CHARGED SURFACES AND FIELD ION EMISSION

Submission for the degree

of

DOCTOR OF SCIENCE

in

The University of Aston in Birmingham

by

RICHARD GORDON FORBES

MA, PhD, FInstP, CEng, MIEE

November 1982

CLASSIFICATION OF PUBLICATIONS

I. FIELD-ION IMAGING

1 Field-ion image formation
R.G. Forbes
Nature Physical Science, 230 (1971) 165-6

2 Reply to comments on 'Field-ion image formation'
R.G. Forbes
Nature Physical Science, 239 (1972) 15-6

3 On why the field-ion microscope works
R.G. Forbes
Proc. 25th. Anniv. Meeting EMAG, Inst. Phys., 1971
(London, Institute of Physics, 1971) pp 248-51

4 A theory of field-ion imaging:
I. A quasi-classical site-current formula
R.G. Forbes
J. Microscopy, 96 (1972) 57-61

5 A theory of field-ion imaging:
II. On the origin of site-current variations
R.G. Forbes
J. Microscopy, 96 (1972) 63-75

6 Field-ion image contrast:
the gas-distribution hypothesis re-examined
J. Duffell and R.G. Forbes
J. Phys. D: Appl. Phys., 11 (1978) L123-5

7 Field-ion imaging theory: On the visibility of adsorbates
R.G. Forbes
Surface Sci., 27 (1971) 659-62

8 Problems in the theory of field-ion imaging
R.G. Forbes
Vacuum, 22 (1972) 517-20

(Reprinted from Nature Physical Science, Vol. 230, No. 16, pp. 165–166, April 19, 1971)

Field-ion Image Formation

RICHARD G. FORBES
Department of Metallurgy, University of Cambridge

It is suggested in this article that the basic physical reason why the field-ion microscope works may be different from that hitherto assumed.

FIELD-ION microscopy is the only established technique capable of resolving single atoms; as such, it is finding increasing application to the problems of metallurgy and surface science[1,2]. The establishment of a secure understanding of how and why the technique works concerns all who wish to use the data that the technique provides. The aim of this article is to direct the theory of image formation, at present somewhat unsatisfactory, towards a new and firmer basis.

The strength of an image spot on the microscope screen depends on the current in the corresponding incident ion beam, which is generated in a small region of space above the corresponding surface site on the specimen[3,4]. The number of ions (J) generated per unit time above any particular surface site is given by[5–7]

$$J = G^{*}\, V\, P^{'}_{\epsilon} \qquad (1)$$

where $P^{'}_{\epsilon}$ is a rate constant called here the characteristic electron transition rate, G^{*} is a characteristic gas concentration, and V is an effective volume, all these parameters being characteristic of the region of space and surface site in question. The product $G^{*}V$ may be interpreted as the probability of finding (the nucleus of) an unionized gas atom in this region of space. J is the "site current".

Consider two surface sites, A and B. The ratio of the corresponding site currents is given by

$$J_A/J_B = (G^{*}_A/G^{*}_B)\,.\,(V_A/V_B)\,.\,(P^{'}_{\epsilon A}/P^{'}_{\epsilon B}) \qquad (2)$$

Up to now, quasiclassical theories of field-ion imaging[6–11] have, in effect, assumed that the dominant factor in this expression would be the ratio $(P^{'}_{\epsilon A}/P^{'}_{\epsilon B})$. Calculations[8,9] have seemed to show that, if the electric field is 1% higher over site A, then $P^{'}_{\epsilon A}$ would be about 30% higher than $P^{'}_{\epsilon B}$.

But variations, as between sites, might also occur in the characteristic gas concentration. If it is assumed that most of the imaging gas atoms trapped near the specimen surface are completely accommodated to the specimen (at temperature T), and that this accommodated population is distributed in accordance with the Maxwell–Boltzmann law, then it can be shown that[5]

$$G^{*}_A/G^{*}_B = \exp[(\tfrac{1}{2}\alpha F_0^2)\,.\,(1/kT)\,.\,(2\delta F/F_0)] \qquad (3)$$

where F_0 is the average of characteristic field strengths above sites A and B (with that above A being the greater), and δF is the difference, k is Boltzmann's constant and α is the gas atom polarizability.

This ratio of gas concentrations is clearly a function of temperature. In the case of helium in a field of 45 V nm^{-1}, it is found that a 1% difference in characteristic field, as between sites, leads to the following ratio values

$$\begin{aligned} &\text{near 80 K,}\quad G^{*}_A/G^{*}_B = 1.2 \\ &\text{near 20 K,}\quad G^{*}_A/G^{*}_B = 2.4 \\ &\text{near 5 K,}\quad G^{*}_A/G^{*}_B = 40 \end{aligned} \qquad (4)$$

These figures should be compared with the value 1.3 quoted earlier for the ratio of characteristic electron transition rates. If this value (1.3) is correct, and if a complete Maxwell–Boltzmann equilibrium at the specimen temperature exists, then near 5 K the gas concentration differences must be the dominant influence on image appearance. They would still be the more important influence near 20 K, and still a significant influence near 80 K.

Thus, at the imaging temperatures now often used in field-ion microscopy, bright image spots may be bright, not because ionization of individual gas atoms takes place more quickly over protruding substrate atoms, but essentially because above these sites there is a greater probability of finding an imaging gas atom in the right place to be ionized. The argument could also apply as between the regions of space above imaged and unimaged sites. It constitutes an essentially new explanation of why the field-ion microscope images the regions of high field above a specimen surface.

The figures also imply that changing the imaging temperature should change the relative strengths of image spots. In fact, a striking characteristic on comparison of near-80 K and near-5 K images of exactly the same tungsten endform is that, within any small region of the image, lowering the temperature makes the bright spots become relatively brighter and the dim spots become relatively dimmer or vanish completely[12]. This experimental evidence cannot be explained on the basis of the traditional explanation of image appearance, based on electron transition rate, because there is no significant temperature dependence in the ratio $P^{'}_{\epsilon A}/P^{'}_{\epsilon B}$ (ref. 5).

The general idea that the behaviour of the imaging gas might affect the appearance of an image at best image voltage is not in itself new. It has been implicitly assumed by many, and explicitly mentioned (in various contexts) by some (refs. 6–9, 13–15, and unpublished work of F. W. Rollgen and H. B. Beckey). Nor am I the first to know about temperature dependent changes in relative spot strengths[15]. The originality lies in the much greater significance attached here to these matters, and in the manner of theoretical approach, particularly the assumption of small scale distribution effects.

There are various complications. For example, it is not certain that most of the trapped population of gas atoms are almost completely accommodated to the specimen temperature, though the temperature dependence of the image spot size[15] tends to suggest they are. More important, if it is assumed that near 5 K the appearance of the image predominantly depends on the distribution of imaging gas, then voltage dependent changes[12] in image appearance must be interpreted as showing that, at voltages in the normal imaging range, a Maxwell–Boltzmann equilibrium does not exist across the whole of the observed part of the specimen surface, but only across small areas of it[5].

It is reasonable that the imaging gas should distribute itself in this way, but it is not yet clear what the precise mechanism is. The most likely supposition is that, at applied voltages in

the normal imaging range: (1) The general brightness of an area of the image is determined by the supply of gas to that area (that is, by the number of atoms per unit time that finish or would tend to finish accommodation within that area), this supply being determined by the processes that occur during accommodation. (2) The distribution of brightness within an area of the image is largely determined by the processes that occur after (or near the end of) accommodation, these processes tending to set up a local Maxwell–Boltzmann equilibrium. Within an area of the image the considerations about image formation outlined earlier would be expected to apply.

This explanation is to be regarded as a provisional working rule. If it were generally applicable, there would be far reaching consequences for field-ion imaging theory. For example, the rule would imply: (1) That the field-ion microscope does not, in any direct sense, image broken bonds or "protruding atomic orbitals"[11], and hence that regional variations of brightness do not result directly from regional variations in the orientation of "protruding orbitals", as has been suggested by Knor and Muller[11]. (2) That it is most unlikely that observation of field-ion images will ever give useful information about the topology of Fermi surfaces, as was suggested by Fonash and Schrenk[16]. (3) That regional brightness differences might well be a consequence of differences, as between different crystallographic regions of the same specimen, in the average depth of the gas atom potential and/or in the average accommodation coefficient. (Further, differences in these parameters, as between different materials, might be a contributory cause of the observed[1] differences in characteristic brightness pattern.)

The working rule is clearly an unproven hypothesis at present. But, equally clearly, the experimental and theoretical evidence described earlier is sufficient to destroy all confidence in the general validity of the usual (electron transition rate) explanation of image appearance. The theoretical problems of a full investigation are massive and fundamental; they will be discussed in depth elsewhere, but there seems little prospect that detailed quantitative theories will become available in the near future. I hope that oulining these ideas now will, however, stimulate further discussion and experiment.

Received November 25, 1970; revised March 29, 1971.

[1] Muller, E. W., and Tsong, T. T., *Field-ion Microscopy: Principles and Applications* (American Elsevier Publishing Co., Inc., New York, 1969).
[2] Hren, J. (edit.), *Surface Sci.* (Spec. Issue on Field-ion, Field Emission Microscopy and Related Topics), 23, 1–258 (1970).
[3] Inghram, M. G., and Gomer, R., *J. Chem. Phys.*, 22, 1279 (1954).
[4] Tsong, T. T., and Muller, E. W., *J. Chem. Phys.*, 41, 3279 (1964).
[5] Forbes, R. G., thesis, Cambridge Univ. (1970).
[6] Gomer, R., in *Field Emission and Field Ionization* (Harvard University Press, 1961).
[7] Hoischer, A. A., thesis, Leiden Univ. (1967).
[8] Muller, E. W., *Adv. Electr. Electron Phys.*, 13, 18 (1960).
[9] Southon, M. J., thesis, Cambridge Univ. (1963).
[10] Tsong, T. T., *Surface Sci.*, 10, 303 (1968).
[11] Knor, Z., and Muller, E. W., *Surface Sci.*, 10, 21 (1968).
[12] Forbes, R. G., and Southon, M. J., in *Thirteenth Field Emission Symp.* (Cornell University, 1966).
[13] Whitmell, D. S., thesis, Cambridge Univ. (1965).
[14] Van Eekelen, H. A. M., *Surface Sci.*, 21, 21 (1970).
[15] Muller, E. W., *J. Appl. Phys.*, 27, 474 (1956).
[16] Fonash, S. J., and Schrenk, G. L., *Phys. Rev.*, 180, 649 (1969).

Printed in Great Britain by Flarepath Printers Limited, Watling Street, Colney Street, St. Albans, Herts.

(*Reprinted from Nature Physical Science, Vol.* 239, *No.* 88, *pp.* 15–16, *September* 4, 1972)

Comments on "Field-Ion Image Formation"

I HAVE several remarks concerning the article on field-ion image formation by Forbes[1]. Forbes says his article directs "the theory of image formation, at present somewhat unsatisfactory, towards a new and firmer basis". There does not, however, appear to be any evidence for this statement. Only the temperature-dependent changes of image brightness were briefly discussed, yet the author said at the end of his article that "the experimental and theoretical evidence described earlier is sufficient to destroy all the confidence in the general validity of the usual explanation of image appearance by electron transition rate".

It seems unnecessary to quote further from the literature to show that the general idea of this article is not new and original and does not represent a firmer basis for field-ion theory; there is enough evidence for this statement in the article itself. In ref. 2, for example, all the problems treated qualitatively by Forbes (the temperature dependence of the total ion current and of the current from individual spots, the concentration of the particles at the tip surface and the extent to which accommodation precedes ionization) are really thoroughly discussed.

The author states that "bright image spots may be bright not because ionization of individual gas atoms takes place more quickly over protruding substrate atoms, but essentially because above these sites there is a greater probability of finding an imaging gas atom in the right place to be ionized". Unfortunately Forbes does not present any explanation (based on his supply–function approach) for this higher probability. It would be interesting to know his reasoning, because it has

already been proposed[3,4] that the mechanism of field-ionization proceeds via the formation of a "transition complex" between a gas atom (ion) and a surface atom; in other words the probability of finding an inert gas atom above the surface atom just before the ionization was expected to be higher.

These errors in Forbes's statements aside, I think it necessary for him to justify his approach by explaining other effects found in field ion microscopes: for example, regional brightness of f.c.c. crystals, alternating visibility of rows of identical atoms on h.c.p. crystals, the hydrogen promotion effect, gas promoted field desorption, the contrast reduction of the regional brightness in the presence of a chemisorbed layer and the formation of ionic complexes between an inert gas atom (ion) and a metal atom (as observed recently by the atom-probe-hole technique). None of these effects is discussed by Forbes, in spite of the fact that these approaches which he rejected without any arguments are able to rationalize all[3,4] or at least some[5] of them.

Finally, it was a complete misunderstanding to classify ref. 3 as a paper dealing with a "classical" explanation of field-ion images. The opposite is, in fact, the case, because the classical explanation was criticized there.

In conclusion, nobody doubts the important role of gas supply and energy accommodation in the field-ionization process. In any surface interaction (like reflexion of atoms on the surface and chemisorption) the first step is always the exchange of energy. If it is not possible, however, to explain all the observed effects in terms of energy accommodation and the gas supply; further steps have to be studied qualitatively (for example, the localized orbitals approximation[3,4]) and quantitatively (for example, the band approach[5]). The introduction of a new theory in physics or chemistry usually leads to the better understanding of several experimentally observed effects, thus proving its usefulness; it is therefore inappropriate to speak only of "far reaching consequences".

Z. KNOR

Institute of Physical Chemistry,
Czechoslovak Academy of Sciences,
Prague

Received October 21, 1971; revised May 24, 1972.

[1] Forbes, R. G., *Nature Physical Science*, **230**, 165 (1971).
[2] Van Eekelen, H. A. M., *Surface Sci.*, **21**, 21 (1970).
[3] Knor, Z., and Müller, E. W., *Surface Sci.*, **10**, 21 (1968).
[4] Knor, Z., *J. Vac. Sci. Technol.*, **8**, 57 (1971).
[5] Sharma, S. P., Fonash, S. J., and Schrenk, G. L., *Surface Sci.*, **23**, 30 (1970).

Reply

A MARKED temperature dependence in relative spot strength suffices to destroy confidence that electron-transition-rate variations are always the dominant influence on image contrast, because no attempted transition-rate calculation contains any significantly temperature-dependent parameter[1,2].

Originality and status were discussed before[1]. There is, for example, no previous proposal of my "provisional working rule"[1,2] that near best image voltage the regional contrast may be primarily determined by supply-and-capture effects, but local contrast by a tendency for local gas equilibria to be established. Van Eekelen[3] gives no results about temperature dependence in image contrast, and clearly states that he excludes "active" distribution effects. Knor[4] is mistaken here, perhaps also in his uncited references.

The origin of the differences in probability $G''V$ is that when imaging-gas atoms bounce on a rough surface during cooling[5] the resulting randomization causes distribution effects, which will be influenced by the details of the potential structure. Investigation suggests the working rule, which fits the observations[2]. Chen and Seidman's recent work[6], when coupled with Van Eekelen's[3] on the degree of accommodation at ionization, now suggests that the local gas distribution may deviate considerably from thermodynamic equilibrium (a thing less certain earlier[1,2]); a temperature-dependent trend similar to the equilibrium one is still to be expected, though[7]. I continue to think that local gas concentration variations will be significant, certainly in some temperature and field ranges, particularly if standard assumptions about the nature of electron-transition-rate variations across the surface are not fully justified[2]. Additional comment appears elsewhere[7].

Other theories have their difficulties too. Knor and Müller[8] assume that "ionization proceeds preferentially in those regions where the fully occupied orbitals of the image gas atom can overlap with those exposed and only partially-occupied orbitals of the surface metal atom". But I know of no clear demonstration that overlap is greater above geometrically protruding surface atoms, if the gas atom is at the critical distance.

Other points arise from Knor's third paragraph[4]. First, he designates the initial part of ref. 1 as a "supply-function approach". In fact this part examined gas concentration variations. A careful distinction should be drawn: a simple gas-supply approach would probably not explain local contrast effects[9].

Second, he confuses the questions of ionization mechanism

("how does ionization occur") and imaging mechanism ("how much ionization occurs where, and why"). The idea of transition-complex formation preceding ionization is first recognizably stated[10] some months after my note[1], but even if complex formation occurred the question of imaging mechanism would still arise. Would transition complexes be formed more often above the most protruding specimen surface atoms? Why?

Third, he interprets transition-complex formation as implying that "the probability of finding an inert gas atom above the surface atom just before ionization was expected to be higher". This seems a detail of complex formation, or trivial. My probabilities $G''V$ relate to the average probability, over a long period of time, of finding an unionized gas atom (or, rather, its nucleus) in a particular ionization zone. This is an average over both the large majority of atom transits through the zone without ionization, and the small minority interrupted by ionization.

The comment[4] about non-"classicality" is itself a misunderstanding. A "quasi-classical" theory applies wave-mechanics to electron behaviour but not to nuclear motion. Ref. 8 seems quasi-classical.

I last deal with general points Knor raises. While the field-ion imaging mechanism is not known, there may be several rational relationships between observed effects and surface geometry and chemistry. Thus he and I may agree that hydrogen promotion effects are probably due to charge transfer; but he may explain them in terms of the effect of changes in orbital orientation and occupancy on (presumably) electron transition rate[8], whereas I might rationalize in terms of the effect of charge distribution changes on the field distribution and hence on the gas behaviour. Similarly, alternative rationalizations might be given for the alternating visibility of rows on h.c.p. crystals, or for the effect on contrast of an adsorbed layer[11]. Because the alternative rationalizations are related via more basic surface geometrical and chemical considerations, advancing one or other of them will not in general provide decisive arguments about why the microscope works. In my earlier communication[1] I concentrated on a potentially decisive observation.

In principle, I disagree with Knor's[4] final point; the modern philosophy of science recognizes that a new theory built around failure points of an older one may not initially be more powerful than the older one[12]. But actually the gas behaviour approach probably meets his criterion of greater usefulness, for it can rationalize[13] the imaging features that the electron-transition-rate approach can, and also the temperature and voltage dependent changes in helium-ion images of tungsten that the

latter apparently cannot[2,7]. Further, the recent results of Schmidt *et al.*[14] concerning "hopping bright spots", and the "dim ring phenomenon" observed by Schubert (private communication) and by Boyes *et al.*[15], are consistent with the first part of the working rule but not with simple transition-rate arguments.

R. G. FORBES*

University of Cambridge,
Department of Metallurgy and Materials Science,
Pembroke Street, Cambridge

Received July 7, 1972.

*Present address: Department of Physics, University of Aston in Birmingham, Gosta Green, Birmingham B4 7ET

[1] Forbes, R. G., *Nature Physical Science*, **230**, 165 (1971).
[2] Forbes, R. G., thesis, Cambridge University (1970).
[3] Van Eekelen, H. A. M., *Surface Sci.*, **21**, 21 (1970).
[4] Knor, Z., *Nature Physical Science* (this issue).
[5] Müller, E. W., *Adv. Electr. Electron Phys.*, **13**, 83 (1960).
[6] Chen, Y. C., and Seidman, D. N., *Surface Sci.*, **26**, 61 (1971).
[7] Forbes, R. G., *J. Microscopy*, **96**, 63 (1972).
[8] Knor, Z., and Müller, E. W., *Surface Sci.*, **10**, 21 (1968).
[9] Holscher, A. A., thesis, Leiden University (1967).
[10] Knor, Z., *J. Vac. Sci. Technol.*, **8**, 57 (1971).
[11] Forbes, R. G., *Surface Sci.*, **27**, 659 (1971).
[12] Kuhn, T. S., *The Structure of Scientific Revolutions*, second ed. (Univ. of Chicago Press, 1970).
[13] Forbes, R. G., in *Eighteenth Field Emission Symposium* (Eindhoven, 1971).
[14] Schmidt, W., Reisner, Th., and Krautz, E., *Surface Sci.*, **26**, 297 (1971).
[15] Boyes, E. D., Turner, P. J., and Southon, M. J., in *Eighteenth Field Emission Symposium* (Eindhoven, 1971).

Printed in Great Britain by Flarepath Printers Ltd., St. Albans, Herts.

PROC. 25th ANNIVERSARY MEETING OF EMAG, INST. PHYSICS, 1971

ON WHY THE FIELD-ION MICROSCOPE WORKS

Richard G. Forbes

University of Cambridge, Department of Metallurgy, Cambridge, England.

The brightness of an image spot on a field-ion microscope screen is determined by the current (J) in the corresponding incident ion beam. J is given by:

$$J = G'' \; V \; P''_e \qquad (1)$$

where P''_e is a rate-constant, V is a volume, and G" is a gas concentration, all these parameters being characteristic of the region of space above the imaged site in question. Gas concentration is here used in the sense of probability per unit volume.

Until recently, it has in effect been assumed that a field-ion micrograph is a map of those places on the specimen surface where the rate-constant for ionization (often treated as an electron tunnelling rate) is relatively high. However, it has recently been suggested (Forbes 1970) that the imaged sites may in fact be those above which the gas concentration is particularly high; this is certainly thought to be the situation for imaging temperatures near 5°K. The original basis for this suggestion was: (a) experimental, in that temperature-dependent changes in image appearance occur but the rate-constant P''_e is not significantly temperature dependent; (b) theoretical, in that preliminary calculations suggested (for example) that a one-percent difference in field, as between two imaged sites, would lead to a 30% difference in the rate-constant but to a 40-fold difference (near 5°K) in the gas concentration (- if the gas is taken to be in thermodynamic equilibrium with the specimen).

Reinforcement for the suggestion came when closer theoretical investigations disclosed faults in the rate-constant calculations: it was impossible to prove, even, that the rate-constant is higher above the imaged sites. Decisive calculations are not yet possible, but, clearly, if it were confirmed that P''_e is lower above the imaged sites, then the existence of field-ion images would in itself show that gas-concentration differences are the cause of images.

Information about the nature of the imaging-gas distribution comes from analysing the dependence of image appearance on the voltage applied to the specimen. If the imaging gas is taken to be in thermodynamic equilibrium with the specimen, then it may be shown (Forbes 1970) that the ratio (G''_A/G''_B) of the gas concentrations above two surface sites A and B is given by:

$$G''_A/G''_B = \exp\left[(\tfrac{1}{2}\alpha F_o^2) \cdot (1/kT) \cdot (2\delta F/F_o)\right] \qquad (2)$$

where F_o is the average of characteristic field strengths above the sites (with that above A being the greater), and δF is the difference; k is Boltzmann's constant; and α is the gas-atom polarizability. The $(\frac{1}{2}\alpha F_o^2)$ factor predicts that, for a given pair of image spots, the ratio of their strengths should increase as the average field strength increases, i.e. that local contrast in the image should get sharper as the voltage applied is increased.

On the other hand, it can be shown (Forbes, to be published) that, if variations in the rate-constant were responsible for image appearance, then local contrast in the image would be expected to become <u>less</u> sharp. An experimental test of imaging mechanism thus appears possible.

Fig. (1) shows a set of micrographs taken, at a specimen temperature near 5°K, with the applied voltage gradually increasing from the point of image detection (near 8kV) to a point just below the near-80°K best image voltage (about 15kV). Through the lower part of the range (fig. 1a to 1d) the image contrast gets markedly sharper, both locally and on an inter-regional scale. This seems to show decisively that, with helium, near 5°K, at low fields, the gas-concentration mechanism of image formation operates.

PROC. 25th ANNIVERSARY MEETING OF EMAG, INST. PHYSICS, 1971

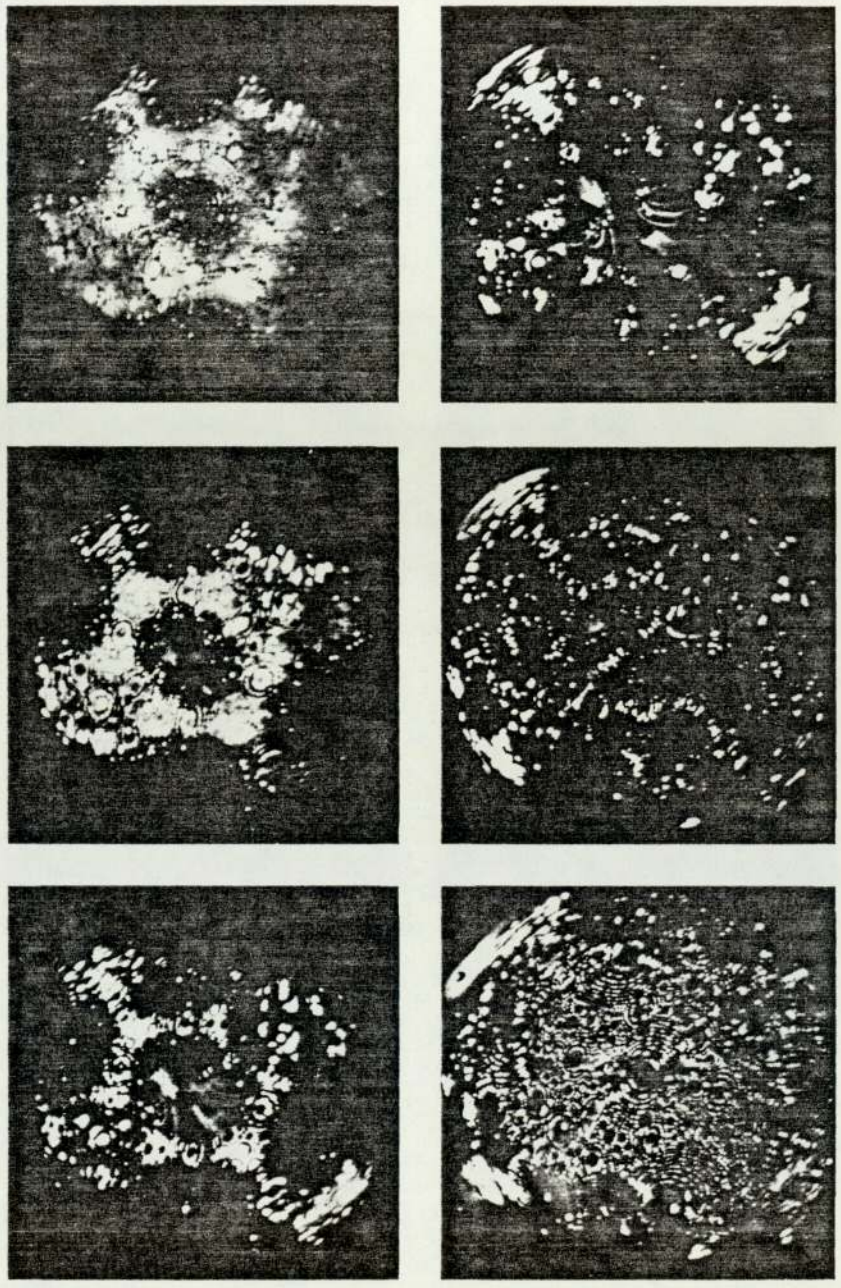


Figure 1: Helium-ion micrographs of Tungsten: a sequence taken at a specimen temperature near $5^{\circ}K$, with the applied voltage increased in steps.
(The black spots in fig. 1a are the result of phosphor damage on the microscope screen, and the curved misty streaks on some micrographs are due to secondary effects: for present purposes both may be ignored.)

However, with further increase in voltage (from fig. 1d onwards) the trend of the change in image appearance is reversed. Closer analysis of this and other evidence (Forbes 1970) strongly suggests that the reversal results, not from a change from a gas-concentration to a rate-constant regime, but from breakdown of the assumption that the gas is in thermodynamic equilibrium with the specimen. More careful enquiry into the statistics of the histories of individual gas atoms then suggests the following.

As a result of the polarization forces that have drawn them towards the specimen's tip, imaging-gas atoms have an energy equivalent to at least several hundred degrees Kelvin when they first hit the tip. The typical gas-atom history can then be divided into two further stages: an accommodation stage, during which the gas cools down towards the tip temperature; and a diffusion stage, during which the accommodated atoms move about on the tip surface. Distribution processes occur during both stages. However, the accommodation stage is completed sufficiently quickly for there to be no chance during it for a Maxwell-Boltzmann concentration equilibrium, at the specimen temperature, to be set up across the whole tip surface. Such a distribution can be set up only if the lifetime of an atom in the diffusion stage is long enough for it to travel long distances across the tip surface.

Statistically, as the duration of the history increases, the concentrations gradually change from near-uniformity to the values characteristic of the endpoint thermodynamic equilibrium. At low fields, it seems, the mean gas-atom lifetime at the tip is long enough for the endpoint nearly to be reached: hence the observed voltage-dependent changes reflect changes in the endpoint gas concentrations.

However, at higher fields ionization occurs before the endpoint is reached: so, although the endpoint gas-concentration differences continue to become larger, the observed image changes now reflect the fact that, with increase in the applied field, ionization occurs increasingly early in the gas-atom history. The effective gas concentrations are now determined by the details of the gas-distribution processes that occur during the accommodation stage, and by the actual lifetime length.

This analysis is a simplified summary of a more careful investigation (Forbes 1970), which is itself only a beginning towards solving a very complex problem. However, it seems justified to formulate a provisional working rule about the brightness distribution in a normal-imaging-range image:

"Image appearance is largely determined by the statistics of the behaviour of the imaging-gas atoms. The general brightness of an area of the image is determined by the supply of gas to that area (i.e. by the number of atoms per unit time that finish or would tend to finish accommodation within that area), this supply being determined by the processes that occur during accommodation. The distribution of intensity within an area of the image is largely determined by the gas-distribution processes that occur after (or near the end of) accommodation, these processes tending to set up a local Maxwell-Boltzmann equilibrium".

Now, although the argument has been based on experiments with helium-ion images of tungsten, there is nothing special about the argument itself that limits its applicability. Far-reaching consequences could follow for imaging theory if (as seems likely) the working rule is essentially correct and generally applicable. For example, the characteristic brightness patterns observed for different materials (see Müller and Tsong 1969) would certainly not be a direct consequence of regional variations in the orientation of "protruding atomic orbitals", as suggested by Knor and Müller (1968). Regional brightness differences, both as between different facets of the same specimen and as between images of different materials, might well be consequences of regional differences in the (average) depth of the gas-atom binding potential, and/or in accommodation coefficients. It would also be implied that observations of field-ion images are most unlikely to ever give useful information about the topology of fermi surfaces.

If, as seems probable, gas-concentration variations are shown to be the dominant influence on image formation at most temperatures now normally used for microscopy, then the basic relationship between surface structure and image appearance may be stated as follows:

"Image spots correspond to those parts of the specimen surface above which the greatest probability exists of finding imaging-gas atoms in the right place and state to be ionized. These are the relatively high field parts of the surface. With an element, each protruding locality on the surface is imaged; in each locality the most protruding surface nucleus will be imaged, and other protruding nuclei may be imaged."

The gas-concentration mechanism does provide a reason why the field-ion microscope should image surface nuclei in virtue of the high fields above them (a thing not unequivocably proved up till now, and, indeed, recently challenged - see Sharma et al. 1970). However, there is not necessarily a monotonic relationship between field strength and image spot strength. So, although one might expect Moore's (1962) model to give a good general impression of image appearance, there is a good theoretical reason why there should not be exact correlation between model prediction and image appearance.

Other aspects and implications are to be found elsewhere (Forbes 1970, and to be published). Many theoretical problems remain. However, it does seem justifiable to claim, firstly, that the evidence available is sufficient to destroy all confidence in the general validity of the traditional electron-transition-rate explanation of image formation; secondly, that considerable progress has been made towards establishing imaging theory on a new and firmer basis. This in turn should make the interpretation of images more secure, and make for a greater confidence in the data the technique can validly supply.

Forbes R G 1970 Ph.D. Thesis, Cambridge University

Knor Z and Müller E W 1968 Surface Sci. 10 21.

Moore A J W 1962 J. Phys. Chem. Solids 23 907.

Müller E W and Tsong T T 1969 Field Ion Microscopy (New York: Elsevier).

Sharma S P, Fonash S J and Schrenk G L 1970 Surface Sci. 23 30.

© *Journal of Microscopy, Vol.* 96, *Pt* 1, *August* 1972, *pp.* 57–61.
Received 26 *September* 1971; *revision received* 23 *November* 1971

A theory of field-ion imaging: I. A quasi-classical site-current formula

by RICHARD G. FORBES, *Department of Metallurgy and Materials Science, University of Cambridge*

SUMMARY

A quasi-classical formulation of field-ion imaging theory is defined and discussed. Concepts and terminology are defined, and a formula is established for the ion current generated above a single surface site.

INTRODUCTION

The field-ion microscope is an instrument that is able to image individual atoms. Within the last few years technological developments have allowed this type of microscopy to be applied to a variety of topics in metallurgy and in surface science (Müller & Tsong, 1969; Hren, 1970). Currently, attempts are being made to extend the technique into the biological area (e.g. Machlin, 1971).

The microscope's mode of operation is as yet incompletely understood. Physical arguments about imaging mechanism follow later; this paper deals with prior logical tasks. It defines concepts and terminology, thereby circumventing any problems due to imprecision in the existing literature. And it formally derives a basic quasi-classical formula for the ion current generated above a specimen surface atom. The contents may be seen as a redefinition and extension of basic ideas put forward long ago by Gomer (e.g. 1961), Müller (e.g. 1960) and others, in order that they may clearly be applicable to surface models that take into account details of surface structure.

SOME BASIC DEFINITIONS

General

The spots on a field-ion microscope screen are each formed by a beam of ionized imaging gas atoms, each beam having originated in a small region of space (the *ionization zone*) above a particular surface site (Tsong & Müller, 1964). The *site current* is the number of ions generated in the ionization zone per unit time. Ionization zones can be defined above every specimen surface atom, but we see only those which produce relatively high site currents.

Electron behaviour

Suppose that an imaging-gas atom could be held stationary. Field ionization of such an atom would be a random process, governed by the usual exponential decay law. In imaging-mode field ionization, the electron makes a transition into a

state of equal energy in the specimen. The symbol P_e is defined to denote the rate constant appropriate to circumstances where all the energetically accessible final states are unoccupied before the transition, and is here called the *electron transition rate*. It may be calculated in three ways: from a barrier penetration coefficient (Müller & Bahadur, 1956; Gomer, 1961; Southon, 1963, 1968; Alferieff & Duke, 1967); from the time-dependent Schrödinger equation (Jason, 1967); or by application of Fermi's 'Golden Rule' or some variant thereof (Boudreaux & Cutler, 1966a; Holscher, 1967; Forbes, unpublished work; Tsong, 1968). The three methods are not entirely equivalent.

In circumstances where the final states are fully or partially occupied (with occupation probability f) the rate constant is not P_e but Q_e, where:

$$Q_e = (1 - f)P_e. \tag{1}$$

Q_e is termed the *electron transfer rate*. For a metal specimen, f may be taken as the Fermi function. f is equal to $\frac{1}{2}$ for all positions of the gas atom nucleus such that the 'energy' (or electrochemical potential) of the transiting electron is equal to the Fermi level of the metal electrons. The surface containing all such nuclear positions is termed the *critical surface*. Geometrically, the critical surface would seem to be a smoothly corrugated ('egg-box shaped') surface roughly parallel to the specimen surface.

Near the Fermi level the Fermi function is a sensitive function of energy. The energy level of the gas atom electron is a sensitive function of the position of the gas atom nucleus. Consequently, Q_e may in practice be taken equal to zero for nuclear positions inside the critical surface, and equal to P_e for positions outside it. A quasi-classical theory thus assumes that no ionization occurs inside the critical surface: this is the three-dimensional analogue of Müller's (e.g. 1960) statement that no ionization occurs inside the critical distance. The idea was originally introduced, in order to explain field-ion energy-analysis measurements, by Inghram & Gomer (1954, 1955).

Gas behaviour

In a macroscopic region of space the rate of ion generation depends both on the rate constant for ionization and on the number (N) of unionized gas atoms present, the latter being determinable from the average gas concentration ($\bar{G}$) by the trivial formula: $N = \bar{G}v$. $\bar{G}$ is a space average taken over the volume (v) of the region. With a very small region of space, an ion can be said to be generated in the small volume dv if the atomic nucleus is within dv at the instant of ionization, and the corresponding formula is:

$$dN = G dv. \tag{2}$$

dN is the probability (the average probability taken over a long period of time) of finding the nucleus of an unionized atom in dv. G is the *probability distribution function* for the imaging gas, but may also be called the *effective gas concentration*. G, as defined mathematically in (2), can vary sharply over small distances.

Characteristic points and values

The parameters P_e and G vary with position in space, and so does the electric field strength, F. In particular, they all vary with position in the critical surface. One may hypothesize that in the critical surface above a surface site there exists some point where the field strength F is greater than at any other nearby point in the critical surface. This point where F has a constrained maximum is termed the *characteristic point* for the surface site in question. For any particular surface

site, the *characteristic value* of any space-varying parameter is its value taken at the characteristic point. Characteristic values will be denoted by adding a double-prime (″) to the parameter in question: thus, P_e'' denotes the characteristic electron transition rate.

Figure 1 illustrates these definitions. In Fig. 1a, S is a protruding surface nucleus, and C is a portion of the critical surface. L is a line that lies in the critical surface and passes through the characteristic point R'' corresponding to surface nucleus S. Figure 1b shows schematically the variation in field strength with position in line L: the field strength has a maximum (of value F'') at the characteristic point R''.

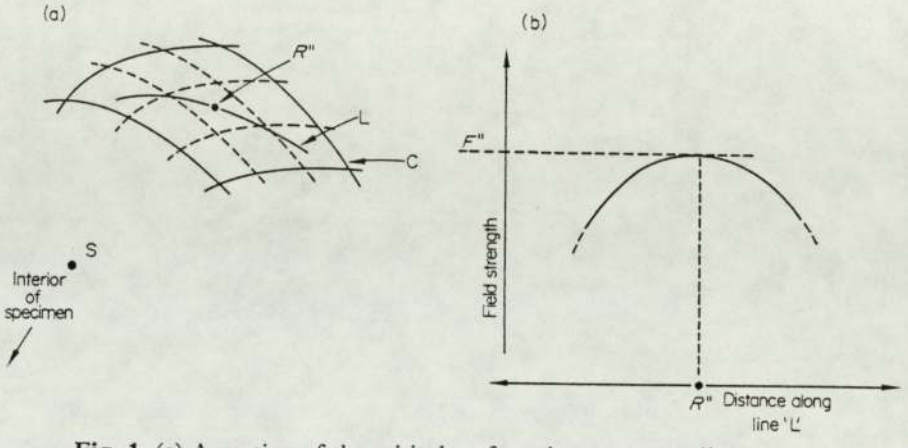


Fig. 1. (a) A portion of the critical surface above a protruding surface nucleus (S) of the specimen. (b) To illustrate how the characteristic point R'' is defined. The field strength has a constrained maximum (of value F'') at the characteristic point.

The defining of characteristic points is necessary because, in establishing equation (4) below, one must think three-dimensionally about the origin of ion current. A gas atom with its nucleus at a characteristic point would in a one-dimensional model (Müller, 1960) be said to be 'at the critical distance'.

In certain circumstances, or for some sites, the above definition of characteristic point may fail (Forbes, 1970). But it is still possible to characterize ion-current generation above any particular site by certain parameter values.

DERIVATION OF FORMULAE

A general expression is now derived for the current, $\mathcal{J}$, from a single surface site.

The number ($d\mathcal{J}$) of ions generated per unit time in a small region of space, of volume dv, is given by:

$$d\mathcal{J} = G dv \,.\, Q_e \qquad (3)$$

—where G is the effective gas concentration, and Q_e the electron transfer rate, for this region. (The r.h.s. of (3) is the product of the rate constant for ionization and the probability of finding a gas atom with its nucleus in the region.)

The site current $\mathcal{J}$ is obtained by integrating (3) over the whole region of space above the relevant surface site. However, Q_e may be taken equal to zero inside the critical surface and equal to P_e outside, so the integration reduces to one taken

over an appropriate region of space, ϵ, outside the critical surface. Further, under normal conditions of imaging, most of this ionization occurs in the near vicinity of the relevant characteristic point; so it is convenient to write the result of integration in the form:

$$\mathcal{J} = \int_{\epsilon} GP_e dv = G''VP_e'' \qquad (4)$$

where G'' is the relevant characteristic gas concentration, P_e'' is the relevant characteristic electron transition rate, and V is an *effective volume.*

The quantity V comes out of the mathematics, being related to the rate of fall-off of (GP_e) with distance from the characteristic point. However, it is convenient to think of V as the volume of a Muller-type ionization zone, with $(G''V)$ the probability of finding a neutral gas-atom nucleus within the zone, and P_e'' a transition rate that is constant throughout the zone. The equivalent one-dimensional model has been widely employed (e.g. Southon, 1963).

A field-ion micrograph shows only the relative values of spot strengths. The most relevant quantity in a discussion of imaging mechanism is the ratio of site currents. For two sites A and B this ratio $(\mathcal{J}_A/\mathcal{J}_B)$ may be written:

$$(\mathcal{J}_A/\mathcal{J}_B) = (G_A''/G_B'') \cdot (V_A/V_B) \cdot (P_{eA}''/P_{eB}''). \qquad (5)$$

Expression (4) may be termed the *quasi-classical site-current formula,* and (5) the *quasi-classical ratio formula.* The description 'quasi-classical' implies that wave-mechanics has not been applied to the motion of the gas-atom nucleus: corresponding approximations in other branches of physics and chemistry are similarly named.

DISCUSSION

Formulæ similar in form to equation (4) have previously been used by Gomer (1961) and by Holscher (1967), in discussion of total ion current behaviour. The symbols that appear in (4) are, however, rather more closely defined, partly in order to show the sense in discussing local electron transition rate and gas concentration variations.

Two basic physical assumptions made deserve recording. Implicit in the defining of characteristic points is the assumption that the field strength at any point is constant in time. In effect, we are disregarding the zero-point motion of the surface nuclei and statistical fluctuations in the state of the surface, and are taking the field distribution to be that derived from some specific or time-averaged surface charge distribution. The present treatment also assumes that the type of field ionization under discussion obeys locally first-order kinetics, in the sense that the local rate of reaction (i.e. ion generation) depends on the probability of finding a *single* body in an appropriate place.

It also deserves stressing that the quantity P_e'' is defined within the context of a quasi-classical approach. Fully wave-mechanical treatments are also possible, and in these there appears a rate constant $P^\star$ (Forbes, 1970) that has in effect been estimated by Gomer & Swanson (1963), by Boudreaux & Cutler (1966b), and by Schrenck and co-workers (Sharma, Fonash & Shrenck, 1970). However, the quantities P_e'' and $P^\star$ are intrinsically different, being defined within the contexts of different world-views, and there is no substantive physical relationship between them. P_e'' is also different in kind from the rate constant k_i that appears in an alternative formulation (Gomer, 1961, p. 78) which treats field ionization as an activated process.

The present treatment is *quasi-classical.* It is a simple, *first-level,* approach. It is *direct,* in the sense that ionization is not regarded as activated. And it is a *space*

formulation, in the sense that equation (4) involves an integral over space, and not the integrals along gas-atom trajectories that appear in one of Gomer's (1961) treatments. These characteristics distinguish the present 'QI–DS' formulation from the many others possible.

To some extent this work is a formal redefinition of ideas implicit in earlier treatments of imaging. But past literature and discussions have not always been consistent in their use of terminology. For example, P_e'' and its reciprocal have been given a variety of names, and the name 'ionization probability' has been applied to P_e'', $P^\star$, k_i, $\mathcal{J}$, and to various kinds of true probabilities; and on occasion the distinctions between the different kinds of rate constant have not been realized, and the difference between a rate constant and a true probability blurred. Careful formulation and naming have perhaps not been necessary in the past, but continuing theoretical development now seems to advise it, in order to make the analysis of differences between different imaging models easier, and the understanding of theoretical arguments less time-consuming for non-specialists. This paper has offered some suggestions.

References

Alferieff, M.A. & Duke, C.B. (1967) Field ionization near nonuniform metal surfaces. *J. chem. Phys.* **46**, 938.

Boudreaux, D.S. & Cutler, P.H. (1966a) Theory of atom–metal interaction. II. Perturbation theory in field ionization calculations. *Surf. Sci.* **5**, 230.

Boudreaux, D.S. & Cutler, P.H. (1966b) Theory of atom–metal interactions. I. Application of quantum-mechanical collision theory to field-ionization processes. *Phys. Rev.* **149**, 170.

Forbes, R.G. (1970) *Field-ion microscopy at very low temperatures.* Ph.D. Thesis, Cambridge University.

Gomer, R. (1961) *Field Emission and Field Ionization.* Harvard University Press, Cambridge, Mass.

Gomer, R. & Swanson, L.W. (1963) Theory of field desorption. *J. chem. Phys.* **38**, 1613.

Holscher, A.A. (1967) *Adsorption studies with the field-emission and field-ion microscope.* Doctorate Thesis, Leiden University.

Hren, J.J. (1970) (Guest Editor) Special issue on field-ion, field emission and related topics. *Surf. Sci.* **23**, 1.

Inghram, M.G. & Gomer, R. (1954) Mass spectrometric analysis of ions from the field-ion microscope. *J. chem. Phys.* **22**, 1279.

Inghram, M.G. & Gomer, R. (1955) Massenspektrometrische untersuchungen der feld-emission positiver ionen. *Z. Naturf.* **10A**, 863.

Jason, A.J. (1967) Field-induced resonance states at a surface. *Phys. Rev.* **156**, 266.

Machlin, E.S. (1971) Low field strength ion microscopy of organic and bio-molecules. *Eighteenth Field Emission Symposium* (Eindhoven).

Müller, E.W. (1950). Field ionization and field-ion microscopy. *Adv. Electronics Electron Phys.* **13**, 83.

Müller, E.W. & Bahadur, K. (1956) Field ionization of gases at a metal surface and the resolution of the field-ion microscope. *Phys. Rev.* **102**, 624.

Müller, E.W. & Tsong, T.T. (1969) *Field-ion Microscopy: Principles and Applications.* American Elsevier Publishing Co., New York.

Sharma, S.P., Fonash, S.J. & Schrenk, G.L. (1970) The electron transfer process in field ionization. *Surf. Sci.* **23**, 30.

Southon, M.J. (1963) *Image formation in the field-ion microscope.* Ph.D. Thesis, Cambridge University.

Southon, M.J. (1968) Field emission and field ionization. In: *Field-ion Microscopy* (Ed. by J. J. Hren and S. Ranganathan). Plenum Press, New York.

Tsong, T.T. (1968) Surface states and field-ion image formation. *Surf. Sci.* **10**, 303.

Tsong, T.T. & Müller, E.W. (1964) Measurement of the energy distribution in field ionization. *J. chem. Phys.* **41**, 3279.

© *Journal of Microscopy, Vol.* 96, *Pt* 1, *August* 1972, *pp.* 63–75.
Received 29 *September* 1971; *revision received* 23 *November* 1971

A theory of field-ion imaging: II. On the origin of site-current variations*

by RICHARD G. FORBES, *Department of Metallurgy and Materials Science, University of Cambridge*

SUMMARY

Experimental evidence demonstrates the incompleteness of the usual rate constant explanation of field-ion image appearance, and suggests that gas distribution effects may be the dominant influence on image contrast, certainly in some circumstances. Further evidence shows that gas concentration variations across the surface would not be given by any simple formal rule. However, significant features of image behaviour in the helium-on-tungsten system can be rationalized by closer examination of gas behaviour during and after accommodation. For the normal imaging range a provisional working rule is suggested, that regional contrast may be primarily determined by supply-and-capture effects, but local contrast by a tendency for local gas equilibria to be established.

INTRODUCTION

This paper contains physical arguments about the nature of the field-ion imaging mechanism. It advances the idea that both the general appearance and the details of a field-ion image may be significantly influenced, and sometimes dominated, by the way in which the imaging gas distributes itself across the specimen surface. The general idea that gas behaviour might affect image appearance is not in itself new (see, for example: Müller, 1956; Gomer, 1961; Southon, 1963; Whitmell, 1965; Holscher, 1967; Müller & Tsong, 1969; Van Eekelen, 1790; Rollgen & Beckey, 1971; Chen & Seidman, 1971a, 1971b). But the suggestion of possible dominance is a new one, and so are aspects of the approach adopted. The procedure is to identify what seem the most significant features of the helium-ion imaging of tungsten, and to develop, via several stages of argument, a sufficiently general empirical explanation.

One motive behind the work is the author's belief that the helium-on-tungsten system is the 'Bohr atom' of field-ion theory; that its proper understanding is a key to further theoretical progress; and that if we do not understand this system then we cannot rightly claim to know why the field-ion microscope works.

THE TEMPERATURE DEPENDENCE OF IMAGE APPEARANCE

It was shown in the preceding paper that the site-current ratio for two sites A and B is given by:

$$\mathcal{J}_A/\mathcal{J}_B = (G_A''/G_B'') \cdot (V_A/V_B) \cdot (P_{eA}''/P_{eB}'') \qquad (1)$$

where the symbols have the meanings defined previously.

* Based on a paper presented at the Royal Microscopical Society Conference, 'Micro-70' in September 1970.

In the past, quasi-classical theories of field-ion imaging have in effect taken (P_{eA}''/P_{eB}'') as the most important factor in equation (1), assuming that protruding specimen atoms are imaged because individual gas atoms get ionized more quickly above these sites. With helium this has seemed justified by calculations (Müller & Bahadur, 1956; Southon, 1963) and ion-current measurements (Southon & Brandon, 1963) that have been taken to show that a 1% increase in field across the surface leads to a 1% increase in P_e.

However, the traditional standpoint is challenged by the experimental evidence illustrated in Figs. 1 and 2. Lowering the specimen temperature causes changes, not only in spot size, but also in relative spot strengths: within many individual small areas of the image the strong spots have become stronger and the weak spots have become weaker or have vanished altogether. The effect is more obvious on the microscope screen than it is on the micrographs.

There are many attempted calculations of electron transition rate; but temperature neither appears explicitly nor significantly affects the parameters that do appear. Therefore, because the reality is that temperature change produces contrast change, the usual imaging explanation is certainly incomplete.

The term (V_A/V_B) in equation (1) should not normally vary significantly with temperature. This leaves only the gas concentration term (G_A''/G_B'') to explain the illustrated changes.

The accepted explanation of the spot-size decrease (Müller, 1956) is that lowering the specimen temperature reduces the mean lateral speed of the imaging-gas atoms. This implies a strong link between gas temperature at ionization and specimen temperature. Incoming gas atoms hit the specimen with appreciable kinetic energies (Müller, 1956). The many collisions that occur during cooling (Müller, 1960) will tend to randomize the motion, and hence *tend* to produce a kinetic equilibrium (see for example, Jeans, 1940; also Holscher, 1967). Let us therefore take the hypothesis that the gas atom population subject to ionization be in thermodynamic equilibrium with the specimen (at temperature T) and, temporarily deferring questions of validity, examine the consequences.

Spatial variations in gas-atom potential (U) would give rise to variations in local gas concentration (G), according to the Boltzmann rule:

$$G \propto \exp(-U/kT). \tag{2}$$

For points in or near the critical surface, the dominant term in the full expression for U is expected to be the 'polarization term' $(-\frac{1}{2}\alpha F^2)$, where α is the gas-atom polarizability and F is the electric field strength at the position of the gas atom nucleus (Forbes, 1970). Whence:

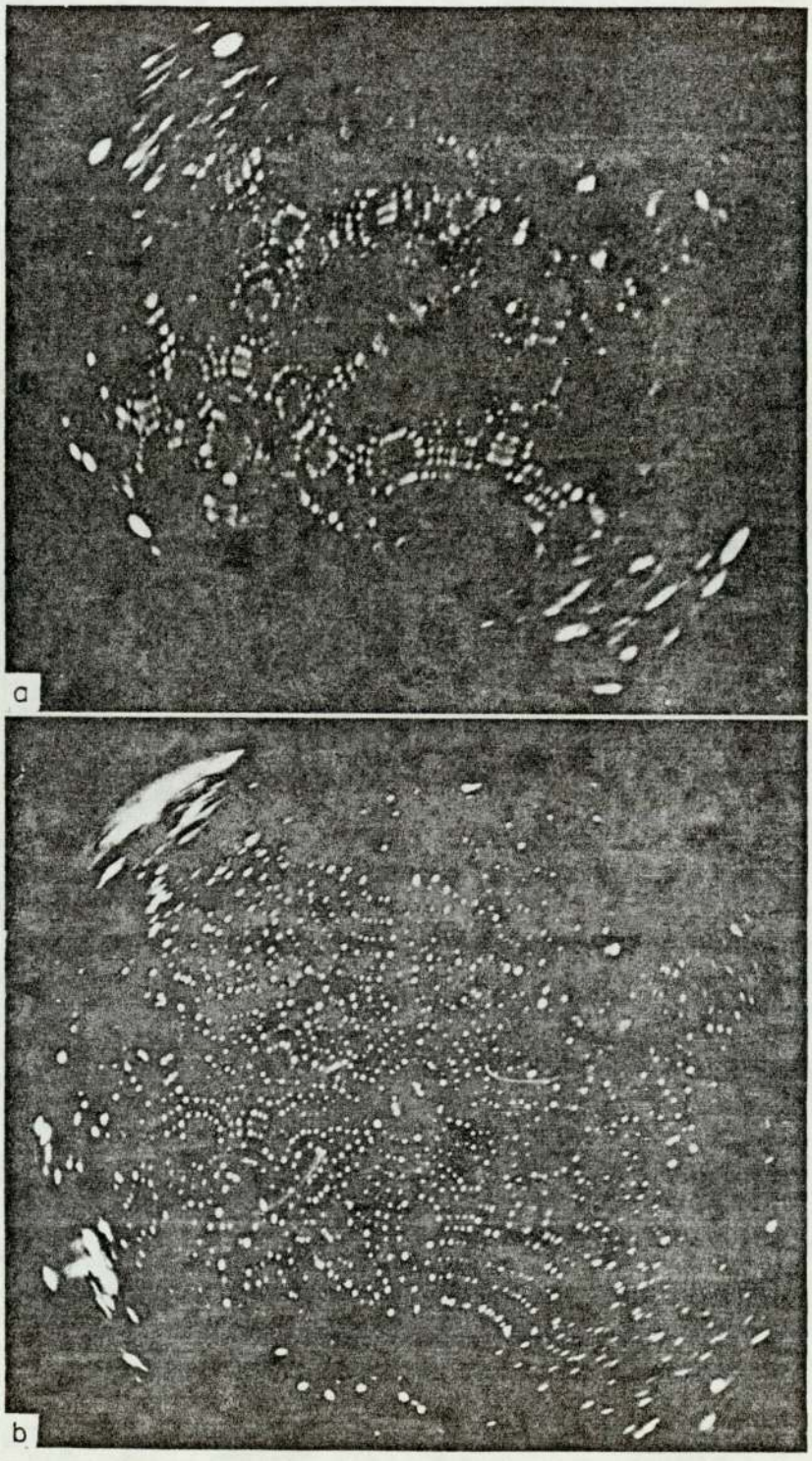
$$G_A''/G_B'' = \exp[(\tfrac{1}{2}\alpha F_0^2) \,.\, (1/kT) \,.\, (2\delta F/F_0)] \tag{3}$$

where F_0 is the mean of characteristic field strengths (F_A'', F_B'') above sites A and B, and δF is the difference: $\delta F = F_A'' - F_B''$.

It perhaps deserves stressing that F_A'' and F_B'' are *characteristic* field strengths as defined in the preceding paper (i.e. values taken at characteristic points); also

---

Fig. 1. Helium-ion images of a tungsten endform, taken near best image voltage and at temperatures (a) near 80°K and (b) near 5°K. (The micrographs in this paper were taken on a large glass field-ion microscope of the conventional double-dewar configuration (Forbes, 1970). The temperature was changed by pouring liquid nitrogen into the inner dewar as the last of some liquid helium boiled away.)


a
b

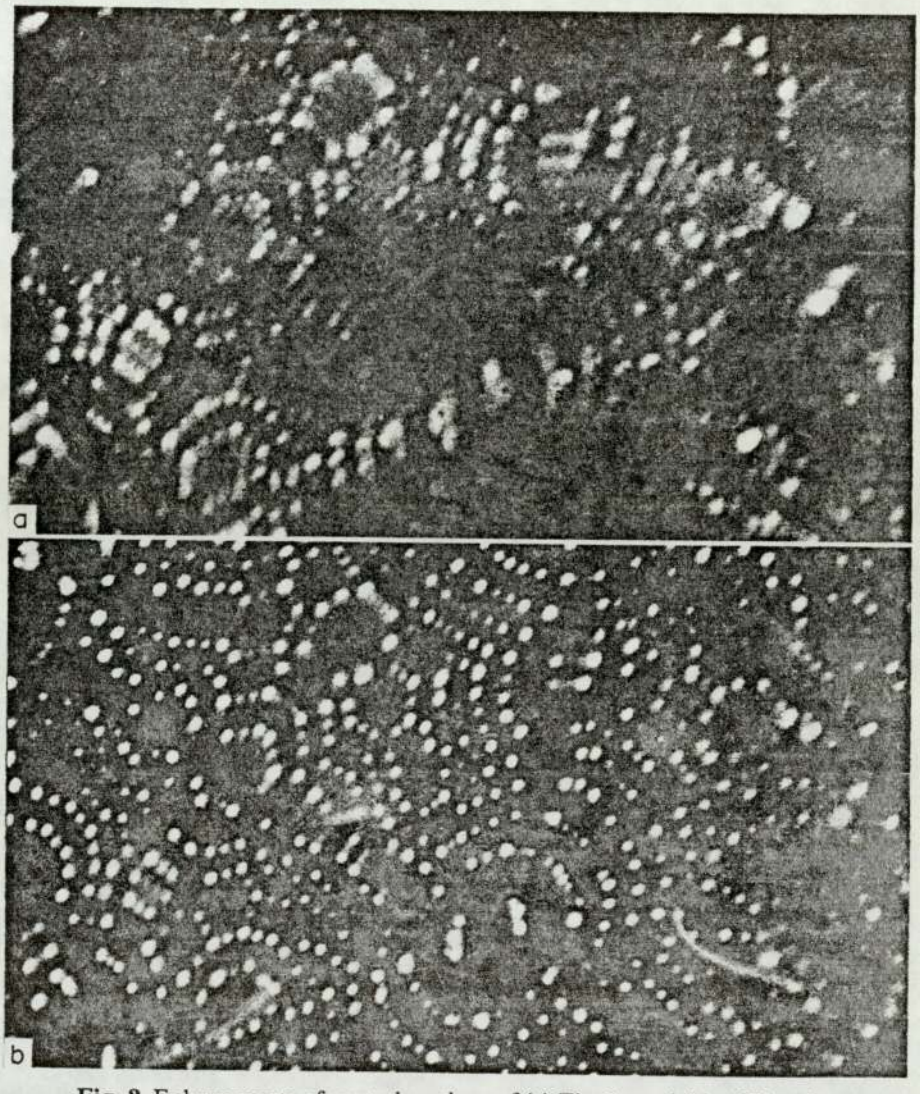


Fig. 2. Enlargements of central sections of (a) Fig. 1a and (b) of Fig. 1b.

that the mode of derivation of equation (3) automatically takes cognizance of effects due to inhomogeneities in the metal's surface charge distribution, including the types recently discussed by Forbes (1970) and by Tsong & Müller (1970). (The F_0 in my work is a different quantity from the F_0 in Tsong and Müller's.)

Substituting into equation (3) numerical values appropriate to helium (α_{He} = 0·0002 nm^3, F_0 = 45 V/nm), and assuming a 1% difference in characteristic field, leads to the following ratio values:

$$\begin{aligned} \text{Near } 80^\circ\text{K} \quad & G_A''/G_B'' = 1{\cdot}2 \\ 20^\circ\text{K} \quad & = 2{\cdot}4 \\ 5^\circ\text{K} \quad & = 40. \end{aligned} \tag{4}$$

If one makes the additional, not unreasonable, assumption that the characteristic field strength is higher for the more protruding specimen surface atoms, it is

evident from (3) and (4) that our hypothesis is *sufficient* to explain local temperature dependent changes in relative spot strength.

The converse does not hold, of course. For a gas concentration explanation to work qualitatively, it is only necessary to assume that the general supply of gas to the tip should partition itself preferentially to the high-field localities, and that lowering the specimen temperature should cause increasingly strong partitioning to these localities. The figures in table (4) represent an upper limit which the actual gas concentration ratios cannot exceed: the ratios may in reality be much less, though the same general sort of temperature-dependent trend should exist.

Nonetheless, when compared with the figure of 1·3 quoted earlier for the corresponding transition rate ratio, table (4) suggests an important possibility: that, at sufficiently low temperatures, the variations in gas concentration may be the dominant influence on image contrast: that sometimes bright image spots may be bright, *not* because individual gas atoms get ionized more quickly above protruding specimen atoms, *but essentially because above such sites there is a greater probability of finding an imaging-gas atom in the right place to be ionized.* When applied as between the regions of space above imaged and unimaged sites, the idea amounts to a possible new explanation of why the field-ion microscope images the high-field regions above a specimen surface.

The fact of local temperature-dependent change tends to suggest that this is the imaging mechanism near 5°K. More generally, one might expect the new explanation to be applicable at low specimen temperatures. At high temperatures, however, the gas concentration variations tend to become negligible and the electron transition-rate explanation should become correct. In between, there might be supposed to exist a 'balance temperature' at which (P_{eA}''/P_{eB}'') would equal (G_A''/G_B''). Its value is of vital interest. But fundamental difficulties arise when one tries to determine it. These problems have been outlined elsewhere (Forbes, 1970); the main points are as follows:

(1) Because of the nature of the potential variation near a charged surface, a 1% *difference* in characteristic field between different surface sites does not necessarily lead to the same variation in characteristic electron transition rate as would a 1% *increase* in the characteristic field at any particular site.

(2) Consequently, Southon & Brandon's (1963) ion-current measurements do *not* directly support any theoretical work concerned with the calculation of (P_{eA}''/P_{eB}'').

(3) In fact it seems impossible at present to prove theoretically that calculated characteristic electron transition rates would be higher over the more protruding sites. (Definitive calculations are not feasible, but preliminary arguments suggest that the result could go in the wrong direction, and in this case the traditional explanation of contrast formation would not work at all. The gist of the matter is that the Hartree potential well surrounding a more protruding nucleus might be broader and deeper than at less protruding sites, and in consequence the distance between surface nucleus and the characteristic point might be greater, the degree of overlap between gas-atom orbitals and unoccupied surface-atom orbitals less, and hence the electron transition rate less.)

(4) Further, it is logically invalid to suppose that isolated observation of a normal field-ion image can in itself prove that the rate constant (or any other) explanation would work in the right direction.

(5) The uncertainty over the behaviour of P_e'' notwithstanding, there are some grounds for believing that the balance temperature would not lie in the range 0°–60°K if gas thermodynamic equilibrium were assumed. (The argument is based on the Southon & Brandon ion-current measurements, but is not direct.)

A separate kind of difficulty arises if the gas distribution deviates significantly from a Maxwell–Boltzmann distribution. The work of Van Eekelen (1970) on the degree of accommodation at ionization, and that of Chen & Seidman (1971a) on the temperature dependence of image spot size, indicates that the effective temperature of the imaging gas (as determined by its velocity distribution) is probably above the temperature of the specimen for field ranges near and above the normal imaging range (though approximations used in their models may counsel caution in accepting their numerical results as quantitatively precise). Thermodynamic equilibrium would certainly not then exist. Revised gas concentration ratios might be obtained by substituting the effective gas temperature into equation (3); but it is far from obvious that such a procedure would be approximately correct.

In this situation few firm theoretical conclusions are possible. For a normal imaging range helium-on-tungsten image it seems that gas distribution effects will almost certainly be the dominant influence on image contrast near 5°K, and may perhaps be significant or dominant at much higher temperatures. Detailed quantitative investigation is urgently required, but it seems likely to be messy and time-consuming.

THE VOLTAGE DEPENDENCE OF IMAGE APPEARANCE NEAR 5°K

Voltage-dependent effects provide further empirical evidence about imaging mechanism. Theoretical predictions, initially based on the thermodynamic equilibrium hypothesis, are compared with observations first presented some years ago (Forbes & Southon, 1966).

Theoretical predictions

The gas concentration mechanism. In equation (3), the $\frac{1}{2}\alpha F_0^2$ factor predicts that, for a given pair of image spots, the ratio of their strengths should increase as the average field strength increases. So, if thermodynamic equilibrium exists and the gas concentration variations are the dominant influence, contrast in the image should get sharper as the voltage applied to the specimen is increased.

The electron transition-rate mechanism. Southon's (1963) calculations, replotted in Fig. 3, suggest that increase in applied field leads to fall-off in the rate of increase of characteristic electron transition rates. Figure 4 shows plots for two separate sites, the horizontal axis representing either the mean field F_0 or the applied field. A set interval on the vertical axis represents a set ratio of characteristic electron transition rates. Clearly, the higher the field, the lower will be this ratio. So, if transition rate variations are the dominant influence, image contrast should get *less* sharp as applied voltage increases.

Observations and discussion

Figure (5) shows a sequence of micrographs taken at a specimen temperature calculated to be less than 4·3°K, with the applied voltage gradually increasing from the point at which an image is just visible (near 8 kV) to a point just below the near-80°K best voltage (at about 15 kV). The micrographs were taken early in the experimental work and their quality leaves something to be desired, but they suffice to illustrate effects observed on the microscope screen and described below.

For discussion, the sequence is split: Figure 5 (a–d) is the 'low field range'; Fig. 5 (e–h) the 'intermediate field range'. The 'high field range', above best image voltage, is not discussed here.

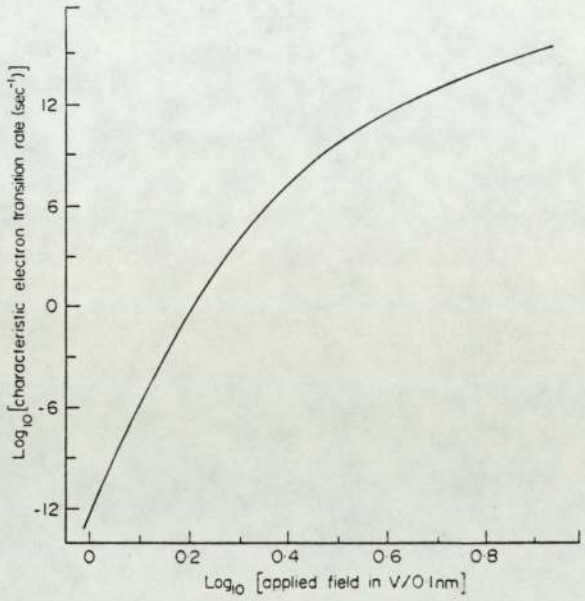


Fig. 3. Logarithmic plot of electron transition rate versus applied field. (Re-plotted from Southon, 1963, Fig. 31.)

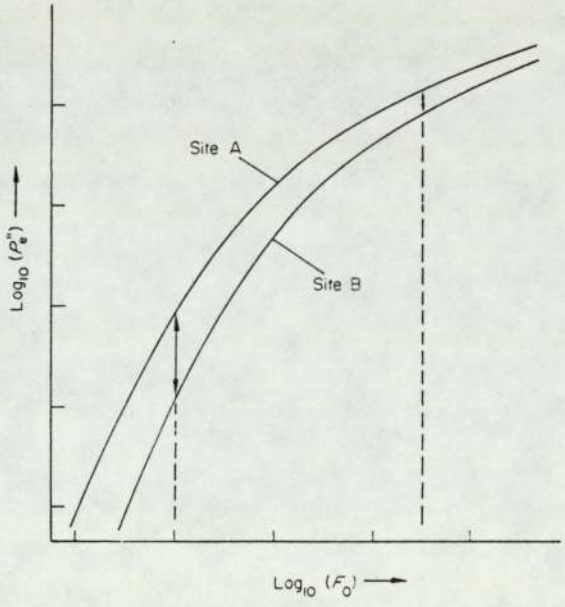


Fig. 4. To show the variation of electron transition-rate ratios with applied field strength (see text).

Low fields. At the lowest voltage (Fig. 5a) the image is generally blurred. With gradual increase in voltage, resolution improves, bright spots get relatively brighter and dim spots relatively dimmer, the amount of image visible shrinks, and the bright regions become very bright indeed (a dramatic effect at 10·6 kV, not well reproduced on the micrograph). (The {111} brightness peak at very low

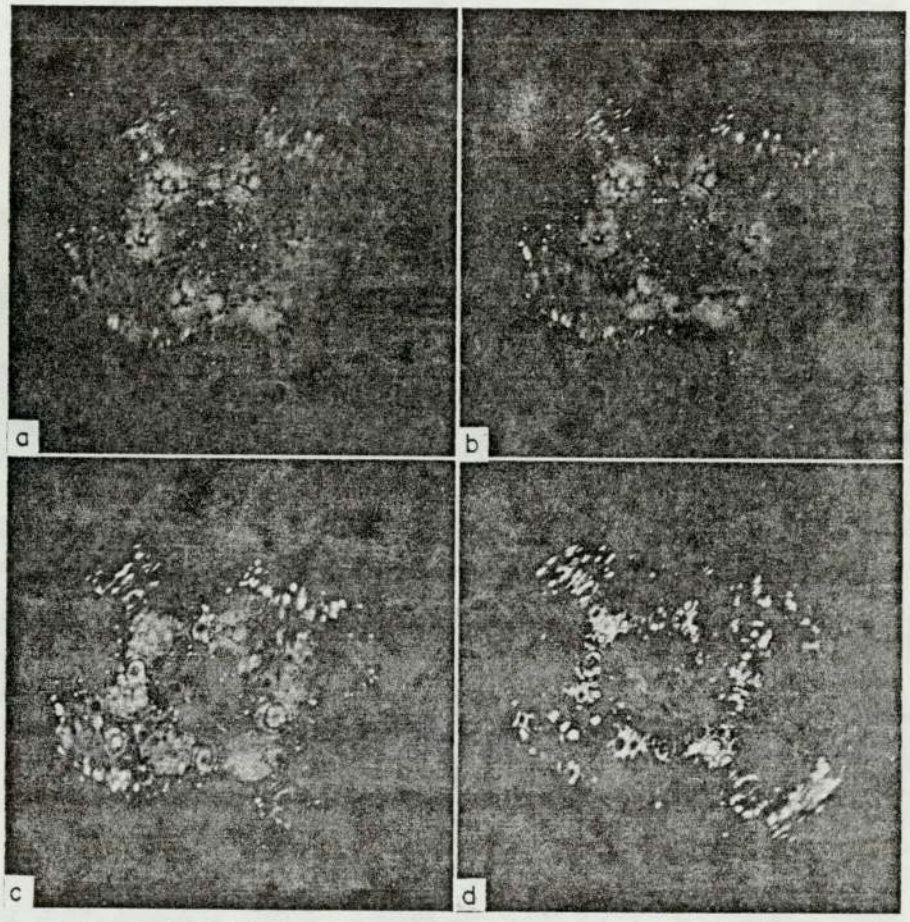


Fig. 5. (a–d). Voltage sequence taken at a specimen temperature near 5°K. (The black spots visible in Figs. 5a to 5c are due to phosphor damage on the microscope screen, and the curved misty streaks on some micrographs are due to secondary effects. These are extraneous effects and can be disregarded here.) (a) 8·0 kV, (b) 8·5 kV, (c) 9·0 kV, (d) 10·0 kV.

temperatures has been noted before, by Müller *et al.*, 1965.) Locally and regionally, the voltage increase leads to sharper contrast. So we can decisively deduce that, with helium, at specimen temperatures near 5°K, at low fields, the gas concentration mechanism operates. The gas concentration variations tend to being those characteristic of a thermodynamic equilibrium (but of course, this does not imply that the endpoint equilibrium situation has in fact been reached).

In an attempt at completeness, a large number of alternative hypotheses have been examined (unpublished work), involving such things as imaging artifacts,

deficiencies in transition rate theory or surface potential theory, scattering, charge exchange, the effects of an adsorbed layer, the co-existence of several ionization mechanisms, two-electron processes, nuclear exchange at the instant of ionization, and gas-phase ionization. In no case could a simple well-established alternative explanation be given for the contrast sharpening just described.

Intermediate fields. Further voltage increase causes a trend reversal: through Fig. 5 (e–h) the amount of image visible increases, and the normal imaging range type of contrast is approached. The reversal shows that even if something approaching thermodynamic equilibrium exists at low fields it does not at intermediate and higher fields—a conclusion that bears some relation (but not a closely-defined relation) to Gomer's (1961) distinguishing of different total ion current regimes.

The 'hot general equilibrium' mechanisms. At first sight, there might seem a simple empirical explanation of the intermediate-field trend, based on the premise that a Maxwell–Boltzmann concentration equilibrium exist across the whole tip surface but that an effective gas temperature (T_e) be above the specimen temperature and vary with the applied field. In this field range T_e might increase

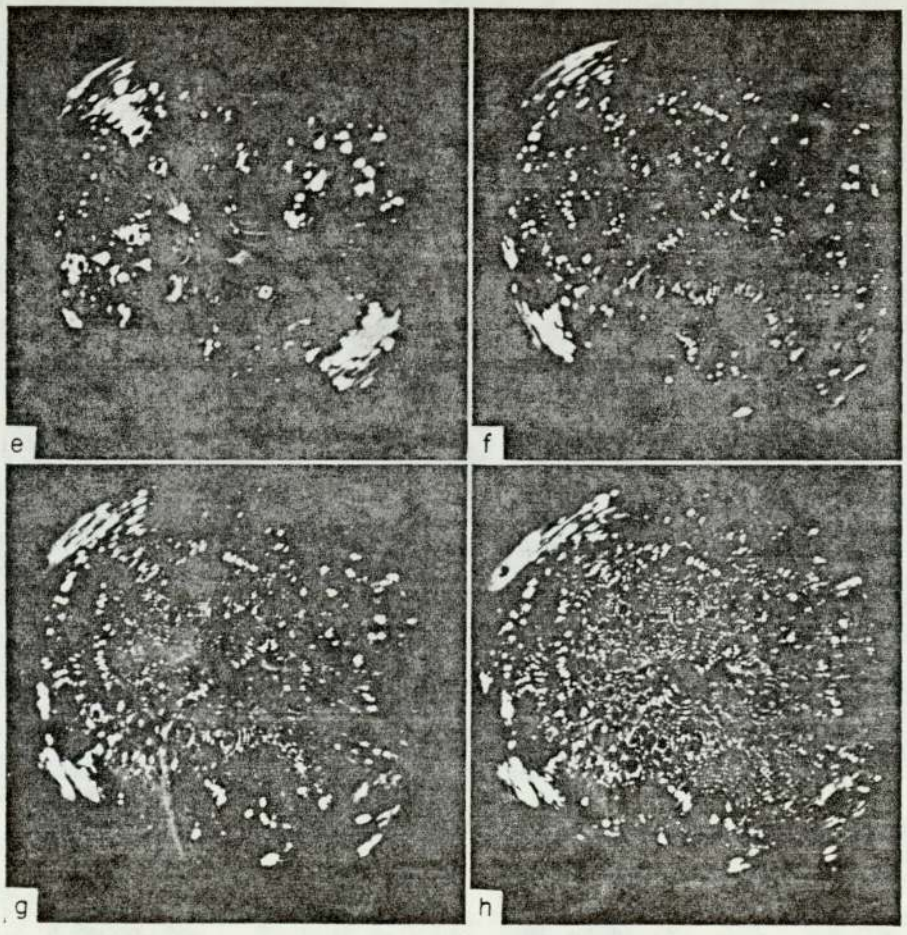


Fig. 5. (e–h). Voltage sequence taken at a specimen temperature near 5°K. (e) 10·6 kV, (f) 12·7 kV, (g) 13·8 kV, (h) 14·5 kV.

so rapidly that $(\frac{1}{2}\alpha F_0^2/kT_e)$ decreased as F_0 increased. Alternatively, if (i) T_e were so high that the gas concentration differences be very small, *and* if (ii) characteristic electron transition rates are in fact higher over the more protruding atoms, then the observed trend could be explained in terms of electron transition rate changes, as in the section on 'Theoretical predictions'. The two mechanisms correspond to reduction of gas concentration differences and electron transition rate differences, respectively. The operation of one or both would make the amount of image visible increase with voltage increase.

But experimental spot-strength comparisons as between different image regions disclose instances where lowering the specimen temperature causes a reversal in relative spot strengths: this shows decisively that 'hot general equilibria' do not exist across the whole tip surface at both specimen temperatures. Irrespective of reservations as to whether the concept is theoretically justifiable, assuming a 'hot general equilibrium' is not empirically sufficient.

The more general inference, from the whole body of evidence, is that gas distribution effects do significantly influence image contrast in a normal imaging range, certainly at very low temperatures, but that the gas concentration variations obey no simple formal rule. A closer investigation of the statistics of gas behaviour thus becomes inevitable.

IMAGING-GAS BEHAVIOUR

Qualitative analysis

Deep statistical analysis would be very complicated, as is evidenced by Van Eekelen's (1970) investigation of accommodation and ionization in a model that assumes spherical symmetry. Reliable and detailed quantitative investigation of lateral distribution effects in the real, structured, situation is virtually impossible, certainly for the present, partly because of the sheer complexity, but also because we lack proper knowledge of the potential in which the gas atoms move, the state of the surface, mechanisms of accommodation and ionization, and of the value of relevant parameters. Nonetheless, one can see in principle (Forbes, 1970) how to extend to the structured situation the primitive accounts of gas behaviour originated long ago by Gomer (1961) and Müller (e.g. 1960) in their discussions of 'spherical tip' and similar structureless models.

If ionization did not occur, the typical gas-atom history could be summarized as follows. As a result of the polarization-induced forces that have drawn them towards the specimen's tip, the imaging-gas atoms have an energy equivalent to at least several hundred degrees Kelvin when they first hit the tip. The typical gas-atom history can then be construed into two further stages: an *accommodation* stage, during which the gas cools down towards the tip temperature; and a *diffusion* stage, during which the accommodated atoms move about on the tip surface. Distribution processes occur during both stages. The accommodation stage is completed sufficiently quickly for there to be no chance during it for a Maxwell–Boltzmann concentration equilibrium at the tip temperature to be set up across the whole tip surface: the details of the distribution processes therefore depend on the details of the potential structure, and may be quite complex: there may be some tendency for *local* Maxwell–Boltzmann concentration equilibria to be set up across individual small areas of the surface (the individual small areas corresponding to 'land-locked valleys', or something similar, in the potential structure). During the diffusion stage the gas-concentration variations existing at the end of the accommodation stage will tend to adjust slowly towards the values characteristics of a thermodynamic equilibrium across the whole surface,

but, because of the nature of the potential structure, the process of adjustment will not necessarily take place uniformly.

In a first approximation, ionization can be regarded as 'cutting off' the gas-atom history at a point equivalent to the mean gas-atom residence time at the tip.

Discussion

If the above analysis is essentially correct, then all the observations discussed may be coherently rationalized.

The voltage-dependent effects (*Fig.* 5). At low fields the gas-atom lifetime at the tip could be long enough for the endpoint thermodynamic equilibrium to be nearly reached: hence, the observed changes would reflect changes in the endpoint gas concentrations. However, at higher fields ionization occurs well before the endpoint is reached: so, although the endpoint gas-concentration differences continue to become larger, the observed image changes now reflect the fact that, with increases in applied field, ionization occurs increasingly early in the gas-atom history. The earlier in the history, the more uniform the imaging-gas distribution.

The local temperature-dependent effects (*p.* 64). Explanation is possible if, near best voltage: (a) the proportion (of the total tip supply) distributed during accommodation to an individual small area of surface depends relatively weakly on temperature, but (b) the distribution within the area depends relatively strongly on temperature, *tending* (see *p.* 67) to obey equilibrium considerations based on equation (3). (Forbes, 1970, suggests that such behaviour is to be expected if nearly-accommodated atoms can move more freely within individual areas than between them, and that the potential structure is probably such as to make this likely.)

The brightness reversal noted on p. 72. Reversal of the relative strengths of image spots in different image regions can again be explained if the distribution as between areas is less temperature sensitive than the distribution within them. A specific case is illustrated in Fig. 6, where spots 1 and 2 are in one area of the image and 3 and 4 in another, and the arrows represent the effect of lowering the temperature. Comparison of the 'hot' strengths 'H' with the 'cold' strengths 'L' shows the effect.

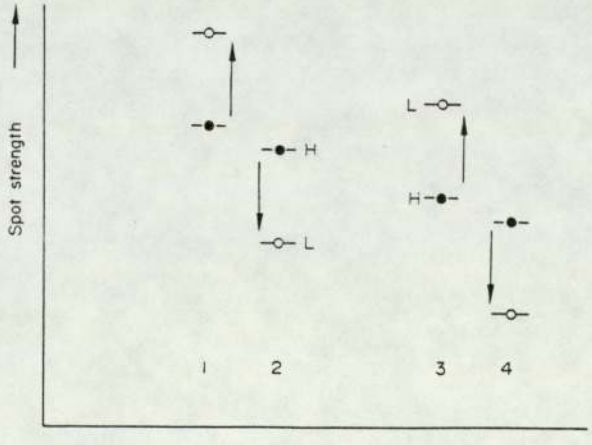


Fig. 6. To illustrate how a temperature-dependent reversal in relative spot strengths could occur (see text).

The effect of a firmly adsorbed layer

So far, the possible effects of a strongly adsorbed gas layer have not been mentioned. The layer's presence on the specimen surface but under the critical surface is strongly suggested by recent results and arguments (Müller *et al.*, 1969; Forbes, 1970; Tsong & Müller, 1970). The layer will undoubtedly influence the behaviour of the more mobile imaging gas atoms bouncing on top, and will certainly affect the numerical details of theories of gas behaviour and ionization mechanism. But, with pure helium, there seems (unpublished work) to be no significant change in the general form and direction of our arguments, for example those about the effect of field and temperature on transition-rate and gas-concentration ratios; the qualitative statistical analyses should therefore still hold good.

CONCLUSIONS

It has been shown that significant contrast features and trends in the helium-ion imaging of tungsten are impossible to explain on the traditional rate constant basis, but can be rationalized by examining how the imaging gas distributes itself prior to ionization.

For normal imaging-range images the arguments can be simplified into a 'provisional working rule':

'Image appearance is largely determined by the statistics of gas distribution. The overall current generated above an area of surface is primarily determined by supply-and-capture considerations, that is by the number of atoms per unit time that finish or tend to finish accommodation within the area. The distribution of current generation across individual small areas of surface is likely to be significantly affected by the gas distribution processes that occur after or near the end of accommodation, these *tending* to set up local Maxwell–Boltzmann concentration equilibria.' The last sentence must be understood in the context of the remarks in the section on the temperature dependence of image appearance. This rule is a slightly modified version of the one announced earlier (Forbes, 1971).

The argument has largely been qualitative and intuitive. These conclusions are therefore to be taken, not as a definitive statement, but as an empirically-sufficient hypothesis offered for further discussion. Nevertheless, the hypothesis seems a significant departure from previous beliefs about the mechanism of contrast formation, and if correct could tend both to alter the emphasis of theoretical field-ion research and to widen the range of possibilities considered by those seeking empirical explanations of new imaging effects. Others apart from the author may care to join in the potentially messy theoretical problems involved in testing its *a-priori* validity.

ACKNOWLEDGMENTS

The author would like to thank Sir Alan Cottrell, F.R.S., for starting him off on the low-temperature project; Dr M. J. Southon, for advice given as Research Supervisor; Professors Sir Alan Cottrell and R. K. W. Honeycombe for provision of facilities for the experimental work; the U.K.A.E.A. (Harwell) for a personal maintenance grant, and for supporting the experimental work; the Master and Fellows of Trinity College, Cambridge, for a Research Scholarship; and many friends and colleagues, in the Cambridge Field-ion Group and elsewhere, for criticism and discussion of the arguments herein.

References

Chen, Y.C. & Seidman, D.N. (1971a) On the atomic resolution of a field-ion microscope. *Surf. Sci.* **26**, 61.

Chen, Y.C. & Seidman, D.N. (1971b) The field ionization characteristics of individual atomic planes. *Surf. Sci.* **27**, 231.

Forbes, R.G. (1970) *Field-ion microscopy at very low temperatures*. Ph.D. Thesis, Cambridge University.

Forbes, R.G. (1971) Field-ion image formation. *Nature Physical Science*, **230**, 165.

Forbes, R.G. & Southon, M.J. (1966) A liquid-helium-cooled field-ion microscope. *Thirteenth Field Emission Symposium* (Cornell University).

Gomer, R. (1961) *Field Emission and Field Ionization*. Harvard University Press, Cambridge, Mass.

Holscher, A.A. (1967) *Adsorption studies with the field emission microscope*. Doctorate Thesis, Leiden University.

Jeans, J. (1940) *An Introduction to the Kinetic Theory of Gases*. Cambridge University Press.

Müller, E.W. (1956) Resolution of the atomic structure of a metal surface by the field-ion microscope. *J. appl. Phys.* **27**, 474.

Müller, E. W. (1960) Field ionization and field-ion microscopy. *Adv. Electronics Electron Phys.* **13**, 83.

Müller E.W. & Bahadur, K. (1956) Field ionization of gases at a metal surface and the resolution of the field-ion microscope. *Phys. Rev.* **102**, 624.

Müller, E.W., McLane, S.B. & Panitz, J.A. (1969) Field adsorption and desorption of helium and neon. *Surf. Sci.* **17**, 430.

Müller, E.W., Nakamura, S., Nishikawa, O. & McLane, S.B. (1965) Gas-surface interactions and field-ion microscopy of refractory metals. *J. appl. Phys.* **36**, 2496.

Müller, E. W. & Tsong, T. T. (1969) *Field-ion Microscopy: Principles and Applications*. American Elsevier Publishing Co., New York.

Röllgen, F.W. & Beckey, H.D. (1971) Field-induced chemisorption and promoted field ionization. *Surf. Sci.* **26**, 100.

Southon, M.J. (1963) *Image formation in the field-ion microscope*. Ph.D. thesis, Cambridge University.

Southon, M.J. & Brandon, D.G. (1963) Current-voltage characteristics of the helium field-ion microscope. *Phil. Mag.* **21**, 23.

Tsong, T.T. & Müller, E.W. (1970) Field adsorption of inert-gas atoms on field-ion emitter surfaces. *Phys. Rev. Lettrs.* **25**, 911.

Van Eekelen, H.A.M. (1970) The behaviour of the field-ion microscope: A gas dynamical calculation. *Surf. Sci.* **21**, 21.

Whitmell, D.S. (1965) *Image intensification in the field-ion microscope*. Ph.D. thesis, Cambridge University.

J. Phys. D: Appl. Phys., Vol. 11, 1978. Printed in Great Britain

LETTER TO THE EDITOR

Field-ion image contrast: the gas distribution hypothesis re-examined

J Duffell and Richard G Forbes
Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET

Received 3 May 1978

Abstract. A correction is reported to the gas concentration ratios calculated by Forbes some years ago, in connection with the gas distribution explanation of field-ion image contrast, and the explanation itself is re-examined in the light of subsequent experimental evidence.

The ratio J_A/J_B of the site-currents from two emitting sites (A and B) on the surface of a field-ion emitter is, in a quasiclassical treatment, given by:

$$J_A/J_B=(C_A''/C_B'').(P_{eA}''/P_{eB}'').(V_A/V_B) \tag{1}$$

where C_A'' is a gas concentration characteristic of site A; P_{eA}'' is an electron transition rate-constant characteristic of site A; V_A is the effective volume of the ionisation zone; and similarly for site B. Equation (1) is a restatement, using post-1971 nomenclature, of a formula previously derived (Forbes 1971, 1972).

Some years ago, one of us (Forbes 1971) suggested that gas distribution effects might play a much greater part in the explanation of field-ion image contrast than had hitherto been thought. To back this suggestion up, a table was presented showing the gas concentration ratios that would exist between two emitting sites if a characteristic-field difference of 1% existed between these two sites and the imaging-gas distribution at the emitter surface were that characteristic of a local thermodynamic equilibrium. These ratio values were based on the formula now written:

$$C_A''/C_B''=\exp\left[(\tfrac{1}{2}\alpha' F_0^2/kT).(2\delta F/F_0)\right] \tag{2}$$

where α' is the SI (absolute) polarisability for the gas atom (as defined by Forbes 1977); F_0 is the average, and δF the difference, of fields characteristic of sites A and B; k is the Boltzmann constant, and T is thermodynamic temperature.

In the course of a final-year undergraduate project (unpublished) by one of us (JD), it has been discovered that the published gas concentration ratios are in error. Corrected ratios for helium imaging gas are shown in table 1, together with the ratios

Table 1. Gas concentration ratios for helium at a best image field of 45 V nm^{-1}, for a characteristic-field difference of 1%

Original values		Corrected values	
Near 80 K	1·2	At 80 K	1·5
20 K	2·4	20 K	5·4
5 K	40	5 K	830

0022-3727/78/0009-9123 $01.00

as originally published. In deriving the corrected figures, the following values have been assumed: $\alpha'(\mathrm{He})=0{\cdot}143$ meV.(V nm^{-1})$^{-2}$ ($\approx 2{\cdot}29\times10^{-41}$ J V^{-2} m^2); $F_0=45$ V nm^{-1}; $\delta F/F_0=0{\cdot}01$; $T=5$ K, 20 K, 80 K. Slightly different temperatures were used in deriving the original ratios, but there was also an error by a factor of 2 in calculating the exponent.

In general terms, the effect of this correction is to make the gas distribution hypothesis relatively more significant (with respect to the rate-constant hypothesis) than the original figures suggested.

The relative influence of the two factors may be expressed in terms of a 'balance temperature' T_b at which (C_A''/C_B'') equals (P_{eA}''/P_{eB}''). Below this temperature the gas distribution effects will certainly be the more important influence if thermodynamic equilibrium exists. If, following Forbes (1971, 1972), we interpret Southon's work (1963) as implying that a field difference of 1% would lead, at typical imaging fields for helium, to an electron transition rate-constant ratio of at most 1·3, then the corrected estimate of T_b is 128 K. (The original estimate was 60 K.) If a field difference of 1% actually leads to a rate-constant ratio of less than 1·3, which, for example, is possible if a reduced local workfunction were associated with a high characteristic field, then the balance temperature estimate could be even higher.

It is now clear experimentally (Chen and Seidman 1971) that during ionisation under 'best image conditions' the imaging gas is not in thermodynamic equilibrium with the emitter. Nevertheless, for adjacent sites *within a single confine* it is a useful approximation to think of the gas as being *locally* in equilibrium, at some effective temperature T_{eff} above that of the emitter. It is T_{eff} that should be substituted into equation (2). The term 'confine', as used here and in earlier work (Forbes 1970), refers to an assemblage of emitting sites within a single relatively-high-field trough in the potential structure in which the imaging-gas atoms move. Geometrically, a confine may perhaps be a row of atoms, or a small crystal net plane, or a net-plane edge that forms a 'ring' in the image, or part of such a ring. In this note 'local contrast' refers to contrast as between emitting sites within a single confine.

If the mechanisms that determine image spot size were completely understood at a quantitative level, then a reasonable estimate of T_{eff} could perhaps be obtained from measurements of the temperature dependence of image-spot size. Chen and Seidman (1971), assuming the principal mechanism to be variation with temperature in the lateral velocity of the gas atom at the instant of ionisation, found T_{eff} to be approximately $7\,T_e$ (T_e being emitter temperature). Their analysis, however, fails to take account of the variation with temperature in the actual emitting area. Because of lack of information about the details of the potential in which the imaging gas moves, it is difficult to treat this second spot-size mechanism quantitatively; the authors' guess is that a revised analysis of their results might lead to the conclusion that under best image conditions, T_{eff} would be approximately $4\,T_e$.

On the basis of these figures one would deduce that, for helium imaging, the gas distribution effects would be highly likely to be the more significant influence on local image contrast under best image conditions at emitter temperatures of near 20 K and below, and that it is quite likely that gas distribution effects are the more significant influence on local contrast at temperatures above this, up to around 30 K or even higher. These theoretical inferences seem compatible with experimental image behaviour, observed qualitatively.

Obviously, these deductions are less definite than one would like. But the lack of detailed knowledge about the potentials in which gas-atom nuclei and orbital electrons

move, at a field-ion emitter surface, and the consequent uncertainties in the prediction of rate-constant and gas distribution effects, make firmer conclusions impossible at the present time.

Finally, we wish to comment on remarks made by the late Professor Müller (1976) and by Rendulic and Krautz (1975). The 'provisional working rule' put forward by Forbes (1972) suggested that the total current emitted from a confine would be determined by supply and capture mechanisms, but that the distribution of current generation within a confine would be strongly influenced by a tendency for a Maxwell–Boltzmann distribution to be set up within the confine. It was assumed that under best imaging conditions most of the ion current was generated from atoms that had lost sufficient energy to become trapped within the confine, and that the distribution arose as a result of the random bounces that each individual gas atom makes with the strongly field-adsorbed layer of gas atoms covering the emitter surface. Provided that there is a randomising influence repeatedly operating in each individual gas atom's history, Liouville's theorem should ensure that a tendency exists for a Maxwell–Boltzmann equilibrium to be established in space close above the emitter surface: it is *not* necessary for inter-gas-atom collisions to occur 'within the narrow confines of the free space above one atomic surface site'. Müller's objection (1976, p 106) is not well-founded.

In their 1975 paper, Rendulic and Krautz assume that most of the ion current is generated from atoms in a mobile, field-adsorbed 'second layer', that moves on top of the strongly-bound, first field-adsorbed layer. There is thus a large measure of agreement between their view of the state of the imaging gas prior to ionisation and ours. The main difference is that we do not assume that complete adsorption (i.e. complete accommodation of the imaging gas to the emitter temperature) occurs prior to ionisation. It appears to us that, if ionisation came from an adsorbed layer, the effective temperature of the ionised gas should be equal to the emitter temperature, which (according to the Chen and Seidman experiments) it is not. We have now given more weight to this experimental evidence than Forbes (1971) originally did.

References

Chen Y C and Seidman D N 1971 *Surface Sci.* **26** 61–84
Forbes R G 1970 *PhD Thesis* University of Cambridge
—— 1971 *Nature Phys. Sci.* **230** 165–6
—— 1972 *J. Microscopy* **96** 57–75
—— 1977 *Surface Sci.* **64** 367–71
Müller E W 1976 *CRC Crit. Rev. Solid St. Sci.* **6** 85–109
Rendulic K D and Krautz E 1975 *Acta Phys. Austr.* **42** 357–66
Southon M J 1963 *PhD Thesis* Cambridge University

SURFACE SCIENCE 27 (1971) 659–662 © North-Holland Publishing Co.

FIELD-ION IMAGING THEORY: ON THE VISIBILITY OF ADSORBATES

Received 5 August 1971

In a recent paper[1]) Lewis and Gomer reported some results of adsorption studies in a microscope operated with argon imaging gas. This note comments on one aspect of their conclusions, namely the visibility of their adsorbates under the argon-ion imaging process, and on adsorbate visibility in general.

Lewis and Gomer conclude that oxygen and carbon monoxide were not visible but that methane may be, and make the general comment that: "The situation for what may loosely be called electronegative adsorbates seems to be quite different from metallic ones, which image, provided they are not field desorbed at the viewing field." Though warning that it is not possible to explain these phenomena at the present time, they undertake some speculation, basing discussion on the implicit assumption that field-ion image appearance is determined by variations in the rate-constant for ionization. They treat field ionization as an electron transition problem, with the rate-constant expressible in terms of overlap integrals between electron orbitals.

However, it has recently been suggested[2, 3]) that variations in gas concentration may play a more significant role in determining image appearance than do variations in rate-constant. The assumption of a gas-concentration imaging mechanism here could, it seems, make significant sense out of the Lewis and Gomer conclusions.

With the gas-concentration mechanism the generation of ion current is generally highest above those localities on the surface where the local field is highest, because the imaging gas tends to partition to those localities because the gas-atom potential is deepest there. If the presence of an adsorbate led to significant charge re-distribution in the surface, and to corresponding field changes above the surface, then this fact in itself could strongly influence the visibility of the adsorbate; it would not be necessary to go into detailed considerations about the existence, positioning, and width of adsorbate electron energy levels, etc., and the effect of these on the electron transition rate. The basic prediction is that when the adsorbate is more electronegative than the substrate the consequent re-distribution of charge should lead to invisibility of the adsorbate species itself; a relatively electropositive adsorbate should be visible, and perhaps relatively brightly imaged. (The terms

electropositive and electronegative are used loosely, as indicating the direction of charge transfer in the absence of any applied field.)

In practice, however, one would also need to take careful account of geometrical factors, and of the combined effect of these and electronegativity differences on the field distribution above the surface, and of the whole statistical behaviour of the imaging gas in the complicated potential structure above the specimen surface. So, although the basic prediction is simple enough, the application of it to the interpretation of adsorbate images may be far from straightforward. Nevertheless, Lewis and Gomer's empirical conclusion about electronegative adsorbate visibility is, on the face of it, in line with theoretical expectation based on a gas-concentration imaging mechanism.

Before discussing methane, it is necessary to note that a *physically*-adsorbed species could also modify the appearance of a field-ion image. Recent work[2,4,5]) strongly suggests that under normal conditions of imaging there is a strongly-bound layer of imaging gas on the surface. The absorbed gas atoms lie above protruding surface atoms but below the corresponding ionization zones. If, in addition to the imaging gas, there is a second physically adsorbable gas species present, and this species has higher polarisability than the imaging gas, then in effect this second species displaces imaging-gas atoms from some or all the strongly-bound sites. Now, the dipole moment of the polarised strongly-bound atom will give rise to a short-range field: and in the ionization zone above the atom this field will produce a contribution to the potential for a mobile imaging-gas atom: and in general this contribution will be greater when the strongly-bound atom has higher polarisability. So, at a given site, replacement of the less polarisable species by the more polarisable species will lead to deepening of the potential in the corresponding ionization zone: and consequent stronger partitioning of the imaging gas to the region of deeper potential would lead to enhancement of the ion generation rate there, and to diminishment of the ion generation rate above neighbouring sites. This effect provides, for example, a ready explanation of the phenomena reported by Schmidt et al. when imaging in a mixture of helium and neon, with a low partial pressure of neon[6,7]); there is no need to suppose, as they tentatively do, that the rate-constant for ionization above an adsorbed neon atom is increased because of an electron transmission resonance, or by some similar mechanism.

The theoretical arguments above show that visibility of adsorbed methane could certainly be understood in principle if the methane were not decomposed and were (under the conditions of argon ion imaging) adsorbed either by means of some charge-transfer mechanism which resulted in an effective positive charge on the methane, or by means of field adsorption. (The polarisability of methane is 2.6 A^3, as compared with 1.6 A^3 for argon.)

In a current report[8]) Madey and Yates review past work and give new results concerning the methane on tungsten system. They conclude that at low temperatures (~ 100 K) methane adsorbs to appreciable coverage as an undissociated molecular species, certainly on the {100} and {111} faces. They also suggest that their measured binding energies are greater than expected from London-type dispersion forces alone, and postulate that charge-transfer contributions may be present. Further, their work-function studies are consistent with the existence of an effective positive charge on the adsorbed methane. So, although the applied field in the argon-ion imaging situation will certainly make some difference to the adsorption behaviour, recent results from other techniques provide grounds for supposing that the methane might be present undecomposed, and that the adsorption mechanism might be one of the two noted earlier[9]).

Even if decomposition does occur, it seems probable that the more electropositive component of any decomposition products might be visible in the field-ion microscope for broadly the same sort of reason as a methane molecule might be. So the visibility of methane (or decomposition product) tentatively reported by Lewis and Gomer is generally in line with theoretical expectation based on a gas-concentration imaging mechanism.

From the theoretician's viewpoint, if gas partitioning effects do have the general significance in field-ion image formation that I have hypothesised[2,3]), then it is necessary to its ultimate success that the hypothesis be able to explain (inter alia) argon-ion images of adsorbates. At a preliminary level of discussion the gas concentration mechanism seems an acceptable means of interpretation for the adsorbate observations, and at least as good as the rate-constant mechanism assumed by Lewis and Gomer.

Obviously, my remarks on adsorbate visibility in Lewis and Gomer's experiments are (like theirs) speculative and largely concerned with generalities; indeed, I believe it would be theoretically premature to attempt detailed explanations of relatively complex images when no general consensus exists as to how ordinary helium-ion images of tungsten are formed. Nevertheless, the gas-concentration imaging hypothesis does imply principles for the imaging of adsorbates, and it seems timely to outline and try to apply them: if attempts to image DNA and other biological molecules in the field-ion microscope lead to the production of reasonable micrographs, then there may develop a pressing need to understand the mechanism of image formation above organic molecules and their decomposition products.

RICHARD G. FORBES

Department of Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, England

References

1) R. T. Lewis and R. Gomer, Surface Sci. **26** (1971) 197.
2) R. G. Forbes, Ph. D. Thesis, Cambridge University (1970).
3) R. G. Forbes, Nature Physical Science **230** (1971) 165.
4) E. W. Muller, S. B. McLane and J. A. Panitz, Surface Sci. **17** (1969) 430.
5) T. T. Tsong and E. W. Muller, Phys. Rev. Letters **25** (1970) 911.
6) W. Schmidt, Th. Reisner and E. Krautz, Surface Sci. **26** (1971) 297.
7) Generally similar effects have also been observed, in the Cambridge field-ion group, by E. D. Boyes (private communication).
8) J. T. Yates and T. E. Madey, submitted to Surface Sci. I am grateful for a preprint of this paper, and for additional unpublished information.
9) I should like to acknowledge helpful discussions with Drs. M. J. Duell and J. T. Yates.

Problems in the theory of field-ion imaging*

received 21 August 1972

R G Forbes†, *Cambridge University, England*

Although field-ion microscopy is a well-established technique[1], there is currently some divergence of opinion about the physical origin of image contrast.[2] However, discussion of this is complicated by other problems, about the nature of the imaged surface and about the nature of the mechanism of ionization, and because there are various ways of formulating imaging theory. Much of the original work on field-ion theory was done by Gomer[3,4] but many others have made contributions[4], notably Müller, Southon, Cutler, Forbes, Hoischer, Knor, Ralph, Schrenck, Seidman, Tsong, Van Eekelen, and their various co-workers. This paper attempts, first, to isolate the main points of various theoretical approaches, and thence show the relationship between them; and, second, to discuss in outline some current theoretical problems.

General

The field-ion microscope is an instrument based on the continuous supply of atoms to, and emission of ions from, a charged surface. But, empirically, field-ion emission differs from other desorption processes in that one can *see* where the emission is coming from, can see which are the most active sites. So much effort has gone into examining why there are ion-current variations across the surface.

Answers have been sought at three levels (cf Figure 1). First, one attempts to correlate spot position and brightness with specimen geometry—which is moderately successful, particularly with pure-metal specimens. At a second level, explanations of contrast attribute it to variations in surface characteristics such as: protruding orbitals, broken bonds, the local work function, the local charge excess, or the local field. At the third level one *also* asks what effect these various characteristics have on the parameters that appear in site-current formulae.

There are various permissible sets of parameters, each set corresponding to a different formulation of imaging theory. Most previous discussions have in effect used quasi-classical language; and, in particular, I have used the "Q1-DS" formulation[4]:

$$J = G'' V P_e'' \qquad (1)$$

where G'' is a characteristic gas concentration, P_e'' is a characteristic electron transition rate, and V is an effective volume.

The basic physical picture behind a quasi-classical formulation is that most of the ionization occurs in a very small region of space. The experimental evidence is: first, it can be seen from micrographs that ionization is well confined *laterally*; second, ion-energy analysis[5] shows a narrow main peak (see Figure 2). Each ion is taken to have an energy characteristic of the place where it was supposedly created. Hence, because in a field of 40–50 V/nm (4–5 V/Å) the observed energy half-width corresponds to a distance half-width of about 0.01 nm (0.1 Å), it is deduced that most of the ionization occurs within a narrow layer of space parallel to the tip surface.

This latter argument is vaguely reminiscient of Bohr's account of the hydrogen atom, in as much as sharp spectral lines were taken to mean that electrons jumped between orbitals at well-defined distances from the nucleus. But, of course, it is now held that electrons jump between states that are well-defined in energy but spread out over space.

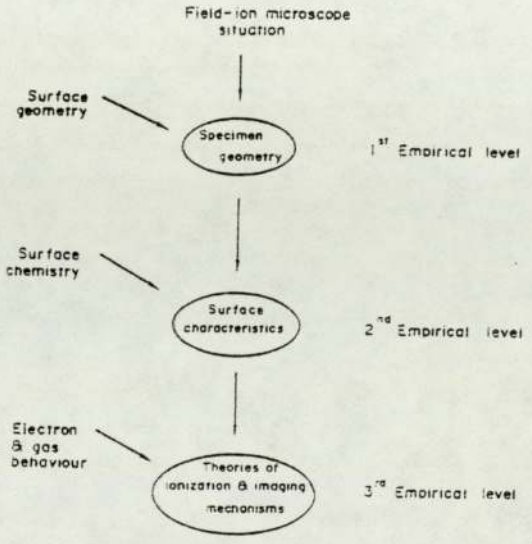


Figure 1. Empirical explanations of field-ion image contrast.

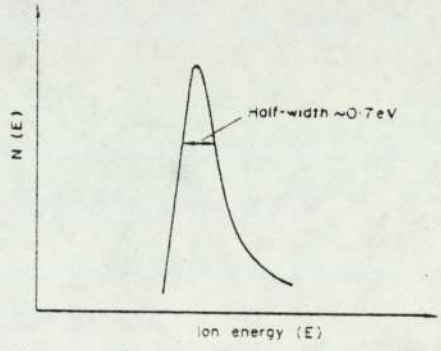


Figure 2. Ion-energy distribution (schematic).

*Paper No 3.

†Present address: University of Aston in Birmingham, Department of Physics, Gosta Green, Birmingham B4 7ET.

A similar change in viewpoint can be made in the field-ion situation in respect of the gas-atom nuclei. All that Figure 2 shows is that most of the ionization takes place into states lying within a closely-defined energy range. Imaging-mode field ionization can be considered as a *molecular* transition between molecular vibrational states defined by applying wave-mechanics to the motion of the gas-atom nucleus. The first people to think of field-ionization like this were Gomer and Swanson, who performed a detailed one-dimensional calculation a long time ago.[4] This paper generalizes their approach slightly, but without going into any of the details.

A chemical formulation of imaging theory

Formally, space above the specimen surface can be divided into cells, perhaps of different sizes, but one for each surface site. In each cell the potential for nuclear motion *normal* to the surface will be generally similar to the U^0 curve in Figure 3, if the imaging gas atom is neutral; or similar to the U^+ curve, in the case of a positively-charged ion. (In fact there are a whole family of ionic terms, depending on what energy the transferred electron has after transfer: U^+ can be considered as a "standard" term, corresponding to the electron ending up at the fermi level.) There will also be potential variations *across* the surface, but in the present approximation the lateral potential variation within a cell is neglected.

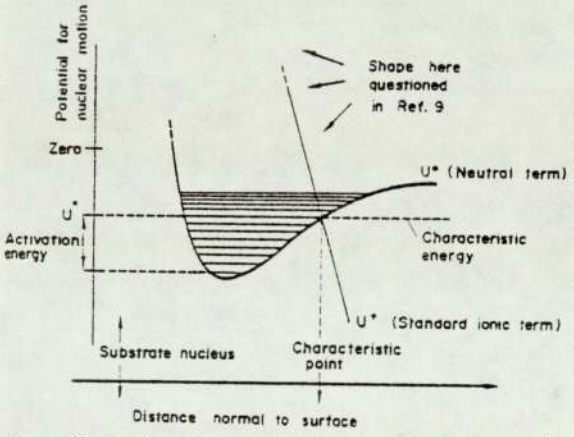


Figure 3. Section through molecular terms (molecular potential distributions) along a line through a protruding substrate nucleus and the characteristic point above it. The surface is assumed clean. The *standard ionic term* is the ionic term arising when electron transfer has taken place at the level of the metal fermi energy. The *critical surface* is the crossing surface in which the neutral term and the standard ionic term have common values: $U^0 = U^+ = U^*$. U^* varies with position in the critical surface. The *characteristic point* is the point at which U^* has a local minimum of value U^*.

Suppose that within each cell the Schrödinger equation can be solved and a set of nuclear vibrational states defined. And further suppose that for each neutral-atom vibrational state one can define an *occupation number* n_i (which is the probability of finding some atom in state i), and a rate-constant (P_i) for transition out of the neutral state i into ionic states. P_i, which is a *molecular* transition rate, can perhaps be given by an overlap integral between a molecular wave-function for the neutral state and one for the ionic state (or in fact by a sum of terms, taken over all molecular terms to which the neutral atom nucleus has reasonable access).

The product $(n_i P_i)$ is the contribution to the site-current from state i; so the total current generated within the cell is given by:

$$J = \sum_{\text{cell}} n_i P_i \qquad (2)$$

Gomer and Swanson have shown that under normal conditions most of the ionization takes place out of certain states (here called *channel states*) whose energies lie near the *characteristic energy* at which neutral and standard-ionic terms cross (see Figures 3 and 4). This is because: below the characteristic energy the nuclei have to tunnel, and the molecular transition rate goes down very rapidly; above the characteristic energy both the transition rate and the occupation number fall off with increase in energy level, and (particularly when the occupation number is essentially determined by a Boltzmann exponential) this fall off can be quite rapid, particularly at low temperatures.

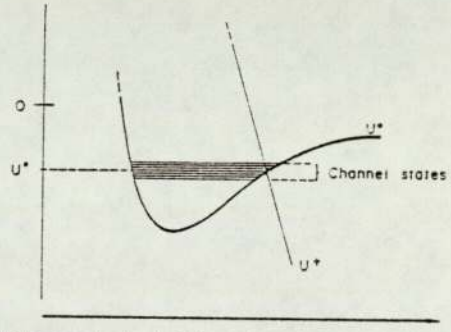


Figure 4. To show the band of channel states.

If most of the ionization in a cell takes place through channel states lying in a closely-defined energy range, the site current can be written in the "W2-DS" form:

$$J = \lambda n^* P^* \qquad (3)$$

where n^* is the occupation number for the channel states in the cell: P^* is the molecular transition rate for these states: and λ is a term which can usefully be regarded as the number of such states (but which actually encompasses other effects).

This is a second way of formulating an equation for the site current: equation (3) is wave-mechanical; equation (1) was quasi-classical.

The relation of quasi-classical and wave-mechanical formulations

The question arises: What is the relationship between the two formulations? In general, this is none too easy to define exactly, but it takes a specially simple form when the imaging gas is in thermodynamic equilibrium. For any particular site, the characteristic energy U^* is given by: $U^* = \frac{1}{2}\alpha(F^*)^2$, where F^* is the local characteristic field. So, if the gas distribution across the surface is that corresponding to a thermodynamic equilibrium, then the field variations across the surface will cause variations in characteristic energy, and these will give rise to variations in occupation number. And for two sites A and B the relevant occupation-number ratio is:

$$n_A^*/n_B^* \simeq \exp\left[\tfrac{1}{2}\alpha F_0^2 \cdot (1/kT) \cdot 2\delta F/F_0\right] = G_A''/G_B'' \qquad (4)$$

(where F_0 is the average of the characteristic fields for sites A and B, and δF is the difference). But this is of course equal to the ratio (G_A^*/G_B^*) of gas concentrations that appears in a quasi-classical theory, if thermodynamic equilibrium exists.

Recently[2], I have proposed that field-ion contrast may be determined more by gas distribution effects prior to ionization than by electron transition behaviour; and have suggested a provisional working rule that, near best image voltage, the regional contrast is basically determined by differentials in gas-supply-and-capture effects, but local contrast by a tendency for local Maxwell–Boltzmann equilibria to be set up.

Previously the arguments about gas distribution have been referred to the gas concentration ratios that appear in the quasi-classical formulation. But most of the argument can be referred equally well to the occupation-number ratios that appear in the wave-mechanical formulation. The gas distribution arguments transfer between formulations.

However, the parameter P_e^* traditionally often used to explain image contrast does not *exist* in a wave-mechanical formulation. The electron-wave-behavioural aspects of the problem are now incorporated, along with nuclear wave behaviour, in the molecular transition rate P^*. Note that P_e^* and P^* are *not* the same, because they are defined within different world-views.

The presence of the bound layer

Further aspects of the field-ion problem are exposed by transforming the direct wave-mechanical formulation above into an activated formulation as commonly used in desorption theory. On multiplying top and bottom of equation (3) by Σn_i, one obtains the "W2–AS" form:

$$J = \left(\sum_{\text{cell}} n_i\right) \cdot \frac{\sum_{\text{cell}} n_i P_i}{\sum_{\text{cell}} n_i} \tag{5a}$$

$$= g \cdot k \tag{5b}$$

k is a "cell rate-constant" similar but not identical to an Arrhenius rate-constant: and g is the cell population (ie, the probability of finding an imaging-gas atom in the cell in question). g less than one means that there is one atom there only part of the time. This activated formulation shows the relationship (and the differences) between field-ion theory and the standard theory of desorption. k in (5b) is defined via a distribution argument not the usual kinetic one, without the assumption of thermodynamic equilibrium, and using transition rates not vibration frequencies.

When thermodynamic equilibrium does exist, k has the familiar form:

$$k = C \exp(-Q/kT) \tag{6}$$

where C is a pre-exponential factor and Q is roughly equal to the activation energy shown in Figure 3.

On making plausible estimates for all relevant quantities, one may compare a maximum predicted ion current with experimentally observed ones. Near 5°K the result is:

$$J_{\text{observed}}/J_{\text{predicted}} \sim \exp(200). \tag{7}$$

This was a very unexpected result when I first discovered it some years ago. In principle, there could be various possible explanations, for example that field ionization not be a thermally-activated process; but it now seems that under normal imaging conditions there must be a strongly-adsorbed layer of imaging gas more-or-less permanently bound to the specimen surface. This layer is assumed to reduce the ionization activation energy needed by a more mobile population of gas atoms bouncing on top, and hence to lead to a reasonable number in (7). The presence of the bound layer has previously been inferred from atom-probe experiments, originally by Müller *et al.*[8]

The layer's presence will clearly affect all field-ion arguments that assumed the surface was clean, and in particular the numerical details of theories of gas distribution and of ionization mechanism.

The charge state of the bound layer

A further complication arises from Röllgen and Beckey's suggestion[9] that the strongly-adsorbed layer might be of positive ions rather than neutrals. In their view this could happen if (as a result of field penetration) the crossing point shown in Figure 3 were in fact rather nearer to the substrate atom and if there could be some "chemical" binding between the ion and the metal that would bring the ionic curve down below the neutral one for positions very near the substrate atom.

This suggestion seems intuitively unlikely. But, because very little is known for certain about the potential variation near a charged surface, it is difficult to prove their potential curve argument wrong in itself. Two counter-arguments are therefore presented here, for the case of helium imaging gas.

In the absence of any helium but with the field on, there is an average charge per surface metal atom, typically some tenths of the proton charge for normal helium imaging fields. The adsorption of a monolayer of specimen-metal atoms on to the original surface would, in effect, reproduce the original situation; so, on adsorption, somewhat less than an electron charge would be transferred from each adsorbate atom to an original surface atom.

Since the ionization potential of helium is much greater (at 24.5 eV) than that of metal atoms (typically 8 eV), it is difficult to believe that helium would be a more effective electron donor than a metal atom, even in the odd situation of the field-ion microscope. Consequently, it seems unlikely that on adsorption of a helium atom there would be a transfer of a whole electron from the helium to a substrate atom. If the adsorbed helium is not fully ionic, then the mechanism of binding invoked by Röllgen and Beckey cannot operate, and their argument falls. (It seems most unlikely that helium could bond, quasi-metallically, in a partially-ionized state, so this possibility can also be ruled out.)

This whole argument of course has less force for the heavier inert gases, with their lower ionization potentials.

A second argument against the Röllgen and Beckey hypothesis is that their energy balance calculation ignores the repulsive reaction between adjacently-adsorbed ions. If it is assumed that there is an ion-image dipole at each site on a tungsten (111)-plane (say), and the total repulsive energy for a dipole near the centre is obtained by summing over the various individual dipole-dipole interactions, a figure of about 2 eV is obtained. This has to be offset against their hypothesized binding energy of 3 eV; the geometrical plausibility of their arguments is thereby somewhat reduced.

It thus seems beyond reasonable doubt that, with helium imaging, the bound layer is there and the bound layer is neutral. The question thus arises: What effect does it have on imaging theory?

The effect of the bound layer on gas distribution and imaging mechanism

The presence of the bound layer is clearly going to affect the numerical values of accommodation coefficients and sticking probabilities—which will be important if the gas is not in thermodynamic equilibrium.

Thermodynamic–equilibrium gas-concentration ratios, or occupation-number ratios, will probably not be affected much in the case of adsorbed helium, because the presence of the strongly-adsorbed helium probably has a relatively small effect on the relative values of characteristic energies. This may be less true for the other imaging gases.

The awkward questions concern the behaviour of the transition rates, because uncertainty now arises about the nature of the ionization mechanism. Assuming that the bound layer is neutral, there seem basically three possibilities, for a helium-operated microscope:

(1) A one-stage one-electron process, involving resonance through or scattering around the strongly bound atom;

(2) A two-stage process, representable by the chemical equations:

$$(M + He) + He \rightarrow (M + He^-) + He^+$$
$$(M + He^-) \rightarrow (M^- + He)$$

(3) A one-stage two-electron process, involving the transitions shown in Figure 5, generally similar to an Auger process.

These possibilities can be discussed in outline.

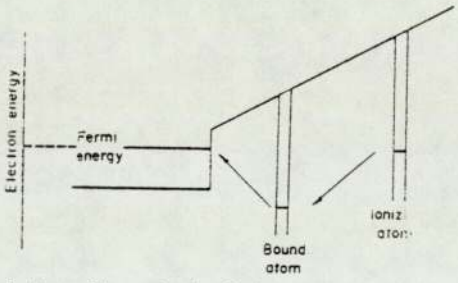


Figure 5. A possible mechanism for imaging-mode field ionization. (Same distance scale as Figure 3, but different energy scale.)

With the two-stage process, energy has to be conserved in each step separately. This means that the electron from the mobile atom has to go into a level of the He^- ion, and it seems that the relevant level is some 18–19 eV positive. This means that in a field of about 45 V/nm (4.5 V/Å) the inter-nuclear separation of the two helium nuclei must be about 0.4 nm (4Å) for the reaction to occur. It is difficult to think that everywhere across the surface one could simultaneously satisfy this condition and also the condition (that follows from energy-analysis results) that the outer nucleus be in the critical surface.

A similar difficulty arises if the mechanism is assumed to be resonance tunnelling through the adsorbed helium: it is by no means obvious that there is a level of the bound atom at the right energy.

Tsong[10] has recently proposed a mechanism of ionization that corresponds to one-or-other of these two possibilities, but his method seems to disregard the requirement that the reaction be energetically convenient, or to be applying a straight WKB integration to a situation where wave-matching is required.

If the field-ionization of helium is a one-electron process, it seems to the present author that the most likely mechanism is scattering of the transiting electron around the bound atom. And in this case it seems intuitively unlikely that the presence of a bound helium layer will have any really significant effect on the electron-transition-rate *ratios* as between different imaged sites.

The other possibility is that field-ionization is the two-electron process shown in Figure 5. But in this case also, if the transition rate is given from a product of overlap integrals, it seems unlikely that the presence of a bound helium layer will have any really significant effect on electron-transition-rate ratios.

Clearly there is an urgent need for detailed numerical investigation of the various possibilities as to ionization mechanism, but this is difficult in itself and handicapped by our lack of good knowledge about potential distributions near a charged surface.

Summary

Thus, in its present state of development, my theoretical argument about field-ion imaging may be summarized as follows:

(1) The traditional explanation of field-ion image contrast is formulated quasi-classically, and is in terms of electron-transition-rate variations. Alternatively, contrast may be explained in terms of gas distribution effects. Certain critical pieces of evidence can only be explained in terms of gas distribution effects, and this led me to formulate a provisional working rule (quoted earlier) about the origin of contrast.

(2) I think a wave-mechanical formulation is better than a quasi-classical one. The gas distribution arguments can be taken over into a wave-mechanical formulation.

(3) The direct wave-mechanical formulation may be transformed into an activated one, and through this and the atom-probe results it becomes clear that there is a strongly-bound layer of helium on the surface during imaging.

(4) For helium, the suggestion that the layer is of positive ions has been discounted. The layer seems almost certain to be neutral.

(5) The effect of the adsorbed layer on imaging parameters, and in particular on the ionization mechanism, has been briefly examined in a largely qualitative investigation; my original *qualitative* arguments about the origin of image contrast still seem to hold good.

Obviously, we have got to look at the question of ionization mechanism in more detail. This is going to be difficult, and may eventually require a many-body theory. However, the next major problem would seem to be a greater understanding of the quantum mechanics of the charged surface itself.

References

[1] The best general introduction is: E W Müller and T T Tsong, *Field Ion Microscopy: Principles and Applications*, Elsevier, New York (1969).
[2] R G Forbes: PhD thesis, Cambridge University, 1970; *Nature Phys Sci*, 230, 1971, 165; *Nature Phys Sci*, 239, 1972, 15; Z Knor, *Nature Phys Sci*, 239, 1972, 15.
[3] R Gomer, *Field Emission and Field Ionization*, Oxford University Press (1961).
[4] R Gomer and L W Swanson, *J Chem Phys*, 38, 1963, 1613.
[5] For detailed lists of references see Refs 1 and 2 above.
[6] R G Forbes, *J Microscopy*, 96, 1972, 57.
[7] T T Tsong and E W Müller, *J Chem Phys*, 41, 1964, 3279; A J Jason, *Phys Rev*, 156, 1967, 266.
[8] E W Müller, S B McLane and J A Panitz, *Surface Sci*, 17, 1969, 430.
[9] F W Röllgen and H D Beckey, *Surface Sci*, 20, 1971, 100.
[10] T T Tsong, *18th Field Emission Symposium*, Eindhoven (1971).

II. THE CONCEPT "AMOUNT OF SUBSTANCE"

9 Some new ideas in field-ion theory
R.G. Forbes
Proc. 7th Intern. Vac. Congr. & 3rd Intern. Conf. Solid Surfaces (Vienna 1977) pp 387-90

10 The mole .... interpreting it
R.G. Forbes
Educ. Chemistry, 14 (1977) 124

11 Confusion over the Avogadro constant
R.G. Forbes
Physics Educ., 12 (1977) 273-4

12 More confusion over the Avogadro constant
R.G. Forbes
Physics Educ., 13 (1978) 5-6

13 Amount of substance: An alternative proposal
R.G. Forbes
Physics Educ., 13 (1978) 269-72

14 In defence of seven dimensions; & Reply to comments
R.G. Forbes
Int. J. Mech. Eng. Educ., 7 (1979) 203-4; 8 (1980) 99-100

15 The seventh SI quantity
R.G. Forbes
Educ. Chemistry, 19 (1982) 102

Proc. 7th Intern. Vac. Congr. & 3rd Intern. Conf. Solid Surfaces (Vienna 1977) 387

SOME NEW IDEAS IN FIELD-ION THEORY

Richard G. Forbes

University of Aston, Dept. of Physics, Gosta Green, Birmingham, U.K.

Abstract: Presentation of the theory of field-ion current formation in a clarified form enables corrections to be made to several items of past work. First, the introduction of an atomic-level unit of amount-of-substance enables a formal distinction to be made between an evaporation current and an evaporation rate-constant, and a numerical error in existing theory to be rectified. Second, the method of elementary thermodynamics can be applied to obtain relationships between different variables; with the evaporation-current field-sensitivity, this enables a correction to be made to past analyses. Third, a generalised formula can be derived for the density-in-energy of a field-ion current.

INTRODUCTION

The aim of this paper is to describe some ideas that, although unorthodox in outward appearance, enable aspects of field-ion theory to be presented in a logically clearer fashion. All the ideas have firm precedents; what is involved is an extension of existing conventions, rather than any deep-seated novelty.

EVAPORATION RATE-CONSTANTS AND CURRENTS

Field evaporation may be regarded as a first-order chemical reaction occurring heterogeneously on a microscopic scale. My objective here is to derive, within the framework of the post-1971 SI system, some relationships between the amount of substance evaporated from a field-ion emitter and the rate-constants at individual high-risk sites.

In 1971 there was introduced into the SI system a seventh dimensionally-independent base-quantity, namely "amount-of-substance". This quantity is represented by the symbol n, and the corresponding dimension by the symbol {N} /1/.

In consequence, various physicochemical symbols had their meanings and dimensions subtlely changed, by international convention /2/. For example, the symbol R now represents the molar gas constant, and has the units J/(K.mol), whereas formerly the symbol R represented a quantity measured in J/K. Other redefinitions can be found in the revised International Standard /2/.

Subtle changes in linguistic habits have also been introduced. For example, in the gas equation:

$$PV = nRT \qquad (1)$$

n now represents "amount of gas" (and has dimensions {N}), whereas formerly it represented "number of gram-molecules" (and was dimensionless).

The main usage of this new convention is in chemistry. As field evaporation is a form of chemical reaction, it seems appropriate to employ the new system in field-ion theory, and to write an equation of the form:

$$J = n k \qquad (2)$$

J is the amount of substance evaporated per unit time (dimensions: $\{NT^{-1}\}$);
n is an amount of substance at risk of evaporation (dimensions: {N});
k is a rate-constant associated with field evaporation (dimensions: $\{T^{-1}\}$)

In past literature /3-5/ there has been inconsistency in the names given to rate-like quantities. The post-1971 SI conventions enable a fresh start, with the nomenclature system:

dimensions $\{NT^{-1}\}$: name - *current*
dimensions $\{T^{-1}\}$: name - *rate-constant*

The SI unit for n is the mole. But this is a most inconvenient unit in the field-ion situation, where one is dealing with relatively few numbers of atoms (or molecules, or complexes). It is better to define a new unit of amount-of-substance, as follows:

"The ordinary substance unit (symbol: osu) is the amount-of-substance of a system which contains one elementary entity. When the osu is used, the elementary entity must be specified and may be an atom, a molecule, an ion, an electron, any other fundamental entity, or a specified group of such entities."

The ordinary substance unit is related to the mole by:

$$1 \text{ mol} = \Lambda \text{ osu} \qquad (3)$$

where Λ is the Avogadro number /2,6/. Eq.(3) is analogous to the relationship between the gram and the unified atomic mass unit u .

In practice, the atoms in the outermost layer of a field-ion emitter are not at equal risk of field evaporation: the more protruding atoms have higher rate-constants. Let the number of atoms at high risk of evaporation be m , and suppose that for each of these atoms the evaporation rate-constant is k_{hr} . If n_{hr} denotes the amount of substance at high risk of evaporation, then we may write the linked equations:

$$J = n_{hr} k_{hr} \; ; \quad n_{hr} = m \text{ osu} \qquad (4)$$

In fact, experimentalists tend to quote the results of current measurements in "layers/s". Clearly, the "layer" must now be regarded as a unit of amount-of-substance, although its size cannot be exactly specified in terms of the mole or the osu. If, on average (i.e. on average during the process of field evaporation), the proportion of a layer at high risk of evaporation is r , then as an alternative to eq.(4) one may write:

$$J = n_{hr} k_{hr} \; ; \quad n_{hr} = r \text{ layers} \qquad (5)$$

If p denotes the average number of atoms in a layer, it follows from eqns (3) and (5) that:

$$1 \text{ layer} = p \text{ osu} \qquad (6)$$

$$m = r p \qquad (7)$$

All of the quantities n_{hr}, m, r, p are in principle dependent on geometrical considerations, and on field and temperature.

The above equations describe field evaporation when it is envisaged as a process involving the evaporation of individual atoms or complexes; they replace the equations given by Forbes /5/.

The real complicating factor is that there is a second method of envisaging field evaporation. This regards it as a process of layer evaporation. At any time there is one layer at general risk of evaporation; and one may define a rate-constant K_{ly} for the process of layer evaporation. With this approach one uses the linked equations:

$$J = n_{ly} K_{ly} \; ; \quad n_{ly} = 1 \text{ layer} \qquad (8)$$

Since J in eqns (5) and (8) denotes the same measurable quantity, it follows that:

$$K_{ly} = r k_{hr} \qquad (9)$$

From the standpoint of the post-1971 SI convention, all previous discussions of evaporation-current theory must be regarded as confused, on questions concerning the meaning and use of symbols printed in equations. Notwithstanding this, it is possible to consider Tsong and Müller's eq.(16) /4/ as corresponding to eq.(9) above, and to identify their symbol c_{hkl} with my r .

Tsong and Müller suggest that c_{hkl} is "probably of the order of the number of atoms in the net plane", (20 to 100, say). However, my r is clearly the proportion of a layer at high risk of evaporation, and has been estimated /5/ as approximately .01 . It thus seems that the suggestion made in Ref./4/ could lead to an error in estimated evaporation current by a factor of up to 10 000 , or so.

FIELD SENSITIVITY OF EVAPORATION CURRENT

Evaporation-current field-sensitivities were first measured by Brandon /3/, who used them to estimate values for the field-evaporation pre-exponential, A , at the high-risk sites. At the time the distinction between a current and a rate-constant was not appreciated. I show here how a correction can be made, and that the correction appears to be small.

The Arrhenius equation normally used in field-evaporation theory /3,4/ can be written in the form:

$$\ln\langle k_{hr}\rangle = M/kT = \ln\langle A\rangle - Q/kT \qquad (10)$$

where: k and T have their usual meanings, Q is the relevant activation energy, $\langle A\rangle$ denotes the numerical value of A when A is expressed in s^{-1} (and likewise for $\langle k_{hr}\rangle$), and M is a quantity defined by this equation. It follows that if F^c denotes the evaporation field defined by the unity-rate-constant criterion /5/ then $k_{hr}(F^c) = 1\ s^{-1}$, and $M(F^c) = 0$.

From eq.(5), we may define two dimensionless quantities j and L by:

$$J = j \text{ layers/s} \qquad (11)$$

$$L = \ln j \qquad (12)$$

Hence one obtains:

$$\ln j = L = \ln r + \ln\langle A\rangle - Q/kT \quad (13)$$

Since all of r, A and Q can be regarded as functions of the independent variables temperature (T) and field (F), eq.(13) in effect has the form:

$$g(L,F,T) = 0 \quad (14)$$

It follows that the same mathematical relationships and procedures apply here as in elementary thermodynamics.

Hence, from eq.(13):

$$(\partial L/\partial T)_F = (\partial \ln r/\partial T)_F + (\partial \ln\langle A\rangle/\partial T)_F - (\partial Q/\partial T)_F/k + Q/kT^2 \quad (15)$$

If the first three terms are negligible, then substitution into eq.(13) gives:

$$\ln\langle A\rangle = \ln j - \ln r + T(\partial L/\partial T)_F \quad (16)$$

This equation holds for any compatible set of L, F, T values. Obviously, it can be simplified if L can be chosen in such a way that L (= $\ln j$) = $\ln r$. This is the case if $k_{hr} = 1\ s^{-1}$: but this is the condition that $F = F^c$. It follows that:

$$\ln\langle A^c\rangle = T(\partial L/\partial T)_{F=F^c} \quad (17)$$

where A^c is the value of A at the evaporation field F^c.

The derivative $(\partial L/\partial T)_F$ is in principle measurable. But Brandon /3/ felt that it would in practice be more accurate to make use of the formula:

$$(\partial L/\partial T)_F = -F(\partial L/\partial F)_T \cdot (1/F)(\partial F/\partial T)_L \quad (18)$$

The first term on the r.h.s. is the logarithmic current field-sensitivity. Both terms can be measured for any given pair of L,T values,

The problem is that one wishes to choose an L-value that corresponds, at a given temperature, to there being a field F^c at the high-risk sites. But one does not precisely know what this L-value is, because r is not well known. The effect of a bad guess can be derived as follows.

Let the L-value actually used be $\ln j^*$ (corresponding to field F^*), rather than $\ln r^c$ (corresponding to field F^c). The error (δL) in L is thus:

$$\delta L = L(F^*) - L(F^c) = \ln j^* - \ln r^c \quad (19)$$

Experimentally, one has measured $(\partial L/\partial T)_{F=F^*}$ rather than the desired $(\partial L/\partial T)_{F=F^c}$. Using eq.(13), it may be shown that if the F and T derivatives of $\ln\langle A\rangle$ and $\ln r$ are small then the L-derivative (at constant temperature) of $T(\partial L/\partial T)_F$ equals -1 , so:

$$T(\partial L/\partial T)_{F=F^*} = T(\partial L/\partial T)_{F=F^c} - \delta L \quad (20)$$

Substituting into eq.(16) gives:

$$\ln\langle A^c\rangle = \ln(j^*/r^c) + (\partial L/\partial T)_{F=F^*} \quad (21)$$

Hence, when the bad guess is discovered, a correction may be made.

This discussion may be applied to the work of Brandon /3/. For Tungsten at a temperature near 80 K he took J^* to be .1 layers/s (i.e. $j^* = .1$). The present author, however, takes the value of r^c for Tungsten near 80 K to be .01 /5/. Hence, on this basis, Brandon's estimate of A^c is too low by a factor of 10 .

In fact, this correction (which is less than one might perhaps have expected intuitively) is not large enough to account for the discrepancy between Brandon's deduced pre-exponential values and those predicted on the assumption that the pre-exponential should be approximately equal to the nuclear vibrational frequency. The discrepancy must have some other cause.

The use of quasithermodynamic ("thermokinematic") relationships in field-evaporation theory is not in itself new. The present account, however, is not subject to certain technical objections that can be made against previous discussions, - and serves to underline the need for care over evaporation onset criteria.

DENSITY-IN-ENERGY OF FIELD-ION CURRENT

A final application, here, of the new convention is to derive a general formula for the current per unit-energy-range generated above a field-ionization site, in the real 3-dimensional situation.

In the post-1971 SI system, the amount (d^3J) of substance in the form of ions generated per unit time in a small volume dv is given by:

$$d^3J = C\,Q_e\,dv \quad (22)$$

where C is the SI quantity now called *concentration* (of neutral gas atoms), and Q_e is the *electron transfer rate-constant* as used in Ref./7/. This equation replaces eq.(3) in Ref./7/.

Consider the space above a given emitting site as divided up by a family of equipotentials of standard ion potential energy U_n /8/, and let z denote the curlinear space coordinate normal to these surfaces. If g is used to denote:

$$g = (\partial U_n/\partial z)^{-1} \qquad (23)$$

then the element of volume between two equipotentials dU_n apart is given by:

$$dv = dz\, dS = g\, dU_n\, dS \qquad (24)$$

where dS is the element of area on an equipotential surface. Consequently, the amount (dJ) of substance (ions) generated per unit time in the energy shell between U_n and $U_n + dU_n$ is:

$$dJ = (dJ/dU_n)\, dU_n = dU_n \int g\, C\, Q_e\, dS \qquad (25)$$

In the performance of the integral above, one must take into account that for points inside the critical surface /7/ $Q_e = 0$, whereas for points outside the critical surface $Q_e = P_e$, where P_e is the *electron transition rate-constant*. The result of integration can be written formally as:

$$dJ/dU_n = P'_e \,.\, (C'\, g'\, S') \qquad (26)$$

where P'_e, C' and g' are values taken on the axis of the ionization zone, and S' is the "effective area of ionization on the surface of energy U_n".

If any ion kinetic energy at the instant of ionization is ignored, then dJ/dU_n may be identified with the measured current per unit-energy-range.

As the fine details of potentials are not well known at the present time, it is difficult to take the analysis further. However, it is clear from Fig. 1 that one can envisage geometrical arrangements of the equipotentials and the critical surface such that S' could be a rapidly varying function of U_n, because of the variation (with U_n) in the area of equipotential surface that lies outside the critical surface.

Clearly, one must consider the possibility that the half-width of the low-energy-deficit side of the energy distribution "main peak" may be affected by the effect just described. Energy dependence of the effective area of ionization can arise only in 3-dimensional models; its existence has not previously been pointed out.

One final comment applies to this paper as a whole. It may seem strange that at the present time one should seek to introduce a new unit. It is possible to give consistent formulations of the theory of field-ion current generation within the framework of the pre-1971 system of dimensions. However, the author finds the approach developed here much clearer, once the nature of the osu has been understood.

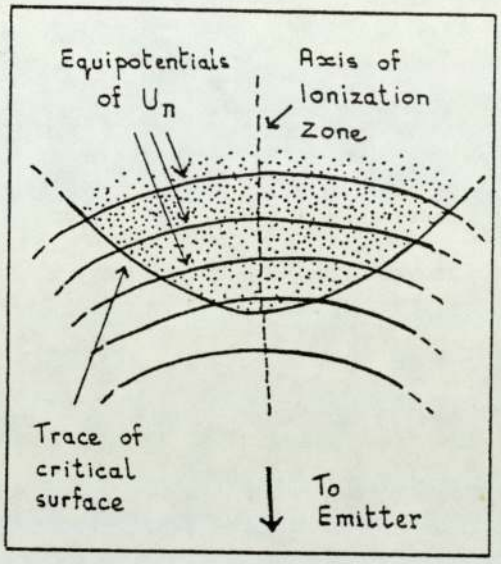


Fig. 1. Schematic diagram of the ionization zone above an emitting site. Ionization can occur only in the dotted region outside the critical surface. (The real shape of the critical surface is not known.)

REFERENCES

/1/ D.A.Lowe: A guide to International Recommendations on Names and Symbols for Quantities and on Units of Measurement (World Health Organisation Geneva, 1975).
/2/ International Standard ISO 31/VIII--1973(E).
/3/ D.G.Brandon: Phil. Mag. 14 (1966) 803.
/4/ T.T.Tsong and E.W.Müller: Phys. Stat. Sol. (a) 1 (1970) 513.
/5/ R.G.Forbes: Surf. Sci. 46 (1974) 577.
/6/ R.G.Forbes: Physics Education, July 1977.
/7/ R.G.Forbes: J. Microscopy 96 (1972) 57.
/8/ As defined in, for example: R.G.Forbes: Surf. Sci. 61 (1976) 221.

EDUCATION IN CHEMISTRY - Letter to the Editor

Educ. Chem., 1977, 14, 124

The mole—defining it . . .

I read with interest Professor McGlashan's comments on the use of the 'mol^{-1}' in thermochemical statements (*Educ. Chem.*, 1977, 14, 17), but I think we need to go one stage further. The definition of the mole, quoted in the article, read in part 'the elementary units must be specified'. In other words we should always state 'per mole of something'.

The statement

$2H_2(g) + O_2(g) = 2H_2O(l)$
and $\Delta H^\ominus_r(298.15\ K) = -571.66\ kJ\ mol^{-1}$

could be considered to be ambiguous. Do we mean 'per mole of $H_2O(l)$ formed' or perhaps 'per mole $O_2(g)$ reacting'? Most of us know that the $\Delta H^\ominus_r(298.15\ K) = -571.66$ kJ refers to the reaction between 2 moles $H_2(g)$ and 1 mole $O_2(g)$ to give 2 moles $H_2O(l)$. I would suggest we use the term 'per mole equation', unless someone can suggest something better.

P. V. Shooter
The Abbot Beyne School,
Evershed Building,
Mill Hill Lane,
Burton upon Trent,
Staffs. DE15 0AZ.

. . . and interpreting it

On comparing my earlier letter (*Educ. Chem.*, 1976, 13, 92) with Dr Hudson's article (*Educ. Chem.*, 1976, 13, 110), it is clear that he and I have different concepts of the mole. I wish to argue that the 'counting unit' interpretation is more fundamental and easier to understand. To aid discussion I shall use the word 'ont' to refer to any form of elementary entity (atoms, molecules, ions, protons, electrons *etc*). 'Ont' comes from the Greek word ὄν (meaning 'being' or 'a being thing') that has been widely used in forming names.

My concept of the mole is that it is simply a large counting unit. Dr Hudson's concept, however, seems closer to what one might call the 'gram-ont'. For him, 'The essential feature of the mole is that it is a quantity or amount of substance *expressed as a mass or volume*'. This idea seems the natural descendant of the concept of a gram-equivalent, which originated in the days before science was really sure that atoms existed. There is, however, nothing in the formal definitions of the SI system that requires the mole to be interpreted as a gram-ont.

It seems to me that my concept of the mole should be fairly easy to grasp. Further, the idea of 'mass per counting unit', and its variation as between different types of object, should be relatively easy to put over. Given adequate numbers of reasonably uniform ball-bearings, of two different sizes, and a chemical balance, it should be possible to determine their 'masses-per-dozen', and to show that getting two dozen of type A *by weighing* requires a different weight from getting two dozen of type B by weighing. From the equation:

mass = mass-per-dozen × quantity (expressed in dozens)

eg 18 g = 9 g dozen^{-1} × 2 dozen

it should be possible to proceed to the equation:

mass = mass-per-mole × quantity (expressed in moles)

eg 32 g = 16 g mol^{-1} × 2 mol

Hence one arrives at a method of measuring out moles, but the apparently difficult concept of relative atomic mass has not been used. No doubt, books of constants will continue to list 'relative atomic mass' rather than 'naturally occurring molar mass', but it seems to me that the latter is easier to understand and use—because it has the units g mol^{-1}.

Dr Hudson writes that it is not correct to think of the mole as analogous to the dozen, because the mole is not a number. In my opinion the dozen is not a number either: both are counting units. The analogies I see between the counting of molecules in moles and (say) ball-bearings in dozens are as in the Table below.

Molecules	Ball-bearings
Mole	Dozen
Avogadro constant 6.022×10^{23} mol^{-1}	Salesmen's constant 12 dozen^{-1}

(Note that in the post-1971 SI system the Avogadro constant has the dimensions [amount-of-substance^{-1}], and is *not* a pure number.[1] In my previous letter the symbol in the equations should have been the new mass-unit ratio Λ_{kg} (1 kg = Λ_{kg} u), not the post-1971 Avogadro constant.)

If there is lack of strict comparability between the mole and the dozen, it is because the post-1971 SI system has got itself into a bizarre logical state whereby a large counting unit is regarded as having a dimension formally called [amount-of-substance], but the ordinary counting unit[2] is not recognised or is treated as dimensionless. It is then an arbitrary choice as to whether the counting unit, the 'dozen', does or does not formally have the SI dimension [amount-of-substance].

There are good historical reasons for the present situation, which is perhaps an inevitable transition stage. But the sooner we get out of it by adopting the revised convention that small counting units all have the same formal dimension as the large counting unit the mole, and by achieving understandings about the meaning of the word 'number' as used in scientific names, then the less confusion there should be.

Returning to the question of alternative concepts of the mole, I wonder if the vital difference between the two interpretations lies in their origins. The 'gram-ont' interpretation comes out of the historical, macroscopic, tradition of chemical exploration. The 'counting unit' interpretation comes (for me) out of research into field-ion microscopy and related techniques. Here, one is accustomed to seeing metal atoms directly, to measuring their masses directly (with a mass spectrometer), and to measuring the rates of ion formation in certain special chemical reactions by directly counting the number of individual ions formed per unit time. In the corresponding theory neither the concept of weighing nor the gram-ont interpretation of the mole have any place whatsoever. But if the mole is interpreted as a counting unit, and an atomic-level counting unit is specified, then one can construct a microscopic reaction-rate theory that is closely analogous to the normal macroscopic reaction-rate theory for unimolecular reactions. Therefore, I believe, the counting unit interpretation must be regarded as the more powerful and more fundamental one.

Richard G. Forbes
University of Aston,
Department of Physics,
Gosta Green,
Birmingham B4 7ET.

References

1. See *International Standard ISO 31/VIII–1973 (E)*, definition 8–4. 1.
2. *i.e.* the 'single-ont', or ordinary substance unit (osu), of my previous letter.

Physics Education 12, 273-4 (1977)

CONFUSION OVER THE AVOGADRO CONSTANT

As a result of some recent experience, I would like to draw wider attention to conceptual errors in the table of fundamental constants in the SI edition of Kaye and Laby (1973).

Before 1971 the names 'Avogadro constant' and 'Avogadro number' were often used interchangeably. The new International Standard (1973) has altered this. The Avogadro constant, L (or N_A), is now a *physical quantity* with the dimensions [amount-of-substance]$^{-1}$. But the Avogadro number is a pure dimensionless number—the numerical value of L when L is expressed in mol^{-1}. There is no internationally recognized symbol for the Avogadro number; for the sake of the present argument I denote it by Λ. Thus, $\Lambda = L/\text{mol}^{-1}$. Kaye and Laby give the Avogadro constant (L) as $6{\cdot}022\,17 \times 10^{26}$ kmol^{-1}. This form is unconventional but not incorrect. However, to avoid confusion (particularly as between physics and chemistry teaching), it may be preferable normally to cite L in mol^{-1}, as below.

Conceptual errors arise because the form of words used in Kaye and Laby to explain the term 'Avogadro constant' is in fact a definition of the number 1000Λ. The symbol N as used in Kaye and Laby sometimes means L, sometimes means 1000Λ and should not be there at all in the citation of electron mass. The compiler of the table has failed to distinguish conceptually between a physical quantity and an associated numerical value (Royal Society 1975).

There are also several other discrepancies between international practice and the tables on pp16 and 17 of Kaye and Laby. I give in table 1 corrected versions of various faulty or ambiguous entries. For consistency, I retain the numerical values used by Kaye and Laby. (More recent values are to be found in Royal Society 1975.) The right-hand column gives the standard error, in parts in 10^6.

The fourth entry here deserves comment. The unified atomic mass unit (symbol: u) is the mass unit equal in size to the unified atomic mass constant m_u. This mass unit is internationally recognized for continued use alongside SI units, and I would regard it as acceptable to cite electron mass in this unit. However, some authorities prefer to use SI units exclusively, in the context of teaching. Hence I have avoided using this unit by giving instead the relative molecular mass (m_e/m_u) for the electron.

It must also be made clear that, although the meaning of the phrase 'Avogadro number' is stated as a remark in the International Standard (1973), no entry concerning it appears in any other authoritative table. (This information was provided by a referee, who strongly warned that the inclusion of the entry concerning Avogadro number might serve to perpetuate confusion. He suggests that the remark in the International Standard was intended only as an historical cross-reference.) It is included here because I personally find the concept necessary in my research, and because the concept may be essential in the potential future development (Deslattes *et al* 1974) of the SI system itself. However, within the framework of the SI system as at present constituted it seems best to use the Avogadro constant rather than the Avogadro number in all normal teaching contexts.

More generally, it seems that the new International Standard forces on us subtle distinctions and changes in linguistic habits and creates considerable potentialities for confusion. I fear that Kaye and Laby's treatment may unconsciously teach students that carelessness about units and dimensions is acceptable in this area.

Richard G Forbes

Department of Physics, University of Aston

REFERENCES

Deslattes R D *et al* 1974 *Phys. Rev. Lett.* 33 463

International Standard ISO 31/viii-1973 (E) *Quantities and Units of Physical Chemistry and Molecular Physics*

Kaye G W C and Laby T H 1973 *Tables of Physical and Chemical Constants* 14th edn (in SI units) (London: Longman) p16

Royal Society 1975 *Quantities, Units and Symbols* 2nd edn (London: Royal Society)

Table 1 *Values of fundamental constants. Numerical values from Kaye and Laby (1973)*

Fundamental constant	*Symbol*	*Value*	*Standard error (in 10^6)*
Elementary charge (of proton)	e	$1{\cdot}602\,192 \times 10^{-19}$ C	4·5
Unified atomic mass constant (1/12 of the mass of an atom of carbon-12)	m_u	$1{\cdot}660\,53 \times 10^{-27}$ kg	7
Mass of electron	m_e	$9{\cdot}109\,56 \times 10^{-31}$ kg	6
Relative molecular mass of electron ($= m_e/m_u$)	$M_r(e)$	$5{\cdot}485\,93 \times 10^{-4}$	7
Avogadro constant (number of molecules/amount-of-substance)	L	$6{\cdot}022\,17 \times 10^{23}$ mol^{-1}	7
Avogadro number (number of atoms in 12 g of carbon-12)	Λ	$6{\cdot}022\,17 \times 10^{23}$	7
Faraday constant ($F = Le$)	F	$9{\cdot}648\,67 \times 10^{4}$ C mol^{-1}	6
Molar gas constant	R	8·3143 J K^{-1} mol^{-1}	42
Boltzmann constant ($k = R/L$)	k	$1{\cdot}380\,62 \times 10^{-23}$ J K^{-1}	43

Physics Education 13, 5-6 (1978)

MORE CONFUSION OVER THE AVOGADRO CONSTANT

In the July issue of *Physics Education* (1977 12 273–4) I warned of the possibility of confusion between the Avogadro constant and the Avogadro number, and suggested the need for subtle changes in linguistic habits. Unfortunately, three of the contributed articles in the same issue did not succeed in completely avoiding this trap.

A trivial example occurs at the top of p283. P M Whelan writes that 'The volume from which a mole of molecules will be excluded is $\frac{1}{2}N_A(\frac{4}{3}\pi r_0^3)$'. This wording is dimensionally inconsistent. He should have written 'The volume per mole from which molecules will . . .' or perhaps 'The volume per unit amount of gas from which molecules will . . .'. I am inclined to think that the advantage of being precise outweighs the disadvantage of seeming pedantic.

A second example is at the bottom of p285, where R T Allsop writes 'The definition given (of the mole) is equivalent to saying that one mole of any substance is that amount which contains the Avogadro constant of its particles as defined on the carbon-12 scale. The Avogadro constant is a constant number of particles per mole of particles for all substances'. The first cited sentence, as it stands, is dimensionally inconsistent and less than clear. An alternative wording less subject to such criticism would be 'The definition is equivalent to saying that for all substances there are the same number of particles per mole. The relationship between number of particles (N) and amount of substance (n) is

$$N = L\,n \qquad (1)$$

where L is the Avogadro constant ($\sim 6.022 \times 10^{23}$ mol^{-1}). L is a universal constant'.

In the third case a fault occurs in precisely the area (development of the SI itself) mentioned in my previous letter as requiring retention of the concept of Avogadro number. On p291 B W Petley writes 'Thus if we fix the value of N_A, by definition, then N_A molecules of a substance of molecular weight M would have a mass of exactly $10^{-3}\,M$ kg '. In the table on p294 it is clear that Dr Petley's N_A is indeed the post-1971 Avogadro constant. However the sentence is dimensionally consistent and scientifically meaningful only if N_A there represents the Avogadro *number* (assuming M denotes relative molecular mass).

At this point I become heretical because I feel that the true nature of the post-1971 Avogadro constant has not been adequately appreciated. To show this I introduce a new unit of amount-of-substance by the following definition:

'The *ordinary substance unit* (symbol: osu) is the amount-of-substance of a system that contains one elementary entity. When the ordinary substance unit is used the nature of the elementary entity must be specified and may be an atom, a molecule, an ion, an electron, any other fundamental entity, or a specified group of such entities'.

The relationship between the mole and the ordinary substance unit is

$$1 \text{ mol} = \Lambda \text{ osu} \qquad (2)$$

where Λ is the Avogadro number, as defined by

$$1 \text{ g} = \Lambda \text{ u}. \qquad (3)$$

The ordinary substance unit thus bears the same relationship to the mole as does the unified atomic mass unit (u) to the gram. One can also define another new unit of amount-of-substance, the 'dozen', by

$$1 \text{ dozen} = 12 \text{ osu}. \qquad (4)$$

These units are of course not 'recognised', scientifically. However, the definitions are precise and are completely compatible, logically and dimensionally, with the existing (post-1971) SI system. The usefulness of the osu in a research context is discussed elsewhere (Forbes 1977).

The Avogadro constant L is a physical quantity with the dimension [amount-of-substance^{-1}]. Consequently it can be meaningfully expressed in any unit having this dimension (Royal Society 1975). In particular, its value can be expressed in any of the forms shown below:

(i) $L \approx 6.022 \times 10^{26}$ kmol^{-1}
(ii) $L \approx 6.022 \times 10^{23}$ mol^{-1}
(iii) $L = 12$ dozen^{-1}
(iv) $L = 1$ osu^{-1}.

Since neither any physical quantity nor the symbol used to denote it should imply a particular choice of unit (Royal Society 1975 p6), it follows irrevocably that Dr Petley's discussion is about the fixing by international convention of the Avogadro *number* rather than the Avogadro constant. It further follows that attempts to de-legitimise the concept of Avogadro number, and to discontinue its usage in scientific literature, may be somewhat short sighted.

Form (iv) above shows the most clearly that L is the symbolic representation of the logical process of replacing '(property) *of* a molecule' by '(property) *per* unit of amount-of-substance'. This is the origin of the post-1971 need to become extremely careful about the use of the words 'of' and 'per' in ordinary English-language discussions of molecular physics and related topics.

That this operator L should be called the 'Avogadro constant' when used in a scientific context seems an unfortunate recent accident of scientific history. L is a more general logical object, associated with the use of counting units. It can find everyday use, in equations such as

Cost *per* counting unit = $L \times$ Cost *of* a single item.

The English language contains many words for counting units (which are not the same concept as numbers), e.g. dozen, score, gross, six-pack, single, half-brace. The mole is basically a large counting unit, used in the counting of objects very small by human standards, that has a size determined by scientific history.

The real significance of the changes made to the SI in 1971, it seems to me, is that they constitute a first step towards a formalisation within science of the process of counting, and imply a declaration that from this time on counting is to be part of science and is not to be left to mathematics and language teaching. All of which seems a natural (though belated) historical consequence of the discovery of the atomic nature of matter.

Richard G Forbes

Department of Physics, University of Aston

REFERENCES

Forbes R G 1977 *Proc. 7th Int. Vac. Congr. and 3rd Int. Conf. Solid Surfaces, Vienna 1977* p387

Royal Society 1975 *Quantities, Units and Symbols* 2nd edn (London: Royal Society)

Physics Education 13, 269-272 (1978)

Phys. Educ., Vol 13, 1978. Printed in Great Britain

AMOUNT OF SUBSTANCE: AN ALTERNATIVE PROPOSAL

In his recent article M L McGlashan (*Phys. Educ.* 1977 12 276-8) suggested abandoning the physical quantity 'amount of substance' and its unit, the mole. This amounts to a change in the international conventions about how to express the physical *property* 'amount of substance' (i.e. how to express how much of a substance is present, when thinking in terms of a count of its elementary entities). This letter demonstrates that there are four possible conventions, and advocates a change to one apparently not so far considered. A necessary preliminary is to look again at the basic nature of the international system of measurement, and the 1971 changes therein, and to sharpen up some definitions and nomenclature.

A clear distinction must be made between the concepts of 'physical property' and 'physical quantity'. Crudely, a physical property is some property of the world that we wish to measure; its existence is a physical fact. A physical quantity used to 'express' a physical property is, by definition (e.g. Royal Society 1975), a logical entity with definite dimensions; a physical quantity is thus a thing of convention, because it has to be defined within the context of a system of dimensions fixed by international agreement. Within a given system of dimensions most physical properties can in principle be 'expressed as' any one of several physical quantities. To take a trivial example, the property 'distance' can be expressed either as the quantity 'length' or as the quantity 'number of metres'. These two are certainly different quantities because the latter is dimensionless whereas the former is not.

Usually, a convention has developed that one particular physical quantity is used to express the physical property. In some cases, however, competing conventions exist. For example, 'magnetisation' M and 'magnetic polarisation' J ($= \mu_0 M$) are two different quantities expressing a single property. The relics of past competitions also lie about the system. Thus the property 'amount of electricity' is now normally expressed as the quantity 'electric charge', q; but the obsolete alternative quantity 'Gaussian electric charge', q_s [$= q/(4\pi\varepsilon_0)^{1/2}$], is still defined in the current International Standard.

In the international system of measurement each designated base quantity is regarded as having its own independent dimension, and the dimensions of all quantities defined in physical science can be stated in terms of these base quantity dimensions. The fundamental change made in 1971 was the introduction of a new dimension. The corresponding base quantity is here represented by ν.

Where a given physical property is normally expressed by only one physical quantity, the same name tends to be used for both property and quantity, without confusion. But when more than one quantity is being used to express a given property, discussion is clearer if all the quantities have names different from that of the property. In pre-1971 literature the phrase 'amount of substance' tends to be used linguistically in much the same way as I have used 'amount of electricity' above. Hence I shall regard 'amount of substance' as the continuing name of a physical *property*. In consequence, for the purposes of this letter, I require an alternative name for the physical quantity ν whose base unit is the mole. A name based on the word 'count' seems most appropriate. Phrases such as 'ion count' are in use in some areas of physics. But it seems preferable to employ a new word, and I shall use '*ont-count*'. (The prefix 'ont-' comes from the Greek word *ov* that has been widely used in forming the names of fundamental entities; 'ont-count' can be translated, roughly, as 'count of being things' or as 'thing-count'.) We are now in a position to define four possible conventions about how the property 'amount of substance' may be expressed.

(1) Prior to 1971 the property 'amount of substance' was normally expressed in terms of one of two dimensionless physical quantities. Small amounts of substance were expressed as a 'number of particles'. Large amounts of substance were expressed as a 'number of gram-molecules' (or of gram-atoms, etc).

(2) Subsequent to 1971 the property 'amount of substance' has been expressed in terms of physical quantities of different dimensions. Small amounts of substance continue to be expressed as a 'number of particles (N)'. But large amounts of substance are now expressed as an 'ont-count (ν)', whose value is given in terms of the unit 'mole'. The relationship between the

Phys. Educ., Vol. 13, 1978. Printed in Great Britain

two quantities is:

$$N = Lv \qquad (1)$$

where L is the (post-1971) Avogadro constant ($\sim 6.022 \times 10^{23}$ mol^{-1}). The quantity L is a fundamental constant of the international system of base quantities and dimensions, as are the electric constant ε_0 and the magnetic constant μ_0, and the above relationship is logically analogous to the relationships between J and M, and between q_s and q, cited earlier.

The recent changes also involved international agreement that there would be changes or modifications in the scientific meanings of various words, phrases, symbols and equations. For example, these included mole, molar, amount of substance, Avogadro constant, van der Waal's equation, n, R, $pV = nRT$. In the author's view, much of the confusion evident in the literature over the SI system results from failures to realise that changes in meaning have been made. This letter tries to minimise difficulties by careful choice of nomenclature.

The linguistic misunderstandings will no doubt diminish with time. But the present underlying conceptual situation is itself unsatisfactory, in that two physical quantities with different dimensions are being used to express a single physical property. Prior to 1960 this kind of situation existed with respect to the property 'amount of electricity' and other electrical properties. The SI system was originally established with the object of eliminating the confusion over the electric and magnetic quantities. But the 1971 changes seem to have *introduced* problems of an essentially similar logical nature with respect to the quantities expressing the property 'amount of substance'.

The problems are most noticeable when attempting to relate molecular physics to macroscopic physical equations, when the Avogadro constant has to be introduced in order to change from one quantity to another with different dimensions, and to some extent they may be permanent. However, there are two further alternative conventions that could ameliorate the situation.

(3) Professor McGlashan suggests that use of the quantity 'ont-count' and the mole should be progressively abandoned. He suggests that we should move progressively to the use of molecular quantities (i.e. to the use of quantities expressing the properties *of* elementary entities), and to the use of equations involving molecular quantities. Both small and large amounts of substance would presumably be expressed as a 'number of particles'.

(4) The present author's suggestion is that use of the quantity 'number of particles' should be progressively abandoned. Both small and large amounts of substance would be expressed as the quantity 'ont-count'. The value of an ont-count would be given in terms of the unit 'mole' for large ont-counts, but for small ont-counts it could alternatively be given in terms of the 'ordinary substance unit' introduced in an earlier letter (R G Forbes *Phys. Educ.* 1978 13 5-6). (The ordinary substance unit is the ont-count of a system that contains one elementary entity. When the ordinary substance unit is used the nature of the elementary entity must be specified and may be an atom, molecule, ion, electron, any other fundamental entity, or a specified group of such entities.)

If one believes that it is part of the fundamental philosophy of the international system of measurement to develop to the position where each base property is conventionally expressed by only one physical quantity, then conventions (3) and (4) must both be considered superior to the official (post-1971) convention (2). In the choice between (3) and (4), several factors support convention (4). The SI unit system has its origins in a need felt for 'practical' units of convenient size. If convention (3) were adopted it seems likely that a need would be experienced for a convenient way of talking about large numbers of elementary entities. Convention (3) thus seems unstable, and likely to revert to convention (1) or (2).

At a more philosophical level one can argue that counting is one of the very basic experimental processes of science, and that one should formalise this process in a manner directly analogous to that used for other basic processes such as time-measuring and distance-measuring. (In the case of distance, a possible analogy to convention (2) is that large distances should be expressed as a length, but small distances as a number of Bohr radii; the analogy to (3) would be that all distances should be expressed as a number of Bohr radii; the analogy to (4) would be that all distances should be expressed as lengths. Clearly, the analogy to (4) is the recognised convention.)

It also seems that the pedagogical advantages of being able to allocate different dimensions (and hence units) to 'amount of substance formed per unit time' and to 'rate constant' are not to be surrendered lightly. In convention (3) both would presumably be measured in s^{-1}. The nuclear physicists, faced with a partially analogous situation, felt obliged to resort to the 'special' procedure of defining a special unit with dimensions $[T^{-1}]$, namely the becquerel, so as to be able to make appropriate discriminations. Convention (3) might well produce a demand for a 'chemists' bequerel'. (The usefulness of convention (4) in clarifying discussion of atomic-scale chemical rate phenomena has been demonstrated elsewhere (Forbes 1977), in connection with aspects of field-ion emission.)

The principal objection to convention (4) will probably be that it envisages the continued use of two different units of ont-count, the 'mole' and (in appropriate contexts) the 'ordinary substance unit' (osu). To the author's mind, however, this duality is a lesser evil than the duality inherent in the current convention (2), where one has two sets of physical

quantities and two sets of equations, for use according to context (e.g. $pV = \nu RT$ for macroscopic use, $pV = NkT$ for atomic-scale use). The principle of duality with respect to units, in appropriate contexts, is already part of the international system of measurement, where the unified atomic mass unit (u) and the electron volt (eV) are recognised for continued use alongside SI units; the principle reflects the realities of modern experimental science, where the behaviour of individual atoms can be investigated. Obviously, the adoption of convention (4) would bring with it certain changes in the teaching of molecular physics, particularly elementary kinetic theory. My impression is, though, that the advantages could more than offset the disadvantages, once teachers' unfamiliarity had been overcome.

In summary, I hope that if nothing else this letter has clarified the concepts associated with the property 'amount of substance', and has exposed some of the current problems and possibilities. Professor McGlashan expressed the view that the physical quantity here called ont-count and its unit the mole are 'redundant lumber' that science would do well to get rid of. On the contrary, I feel that the 1971 decision, which in effect introduces a dimension formally associated with the *process of counting*, could come to be seen as an important and wide-reaching scientific and intellectual development. This I now wish to build on, by having the ordinary substance unit (under this or some other name) recognised for continued use alongside SI units. If the ideas expressed here found wider acceptance, then a slow progressive change to convention (4) could take place.

It must, of course, be clearly understood that the views expressed here have no official status whatsoever. Nevertheless, I believe that this letter represents a coherent logical analysis of the situation as it currently exists. My hope is that it may serve to stimulate fundamental discussion of the alternatives.

Richard G Forbes

Department of Physics, University of Aston

REFERENCES

Forbes R G 1977 *Proc. 7th Int. Vac. Congr. and 3rd Int. Conf. Solid Surfaces, Vienna 1977*, pp387-91

Royal Society 1975 *Quantities, Units and Symbols* 2nd edn (London: Royal Society)

International Journal of Mechanical Engineering Education, 7, 203–204 (1979)

Letters to the Editor

IN DEFENCE OF SEVEN DIMENSIONS

Dear Sir,

Mr Clayton draws attention (1) to misunderstandings concerning the mole, and Professor Walshaw (2) wants to know why the mole should be considered an SI base unit. In taking up his challenge I am not defending either the half-finished job so far made of introducing the underlying concept of 'amount of substance', or the massive confusion that exists in current literature. But I think I can justify the introduction of a seventh independent base-quantity into the International System of Measurement, by making a series of points.

The mole as a counting unit

1. Mr Clayton is right in pointing out that several different concepts of the mole are in use. These date from the pre-1971 era. It is abundantly clear, however, that the new international standard (3) lays down a 'counting unit' definition of the mole. This new mole is a different concept from the old 'gram-molecule' (4). Realization of this seems to be rather slow in the United Kingdom. Some people have changed the name but not their concepts.
2. A distinction must now be made between 'numbering' and 'counting', at least in scientific contexts. Numbering is going 'one, two, three', counting is going 'one thing, two things, three things'. Counting is an experimental physical process; numbering is something more abstract. The distinction between counting and numbering is deep-seated in the English language, although not total. There are many words (nouns) that are used only in connection with counting, for example: brace, six-pack, dozen, score, gross. They are different in usage from the corresponding numbers, and we may call them 'counting units'.
3. The algebraic formulae by which we write down the rules of engineering and science are based on the concept of 'physical quantity', for example—length. The symbol, L, representing the length of a pin, stands for 1 in *or* (1/12) ft *or* 2·54 cm, and *not* for any of the numbers 1, (1/12) or 2·54. This is specified by the International Standard that defines the SI system, and also by such bodies as the Royal Society Symbols Committee. In general, we write the value of a physical quantity in the form:

$$\text{value} = \text{numerical value} \times \text{unit}$$

4. The base-quantities of the international system of measurement correspond to basic experimental processes, such as 'distance measuring', or 'duration measuring'.
5. Counting things is also a basic experimental process. What happened in 1971 was that the CGPM decided, in effect, to put the process of counting onto the same formal symbolic and mathematical basis as the other basic physical processes. That is, to create a system whereby we write:

$$\text{value of a 'count'} = \text{numerical value} \times \text{counting unit}$$

In Professor Walshaw's terminology (5), counting has now been formalized in terms of a 'quantity algebra'. The new mole is a large counting unit, useful for counting objects that are very small by human standards.

A historical perspective

Why has this change been introduced? I think that it is just an historically inevitable consequence of the discovery of the atomic nature of matter. For thousands of years men have known that measuring length and time is important in what we now call science, and hence the system formalizes their reckoning. But it is only within the last hundred years that we have become *convinced* that in understanding chemistry and molecular physics it is essential to have a count of the fundamental entities involved, and a knowledge of their nature. And it is only now that this knowledge is beginning to be incorporated into the system of scientific measurement itself. (Commerce, of course, has used counting units for many, many centuries.)

Unnecessary linguistic confusion

Strictly, in the above equation I should have written: value of an amount-of-substance. But my experience is that this term is a major source of confusion. When the new base-quantity whose base-unit is the mole was introduced, it would have been far better, I believe, to have introduced a new scientific name for it. 'Entity-count' would be one possibility. But my preference is to invent the word 'ont-count' (literally: count of being things), which is derived from the Greek word *ov* that has been widely used in forming the names of elementary entities. One would then be able to make statements like: 'An amount of substance can be expressed either as an ont-count or as a mass'. As things currently stand, the words 'amount of substance' have both a general everyday meaning and a new (post-1971) technical meaning; it is lack of awareness of this new, confusing, multiple scientific usage that causes many misunderstandings. I have discussed this in more general terms elsewhere (6).

The mole but no mass units

Professor Walshaw (2) asks for an example in which the use of the mole does not involve a link with a unit of mass. Consider my last year's examination question: 'Estimate what amount of gas hits unit area of your finger per unit time'. In the post-1971 system this quantity (here denoted by J) is given by:

$$J = \tfrac{1}{4}C\bar{u}$$

where $\bar{u}$ is the mean speed of a gas molecule, and C is the gas concentration ($\sim 44{\cdot}6$ mol/m^3 for gases under standard conditions). There are no mass units visible here.

Obviously, the size of the base-unit of amount-of-substance is, for general convenience, linked in with the size of the kilogram. But this does not render the concept of amount-of-substance (ont-count) redundant. Similarly, the use of the 'light-year' in astronomy does not render the concept of length redundant there.

Some work of mine (7), albeit not in the context of mechanical engineering, provides an example of a situation where greater clarity of discussion is obtainable in the post-1971 system; and also demonstrates that a scientific discussion can use the concept of amount-of-substance without implying any link whatsoever with the concept of mass.

The need for a long-term view

Finally, I wish to suggest that the real issue is one of *long-term* convenience and clarity. By international agreement, and with this aim in mind, there has been a change in the language and in the style of thinking used to discuss chemistry and molecular physics. The full nature of the change is, as yet, poorly publicised, inadequately understood, and incompletely worked through. There will inevitably be a period of some years during which parallel systems operate and confusion occurs between them. However, it took many years for electrical rationalization to become accepted, yet now the whole business of electrostatic and electromagnetic unit systems looks markedly archaic. In like manner, I believe, will the old systems of gram-molecules, pound-moles, etc, gradually fade into obscurity, because the new mole-based system (once understood) is conceptually simpler.

R G Forbes
The University of Aston in Birmingham
9 July 1979

References

(1) CLAYTON, D G. The unwanted SI mole. *IJMEE*, 1978, 6 No 3, 153
(2) WALSHAW, A C. The mole—a base unit of the SI? *IJMEE*, 1973, 1 No 1, 77-81
(3) International Standard ISO 31/viii-1973 (E) *Quantities and units of physical chemistry and molecular physics*
(4) FORBES, R G. The mole—interpreting it. *Educ Chem*, 1977, 14, 124
(5) WALSHAW, A C. Different algebras and conventions in predicting magnitudes. *IJMEE*, 1978, 6 No 4, 181-185
(6) FORBES, R G. Amount of Substance: an alternative proposal. *Physics Educ*, 1978, 13, 269-272
(7) FORBES, R G. Some new ideas in field-ion theory. *Proc 7th Intern Vac Congr and Third Intern Conf Solid Surfaces*, Vienna, 1977, pp 387-390

International Journal of Mechanical Engineering Education, 8, 99–100 (1980)

Letters to the Editor

IN DEFENCE OF SEVEN DIMENSIONS

Dear Sir,

Dr Forbes' attempt (1) to justify the 1971 mole as an *independent* base unit of the SI and the consequent demise of the lb-mol and g-mol in favour of a 'numerical' mole based on Avogadro's number does not succeed completely because the definition of the 1971 mole is based on a particular mass—0·012 kg of 1 carbon 12; and I doubt if this definition will be modified to include his new word 'ont-count' or even 'entity-count' since the latter is implied in the 1971 definition (2). He quotes a gas concentration of 44·6 mol/m^3 followed by 'There are no mass units visible here' (1). Quite so! but 'not visible here' does not mean that they are not involved—i.e. 'hiding' within the mole in a similar way that a unit of length is 'not visible' in circular measurement of plane angle, but a unit of length is, nevertheless, always involved in the radian.

Dr Forbes does not include a reference (inadvertently omitted, no doubt) to my recent letter (3) which gives the links between mass m, moles n and number B of entities, and poses queries on independence. The usefulness of the mole is not in question but merely its 'masquerading as a *base* unit of the SI' (3). However, even this cannot now be proceeded with, since I am informed by the BIPM that 'The base units of the SI system are simply those units that have been chosen, arbitrarily, to form the most convenient system. It is not necessary that the base units, so chosen, be independent of one another. The important requisites are that they form a coherent system and most conveniently serve the needs of those who use them'.

So, now we know!—base units are not all that some of us thought they were intended to be—i.e. they are not so 'fundamental' as can be defined without recourse to any other unit. This, I suppose, complies with a 1948 resolution (2) of the CGPM re: 'the establishment of a practical system of units of measurement'. Hence, the mole as a base unit!

Finis.

Professor A C Walshaw
The University of Aston in Birmingham
4 December 1979

References

(1) FORBES, R G. In defence of seven dimensions. *IJMEE*, 1979, 7, No 4, 203
(2) HMSO. The International system of units, 1977, pp 1 and 40
(3) WALSHAW, A C. The unwanted mole. *IJMEE*, 1979, 7, No 2, 107

Dr Forbes replies

In the construction of a system of measurement, decisions have to be taken at three levels.

First, it has to be decided what properties/processes in the world are to be regarded as basic and independent. There is some degree of arbitrariness about this: for example, one could in principle take rate of motion, rather than distance, as basic. However, it is surely legitimate to regard counting as a basic scientific process, particularly now that techniques exist (based on field-ion emission) whereby individual atoms can be seen. Surely, also, counting is a process that is conceptually distinguishable from the measurement of inertia.

Second, it has to be decided how each physical property is to be expressed as a physical quantity. For each basic property/process there should be a single corresponding base-quantity. Before 1971, in the context of chemistry and gas physics, the process of counting was expressed in terms of the dimensionless quantity 'number of gram-molecules'. The 1971 change was to replace this by a physical quantity 'amount-of-substance' (my 'entity-count' or 'ont-count') defined as having its own dimension and units. This new physical quantity is conceptually independent of the physical quantity 'mass', and this independence is not affected by the way in which its SI base unit is currently defined, nor by the existence of an experimental relationship between mass and entity-count for any given material.

Third, for each base quantity a corresponding base unit has to be chosen and defined. *A priori*, many alternatives exist. The SI base unit the mole is currently defined by reference to a mass unit; logically, it could be defined by reference to a charge unit; in twenty years time I think its definition will have the form:

> 'The mole is the amount of substance (entity-count) of a system that contains 6·022 045 × 10^{23} elementary entities. When the mole is used the nature of the elementary entity must be specified.'

For an account of the experiments, already in progress, that have this objective in mind, see the article by Deslattes *et al* (4). The mole as just defined would be conceptually pure.

With respect to Professor Walshaw's comments on my earlier letter (1) and his previous letter (3)—reference to which I indeed omitted inadvertently, may I make four points:

1. We are united about the convenience of the mole in many contexts, and on the point that it should be taught and used correctly; we differ only about what its status should be. This difference comes from divergences in opinion as to the status of the base units in a system of measurement.
2. The logical construction of a system of measurement is a more subtle and complicated procedure than the definition of its base units. (British textbooks and official publications almost never discuss all the issues involved. The International Standard (5) is somewhat clearer.)
3. In my view, the essential point is that the process of counting is conceptually independent and basic, and the quantity 'amount-of-substance' (my 'ont-count') is independent and basic; hence I see no difficulty in justifying the mole as an SI base unit.

4. The ampere, like the mole, is currently defined by making reference to other base units. If the mole is masquerading then so—logically—is the ampere. It seems better to believe that neither is.

Finis, peradventure, to this dialogue; but a large task of education remains, I feel.

R G Forbes
The University of Aston in Birmingham
17 December 1979

References

(4) DESLATTES, R D, *et al*, Determination of the Avogadro constant, *Phys Rev Lettrs*, 1974, 33, 463–466
(5) International Standard ISO 31/0–1974(E), General introduction to ISO 31: general principles concerning quantities, units and symbols

EDUCATION IN CHEMISTRY - Letter to the Editor

Educ. Chem., 1982, 19, 102

The seventh SI quantity

The post-1971 changes to the international system of measurement, and their linguistic implications, continue to trouble me and others.[1]

R. E. Lee sees a problem in the statement: 'The mole is the unit of the quantity "amount of substance".' He comments: 'It is not, it is a number of particles'. In everyday language I would agree with him; but in official scientific language he is wrong: the official international and national bodies have *defined* 'amount of substance' to be the English proper name of the quantity for which the mole is the SI unit. I see much sense in persuading them to introduce a second, optional, more obvious in its meaning, name for this physical quantity. I see a partial analogy in electricity, where the Royal Society Symbols Committee [2] lists both 'quantity of electricity' and 'electric charge' as names for the physical quantity whose SI unit is the coulomb.

The mole (post-1971) is manifestly a unit associated with specifying how much is present, in terms of a count of the elementary entities in question. But *linguistically* it is *not* a number. Consider the sentence: 'There were four atoms present'. Now replace the word 'four' by the word 'mole', and then by the words 'dozen' and 'pair'. In each case, the result is not linguistically correct English. Hence 'mole' and 'dozen' and 'pair' are not numbers in the sense that 'four' is. I call them 'counting units', and believe the distinction is not only useful but crucial: in the quantity calculus of the revised SI system, 'numbers' are dimensionless, but 'counting units' are defined to have a dimension of their own, symbolised by [N].

The other basic question with the new post-1971 'mole' is whether it is the name of an *abstract* thing, a pure unit of measurement, or the name of a *concrete* thing, the pile of powder or whatever that contains the Avogadro number of basic entities, or both. I think this is really what Messrs Lee and McManus were discussing. Before 1971 'mole' probably was both; but in my view the formal adoption of this word as the name of an SI unit has de-legitimised its use as a concrete name. (My analogy is: the sticks of wood that children use to measure length in laboratories are not called 'metres', but 'metre-rules'.) The need for a consensual name for the pile of powder or whatever also exists: perhaps we should invent one: for example, 'mole-measure' or 'normal measure'.

(The more radical, but not recommended, alternative would be to let 'mole' revert to being the concrete name, and think up a new name – perhaps 'milliol' – for the present SI unit.)

On the question of a second name for the physical quantity, my requirements are: it should be a word manifestly associated with the process of counting; it should have not more than three syllables; it should not cause confusion by having other uses or associations in science; it should not offend basic rules of the international system of measurement. I thus reject most existing suggestions, including: moleage, quantity, amount, population, count, and entity-count. My personal preference remains the invented word 'ontcount' (meaning 'count of being things') – with my best normal-English suggestion being 'content' (perhaps 'chemical content' in school chemistry).

Can I urge some official action before not too long, if only a commissioned study of the deep and involved *linguistic* problems that seem to exist.

Richard G. Forbes,
Department of Physics,
University of Aston,
Gosta Green,
Birmingham B4 7ET.

References

1. D. G. Clayton, *Educ. Chem.*, 1981, **18**, 164; R. E. Lee, *ibid*, 1982, **19**, 6; F. R. McManus, *ibid*, 1982, **19**, 6; W. Dierks, *Eur. J. Sci. Educ.*, 1981, **3**, 145.
2. *Quantities, units and symbols*. London: Royal Society, 1975.

III. CHARGED-SURFACE PROPERTIES

IIIA THE COEFFICIENT OF THE F^2 ENERGY TERM

16 Atomic polarisability values in the SI system
R.G. Forbes
Surface Sci., 64 (1977) 367-71

17 The influence of depolarisation on surface-atom polarisation energy
R.G. Forbes
Surface Sci., 82 (1979) L620-4

18 Derivation of surface-atom polarisability from field-ion energy deficits
R.G. Forbes
Appl. Phys. Lettrs., 36 (1980) 739-40

19 A fresh look at the electric-field dependence of surface-atom binding energy
R.G. Forbes and K. Chibane
Surface Sci., 122 (1982) (in press)

IIIB WORK-FUNCTION AND RELATED TOPICS

20 Negative work-function correction at a positively-charged surface
R.G. Forbes
J. Phys. D: Appl. Phys., 11 (1978) L161-4

21 Re-exploration of the concepts of work-function and ionization energy
R.G. Forbes
Proc. 7th Intern. Vac. Congr. & 3rd Intern. Conf. Solid Surfaces (Vienna 1977) pp 433-6

Surface Science 64 (1977) 367–371

ATOMIC POLARISABILITY VALUES IN THE SI SYSTEM

Received 23 September 1976; manuscript received in final form 11 January 1977

The object of this note is to present a tabulation of values of the SI quantity "free-space atomic polarisability", for selected chemical species of particular interest to field-ion theory [1–4]. This quantity is of general interest to discussions of atomic and molecular phenomena occurring in high electric fields, but values are not tabulated in the more readily accessible data handbooks now published in SI units. A subsidiary aim is to clarify and simplify the procedure by which polarisation energies are calculated in current field-ion literature.

Principles. In the cgs esu and Gaussian unit systems the property of polarisability is represented by a quantity α_s now called *Gaussian polarisability* [5–7], defined by:

$$p_s = \alpha_s E_s^{loc}, \tag{1}$$

where p_s is Gaussian electric dipole moment and E_s^{loc} is local Gaussian electric field strength. From ref. [6], p_s and E_s^{loc} are related to the corresponding SI quantities p and E^{loc} by: $p_s = (4\pi\epsilon_0)^{-1/2} p$; $E_s^{loc} = (4\pi\epsilon_0)^{1/2} E^{loc}$: hence the defining relationship for polarisability in the SI system can be written:

$$P = \alpha' E^{loc} = \alpha_r \epsilon_0 E^{loc} = 4\pi\alpha_s \epsilon_0 E^{loc}, \tag{2}$$

where ϵ_0 is the electric constant, and α' and α_r are quantities defined by this equation.

By international convention [7], in the SI system the quantity α' ($= 4\pi\epsilon_0\alpha_s$) is called "polarisability" (strictly – electric polarisability of a molecule). The quantity α_r (which could be called the relative, or Lorentzian, polarisability) is not internationally recognised, although it has been used in some widely-distributed current textbooks [8].

All three of the quantities α', α_r, α_s are normally denoted by the simple symbol α, and are normally called simply "polarisability", in the contexts in which they are used. To obviate possible confusion, it is suggested that in future publications the prime might be retained on the SI quantity, at least for the time being, and that when a distinctive name is needed α' should be called "SI (absolute) polarisability".

Units of SI polarisability. The official SI unit for α' is J V^{-2} m^2, as may be seen from the expression for the polarisation energy U_{pol} (unit: J) for an atom in an

electric field E^{loc} (unit: V m^{-1}):

$$U_{pol} = -1/2\ \alpha'\ (E^{loc})^2. \quad (3)$$

However, as can be seen from the tables below, the SI unit is inconveniently large. A much more convenient unit is the meV V^{-2} nm^2 (1 meV V^{-2} nm^2 $\approx$ 1.602 189 $\times$ 10^{-40} J V^{-2} m^2). With α' in this unit it is particularly simple to calculate U_{pol} in eV when E^{loc} is expressed in V/nm: in field-ion theory this calculation is the commonest use of atomic polarisability. Conversion of α_s expressed in Å^3 to α' expressed in meV V^{-2} nm^2, via eq. (7), is also simple, and the numerical values so obtained are of a convenient size. In the tables α' is thus given in both forms of unit.

It is suggested that α' might normally be cited in meV V^{-2} nm^2, but that the conversion factor should be recorded once in each publication.

Data. Values of Gaussian and SI free-space atomic/molecular polarisabilities are shown in the tables. Table 1 gives values for some particularly relevant gas species. With the exception of Xe, the data are in fact derived from experimental measurements [9–11], at room temperature, of the First Dielectric Virial Coefficient A_ϵ. α_s and α' are related to this by:

$$A_\epsilon = (4\pi/3)\, N_A\, \alpha_s = (N_A/3\epsilon_0)\, \alpha', \quad (4)$$

where N_A is the Avogrado constant (6.022 045 $\times$ 10^{23} mol^{-1}). The value of α_s for Xe is taken from ref. [13].

With the exception of O_2 and Xe, the values are based on what Sutter, in his review article [12], believes are the most accurate values of the dielectric virial coefficients to date: cited errors are all much less than 1%. Errors for O_2 and Xe are not cited in the sources used. In some cases the α_s values in table 1 differ slightly (by at most a few percent) from the values cited in earlier tabulations [13–16].

Note that the cited values in effect refer to gas at the limit of infinite dilution, at room temperature. If temperature affects the distribution characteristics as between different internal electronic or vibrational states, then temperature dependence in the average polarisability is to be expected. Also, effective values when the atom or molecule is close to the surface may be slightly or somewhat different [12]. It is beyond the scope of this note to go into details of such variations. However, the author believes that the tabulated values yield a fair indication of the relative magnitudes of polarisation energies in the operating conditions of a field-ion microscope.

Table 2 gives polarisabilities for some relevant metal species. The values of α_s are extracted from a longer table given by Thorhallson et al. [17], obtained from calculations with SCF functions built up from minimal basis sets, and with optimised orbital exponents. In the few cases where experimental data is available for comparison, there is agreement to within a factor of two or better [17].

With metal species particular care must be exercised in using free-space polarisa-

Table 1
Free-space polarisabilities for selected gas species, with origin of data shown; α_s is Gaussian polarisability; α' is SI (absolute) polarisability ($1\ \text{Å}^3 = 10^{-30}\ \text{m}^3$; $1\ \text{meV V}^{-2}\ \text{nm}^2 \approx 1.602189 \times 10^{-40}\ \text{J V}^{-2}\ \text{m}^2$)

Species	$\alpha_s/(\text{Å}^3)$	$\alpha'/(\text{J V}^{-2}\ \text{m}^2)$	$\alpha'/(\text{meV V}^{-2}\ \text{nm}^2)$
He [10]	0.206	2.29×10^{-41}	0.143
Ne [10]	0.396	4.40×10^{-41}	0.275
Ar [10]	1.64	1.83×10^{-40}	1.14
Kr [10]	2.48	2.76×10^{-40}	1.73
Xe [13]	4.01	4.46×10^{-40}	2.78
H_2 [10]	0.806	8.97×10^{-41}	0.560
N_2 [10]	1.74	1.94×10^{-40}	1.21
O_2 [9]	1.57	1.75×10^{-40}	1.09
CO_2 [11]	2.91	3.24×10^{-40}	2.02
CH_4 [9]	2.59	2.89×10^{-40}	1.80

bilities to calculate the binding-energy changes experienced by atoms very close to a highly-charged surface. As has been shown by Tsong and co-workers [18], charge-transfer effects may play a significant role; second-order field-dependences in the binding energy might be associated with variations in the amount of charge transfer rather than with polarisation of atomic orbitals in the adatom.

The numerical expressions used in preparing the tables are:

$$\alpha_s/(\text{Å}^3) = 0.396431\ A_\epsilon/(\text{cm}^3\ \text{mol}^{-1}), \tag{5}$$

$$\alpha'/(\text{J V}^{-2}\ \text{m}^2) = 1.11265 \times 10^{-40}\ \alpha_s/(\text{Å}^3), \tag{6}$$

$$\alpha'/(\text{meV V}^{-2}\ \text{nm}^2) = 0.694456\ \alpha_s/(\text{Å}^3). \tag{7}$$

The numerical factor in eq. (7) is closely equal to 1/1.44; this form is sometimes more convenient, and is more easily memorised.

Comments on current practice. In current field-ion literature the symbol α almost always signifies Gaussian polarisability (though it may be expressed in nm^3); the symbol F, however, denotes a quantity with the dimensions of SI electric field. Thus in present nomenclature the polarisation expression often used would be written:

$$U_{\text{pol}} = -\,1/2\ \alpha_s F^2. \tag{8}$$

Comparison with eqs. (2) and (3) shows that eq. (8) is dimensionally inconsistent. The missing factor is $4\pi\epsilon_0$. This factor can be expressed in the form: $(14.4)^{-1}$ eV $\text{Å}^{-1}\ \text{V}^{-2}$. This is the origin of the "divide by 14.4" rule used in practice to get U_{pol} in eV when F is in V/Å and α_s in Å^3.

Moreover, the quantity F is frequently used as if it were a parameter associated

Table 2
Theoretical free-space polarisabilities for selected metal atom species (nomenclature as in table 1)

Species	$\alpha_s/(\text{Å}^3)$	$\alpha'/(\text{J V}^{-2}\text{ m}^2)$	$\alpha'/(\text{meV V}^{-2}\text{ nm}^2)$
Mo	18.3	2.04×10^{-39}	12.7
Ru	13.9	1.55×10^{-39}	9.65
Rh	12.4	1.38×10^{-39}	8.61
Ag	9.22	1.03×10^{-39}	6.40
Cs	46.2	5.14×10^{-39}	32.1
Ba	49.7	5.53×10^{-39}	34.5
Hf	22.7	2.53×10^{-39}	15.8
Ta	19.4	2.16×10^{-39}	13.5
W	16.8	1.86×10^{-39}	11.7
Re	15.0	1.67×10^{-39}	10.4
Os	13.8	1.54×10^{-39}	9.58
Ir	12.5	1.39×10^{-39}	8.68
Pt	11.5	1.47×10^{-39}	7.99
Au	10.7	1.19×10^{-39}	7.43
Hg	9.78	1.09×10^{-39}	6.79

with a binding site as a whole. Position-dependent variations in E^{loc} (and hence U_{pol}) are then represented as a variation with position in an "effective polarisability" (α_s^{eff} here, but usually just α). The name "effective polarisability" and the symbol α are also sometimes applied to the difference between α_s^{eff} for a neutral atom and α_s^{eff} for the appropriate corresponding ion, even when the atom and ion are at slightly different positions.

The author's impression is that these procedures have sometimes been responsible for confusion concerning polarisability and polarisation energies. I emphasize that E^{loc}, as used here, denotes the local or effective field as normally used in textbook discussions of electricity and magnetism [19].

Richard G. FORBES

Department of Physics, University of Aston,
Gosta Green, Birmingham B4 7ET, England

References

[1] E.W. Müller and T.T. Tsong, Field Ion Microscopy, Principles and Applications (Elsevier, New York, 1969).
[2] R.G. Forbes, J. Microscopy 96 (1972) 63.
[3] T.T. Tsong, J. Chem. Phys. 54 (1971) 4205.
[4] R.G. Forbes, Surface Sci. 61 (1976) 221.

[5] I thank Professor L.F. Bates for pointing out that this nomenclature can be inferred if refs. [6] and [7] are read in conjunction.
[6] Draft International Standard ISO/DIS 31/V, entitled: Quantities and units of electricity and magnetism.
[7] International Standard ISO 31/VIII-1973(E), entitled: Quantities and units of physical chemistry and molecular physics.
[8] For example, in ref. [19] and in: R.P. Feynman, R.B. Leighton and M. Sand, Lectures on Physics, Vol II (Addison-Wesley, London, 1964).
[9] D.R. Johnston, G.J. Oudemans and R.H. Cole, J. Chem. Phys. 33 (1960) 1310.
[10] R.H. Orcutt and R.H. Cole, J. Chem. Phys. 46 (1967) 696.
[11] T.K. Bose and R.H. Cole, J. Chem. Phys. 52 (1970) 140.
[12] H.E. Sutter, in: Dielectric and Related Molecular Processes, Vol. I (Specialist Periodical Report, Chemical Society, London, 1972).
[13] Landolt-Bornstein Tables (6. Auflage); 1. Band, Atom- und-Molecular-Physik; 1. Teil, Atome und Ionen (Springer, Berlin, 1950) p. 401.
[14] H. Margenau, Rev. Mod. Phys. 11 (1939) 1.
[15] Handbuch der Physik, Vol. XXXVI, Atome II (Springer, Berlin, 1956) p. 192.
[16] M. Born, Atomic Physics, 7th ed. (Blackie, London, 1961).
[17] J. Thorhallsson, C. Fisk and S. Fraga, J. Chem. Phys. 49 (1968) 1987.
[18] For example, T.T. Tsong and G. Kellogg, Phys. Rev. B12 (1975) 1343.
[19] For example, W.K.H. Panofsky and M. Phillips, Classical Electricity and Magnetism (Addison-Wesley, London, 1962).

Surface Science 82 (1979) L620–L624

SURFACE SCIENCE LETTERS

THE INFLUENCE OF DEPOLARISATION ON SURFACE-ATOM POLARISATION ENERGY

Richard G. FORBES
Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET, UK

Received 5 June 1978; manuscript received in final form 22 November 1978

The total binding energy of a metal atom to a charged surface contains a contribution resulting from polarisation of the metal atom orbitals. The magnitude of this contribution is of interest in field-ion theory, particularly in the context of field evaporation appearance energies and surface diffusion in high electric fields. This note points out that, for a given external field, if one considers how the magnitude $|U^{\rm pol}|$ of the contribution depends on the proper polarisability of the surface metal atoms, then there may be a theoretical maximum value for $|U^{\rm pol}|$. The maximum value is *independent* of the proper polarisability of the surface atom, but will be a function of its crystallographic environment.

This effect is a consequence of the self-depolarising behaviour of a planar (or nearly planar) array of polarisable atoms. For a given local field acting on a particular atom, the polarisation energy is directly proportional to atomic polarisability. However, the greater the polarisability of the other atoms in the array, the greater will be the depolarising field acting on the original atom, and hence the lower will be the (total) local field acting on it. Atomic polarisability thus affects polarisation energy both directly and indirectly. In consequence, a theoretical polarisation-energy maximum may exist.

This effect can be illustrated by analysing the behaviour of a simplified surface model. This represents a charged surface by an infinite square array of superimposed monopoles and dipoles in the plane of the surface, together with a corresponding distant array of monopoles of the opposite sign. This model is a development of models previously used in the context of field-ion theory [1–3], and will be discussed in more detail elsewhere.

Because considerable linguistic confusion exists [4] in the literature, some definitions will be helpful. If $F^{\rm loc}$ is the *local* SI electric field at the nucleus of a surface atom, then the polarisation energy ($U^{\rm pol}$) of the atom is given by:

$$U^{\rm pol} = -\tfrac{1}{2} b' (F^{\rm loc})^2 ; \qquad (1)$$

b' is the quantity called "electric polarisability" in the current International Standard [5]; for clarity here I call b' the *proper* SI polarisability of the surface atom.

The local field F^{loc} is the total field that would exist at the atomic nucleus, in the absence of the charge distribution associated with the polarised atom in question, but in the presence of any charge rearrangements induced in neighbouring atoms when the original atom is present.

Two alternative expressions for polarisation energy are:

$$U^{pol} = -\tfrac{1}{2}c'(F^{ext})^2 = -\tfrac{1}{2}(4\pi\epsilon_0)c_s(F^{ext})^2. \tag{2}$$

In the context of an array model, F^{ext} represents the SI electric field at a point well away from the array. F^{ext} can also be identified with the "applied field" as used in field-ion literature. In field-ion literature the coefficient c_s has often been called "surface-atom polarisability", or similar names, although it is not a polarisability (*or* a Gaussian polarisability) in the textbook sense. Perhaps a better name would be "Gaussian polarisation-energy coefficient".

Since cited applied-field values are in effect values of F^{ext}, the basic problem is to determine an expression for c' (or c_s). An attempt was made by Müller [6], but the validity of his approach is highly questionable. The present approach assumes merely that the quantity b' is fairly well defined, and evaluates the coefficient β defined by:

$$F^{loc} = \beta F^{ext}. \tag{3}$$

The polarisation-energy contribution can then be determined from the equation:

$$U^{pol} = -\tfrac{1}{2}b'\beta^2(F^{ext})^2 \tag{4}$$

I shall work with the international (SI) system of electrical quantities and equations [5], rather than the mixed, dimensionally-inconsistent system often used in field-ion literature in this context.

Consider the infinite square array model. If the field acting on each atom is F^{loc}, then the total dipole moment p' of each atom is given by:

$$p' = \mu' + b'F^{loc}, \tag{5}$$

where μ' is the zero-field SI dipole moment of the atom, the convention being that μ' and p' are positive if the positive end of the dipole is directed outwards from the surface (i.e. towards the distant array). When the electrostatic contribution [7] to the work-function is positive, it is necessary to regard μ' as negative for the surface atoms.

The field F^{loc} contains two main contributions, due to the monopoles (F^m) and to the dipoles (F^d), thus:

$$F^{loc} = F^m + F^d. \tag{6}$$

It is well-established from classical electrostatics [8] that at an atom in the array:

$$F^m = \tfrac{1}{2}F^{ext}. \tag{7}$$

At some atom, considered to be at the centre of the array, the field F^d is obtained

by summing over the fields due to all the other dipoles, thus:

$$F^{\mathrm{d}} = \sum_{ij}{}' [-p'/\{4\pi\epsilon_0 (i^2a^2 + j^2a^2)^{3/2}\}] = -K_1 p'/4\pi\epsilon_0 a^3 , \tag{8}$$

where a is the lattice spacing in the array; i and j are integers labelling the dipole positions; and the prime indicates that the summation is over all pairs of values except (0,0). K_1 is a "structure factor" [3] defined by:

$$K_1 = \sum_{ij}{}' (i^2 + j^2)^{-3/2} . \tag{9}$$

Combining eqs. (3) to (8) leads, after some rearrangement, to the formula:

$$\beta = \tfrac{1}{2}(1 - 2\mu' K_1/4\pi\epsilon_0 a^3 F^{\mathrm{ext}})/(1 + K_1 b'/4\pi\epsilon_0 a^3). \tag{10}$$

This result can be regarded as the correct form of an erroneous result given in ref. [3]. Note, in particular, the additional factor of $\frac{1}{2}$ present here.

Eq. (10) has the form:

$$\beta = n/(1 + mb'), \tag{11}$$

where n and m represent the relevant expressions in eq. (10). It follows from eq. (4) that $|U^{\mathrm{pol}}|$ is given by an expression of the form:

$$|U^{\mathrm{pol}}| = \tfrac{1}{2} n^2 \{b'/(1 + mb')^2\}(F^{\mathrm{ext}})^2 . \tag{12}$$

Considered as a function of b', the expression in curly brackets clearly passes through a maximum. It is easily shown that the condition for the maximum is:

$$b' = 1/m = 4\pi\epsilon_0 a^3/K_1 . \tag{13}$$

Whence it follows that:

$$|U^{\mathrm{pol}}|_{\max} = (4\pi\epsilon_0 a^3/32K_1)(F^{\mathrm{ext}} - 2K_1\mu'/4\pi\epsilon_0 a^3)^2 . \tag{14}$$

The second bracket is this equation can be simplified by writing μ' in terms of the corresponding electrostatic work-function contribution ψ. For a square array:

$$\psi = -e\mu'/a^2\epsilon_0 , \tag{15}$$

where e is the elementary (proton) charge. Hence we obtain:

$$|U^{\mathrm{pol}}|_{\max} = (4\pi\epsilon_0 a^3/32K_1)(F^{\mathrm{ext}} + K_1\psi/2\pi ae)^2 . \tag{16}$$

Typically, $|\psi|$ is of order a few tenths of an eV. Taking $|\psi|$ as 0.5 eV, and a as 0.5 nm, and using Topping's [9] value for K_1, namely 9.034, leads to an estimate for the size of the zero-external-field term:

$$|K_1\psi/2\pi ae| \sim 1.5 \text{ V/nm}. \tag{17}$$

At the field strengths relevant to field-ion theory this term can be disregarded, certainly in the context of the present planar-array model.

To form an estimate of the likely magnitude of $|U^{pol}|_{max}$, let us set a equal to the interatom spacing on the tungsten (111) plane (namely 0.4476 nm), and F^{ext} equal to the conventional evaporation-field value for tungsten (57 V/nm). This leads to the estimate:

$$|U^{pol}|_{max} \sim 0.7 \text{ eV} \quad (\text{for } F^{ext} = 57 \text{ V/nm}). \tag{18}$$

This would be increased to 1.2 eV if one assumed for F^{ext} the value 74.5 V/nm cited by Müller and Tsong [10] for field evaporation from the (111) plane.

For comparison, Müller's theoretical method [6] leads to a polarisation energy of 2.7 eV for atoms in the (111) plane at 74.5 V/nm [10]; a field-dependent energy contribution of over 10 eV would be deduced at this field if one assumed the c_s value derived by Tsong and Kellogg [11], from analysis of the diffusion of tungsten atoms on the tungsten (110) plane (at somewhat lower field strengths).

From eq. (13), the b' value for which the maximum occurs is 6.9 meV $\cdot$ (V/nm)$^{-2}$. (1 meV $\cdot$ (V/nm)$^{-2}$ $\approx$ 1.602 189 $\times$ 10^{-40} J V^{-2} m^{-2}.) For comparison, the theoretical estimate of b' for tungsten is 11.7 meV $\cdot$ (V/nm)$^{-2}$ [12,4].

Discussion. The important result from this analysis is not so much the numerical values derived above, but the fact that a theoretical maximum does exist – and may also exist for models employing more realistic surface geometries. A major problem in estimating surface-atom polarisation energies has always been the unreliability of both theoretical and experimental estimates of coefficients such as b' and c_s. (The unreliability of the experimental estimates comes from conceptual inadequacy in the models used to analyse the data, rather than defects in the data.) The result here suggests that a limit can perhaps be placed on the polarisation-energy contribution without having values for b' and c_s.

Knowing such a limit would, for example, be very useful in connection with the atomic-jug formalism in field evaporation theory [13]. It would enable a limit to be put to the size of one of the correction factors (the second bracket in eq. (40) of ref. [13]) that could not be satisfactorily estimated. Having a limit for the polarisation-energy contribution would also help in resolving the discrepancy between measured and predicted values of field-evaporation appearance energies [14,15]. If the predicted maximum were genuinely as low as the above figures suggest, then the polarisation-energy contribution could be dismissed as of little significance in either of the above situations.

However, in reality, an atom about to field evaporate is in a kink-site, and a diffusing atom is always above the plane of the array, so the simplified model used above is of limited relevance. In reality, for a given value of F^{ext}, one might expect $|F^m|$ to be greater and $|F^d|$ to be less, with the result that β and F^{loc} would be greater than in the simple model used here (all these at the position of the atom in question). Hence $|U^{pol}|$ would be greater.

Because real situations do not have the symmetry and uniformity of the simple array model, derivation of analytical expressions for β in real situations is far from straightforward. Nevertheless, it seems that theoretical maxima for $|U^{pol}|$ (for a

given F^{ext}) shoult exist, certainly for some real geometries. Attempts to determine maximum values for more realistic model situations would be of considerable interest.

References

[1] R.G. Forbes, in: 21st Intern. Field Emission Symposium, Marseilles, 1974.
[2] M.K. Wafi, M.Sc. thesis, University of Aston (1977).
[3] T.T. Tsong, Surface Sci. 70 (1978) 211.
[4] R.G. Forbes, Surface Sci. 64 (1977) 367.
[5] International Standard ISO 31/VIII-1973(E), entitled: Quantities and Units of Physical Chemistry and Molecular Physics; International Standard ISO/DIS 31/V, entitled: Quantities and Units of Electricity and Magnetism.
[6] E.W. Müller, Surface Sci. 2 (1964) 484.
[7] R. Smoluchowski, Phys. Rev. 60 (1941) 661.
[8] For example, R.P. Feynman, R.B. Leighton and M. Sands, Lectures on Physics, Vol. II (Addison-Wesley, 1964).
[9] J. Topping, Proc. Roy. Soc. (London) A114 (1927) 67.
[10] E.W. Müller and T.T. Tsong, Field Ion Microscopy: Principles and Applications (Elsevier, Amsterdam, 1969).
[11] T.T. Tsong and G. Kellogg, Phys. Rev. B12 (1975) 1343.
[12] J. Thorhallsson, C. Fisk and S. Fraga, J. Chem. Phys. 49 (1968) 1987).
[13] R.G. Forbes, Surface Sci. 70 (1978) 239.
[14] A.R. Waugh and M.J. Southon, J. Phys. D (Appl. Phys.) 9 (1976) 1017.
[15] R.G. Forbes, Surface Sci. 61 (1976) 221.

Derivation of surface-atom polarizability from field-ion energy deficits

Richard G. Forbes
University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET, United Kingdom

(Received 4 December 1979; accepted for publication 4 February 1980)

The field dependence of the difference in standarized energy between helium ions produced from the field-adsorbed position and from the critical surface has recently been interpreted in terms of field penetration into a jellium surface. A model that takes the atomic structure of the real surface into account interprets these results in terms of the proper polarizability of emitter surface atoms. Values obtained for tungsten are 2.01 and 2.07 meV V^{-2} nm^2.

PACS numbers: 79.70. + q, 73.3.0 + y, 73.20. − r

It is now accepted that, if an auxiliary gas of low ionization potential is present, spectral lines can be detected in an atom probe that correspond to the ionization of helium atoms field adsorbed directly onto the emitter substrate.[1–4] Sakurai and co-workers[4] have recently investigated the field dependence of the difference $D(F)$ in standardized energy between these ions and those that originate in the critical surface. For the adsorption of helium on tungsten it is found that this energy difference is a linearly decreasing function of the external field F for all three planes investigated.

Using a model[5] that supposes field ionization to be taking place in a uniform field above a flat jellium surface, into which the field penetrates by a distance λ^{-1}, they show that $D(F)$ is given by

$$D(F) \cong I_1 - \phi^e - (z_{ad} + \lambda^{-1})eF + \Delta_{pol} + U_{i-s}(z_{ad}), \tag{1}$$

where I_1 is the ionization energy of the adsorbed gas; ϕ^e is the relevant local work function of the emitter; z_{ad} is the distance of the bonding site of the field-adsorbed atom from the surface of the jellium; e is the elementary charge; Δ_{pol} is a term expressed as a difference in atom and ion polarizabilities, said to be small[4]; and $U_{i-s}(z_{ad})$ is an ion-surface interaction term treated as independent of field.

For the (111) face, $z_{ad}+\lambda^{-1}$ is empirically determined to be 180 pm (1 pm = 10^{-12} m). Values are slightly less for the (112) and (110) planes, as shown in Table I. Culbertson *et al.*[4] offer no particular explanation of their numerical values or the trend in them. Within the framework of a jellium model, however, if the helium atom radius is taken to be 122 pm (as Tsong and Müller assume[5]), and the helium atom is assumed to be in direct contact with the jellium, then one would conclude that $z_{ad} = r_{He} = 122$ pm, and $\lambda^{-1} = 58$ pm, for the tungsten (111) face, and somewhat less for the (112) and (110) faces. Conventional thinking[5] is that λ^{-1} should be of the order of 100 pm, so the empirical values are on the low side, but not disturbingly low.

There are, however, certain conceptual difficulties in the application of the jellium surface model to field-ion phenomena. The central fact of field-ion microscopy is that individual emitter atoms can be observed. In particular, each atom in a tungsten (111) facet can be seen, if imaging conditions are chosen appropriately. Therefore, to apply a conceptual model in which the atomic structure of the surface is more or less totally disregarded is to create a mismatch between model and reality. Another problem with a jellium model is that there is some uncertainty over precisely where to locate the jellium surface, relative to some definable feature of a real surface—so it may not really be legitimate to take z_{ad} equal to r_{He}.

An alternative is to apply to the analysis of the above results the charged-surface model being developed by the present author.[6–8] This model represents a charged surface by an infinite, regular, flat array of superimposed charges and classical dipoles, together with a distant layer of charge of the opposite polarity, necessary for electrostatic self-consistency. The relevance of these field-induced emitter – surface-atom dipoles was first pointed out by Tsong and Müller.[9] And it now seems to be established[6] that, in order to explain the experimental fact of field adsorption, it is *necessary* to postulate that there is a dipole moment (as well as an excess charge) associated with every surface atom. A model of this type might be expected to apply adequately to the (111) and (112) planes of tungsten, because they are relatively close packed, and because measurements can be performed in the interiors of these facets.

Associated with a polarized layer of atoms there is a work-function correction (negative if the external field is positive) given by

$$\delta\phi^e = -\tfrac{1}{2}beF/A\epsilon_0 M, \tag{2}$$

where F is the *external* field, some distance above the array; b is the *proper* SI polarizability[10] for the atoms, discussed further below; ϵ_0 is the electric constant; A is the area per atom in the layer; and M is a parameter with the role of a relative permittivity for the array, given by

$$M = 1 + Tb/4\pi\epsilon_0 c^3, \tag{3}$$

where c is the lattice parameter of the *space* lattice (c = 316.5

TABLE I. Parameters relating to a tunsten emitter ($b^\Phi = 1$ meV V^{-2} $nm^2 \simeq 1.602\,189 \times 10^{-40}$ J V^{-2} m^2, $b^\Phi \simeq 4\pi\epsilon_0 \times 1.439\,976$ Å^3).

Facet	$z_{ad} + \lambda^{-1}$ (pm)	d (pm)	A/c_w^2	T	b/b^Φ
(111)	180	79	1.7321	3.9011	2.07
(112)	171	88	1.2247	7.5009	2.01
(011)	164	95	—	—	—

pm for tungsten), and T is a "structure factor" specific to a particular lattice structure and crystallographic face, T is defined by

$$T = \sum_m (c/\rho_m)^3, \tag{4}$$

where one atom in the layer is chosen as the "central" atom and the summation is performed over all the other atoms, ρ_m being the distance of the mth atom from the central atom. The value of T is obtained numerically. These equations represent a corrected and generalized version of a formula given earlier.[7]

It can be shown[11] that this work-function correction can alternatively be thought of as a shift outwards in the effective electrical surface, from the plane of the superimposed charges and dipoles, by a distance d given by

$$d = -\delta\phi^e/eF. \tag{5}$$

Consequently, if r_w denotes the radius of a tungsten atom, and the field-adsorbed helium atom is considered to be in direct contact with a surface tungsten atom, application of this model to the results of Culbertson *et al.*, gives

$$d = r_w + r_{He} - (z_{ad} + \lambda^{-1}) = \tfrac{1}{2}b/A\epsilon_0 M. \tag{6}$$

Values of d derived empirically in this way are shown in Table I.

Although the parameter b enters into the definition of M, as well as appearing directly, it is the only unknown in Eq. (6). Hence, using the structure-factor values tabulated, values of b may be derived for the (111) and (112) planes, and are shown in Table I.

It is not possible at present to know whether the difference between the values derived for the two planes is significant. Polarizability values for atoms in different crystallographic environments are not necessarily expected to be equal; but a discrepancy could also be explained by oversimplicity in the model, and/or imprecision in basic assumptions, and/or inaccuracies in the field calibration used by Culbertson *et al.*

However, there is a marked difference between the empirical values and the value 7 meV V^{-2} nm^2 recommended by Miller and Bederson[12] for a tungsten atom in free space, on the basis of the theoretical calculations of Thorhallsson *et al.*[13] As there is no compelling reason to suppose that a partially ionized surface atom would have the same proper SI polarizability as a free atom, the existence of this difference in not surprising, although the empirical values are somewhat smaller than the present author felt the true value would be.

The importance of the present analysis is that it is free from the conceptual objections that can be raised against previous discussions of parameters called "surface-atom polarizability." The parameter b used here is defined by

$$b = p/F^{loc}, \tag{7}$$

where p is the SI dipole moment associated with a surface atom and F^{loc} is the *local* (SI) electric field acting at the nucleus of the atom. (The local field F^{loc} is, by definition, the field that would act at the position of the nucleus of the atom, in the absence of the atom itself, but in the presence of any effects induced by the atom when itself present.) And the theory is formulated self-consistency in terms of this parameter. The parameters called "polarizability" in much of field-ion literature are fomulated in terms of the external field F (which may differ from F^{loc} by a factor of between 2 and about 5) and cannot be identified with the proper polarizability b. Further, the derivation of a "polarizability" from rate-constant sensitivities[14] is mathematically invalid,[15] the derivation from field-adsorption experiments[16] neglects mutual depolarization,[17] and the derivation from diffusion experiments[18] leads to a parameter associated with charge transfer rather than with the polarization of atomic orbitals.

Obviously, there remain some difficulties over the concept of surface-atom polarizability, and the differences reported[4] between the results for tungsten and those for other materials remain to be explained. Nevertheless, I wish to suggest that measurements of the type reported by Sakurai and co-workers may be of greater theoretical importance than they suggest, and that it would be of interest to have such experiments repeated and refined. For theoretical purposes it would also be helpful to have a micrograph showing the precise area from which measurements were taken and the surrounding crystallographic features, and *in situ* field calibrations.

The proper SI polarizability b appears in various aspects of the theory of highly charged surfaces. It has been a continuing problem to know what value to choose for it. In the case of tungsten, the present analysis suggests that 2 meV V^{-2} nm^2 might be an adequate choice for a "provisional working value."

[1] E.W. Müller and S.V. Krishnaswamy, Surf. Sci. 36, 29 (1973).
[2] T. Sakurai and E.W. Müller, Surf. Sci. 49, 497 (1975).
[3] A.R. Waugh and M.J. Southon, J. Phys. 9, 1017 (1976).
[4] R.J. Culbertson, T. Sakurai, and G.H. Robertson, Phys. Rev. B 19, 4427 (1979).
[5] E.W. Müller and T.T. Tsong, Prog. Surf. Sci. 4, 1 (1973).
[6] R.G. Forbes, Surf. Sci. 78, L504 (1978).
[7] R.G. Forbes, J. Phys. D 11, L161 (1978).
[8] R.G. Forbes, Surf. Sci. 82, L620 (1979).
[9] T.T. Tsong and E.W. Müller, Phys. Rev. Lett. 25, 911 (1970).
[10] R.G. Forbes, Surf. Sci. 64, 367 (1977).
[11] R.G. Forbes (unpublished).
[12] T.M. Miller and B. Bederson, in *Advances in Atomic and Molecular Physics* (Academic, London, 1977), Vol. 13, 1.
[13] J. Thorhallsson, C. Fisk, and S. Fraga, J. Chem. Phys. 49, 1987 (1968).
[14] T.T. Tsong, J. Chem. Phys. 54, 4205 (1971).
[15] R.G. Forbes, Surf. Sci. 70, 239 (1978).
[16] T.T. Tsong and E.W. Müller, J. Chem. Phys. 55, 2884 (1971).
[17] R.G. Forbes, Surf. Sci. (to be published).
[18] T.T. Tsong and G. Kellogg, Phys. Rev. B 12, 1343 (1975).

Surface Science 122 (1982) SSC 879
North-Holland Publishing Company

A FRESH LOOK AT THE ELECTRIC-FIELD DEPENDENCE OF SURFACE-ATOM BINDING ENERGY

Richard G. FORBES and K. CHIBANE

Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET, UK

Received 6 may 1982

This paper aims to clarify the theoretical concepts associated with the field dependence of surface-atom binding energy. The various coefficients involved are defined, and orbital-polarization and charge-transfer effects are distinguished. It is demonstrated that the latter, as well as the former, may have an F^2 form. A new empirical method of estimating the polarization-energy coefficient for a field-evaporating surface atom is described: this is based on the joint measurements of activation energy and onset appearance energy made by Ernst. His data for rhodium are reanalysed to give the value 1.05 ± 0.3 meV V^{-2} nm^2. This value is then used to show that polarization-type effects will influence the interpretation of field-sensitivity data. Further experiments are called for.

1. Introduction

The *binding energy* of a surface atom is defined as the work necessary to detach it from the surface as a neutral and remove it to a point in remote field-free space. Applying a high electric field causes a change $\Delta\Lambda$ in surface–atom binding energy, from its zero-field value Λ^0 to the value Λ.

These quantities and the derivative $d\Lambda/dF$ are used in field-evaporation theory [1]. A specific approximation for $\Delta\Lambda$ is sometimes needed, and an F^2 form is normally assumed. Thus we write:

$$\Delta\Lambda = \tfrac{1}{2} c_\alpha F^2, \tag{1}$$

where F denotes the *external field* (i.e. the electric field somewhat above the surface atom in question); c_α is a parameter, *defined by this equation*, that we call the *polarization-energy coefficient* [2] for the surface atom in question. The suffix α is used to indicate that we are referring to the atom in its as-bound electronic state, that has to be partially ionic in character in order to sustain the external electric field [3]. Tsong [4] has suggested the inclusion of a term linear in F in eq. (1), to allow for a field-induced reduction in the interaction between the electrons of the surface atom and the substrate, but this suggestion has subsequently been disregarded.

0039-6028/82/0000–0000/$02.75

The main stimulus for this note comes from a perceived need to have a numerical estimate for $d\Lambda/dF$ at the zero-Q evaporation field F^e; that is, if eq. (1) is valid, we need to know the value of $c_\alpha F^e$. In the new activation -energy formulae for field evaporation put forward in ref. [1], we need this in order to evaluate the likely influence of "polarization-type" effects on field-evaporation flux and rate-constant field-sensitivities. Since the question is one of the relative sizes of $d\Lambda/dF$ (or $c_\alpha F^e$) in comparison with *nea*, where *ne* is the charge on the escaping ion and *a* is the distance of the surface-atom bonding point from the emitter's electrical surface, it seems likely that even a rough estimate of $c_\alpha F^e$ would be quite useful. F^e can already be estimated [5–8], so we now need a value for c_α.

We shall work with the international (four electric dimensions) system of measurement and the units convention proposed in ref. [9].

The structure of the paper is as follows. Section 2 deals with the definitions of coefficients used to describe the field-dependence of binding energy, and previous attempts to estimate them. Section 3 then describes some basic theory associated with a new method of estimating c_α, based on the measurement of activation energies and appearance energies. Section 4 describes the re-analysis of the relevant experimental results. Section 5 contains some theoretical considerations, about the form of $\Delta\Lambda$ and the value of c_α. Finally, section 6 discusses the paper's achievements and draws some general conclusions.

Finally, here, we emphasize that in this paper Λ denotes the total field-dependent binding energy, i.e. the total energy difference between the ground-state vibrational level and the potential reference zero (the situation of a stationary neutral in remote field-free space). Recent literature employs the symbol Λ with four slightly different meanings, and care is necessary to avoid misunderstandings.

2. Definitions and historical background

The coefficient c_α mainly relates, we assume, to two physical effects: partial electron transfer to the metal interior, to create the fractional surface-atom charge; and polarization of the atomic orbitals of the resulting entity. In a first approximation we can perhaps separate the effects and write:

$$c_\alpha = c_\alpha(\mathrm{ct}) + c_\alpha(\mathrm{orb}), \tag{2}$$

where ct and orb label the charge-transfer and orbital-polarization components. Because c_α is defined with respect to the external field (rather than a local field acting on the atom in question), all quantities in eq. (2) will depend on the surface-atom crystallographic environment; they may also be field dependent.

Because the values of polarization-energy coefficients (and their compo-

nents) are environment dependent, it is convenient to use a suffix to label the different types of bonding situation. For the present we distinguish only two types of situation: (1) an atom participating in normal field evaporation from a kink site (the α-situation); (2) an atom diffusing on top of a crystal plane (the δ-situation). Finer distinctions may be needed elsewhere.

If the self-consistent local field F_α^{loc} acting on a surface atom can be defined, then the orbital-polarization component $\Delta\Lambda(\text{orb})$ of the binding-energy increase for this atom can alternatively be written:

$$\Delta\Lambda(\text{orb}) = \tfrac{1}{2} b_\alpha(\text{orb}) \left(F_\alpha^{\text{loc}} \right)^2, \tag{3}$$

where $b_\alpha(\text{orb})$ is the so-called *proper polarizability* of the surface atom. Defining the field ratio β_α by:

$$\beta_\alpha = F_\alpha^{\text{loc}} / F, \tag{4}$$

we have the following relationship:

$$c_\alpha(\text{orb}) = \beta_\alpha^2 b_\alpha(\text{orb}). \tag{5}$$

The proper polarizability $b_\alpha(\text{orb})$ is expected to be much less dependent on crystallographic environment than is the polarization-energy-coefficient component $c_\alpha(\text{orb})$. Obviously, similar relationships apply to atoms in other bonding situations.

Within the framework of a charged-surface model such as that adopted by Forbes and Wafi [10], β_α is roughly about 0.3, partly because the applied field acting on the surface atoms is only half the external field [11], partly because the effective field acting is even less, as a result of surface-layer depolarization effects. It follows that $c_\alpha(\text{orb})$ will be substantially less than $b_\alpha(\text{orb})$.

Past theoretical and experimental investigations of field-dependent effects have been fraught with difficulties and confusions, not the least of which is use of the name "polarizability" for many conceptually-different quantities. On the theoretical side, the semi-empirical formulae of Brandon for $b_\alpha(\text{orb})$ [12] and of Müller for c_α [13] are of uncertain status; and more-recent theoretical discussions [14–16], based on a "broadened level" picture, have not led to numerical results. Estimates of free-space atomic polarizability [17] can be used as estimates of $b_\alpha(\text{orb})$, but their relevance to the situation of a partially-ionized surface atom can legitimately be challenged, as could the accuracy of any depolarization model used to then get the corresponding value of $c_\alpha(\text{orb})$ (particularly in the case of kink-site atoms).

On the experimental side, the two early methods, involving measurements of field evaporation rates [4,18], and of field-adsorption binding energy [19], are invalid because the formulae used to interpret the data are mathematically or physically invalid [10,20–22]. The same flaw exists in a method of analysis recently used [23] to analyse data [24] about the temperature dependence of evaporation field [8]. Difficulties over the correct estimation of depolarization

effects make it problematical to derive an accurate value of c_α(orb) from the experiments of Culbertson et al. [25] and the discussion of Forbes [26]. The method based on the migration of adatoms in the presence of the field, introduced by Tsong and co-workers [14,16,27], would on the face of it seem to give a fair estimate of c_δ for atoms diffusing on the top of a surface plane; but we cannot expect this approach to produce a value directly relevant to field-evaporating kink-site atoms.

Against this very problematical background, and not without certain hesitations, we propose to briefly explore a further experimental method of deriving estimates for c_α. This is based on the measurement of activation energies and onset appearance energies in the actual circumstances of field evaporation, and is implicit in the sophisticated experimental work of Ernst [28].

3. Appearance energy, activation energy and binding energy

The concept of an appearance energy (A) emerges when the First Law of Thermodynamics is formally applied to surface field ionization [29]. A is in the nature of the electrical work done, by the emitter + field system, in order to produce an ion from the corresponding neutral. A has a quasi-thermodynamic nature, and relates only to the surface ionization process, so it is more fundamentally useful than the measured voltage deficit δ for the ion, to which it is related by:

$$A = re(\delta + \phi^r), \tag{6}$$

where e is the elementary charge and re the charge on the ion when it arrives at the retarding electrode of local work-function ϕ^r.

The value of A depends on the details of the ion emission process. The so-called *standard appearance energy* [29] relates to an ion that just escapes over the top of the relevant activation-energy barrier, with no kinetic energy to spare, and which has the electrons resulting from any ionization processes it may undergo transferred directly to the emitter Fermi level. For an ion that was initially bound in a partially ionic state α, escapes into an n-fold-charged state, and arrives at the energy analyser as an r-fold ion, we denote the standard appearance energy by $A^{\text{stand}}_{\alpha nr}$ and the activation energy for escape by $Q_{\alpha nr}$ [30].

The experimentally measured *onset appearance energy* $A^{\text{onset}}_{\alpha nr}$ for ions of this class, however, is less than the standard appearance energy, because many ions in effect cross the activation-energy barrier with finite kinetic energy. This has been clearly established by Block and co-workers [31–33], who in the case of metals now put:

$$A^{\text{onset}}_{\alpha nr} = A^{\text{stand}}_{\alpha nr} - zkT, \tag{7}$$

where k is the Boltzmann constant, T the thermodynamic temperature and z a number of the order of 10. or less. The correction is really due partly to the statistical distribution of the bound-state vibrational energy, partly to electron transfer (during ionization) into temporarily-unoccupied states below the emitter Fermi level [33]. This last effect has long been recognized in the context of surface ionization [34].

The initial-state binding energy Λ is related to the activation energy and the standard appearance energy by [28,29,35]:

$$\Lambda = A_{\alpha nr}^{\text{stand}} + Q_{\alpha nr} - H_r, \tag{8}$$

where H_r is the energy needed to form the r-fold ion in *remote field-free space*, from the neutral, and is given by the sum of the first r free-space ionization energies. Hence we obtain:

$$\Lambda = (A_{\alpha nr}^{\text{onset}} + Q_{\alpha nr}) - H_r + zkT. \tag{9}$$

The relationships implicit in eqs. (8) and (9) are illustrated in fig. 1.

The activation energy that appears in eqs. (8) and (9) is the energy difference marked in fig. 1. In a one-dimensional model this quantity can be identified with the activation energy obtained from an Arrhenius plot [36]. But, it seems to us, in the real situation the "observed" Arrhenius-plot activation-energy could differ from $Q_{\alpha nr}$ by a very small amount, perhaps about $1\,kT$. Any discrepancies of this size will get "hidden" in the uncertainty over the value of z in eq. (7), so we shall neglect them in what follows.

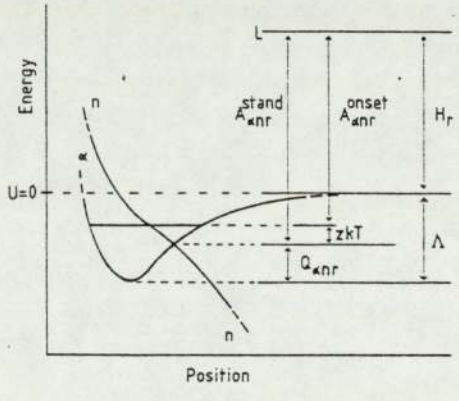


Fig. 1. Schematic diagram illustrating the relationships implicit in eqs. (8) and (9), for an ion that escapes in an n-fold-charged state and arrives at the analyser in an r-fold-charged state. The diagram is constructed by drawing in a level L that is above the energy reference level (at $U=0$) by an amount H_r. (The diagram is not a proof of these relationships, is not to scale, and deliberately omits detail concerning the energy changes in the ionization processes.)

4. Analysis of Ernst's experiments

Ernst [28] was the first to realize that these equations could in practice be used to derive values of binding energies from independent measurements of activation energy and onset appearance energy, but has not carried the argument through to the deduction of a value for c_α. Our objective here is to develop the analysis of his results, as listed in his table 1.

The most useful set are those for rhodium ions arriving in a 2^+ state (i.e. $r = 2$). Ernst established, and Haydock and Kingham helped to confirm [37,38], that these ions are derived by post-ionization of Rh^+ ions (i.e. $n = 1$). The bracketed quantity in our eq. (9) is listed as a function of field strength in table 1 in ref. [28], and is reproduced in column 3 of table 1 here. There is a noticeable field-dependent trend, notwithstanding the instrinsic experimental uncertainties in the individual data points.

Taking the energy H_2 for rhodium as 25.54 eV [15], and temporarily ignoring the zkT term, we show in fig. 2 values of Λ derived from eq. (9), plotted as a function of external field F. Each data point is surrounded by a "box of measurement uncertainty" derived from Ernst's stated error limits. The known zero-field binding energy for rhodium, $\Lambda^0 = 5.75$ eV [15], is also marked.

It is not possible to immediately establish from fig. 2 what the form of the field dependence of Λ is. But it is clear that extrapolating back to zero field, assuming either a simple linear or a simple quadratic dependence on F, would lead to a zero-field binding energy less than the known value [39]. For example, a weighted linear regression gives: $\Lambda^0 = 4.42 \pm 0.12$ eV.

Ernst [39] has tentatively suggested that this may indicate that a linear term of the form $-\theta F$, as proposed by Tsong [4], should be included in eq. (1). A possible field dependence, including both linear and quadratic terms, is shown

Table 1

Data on which fig. 3 is based: the original data are taken from ref. [28], and the values of Λ are derived from eq. (9), using $H_2 = 25.54$ eV; The column headed $3kT$ gives an estimate of the additional error in the corrected ($z = 10.5$) values of Λ due to the correction procedure (this additional error is ignored in the regression, as it is much less than the measurement error)

T (K)	F (V/nm)	$(A_{\alpha 12}^{\text{onset}} + Q_{\alpha 12})$ (eV)	$\Lambda(z=0)$ (eV)	$\Lambda(z=10.5)$ (eV)	$3kT$ (eV)
600	17.0±2.6	30.7±0.7	5.16	5.70	0.16
510	21.0±3.2	31.3±0.5	5.76	6.22	0.13
430	24.0±3.9	31.1±0.4	5.56	5.95	0.11
350	28.0±4.2	31.2±0.4	5.66	5.98	0.09
250	33.0±5.0	31.8±0.3	6.26	6.49	0.06
100	41.0±6.2	32.05±0.22	6.51	6.60	0.03

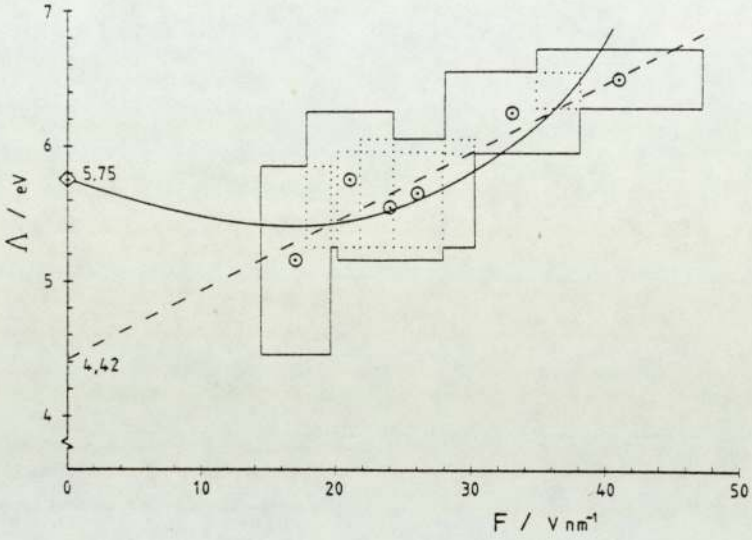


Fig. 2. A plot of "derived binding energy" (from eq. (9), with $z = 0$) against external field. Each point is surrounded by a "box of uncertainty" obtained from the error limits given by Ernst [28] (see table 1). The dashed line represents a weighted linear regression. The known zero-field binding energy for rhodium is marked with a diamond, and the continuous line represents a hypothetical field dependence of Λ involving terms linear and quadratic in F, as discussed in the text.

schematically in fig. 2, but is not convincing. The Tsong proposal, although not unreasonable in itself, was an ad-hoc hypothesis that is without any detailed theoretical justification. For reasons that will become clear later, we prefer to assume a simple F^2 form. The discrepancy then has to be attributed to the zkT correction,term.

Ernst [28] estimated the effect of the zkT term to be less than 0.3 eV, and thus neglected it. Certainly an uniform upwards shift of this amount could not explain the whole of the discrepancy. However, the lower-field data points were in fact taken at higher temperature, and thus a "lever" effect operates. This is illustrated in fig. 3, which shows two sets of Λ-values plotted against F^2.

The original data points are shown as circles. The "corrected" set of data points, marked with squares, were obtained as follows. For a given value of z, we may use the temperature values listed in table 1 of ref. [28] (and reproduced in table 1 here) to produce a modified set of Λ-values via eq. (9). We then carry out a linear regression (linear in F^2) to obtain an "empirical" value of Λ^0. The value of z is then changed, and the process repeated, until the derived value of Λ^0 is adequately close to the known zero-field binding energy. The "corrected" data points in fig. 3 correspond to taking $z = 10.5$; actual values are listed in

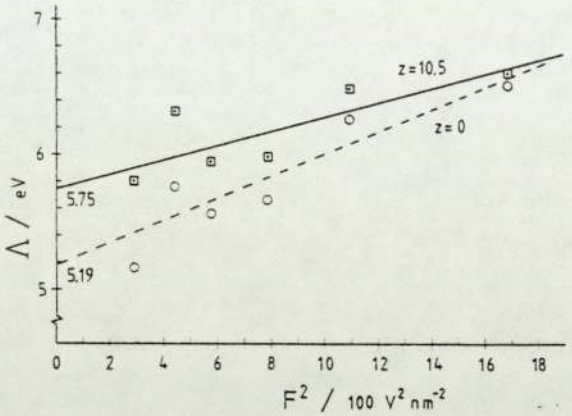


Fig. 3. A plot of binding energy, as derived from eq. (9), against F^2 where F is the external field. The circles and the dashed regression line are obtained by taking $z = 0$. The boxed points and the continuous regression line are obtained by taking $z = 10.5$, which gives an intercept equal to the known zero-field binding energy. Note that the zkT correction is greater for the data points at lower field (= higher temperature, see table 1). Regression results from these plots are given in table 2.

table 1 here, for future reference. The continuous and dashed lines in fig. 3 are the regression fits for $z = 10.5$ and $z = 0$, respectively, and are described by the parameters listed in table 2.

An estimate of the error in z can be obtained as follows. From table 2 we note that the standard deviation of the $z = 10.5$ intercept is $+0.17$ eV. We now repeat the process of varying z and carrying out regressions until a value of z is found for which the intercept is $(\Lambda^0 + 0.17$ eV). This occurs for $z = 13.7$, as shown in table 2. The difference between this and the previous z-value is taken as an estimate of the error in z; so we obtain:

$$z = 10.5 \pm 3. \tag{10}$$

Table 2
Values of the zero-field binding energy Λ^0(der.) and slope $\frac{1}{2}c_\alpha$ derived from weighted regression analysis of fig. 3; the true value of zero-field binding energy is assumed to be 5.75 eV

z	Λ^0(der.) − Λ^0(true) (eV)	Λ^0(der.) (eV)	$\frac{1}{2}c_\alpha$ (meV V^{-2} nm^2)
0	−0.56	5.19±0.18	0.810±0.14
10.5	0.00	5.75±0.17	0.524±0.14
13.7	0.17	5.92±0.17	0.438±0.14

This value of z seems reasonable. In their work on the appearance energies of gas-phase ions, Block and co-workers [33] found a z-value of approximately 10, for a typical signal-to-noise ratio of 10^4. Thermal effects on appearance-energy measurements were also observed in their work on silver [32]. Thus result (10) seems in general accord with the earlier work. We note, however, that the value of zkT we have deduced empirically for 600 K, namely 0.54 eV, is somewhat larger than the maximum likely value estimated by Ernst, namely 0.3 eV. The discrepancy is not large enough for us to be disturbed by it; we are not in a position to establish why Ernst's estimate is apparently slightly low.

For interest, we repeated the procedures described above on the data given by Ernst [28] for singly-charged ions, obtaining $z = 14 \pm 3$. However, the Rh^- data comprise only four points, over a limited field range, and are statistically rather a poor data set for our purpose. So we regard this result as of limited significance. (Possibly, it would be interesting to have experiments performed for a material that produces adequate fluxes of differently-charged ions over a good range in field strength.)

Returning to the estimation of c_α, we note that in fig. 3 there must be an uncertainty in the slope $\frac{1}{2}c_\alpha$ associated with the uncertainty in z. We estimate the total uncertainty in $\frac{1}{2}c_\alpha$ as follows. The difference in the value of the $z = 10.5$ and $z = 13.7$ regressions is 0.087 meV V^{-2}. nm^2. Combining this in quadrature with the standard deviation on the $z = 10.5$ slope, we obtain: $\sigma(\frac{1}{2}c_\alpha) \sim 0.16$ meV V^{-2} nm^2. The final estimate of c_α thus becomes:

$$c_\alpha \sim 1.05 \pm 0.3 \text{ meV V}^{-2} \text{ nm}^2 \quad \text{(experimental estimate)}; \tag{11}$$

the value expressed in SI base units is: $(1.68 \pm 0.5) \times 10^{-40}$ J V^{-2} m^2; the corresponding Gaussian coefficient, $c_\alpha/4\pi\epsilon_0$, has the value: 1.5 ± 0.5 Å^3.

5. Some theoretical considerations

5.1. The form of $\Delta\Lambda$

That an F^2 form is chosen for $\Delta\Lambda$, as in eq. (1), is largely a matter of historical precedent, in that the F^2 form is known to be the correct first approximation for a binding-energy change due to orbital polarization, certainly for a neutral atom in free space. We may reasonably suppose that this is an adequate form for the orbital-polarization contribution in a surface atom, even when partially ionized. But, to the best of our knowledge, it has never been clearly demonstrated that the F^2 form is adequate for the binding-energy contribution $\Delta\Lambda$(ct) associated with the partial ionization of the surface atoms. Thus we shall now present a simple argument that suggests that the F^2 form could be an acceptable first approximation for $\Delta\Lambda$(ct).

As a simple approximation we may suppose that the work necessary to

remove a surface atom, as a neutral, from a charged surface to remote field-free space may be notionally split into three parts:
(1) The work necessary to depolarize the partially-charged surface entity back into a spherically-symmetric state, which is assumed to be $\frac{1}{2}c_\alpha(\mathrm{orb})F^2$.
(2) The work necessary to transfer charge back, from the rest of the emitter, into the atomic orbitals, to make the surface atom neutral; we identify this work as $\Delta\Lambda(\mathrm{ct})$.
(3) The work necessary to remove the neutral atom to remote field-free space. We assume this to be the same as the zero-field binding energy Λ^0 (but recognise that there could be a small field-dependent correction).

A rough estimate of $\Delta\Lambda(\mathrm{ct})$ can be obtained as follows. We assume that the displaced charge (i.e. that charge q necessary to make the surface atom neutral again) has to be moved a distance d, where d is the spacing of the atomic layers normal to the surface. And we assume that the *mean* field in which this charge has to move is some fraction $\bar{\beta}$ of the external field F, with $\bar{\beta}$ a parameter that is independent of F. It follows that:

$$\Delta\Lambda(\mathrm{ct}) = q\bar{\beta}Fd. \tag{12}$$

The charge q on the surface atom is related to the surface charge density σ by: $q = \sigma S$, where S is the surface area per surface atom. And the surface charge density is related to the external field by the Gauss relationship: $\sigma = \epsilon_0 F$, where ϵ_0 is the electric constant (permittivity of free space). Combining these relationships gives:

$$\Delta\Lambda(\mathrm{ct}) = \epsilon_0\bar{\beta}V_{\mathrm{at}}F^2, \tag{13}$$

where V_{at} is the atomic volume for the lattice in question, being given by Sd.

It follows that $\Delta\Lambda(\mathrm{ct})$ has the desired F^2 form, and that an estimate of $c_\alpha(\mathrm{ct})$ may be given (very approximately) by:

$$c_\alpha(\mathrm{ct}) \sim 2\epsilon_0\bar{\beta}V_{\mathrm{at}}. \tag{14}$$

Since $\Delta\Lambda(\mathrm{orb})$ is already assumed to have an F^2 form, it follows that eq. (1) is satisfactory.

5.2. *A rough estimate of* c_α

Eq. (14) enables a rough estimate of $c_\alpha(\mathrm{ct})$ to be made. Whilst this is of little use for predictive purposes, it may be of some help in deciding whether the value of c_α deduced from Ernst's experiments "looks reasonable".

The applied field acting at the surface atom nucleus is, from general electrostatic consideratons, $\frac{1}{2}F$. Surface depolarization effects will decrease this. And the mean field inside the outermost layer is bound to be even less. Let us thus take $\bar{\beta}$ as 0.2. The atomic volume for the rhodium lattice is 0.0138 nm^3. We thus obtain: $c_\alpha(\mathrm{ct}) \sim 0.3$ meV V^{-2} nm^2 [40].

Data are not available to make a realistic estimate of c_α(orb) for rhodium, but the following approach may give some indication. Ref. [26] gives data for the (111) face of tungsten derived from the Culbertson et al. experiments [25]. Using eq. (3) in ref. [26], we may derive a value for the parameter M used there, namely $M = 1.367$. For a clean surface, the field ratio β applicable to a surface atom is given by $\frac{1}{2}M^{-1}$; so for a tungsten atom in the (111) plane we obtain $\beta_* = 0.366$. From ref. [26], the relevant value of b_α(orb) is 2.07 meV V^{-2} nm^2; so from eq. (5) we obtain c_α(orb) = 0.277 meV V^{-2} nm^2. This is a value applicable to tungsten; to obtain a value relevant to rhodium, we ignore the difference in lattice structures [41] and scale the tungsten value by the ratio of the tabulated free-space Gaussian polarizabilities g_S as given by Miller and Bederson [17]. These are: $g_S(Rh) = 7.6$ Å^3; $g_S(W) = 10$ Å^3. Thus we obtain, as a notional value for rhodium, c_α(orb) ~ 0.2 meV V^{-2} nm^2.

Combining this with the estimate of c_α(ct) made above gives:

$$c_\alpha \sim 0.5 \text{ meV } V^{-2} \text{ nm}^2 \quad \text{(rough theoretical estimate)}. \tag{15}$$

This can now be compared with the experimental estimate (11).

6. Discussion

6.1. The value of c_α

The experimental and theoretical estimates of c_α for field-evaporating rhodium atoms agree to within a factor of about 2. Considering the rough and ready nature of the theoretical estimate just made, this agreement is as good as could be expected and is an adequate (albeit weak) demonstration of overall consistency.

To make a more reliable theoretical estimate of c_α would require a rather better theoretical understanding of real charged surfaces than we currently have (or are likely to have for some time). It thus seems that, for the immediately foreseeable future, estimates of c_α are best made by experiments. It would be helpful to have experiments of the type described in ref. [28] extended to a wider range of materials.

Our experimental estimate of c_α for field-evaporating rhodium can also be compared with the value of c_δ for an iridium atom performing a field-induced directional walk on a tungsten (110) plane. Tsong and Kellogg [14] have measured the latter to be 2.28 ± 0.26 meV V^{-2} nm^2 [40]. (We choose iridium for the comparison because, at $g_S = 7$ Å^3, it has free-space Gaussian polarizability [17] closest to that for rhodium, amongst the species used in ref. [14].) This iridium estimate is somewhat higher than the rhodium field-evaporation estimate. This could be a consequence of the diffusing atom being in a more field-exposed situation; or it could be something connected with the chem-

isorption characteristics of iridium on tungsten. Further speculation is hardly justified, but the comparison does bring out a possible value in performing directional self-diffusion experiments for materials other than tungsten.

6.2. *"Polarization-type" effects in the Q-formula*

We now take up the question of whether "polarization-type" effects might significantly influence the prediction of field-sensitivities (or the deduction of information from them). According to ref. [1], for a Gomer-type escape mechanism the activation energy Q_{anr} is given by:

$$Q_{anr} = \tfrac{1}{2}\kappa(a - \Pi^e/ne)^2(1 - \eta_n'/neF)^{-2}(F^e/F - 1)^2, \tag{16}$$

where κ is the force-constant for the surface atom in its as-bound state, Π^e a parameter with the dimensions of SI dipole moment, discussed below, η_n the "purely-chemical" component of the ionic potential energy [42], and η_n' its derivative with respect to distance (evaluated for $F = F^e$ and at the bottom of the bonding well); the other symbols have the same meaning as earlier.

The first bracket contains the correction term due to field-dependent energy changes. If an F^2 form is used both for $\Delta\Lambda$ and for the polarization energy of the escaping ion, then Π^e is given by:

$$\Pi^e = (c_a - c_n)\, F^e \equiv \delta c F^e, \tag{17}$$

where c_n is the polarization-energy coefficient for the n-fold ion, and δc is defined here.

To proceed further we need an estimate of c_1 (since it is known that $n = 1$ for rhodium). An adequate estimate can be obtained from Brandon's discussion [12]. Using his tabulated value of the Rh^+ ion radius, and his fig. 2 relationship between univalent ionic radius and Gaussian polarizability, and note [40], we obtain an estimate of c_1 as ~ 0.15 meV V^{-2} nm^2. Using value (11) for c_a, we thus take δc as ~ 0.9 meV V^{-2} nm^2.

The measured evaporation field for rhodium at 100 K is 41 ± 6 V/nm [28]. For reasons that will be described elsewhere [8,43], we take the zero-Q evaporation field F^e as somewhat higher than this, arbitrarily estimated as 50 V/nm. This gives Π^e/ne as ~ 0.045 nm.

The value of the electrical bonding distance a for the as-bound rhodium atom is expected to be somewhat greater than the neutral-atom radius (0.134 nm in the rhodium lattice), but less than twice this. Thus in eq. (16) the term Π^e/ne is less than a, but is not negligible in comparison with a. It follows that, for rhodium, a significant "polarization correction" is inherent in eq. (1). And, presumably, this may also be the case for other materials.

The implication of this finding is that, if field evaporation experimnts are to be used to obtain accurate estimates of surface atomic parameters such as κ,

then it may be necessary to perform these experiments in such a way that a reasonably accurate value of c_α can be dedued from them, as a preliminary.

6.3. Summary and conclusions

The achievements and conclusions of this paper are as follows:
(1) We have used Ernst's results to make an experimental estimate of the polarization-energy coefficient c_α for a bound, partially-ionized, surface rhodium atom in the circumstances of field evaporation. This estimate is: $c_\alpha = 1.05 \pm 0.3$ meV V^{-2} nm^2.
(2) To do this we have assumed an F^2 form for the field-dependent binding-energy increase $\Delta\Lambda$, and have given a theoretical justification for this choice.
(3) It has also been necessary to make zkT-type corrections to Ernst's appearance energy data. The necessary value for z is 10.5 ± 3. This value seems generally plausible.
(4) A rough theoretical estimate of c_α for rhodium has also been made. This is adequately compatible with the experimental estimate, but it is clear that experiment is currently the better way of getting at values for c_α.
(5) Using the above experimental estimate, it has been demonstrated that "polarization-type" corrections will be significant in the interpretation of field-sensitivity data using Q-formula (16), certainly for rhodium and presumably also for other materials.

The least secure part of this paper is probably the argument concerning the form of $\Delta\Lambda(ct)$, and we would not wish this to be seen as anything other than a classical first attempt. The essential features of the argument are that, in moving electron charge, the work done per unit of charge moved is proportional to the external field, and the amount of charge moved per surface atom is also proportional to the external field; hence an F^2 form. These propositions seem generally plausible. But, clearly, a properly-formulated (electrostatically consistent) quantum-mechanical discussion will eventually be necessary, particularly if a reliable theoretical expression is required for the coefficient c_α.

Summarizing more generally, we note that past literature contains many proposed empirical methods for obtaining the values of coefficients associated with the field dependence of surface-atom binding energy. From work by ourselves and by Wafi [44], not reported in detail here, we conclude that of these proposed methods only three are both valid and useful:
(A) The method described here, based on the combined measurement of activation energies and onset appearance energies, that gives the coefficient c_α for the atoms involved in the process of field evaporation.
(B) The method, introduced by Tsong and coworkers, based on the directed random walk of adatoms in the presence of a field, which gives the coefficient c_δ for the diffusing adatom.
(C) The method, based on re-analysis of anomalous appearance energy data

from the experiments of Sakurai and coworkers [25], introduced by Forbes [26], that gives the coefficient b_α(orb) for the atoms in the surface plane from which the appearance-energy data are taken (provided surface depolarization effects can be modelled correctly [45]).

These three methods seem to be measuring different parameters. So it would be interesting to have all three procedures applied to a selection of materials.

Finally, we hope that in this paper we have demonstrated that past confusions can be put aside, and that – at least at an elementary level – a fairly straightforward account can now be given of the electric-field dependence of surface-atom binding energy.

Acknowledgements

This work forms part of a research project funded by the UK Science and Engineering Research Council. However, the ideas in it were much stimulated from visits to the Fritz-Haber-Institut of the Max-Planck-Gesellschaft, in Berlin, by one of us (R.G.F.). I wish to thank the Max-Planck-Gesellschaft, the German Academic Exchange Service and the British Council for financial support, and to thank Professor J.H. Block, Dr. N. Ernst and their co-workers for stimulating discussions. One of us (K.C.) wishes to thank the Ministry of Higher Education and Scientific Research of the Republic of Algeria for personal financial support.

References

[1] In particular: R.G. Forbes, Surface Sci. 116 (1982) L195.
[2] This coefficient is not in any direct sense a textbook polarizability. We urge that the practice of calling it a polarizability be discontinued, as this seems to contribute to confusion.
[3] c_α has been denoted elsewhere by c_0 and by α_a. The revised notation is thought to be clearer.
[4] T.T. Tsong, J. Chem. Phys. 54 (1971) 4205.
[5] E.W. Müller and T.T. Tsong, Field Ion Microscopy: Principles and Applications (Elsevier, New York, 1969).
[6] R.G. Forbes, Appl. Phys. Letters 40 (1982) 277.
[7] R.G. Forbes, J. Phys. D (Appl. Phys.), 15 (1982) L75.
[8] K. Chibane and R.G. Forbes, Surface Sci, to be published
[9] R.G. Forbes, Surface Sci. 64 (1977) 367.
[10] R.G. Forbes and M.K. Wafi, Surface Sci. 93 (1980) 192.
[11] For example, R.P. Feynman, R.B. Leighton and M. Sands, The Feynman Lectures, Vol. 2 (Addison-Wesley, London, 1964).
[12] D.G. Brandon, Surface Sci. 3 (1964) 1.
[13] E.W. Müller, Surface Sci. 2 (1964) 484.
[14] T.T. Tsong and G. Kellogg, Phys. Rev. B12 (1975) 1343.
[15] T.T. Tsong, Surface Sci. 70 (1978) 211.
[16] T.T. Tsong, Progr. Surface Sci. 10 (1980) 165.

[17] For example, T.M. Miller and B. Bederson, in: Advances in Atomic and Molecular Physics, Vol. 13 (Academic Press, London, 1977).
[18] T.T. Tsong and E.W. Müller, Phys. Status Solidi (a) 1 (1970) 513.
[19] T.T. Tsong and E.W. Müller, J. Chem. Phys. 55 (1971) 2884.
[20] D. McKinstry, Surface Sci. 39 (1972) 37.
[21] R.G. Forbes, Surface Sci. 70 (1978) 239.
[22] R.G. Forbes, R.K. Biswas and K. Chibane, Surface Sci. 114 (1982) 498.
[23] M. Konishi, M. Wada and O. Nishikawa, Surface Sci. 107 (1981) 63.
[24] M. Wada, M. Konishi and O. Nishikawa, Surface Sci. 100 (1980) 439.
[25] R.J. Culbertson, T. Sakurai and G.H. Robertson, Phys. Rev. B19 (1979) 4427.
[26] R.G. Forbes, Appl. Phys. Letters 36 (1980) 739.
[27] T.T. Tsong and R.J. Walko, Phys. Status Solidi (a) 12 (1972) 111.
[28] N. Ernst, Surface Sci. 87 (1979) 469.
[29] R.G. Forbes, Surface Sci. 61 (1976) 221.
[30] Previous papers used a different notation. Now that post-ionization is an established experimental fact, it seems clearest to add all three charge-state suffices to the symbols for activation energy and appearance energy.
[31] T.T. Tsong, W.A. Schmidt and O. Frank, Surface Sci. 65 (1977) 109.
[32] E. Hummel, M. Domke and J.H. Block, Z. Naturforsch. 34a (1978) 46.
[33] M. Domke, E. Hummel and J.H. Block, Surface Sci. 78 (1978) 307.
[34] For example, L.N. Dobretsov and M.V. Gomoyunova, Emission Electronics (Moscow, 1966) (English translation: Israel Program for Scientific Translations, Jerusalem, 1971).
[35] R.G. Forbes, Surface Sci. 102 (1981) 255.
[36] R. Gomer and L.W. Swanson, J. Chem. Phys. 38 (1963) 1613.
[37] R. Haydock and D.R. Kingham, Phys. Rev. Letters 44 (1980) 1520.
[38] R. Haydock and D.R. Kingham, Surface Sci. 104 (1981) L194.
[39] N. Ernst, private communication.
[40] 1 meV V^{-2} $nm^2 \approx 1.602189 \times 10^{-40}$ J V^{-2} $m^2 \approx 4\pi\epsilon_0 \times 1.439976$ Å^3.
[41] Strictly, it would be more appropriate to scale b_a(orb).
[42] R.G. Forbes, J. Phys. D (Appl. Phys.), 15 (1982) 1301.
[43] R.G. Forbes, J. Phys. D (Appl. Phys.), to be published
[44] M.K. Wafi, PhD Thesis, University of Aston in Birmingham (1981).
[45] In principle there is a question as to whether proper polarizability has the same value for a kink-site atom as for an atom in the interior of a surface facet. We neglect this here, partly because the Culbertson et al. data are not sufficiently well-defined to make detailed discussion worthwhile.

J. Phys. D: Appl. Phys., Vol. 11, 1978. Printed in Great Britain

LETTER TO THE EDITOR

Negative work-function correction at a positively-charged surface

Richard G Forbes
Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET

Received 10 July 1978

Abstract. It is shown that if a very high positive electric field is present at a metal surface then the dipole moments electrically induced in the surface atoms must give rise to a field-related negative work-function correction. At the high fields encountered at a field-ion emitter, approximately 50 V nm^{-1}, the correction may be several electron-volts or more. It is also suggested that, in consequence of the Hellman–Feynman theorem, a discrete-charge model may in principle be a better classical model for a positively charged surface than a representation in terms of a continuous positive charge distribution.

Following the experimental discovery of field adsorption of noble-gas atoms on to the surface of an operating field-ion emitter (Müller *et al* 1969, Tsong and Müller 1970), and the calculations by Tsong and Müller (1970), it has been widely believed that the basic cause of the adsorption was the field due to a field-induced dipole moment in the adjacent substrate atom. Tsong (unpublished work cited in Tsong 1978) has recently examined the interaction between an infinite square array of substrate-atom dipoles and a corresponding infinite square array of gas-atom dipoles, and claims to have confirmed theoretically that field adsorption would exist in this case too.

Another theoretical possibility also has to be considered, at least in principle. Forbes (1970) independently argued that a real charged surface ought to be represented by an array of discrete charges, located at the atomic nuclei, and that field variations resulting from the localisation of the positive charge might cause field adsorption. However, investigation (Forbes 1978a) of the field variations above an infinite square array of super imposed monopoles and dipoles strongly suggests that the dominant influence on field adsorption is indeed the substrate-dipole field, although the monopole-induced field variations should not be neglected in a detailed treatment.

Thus, since field adsorption is a real effect, and since substrate-atom dipole moments are needed in order to cause it, one is forced to conclude that the existence of substrate-atom dipole moments at a positively charged surface is a general effect, applicable in principle to all positively charged surfaces.

An important consequence of this seems to have escaped notice: namely that an array of electrically-induced surface dipoles will lead to a field-related work-function correction, in much the same way that local work-function variations are attributed to differences in chemically-induced surface dipole moments (Smoluchowski 1941).

The possible magnitude of this effect can be explored in the context of an infinite-square-array model. By arguments broadly similar to those of Tsong (1978), it may be

0022–3727/78/0014–9161 $01.00

shown that the field-induced SI dipole moment (p') of each individual substrate-atom dipole is given by

$$p' = \tfrac{1}{2}b'F^{\text{ext}}[1 - 2K_1\mu'/(4\pi\epsilon_0 a^3 F^{\text{ext}})]/[1 + K_1 b'/(4\pi\epsilon_0 a^3)] \tag{1}$$

where F^{ext} is the 'external' field in space well above the array, K_1 is a 'structure factor' approximately equal to 9·03, μ' is the zero-field SI dipole moment, ϵ_0 is the electric constant, a is the array lattice spacing, and b' is the *proper* SI polarisability (Forbes 1977) of the substrate atom. By 'proper' polarisability I mean the quantity called polarisability in SI textbooks on electricity and defined by

$$p' = \mu' + b'F^{\text{loc}} \tag{2}$$

where F^{loc} is the *local* SI electric field as conventionally defined in textbooks on electricity.

By comparison with formula (12) in Tsong (1978) for the local field acting on a substrate atom, there are two differences about the treatment here: the present formula is written using the SI system of electrical quantities and equations; and there is an additional factor of $\frac{1}{2}$ here, because it has been realised that the field acting on a layer of charge is half the field above it when dipole-induced effects are ignored (for example, Feynman *et al* 1964).

In terms of p' the correction $\delta\phi$ to the local work-function is given by

$$\delta\phi = -ep'/(a^2\epsilon_0) \tag{3}$$

where e is the elementary (proton) charge. For a positive external field the work-function correction is negative.

To form an estimate of $\delta\phi$ we may: take a equal to the interatom spacing in the tungsten (111) plane, which is 0·4476 nm; regard μ' as negligible; and set b' equal to the theoretical free-space polarisability of tungsten (Thorhallson *et al* 1968, Forbes 1977), namely 11·7 meV V^{-2} nm^2 (1 meV V^{-2} nm$^2 \approx 1{\cdot}602189 \times 10^{-40}$ J V^{-2} m^2). Setting F^{ext} in turn equal to the helium imaging field and the tungsten evaporation field gives:

$$\delta\phi = -8{\cdot}8 \text{ eV } (F^{\text{ext}} = 45 \text{ V nm}^{-1}\text{: helium field-ion imaging}) \tag{3a}$$

$$= -11{\cdot}2 \text{ eV } (F^{\text{ext}} = 57 \text{ V nm}^{-1}\text{: tungsten field evaporation}). \tag{3b}$$

The corrections predicted using this b' value are in fact much larger than one might expect.

In interpreting these figures it must be clearly understood that the Thorhallson *et al* theoretical value for b' is somewhat unreliable, even as an estimate of free-space polarisability. It is also possible that the value appropriate to a surface atom may be different from that appropriate to an isolated atom. Thus it is entirely possible that the corrections predicted in this way are too large by a factor of two or more, although they may not be.

Ideally, one would like to use an empirically-derived estimate of the proper polarisability of a surface atom. Unfortunately, the various empirical quantities called 'polarisability' or 'surface polarisability' in field-ion literature are not in fact polarisabilities in the textbook sense. This is because the definition of the empirical field-ion quantities is via the equation

$$U_{\text{pol}} = -\tfrac{1}{2}c'(F^{\text{ext}})^2 = -\tfrac{1}{2}c_s(4\pi\epsilon_0)(F^{\text{ext}})^2 \tag{4}$$

where U_{pol} is the polarisation contribution to the atomic potential energy and F^{ext} is the external (or 'applied') field above the surface. c_s is the Gaussian quantity usually expressed in Å^3, and c' is its equivalent in the international (SI) system of measurement

(see Forbes 1977); $c'/(\text{meV V}^{-2}\text{ nm}^2) = 0{\cdot}694456\, c_s/(\text{Å}^3)$. The quantity F^{loc} appearing in the official definition of polarisability, that is equation (2), may be smaller than F^{ext} by a factor of as much as 3 or more (Forbes 1978b).

Regrettably, our present theoretical knowledge about charged surfaces is not sufficient for exact numerical relationships to be derived between the proper polarisability b' and the empirical c' (or c_s) values appearing in the literature. However, it is instructive to make an attempt to use an empirical quantity to estimate $\delta\phi$. At a naive level of argument we may suppose that the dipole moment of a surface atom is given approximately by

$$p' \simeq c' F^{\text{ext}}. \tag{5}$$

An estimate of c' for tungsten can be derived from the work of Tsong and Kellogg (1975), and is 4·9 meV V^{-2} nm^2. Substitution into equations (5) and (3) gives

$$\delta\phi = -20 \text{ eV } (F^{\text{ext}} = 45 \text{ V nm}^{-1}) \tag{6a}$$

$$= -25 \text{ eV } (F^{\text{ext}} = 57 \text{ V nm}^{-1}). \tag{6b}$$

On the face of things the empirically derived estimate is even higher than the *a priori* theoretical estimate. A possible reason, however, is that the Tsong and Kellogg quantity represents effects due to charge transfer (which is what they suggest), as well as effects due to the polarisation of metal-atom orbitals. Charge-transfer effects do not occur in the context of a planar-array model, so the empirical c' value has no immediate relevance to the present discussion.

Returning to the original estimates, equation (3*a*,*b*), it is necessary to express some further reservations. An infinite planar square array is, of course, not a particularly realistic model of surface geometry. And the predictions would, obviously, be smaller if a bigger lattice spacing had been assumed. Nevertheless, taking all doubts into account, it seems to the author that at helium imaging fields and higher a field-related work-function correction negative in sign and of magnitude at least several electron-volts is entirely plausible.

Corrections as large as this could have significant repercussions in field-ion theory. In particular they would cause upwards revision of the critical distance values currently accepted, and would affect estimates of field evaporation activation energies obtained from arguments based on a thermionic cycle. Fundamental questions could also arise about the nature of electron states at a charged surface, and the nature of the potential seen by an incoming electron. There thus seems an urgent need to develop better models of a charged emitter surface and the field and potential variations near it.

There also seems to be an important general implication of the arguments here. The field-ion situation is in one sense just an extreme case of an ordinary charged surface. If we are forced to conclude here that the surface of an operating field-ion emitter must be represented by an array of monopoles and dipoles of appropriate strengths, rather than by a continuous classical charge distribution, then in principle all charged surfaces should be represented in this manner. The necessity of such a representation seems to me to be an unnoticed consequence of the Hellman–Feynman theorem, if one thinks in terms of localised-orbital models of a charged surface. Perhaps such a representation would be an important first step in the development of a proper quantum-mechanical theory of charged surfaces, based on LCAO type methods rather than delocalised plane-wave electron wave-functions.

References

Feynman R P, Leighton R B and Sands M 1964 *Lectures on Physics* Vol. II (New York: Addison-Wesley)
Forbes R G 1970 *PhD Thesis* University of Cambridge
—— 1977 *Surface Sci.* **64** 367–71
—— 1978a *Surface Sci* to be published
—— 1978b *25th Int. Field Emission Symp., Albuquerque, 1978* (unpublished)
Müller E W, McLane S B and Panitz J A 1969 *Surface Sci.* **17** 430–8
Smoluchowski R 1941 *Phys. Rev.* **60** 661–74
Thorhallson J, Fisk C and Fraga S 1968 *J. Chem. Phys.* **49** 1987–8
Tsong T T 1978 *Surface Sci.* **70** 211–33
Tsong T T and Kellogg G 1975 *Phys. Rev.* B **12** 1343–53
Tsong T T and Müller E W 1970 *Phys. Rev. Lett.* **25** 911–3

Proc. 7th Intern. Vac. Congr. & 3rd Intern. Conf. Solid Surfaces (Vienna 1977) 433 - 436

RE-EXPLORATION OF THE CONCEPTS OF WORK FUNCTION AND IONIZATION ENERGY

Richard G. Forbes

University of Aston, Dept. of Physics, Gosta Green, Birmingham, U.K.

Abstract: Work functions and ionization energies may in principle be defined by thermionic-type cycles. Some consequences of using this form of definition are explored. It is suggested that the concept of "absolute" or "total" work-function is unsatisfactory, and some replacement concepts are established. Similarly, the concept of an "absolute" or "total" binding energy for an electron orbital in an atom adsorbed on an emitter substrate is unsatisfactory; it is replaced by several types of "work-of-ionization".

These ideas can be used to resolve a longstanding controversy in field evaporation theory over work-functions. Finally, a resemblance between field evaporation theory and the theory of the binding energies observed in photoelectron spectroscopy provides a new viewpoint on the latter theory.

INTRODUCTION

The intentions of this paper are to look again at the basic definitions of work function and ionization energy, and to apply to the question of the referencing of photoelectron energies the type of thermodynamic argumentation used in field ion theory.

Several streams of thought have combined to produce the suggestions here. But particularly fruitful has been the realization that there is a theoretical connection between a long-standing controversy in field evaporation theory over whether it is the local or the total work-function that should appear in a certain formula (see p.92 of Ref./1/), and recent discussions /2,3/ in ESCA literature over the meaning of the term "absolute binding energy" and the referencing of electron energy levels.

DEFINITIONS OF WORK FUNCTION

Though there are various empirical parameters called "work-function" /4/, basically only two types of theoretical parameter have been generally recognised: the "absolute" or "total" work-function, usually defined as the work needed to remove an electron from an emitter fermi level to infinity; and the "local" work-function, which is the commonly used parameter associated with a surface of specific composition and orientation.

The usual definitions of local work function define a parameter associated with a surface of large extent. But in the context of a field-ion emitter it is useful to have a definition that relates to a point outside the emitter. Using a procedure that has some resemblance to Herring's /3/, we may define the *generalised work-function* ϕ^r associated with a point r by:

$$\phi^r = V_z(r) - \omega^r \qquad (1)$$

where: $V_z(r)$ is the work needed to remove an electron from the emitter fermi level to point r, in the absence of any applied electric field, by means of a thermodynamically-slow process; and ω^r is the correlation-and-exchange component of the electron potential energy, at point r. ($\omega^\infty = 0$).

In a "flat-surface" theoretical model, if r is taken to be a point (e) "just outside the emitter surface", then the quantity ϕ^e is identical with the *local work-function* as usually defined, as "the difference between the electrostatic potential energy of an electron just outside the emitter and the electrochemical potential inside the emitter".

An elementary "proof" of this can be made as follows. As well as the correlation-and-exchange forces, an electron is subject to "purely electric" forces that give rise to an "electrostatic" potential energy v. Hence the work done (w) in moving an electron from rest at some point i to rest at some point r is:

$$w = (v^r - v^i) + (\omega^r - \omega^i) \qquad (2)$$

Inside the metal, the total energy (μ) of an electron at the fermi level is given by:

$$\mu = \langle \omega^{int} \rangle + \langle v^{int} \rangle + \epsilon_F \qquad (3)$$

where ϵ_F is the fermi energy, and the brackets denote an appropriate average.

Proc. 7th Intern. Vac. Congr. & 3rd Intern. Conf. Solid Surfaces (Vienna 1977) 434

Further, $V_z(r)$ is given by:

$$V_z(r) = \omega^r + v^r - \langle v^{int} \rangle - \langle \omega^{int} \rangle - \varepsilon_F \quad (4)$$

So combining the above equations leads to:

$$\phi^r = V_z(r) - \omega^r = v^r - \mu \quad (5)$$

Setting $r = e$ gives the required result.

With a crystallographically-flat real emitter of large extent, v^r will be much the same at all points in a surface two or so atom-diameters above the emitter, - so ϕ^e is well-defined. But with an emitter on which there are isolated atoms adsorbed, or with a field-emitter tip, there may be significant point-to-point variations in v^r in space above the emitter surface. It may then be useful to define an *adsorption-site work-function* ϕ^a, by taking r in equation (5) equal to some appropriate point a above the site in question.

If r is taken at any position t beyond the effective range of correlation-and-exchange forces, then $\omega^t = 0$, and:

$$\phi^t = V_z(t) \quad (6)$$

In these circumstances, the work-function (ϕ^t) actually is equal to the thermodynamic work done in extracting the electron. I have previously called ϕ^t an *electrothermodynamic work-function* /5/.

ϕ^t is a parameter defined in terms of a removal process. In general its value depends both on the nature and circumstances of the body concerned and on the location of t. A special case arises if a chemically well-defined body is taken to be isolated in an otherwise empty universe, and t is taken at a remote position R, a very large distance from the body. The work-function thus defined is characteristic of the body in question, has been denoted by ϕ^*, and has been called the *specific work-function* of the body /5/.

An important point is that there seems no good reason to suppose that ϕ^* would be the same for a cube (showing faces of one orientation and local work-function) as it would be for a bipyramid showing only faces of a different orientation and local work-function. Consequently, one must assume that in general the specific work-function ϕ^* is a function of the shape of the emitter.

A further point is that a fermi level is a property of the relevant thermodynamic system. Consequently, if an electron is removed from an emitter that is connected electrically and thermally to a very much larger body, then the work done in removing the electron to a point remote from the emitter will be approximately equal to the specific work-function for the large body, not to that for the emitter. In the laboratory situation, the walls of a vacuum system could be "large" with respect to the emitter, if (say) both were earthed.

One may conclude that the traditional concept of "absolute work-function" is inadequate to describe the real complexity of the situation, and should be replaced by a set of work-function concepts. This conclusion has no significant effect on most experimental work, as experiments nearly always concern themselves with local work-functions. But the refined concepts presented here do enable arguments involving thermodynamic cycles to be formulated in a more rigorous fashion.

Finally, I derive a relationship that will be needed later. For two points r and s, it follows from eq.(2) and the definition of V_z that:

$$V_z(r) - V_z(s) = v^r - v^s + \omega^r - \omega^s \quad (7)$$

and hence from eq.(5) that:

$$\phi^r - \phi^s = v^r - v^s \quad (8)$$

WORK OF IONIZATION

Now consider the situation in which an electron is extracted, by means of a thermodynamically-slow process, from some specified orbital β in an atom adsorbed on an emitter surface.

The removal process may be considered to take place in two stages. First the electron is removed from orbital β to the emitter fermi level; then the electron is removed from the fermi level to its intended destination.

The work B_β involved in the first step is equal to the "binding energy of the orbital relative to the fermi level". If photoemission is in fact a thermodynamically-slow process, then it is this quantity that is in effect measured in a photoelectron spectrometer.

Assuming that the final destination of the electron is outside the range of correlation-and-exchange forces, the work involved in the second step is the relevant electrothermodynamic work function ϕ^t. Consequently the total work L_β^t needed to remove the electron from orbital β to point t is given by:

$$L_\beta^t = B_\beta + \phi^t \qquad (9)$$

L_β^t could perhaps be termed an *electrothermodynamic work-of-ionization*.

From eq.(9) it is clear that L_β^t must depend on the same factors as does ϕ^t. Just as there are different categories of work-function, so there are (analogous) different categories of work-of-ionization. In the same way that the concept of absolute work-function is inadequate, the concept of "absolute" work-of-ionization (i.e. "total" or "absolute" binding energy) is inadequate.

The quantity that comes closest to an absolute binding energy is, perhaps, the *specific work-of-ionization* (L_β^*) for the atom-plus-emitter system, given by:

$$L_\beta^* = B_\beta + \phi_E^* \qquad (10)$$

where ϕ_E^* is the specific work-function for the emitter.

But, from previous remarks, it is clear that this quantity refers to the atom-plus-emitter system when situated in an otherwise empty universe, not to this system when situated in a real vacuum system, and in any case the quantity may depend on the shape of the emitter. Such a parameter is of doubtful experimental use.

The quantity L_β^t is a genuine thermodynamic work done. However, by analogy with the discussion on work-functions and with eq.(9) it is possible to define two quantities:

$$L_\beta^e = B_\beta + \phi^e \qquad (11)$$

$$L_\beta^a = B_\beta + \phi^a \qquad (12)$$

neither of which are genuine works done. L_β^e may be called a *local* work-of-ionization, and L_β^a an *adsorption-site* work-of-ionization. The latter quantity has some relationship with the "ionization limit" E_{IL}^ϕ discussed by Hagstrum /3/ (albeit in the context of a different energy-referencing system).

In their pioneering book on photoelectron spectroscopy , Siegbahn et al. stated (cf. Appendix 1 in Ref./6/) that:

> "The tabulated binding energies are given with zero binding energy at the fermi level. The total binding energy is therefore greater by an amount equal to the work-function."

This statement is a potential source of confusion. To many people the phrase "work-function" automatically implies the local work-function ϕ^e, - but L_β^e is not a binding energy (in the normal sense of the term). L_β^* is a genuine binding energy, - but the work-function ϕ_E^* is not an easily measurable quantity, - and L_β^* is not a quantity of any significant scientific use.

What one actually wants to know is the relationship between B_β and the free-space binding-energy associated with this orbital. I return to this question below.

THE WORK FUNCTION CONTROVERSY

In field evaporation theory /1/ one needs an expression for the work B_{On} needed to convert a bound atom into an n-fold-charged ion at the same position (b), the electrons being placed at the emitter fermi-level by means of a thermodynamically-slow process.

This work can be obtained by considering a thermionic cycle /1,7/. In what follows, H_n^∞ denotes the free-space heat of formation, for the n-fold ion from the neutral, the electrons also being left in remote field-free space, distant from the ion. The work w needed to move the n-fold ion from point i to point r can be written in the form:

$$w = (\eta_n^r - \eta_n^i) + n(v^i - v^r) \qquad (11)$$

where v is the electron potential energy defined earlier. $(-nv)$ is the "purely electric" component of the ion potential energy. η_n is the sum the components of potential energy due to all other ("purely chemical") effects. (At any position R in remote field-free space, $\eta^R = 0$).

Formally, the ion may be created by means of a sequence of steps. The work done by a hypothetical external agent is shown in brackets after each step: *1.* Remove atom from bonding point to remote field-free space (Λ_0). *2.* Create ion in remote field-free space (Ξ_n^∞). *3.* Return electrons to emitter fermi level ($-n\phi^R$). *4.* Return ion to point b ($nv^R - nv^b + \eta_n^b - \eta_n^R$). Summing gives the total work:

$$B_{On} = H_n^\infty - n(\phi^R - v^R + v^b) + \Lambda_0 + \{\eta_n^b - \eta_n^R\} \qquad (12)$$

Λ_0 is the zero-field heat of adsorption for the neutral atom.

With the aid of eq.(8), the term in round brackets reduces to ϕ^b. The term in curly brackets may be written as $(-\Xi_n)$,

where Ξ_n is the "purely chemical" part of the binding energy between the n-fold ion and the emitter. Thus eq.(12) reduces to:

$$B_{On} = H_n^{\infty} - n\phi^b + \Lambda_0 - \Xi_n \quad (13)$$

From this result it is clear that the work-function appearing is indeed a form of local work-function, as has been maintained by Müller and Tsong /1/. However, Brandon /7/ was correct in stating that the work-function that appears in step 3 above is a form of "total" work-function. What is shown here, for the first time, is that the reason why a local work-function appears in eq.(13) is that in step 4 above the ion has to be worked against patch fields in returning it from a remote point to the bonding point b .

APPLICATION TO ENERGY-LEVEL REFERENCING

There is a close formal analogy between the situation just discussed and photoemission theory. If the suffix β is used to label quantities applying to an ion that has been formed by removing an electron from orbital β in a neutral atom, then:

$$B_\beta = H_\beta^{\infty} - \phi^b + \Lambda_0 - \Xi_\beta \quad (14)$$

This is the relationship indicated earlier.

Note that the quantity ϕ^b that appears is a type of local or adsortion-site work-function, but it is a work function associated with the surface in the absence of the atom from which photoemission occurs. The point b is to be taken at the centre-of-charge of the ion that has resulted from photoemission.

If photoemission occurs from an atom considered to be adsorbed on top of a clean surface or on top of an adsorbed layer, then the interpretation of all the parameters in eq. (14) is straightforward, with ϕ^b in effect the local work function of the surface.

However, if the photoemitting atom is considered as adsorbed within a surface layer, then, although all the parameters in eq.(14) remain defined, their interpretation is less obvious.

One of the factors that stimulated the present work was the discovery by Joyner and Roberts /8/ that the fermi-referenced binding energy for the 1s orbital of Oxygen was the same for a number of cases where Oxygen adsorbed on different metal substrates, even though the clean-surface local work-functions for the substrates varied between 4 eV and 5 eV. The theory here provides additional confirmation for their view that their results are chemically significant, - indicating that in all the cases in question the chemical environment of the Oxygen atom must be essentially the same. (Perhaps Oxygen is, in some sense, "exerting its own work-function".)

As compared with the treatment given by Hagstrum /3/, the present approach lays emphasis on atomic and ionic potential energies, and on thermodynamic cycles, rather than on electron energy levels. In our approach it is natural that binding energies should be referred to the emitter fermi level, because this is the obvious reference level for a thermodynamically-based argument. The concept of "vacuum level", and the complications associated with its point-to-point variation throughout space in the real world, are thereby avoided.

Another feature of the present treatment is that it assumes that photoemission is a thermodynamically slow process, not a "sudden" process. Full relaxation is thus implicit in the formulation, as presented above. If photoemission is not, in fact, a "slow" process, then correction terms can be introduced.

REFERENCES

/1/ E.W.Müller and T.T.Tsong: Prog. Surf. Sci. 4 (1971) 1 .
/2/ A.F.Carley, R.W.Joyner, and M.W. Roberts: Chem. Phys. Lettrs. 27 (1974) 580.
/3/ H.D.Hagstrum: Surf. Sci. 54 (1976) 197.
/4/ J.C.Riviere, in: Solid State Surface Science, Vol. III (Ed. M.Green) (Dekker, New York, 1967).
/5/ R.G.Forbes, in: 23rd International Field Emission Symposium, Pennsylvania State University, 1976.
/6/ K.Siegbahn et al.: ESCA - Atomic, Molecular, and Solid State Structure studied by means of Electron Spectroscopy (Almquist and Wiksells, Uppsala, 1967).
/7/ D.G.Brandon, Surf. Sci. 3 (1964) 1.
/8/ R.W.Joyner and M.W.Roberts, Chem. Phys. Lettrs. 28 (1974) 246.

IV. FIELD ADSORPTION

22 Field adsorption: The monopole-dipole interaction
R.G. Forbes
Surface Sci., 78 (1978) L504-7

23 Conceptual errors in the theory of field adsorption
R.G. Forbes
Surface Sci., 87 (1979) L278-84

24 An array model for the field adsorption of helium on tungsten(111)
R.G. Forbes and M.K. Wafi
Surface Sci., 93 (1980) 192-212

25 The influence of hyperpolarisability and field-gradient polarisability on field adsorption binding energies for He on W(111)
R.G. Forbes
Surface Sci., 108 (1981) 311-28

26 Progress with the theory of noble-gas field adsorption
R.G. Forbes
Vacuum, 31 (1981) 567-70

Surface Science 78 (1978) L504–L507

SURFACE SCIENCE LETTERS

FIELD ADSORPTION – THE MONOPOLE–DIPOLE INTERACTION

Richard G. FORBES
University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET, UK

Received 14 April 1978; manuscript received in final form 28 June 1978

The existence of imaging-gas atoms locally adsorbed on the surface of an operating field-ion emitter was discovered experimentally by Müller and coworkers [1,2], who termed the interaction *field adsorption.* The gas atoms are within the critical surface, field ionization takes place above them, and it is thought that the presence of these atoms is essential to the field-ion imaging process [3–5].

Following the calculations by Tsong and Müller [6], it is widely believed that localised field adsorption is basically due to polarisation of the imaging-gas atom by the field resulting from the dipole moment of the underlying emitter substrate atom, the existence of this dipole moment itself being a field-induced effect. Within the subject area, this mechanism is termed "field-induced-dipole–dipole interaction".

An alternative mechanism had in fact been conceived by the present author, who argued [7] that a real positively-charged surface should be represented by an array of discrete positive charges, and suggested that localised field adsorption could result from position-dependent variations in the field resulting from these discrete charges. This mechanism can be called the "monopole–dipole interaction".

With the appearance of ref. [6] it became clear that the total "primary" polarising field should contain monopole and dipole contributions [7]. There are also "secondary" contributions due to the image of the field-adsorbed atom, the presence of other field-adsorbed atoms (and their images), and to the depolarising effect of the field-adsorbed atoms on the substrate atoms, but these contributions are small and can be neglected in a first approximation.

Calculations on the dipole–dipole interaction model initially considered an isolated substrate-atom/gas-atom pair [6]. This treatment may be suitable for a kink-site, but looks inappropriate for a net plane in which every substrate atom is imaged. Thus, recently [8], Tsong has reported the main results of a new calculation of the polarising field on an individual field-adsorbed atom, that takes into account the contributions of an infinite square array of dipoles representing substrate atoms, and of a like array of dipoles representing field-adsorbed atoms.

With the monopole–dipole model, the present author's original calculations [7] (which correspond more-or-less to the kink-site situation) gave an estimated

localised binding energy of ~0.1 eV, for an assumed local charge of $1e$, but it is difficult to justify the approximations used. However, with an infinite square array of charge the calculation of the local field above a symmetry position is straightforward. This note reports the result of such a calculation.

Suppose that at every point there is a charge q whose size is determined by the Gauss relationship:

$$F^{\text{ext}} = \sigma/\epsilon_0 = q/a^2\epsilon_0 \ , \tag{1}$$

where a is the lattice spacing, σ is the average surface charge density, ϵ_0 is the electric constant, and F^{ext} is the external field well above the charged surface.

At a point P_1 a distance (νa) above the central lattice point, the field-contribution $F^{\text{m,pos}}$ due to the positive charges in the array is given by:

$$F^{\text{m,pos}} = (q/4\pi\epsilon_0) \sum_{ij} \nu a/(i^2a^2 + j^2a^2 + \nu^2a^2)^{3/2} \ , \tag{2}$$

where the integers i and j label the lattice points, and the summation is over all lattice points.

In the limit of large ν, $F^{\text{m,pos}}$ tends to $\frac{1}{2} F^{\text{ext}}$. To get the correct result for the whole field-contribution (F^{m}) due to the existence of monopoles, it is necessary to add in the contribution $F^{\text{m,neg}}$ ($= \frac{1}{2} F^{\text{ext}}$) due to an infinite square array of negative charges parallel to the positive charge array and situated a long distance from it:

$$F^{\text{m}} = F^{\text{m,pos}} + F^{\text{m,neg}} \ . \tag{3}$$

This result seems to be not well known, and will be discussed at greater length elsewhere. Combining the above equation, we may define a dimensionless factor β^{m} by:

$$F^{\text{m}}/F^{\text{ext}} = \beta^{\text{m}}(\nu) = \tfrac{1}{2} + (\nu/4\pi) \sum_{ij} (i^2 + j^2 + \nu^2)^{-3/2} \ . \tag{4}$$

In the context of field adsorption, we must assume that the "applied field" acting on the field-adsorbed atoms is the field F^{m} rather than the external field F^{ext}. In principle, therefore, the factor β^{m} should be inserted into Tsong's [8] equations, in appropriate places: for example, his eq. (12) should read:

$$F = \beta^{\text{m}} F_0 - K_1 p/a^3 \ . \tag{5}$$

To compare the relative influences of the monopole (F^{m}) and substrate-dipole (F^{d}) contributions to the polarising field F^{loc}, it is useful to define a second dimensionless factor β^{d} by the ratio $F^{\text{d}}/F^{\text{ext}}$. If "secondary" field contributions are neglected, then the polarising field is given by:

$$F^{\text{loc}} = F^{\text{m}} + F^{\text{d}} = (\beta^{\text{m}} + \beta^{\text{d}}) F^{\text{ext}} \ . \tag{6}$$

The localised polarisation-binding energy, ΔU_{pol}, with respect to an atom in the external field, is thus given by:

$$\Delta U_{\text{pol}} = \tfrac{1}{2} \alpha' [(\beta^{\text{m}} + \beta^{\text{d}})^2 - 1](F^{\text{ext}})^2 \ , \tag{7}$$

where α' is the SI (absolute) polarisability [9] for the imaging-gas atom.

At the level of approximation inherent in the present discussion, $(1+\beta^d)^2$ may be identified with the parameter f_A used in refs. [6] and [8]. For the case of a tungsten substrate, values of f_A are tabulated in ref. [8]. On the basis of these figures it would be deduced that β^d varies between 0.26 (for xenon imaging gas), and 0.55 (for helium imaging gas).

To estimate comparable values for β^m, I take the lattice spacing a as equal to the interatomic spacing on the (111) plane of tungsten, namely 0.4476 nm, and use Tsong's values [8] for the separation between the nuclei of the substrate atom and imaging-gas atom. These are 0.355 nm for xenon, 0.259 nm for helium. The corresponding values of the distance-fraction (ν) are: 0.793 for xenon; 0.579 for helium.

The evaluation of the summation in eq. (4) can be achieved by a combination of numerical and analytical methods, and will be described elsewhere. The results for β^m are: 1.016 for xenon; 1.067 for helium.

Tsong's calculations represent an approximation in which β^m in eqs. (6) and (7) here is taken as unity. On the basis of the figures presented here, we may deduce that this approximation has led to underestimates of the polarisation-energy contribution to binding energy, by roughly 7% in the case of neon, roughly 13% in the case of helium.

These figures should be treated with some circumspection, for several reasons. Firstly, β^d is in some circumstances a function of β^m, albeit a weak one. Second, it is by no means clear to me that the f_A-values cited in ref. [8] are reliable. Third, the procedure of representing a charged metal surface by a superimposed array of point charges and dipoles, although an advance on previous treatments, is still only an approximation to the real situation [10]. Fourth, the surface lattice on the Tungsten (111) plane is hexagonal rather than square, and calculations are clearly needed that represent the arrangement of substrate atoms more realistically.

Nevertheless, the results of more detailed investigations [11] tend to support the basic point established by the simplified treatment here. Namely that localised field adsorption results chiefly from the dipole field of the substrate atoms, but that the discrete nature of positive charge produces significant effects that should not be neglected in detailed treatments.

References

[1] E.W. Müller, S.B. McLane and J.A. Panitz, Surface Sci. 17 (1969) 430.
[2] T.T. Tsong and E.W. Müller, Phys. Rev. Letters 25 (1970) 1911.
[3] R.G. Forbes, in: 21st Intern. Field Emission Symposium, Marseilles, 1974.
[4] H. Iwasaki and S. Nakamura, Surface Sci. 33 (1972) 525.
[5] E.W. Müller, Crit. Rev. Solid State Sci. 6 (1976) 85.
[6] T.T. Tsong and E.W. Müller, Phys. Rev. Letters 25 (1970) 911.
[7] R.G. Forbes, Ph.D. thesis, Cambridge University, 1970.
[8] T.T. Tsong, Surface Sci. 79 (1978) 211.

[9] R.G. Forbes, Surface Sci. 64 (1977) 367.
[10] One referee feels that use of the point-charge approximation will overestimate effects due to the local distribution of excess charge. The author feels that underestimation is also a possibility, if one applies the Hellman–Feynman theorem in the context of a LCAO-type approach to the construction of surface electron wave-functions. Certainly there is a need for a better quantum-mechanical theory of charged surfaces.
[11] R.G. Forbes, unpublished work.

Surface Science 87 (1979) L278–L284

SURFACE SCIENCE LETTERS

CONCEPTUAL ERRORS IN THE THEORY OF FIELD ADSORPTION

Richard G. FORBES
University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET, UK

Received 9 February 1979

Serious conceptual errors are shown to exist in a new theoretical treatment of field adsorption proposed by Tsong in a recent review. A conceptual error also exists in the standard approach, in that the influence of the adsorbate on the internal energy of the source of the electric field has been neglected. A numerical estimate of the size of the correction to the short-range field-adsorption binding energy, made by means of the isolated-dipole-pair approximation, shows that the correction is small but should not be neglected in detailed treatments.

There is, at the surface of an operating field-ion emitter, a firmly-bound monolayer of imaging gas field-adsorbed onto the emitter surface [1]. If, somewhat above this layer the electric field is nearly uniform, and has the value F^{ext}, then the binding energy B of an individual field-adsorbed gas atom can usefully be split into two parts: a long-range part, equal to the electric-field-induced binding energy in the field F^{ext}, and a "localised binding energy" ΔB. Thus:

$$B = \tfrac{1}{2} b_A (F^{ext})^2 + \Delta B, \tag{1}$$

where b_A is the SI polarisability of the gas atom [2,3]. ΔB is equal to the work that needs to be done (by a hypothetical external agent acting on the adsorbate gas atom) to move the atom from its binding site to a position in the "external" field F^{ext}

Although there are components in ΔB resulting from London dispersion forces and repulsive "interpenetration" forces [4], it is normally assumed that the dominant component is the electric-field-induced part ΔB^{el}.

Most attempts to calculate ΔB^{el} have assumed that it is equal to the quantity ΔB(conv.) given by the difference in adsorbate-atom polarisation energy between the binding site and a position in the field F^{ext}:

$$\Delta B(\text{conv.}) = \tfrac{1}{2} b_A (F_A^{loc})^2 - \tfrac{1}{2} b_A (F^{ext})^2 = \tfrac{1}{2} b_A (\beta_A^2 - 1)(F^{ext})^2, \tag{2}$$

where F_A^{loc} denotes the local field acting on the gas atom when it is at the binding site, and β_A denotes the ratio:

$$\beta_A = F_A^{loc}/F^{ext}. \tag{3}$$

As will be shown later, ΔB^{el} is strictly equal to ΔB(conv.) only if there are no mutual induction effects between the adsorbate atom in question and its surroundings (so the local field acting on the gas atom is unaffected by its presence). But in all physically-realistic circumstances it is a reasonable first approximation to set ΔB^{el} equal to ΔB(conv.).

The problem, therefore, has been seen as one of calculating β_A, or the equivalent parameter $f_A(=\beta_A^2)$. At an adsorption site β_A is somewhat greater than unity. Partly this is a consequence of effects due to the localisation of the excess positive charge associated with the surface atoms in a highly-charged surface [5,6]. But, at the time of writing, the more significant influence on β_A seems to be the field component due to a field-induced dipole moment in the surface emitter atoms [7,8].

The first published attempts [7,8] to calculate β_A were made in the context of the "isolated-dipole-pair" approximation. This method ignores all interactions except that between the given adsorbate atom and the dipole moment in the nearest substrate atom, and assumes that both atoms are in an applied field equal to F^{ext}. This leads to the result:

$$\beta_A = \frac{1 + 2b_E/4\pi\epsilon_0 s^3}{1 - 4b_A b_E/(4\pi\epsilon_0)^2 s^6}, \tag{4}$$

where b_E is the proper SI polarisability of the emitter surface atom, as defined in ref. [9]; ϵ_0 is the electric constant (permittivity of free space); and s is the separation of the dipoles representing the two atomic charge distributions.

In his recent review article [10], however, Tsong has proposed a basically different method for calculating the binding energy component ΔB^{el}. His new ("more general") method identifies ΔB^{el} with the sum of electrostatic interaction energies between the dipole in the given adsorbate atom and all other dipoles present. In the context of an isolated dipole pair type of approximation (and ignoring any zero-field dipole moment in the substrate atom), he identifies ΔB^{el} with the quantity $\Delta\epsilon$(T78) given by:

$$\Delta\epsilon(\text{T78}) = 2p_E(\text{ad})\, p_A(\text{ad})/4\pi\epsilon_0 s^3, \tag{5}$$

where p_E(ad) and p_A(ad) are the SI dipole moments [2,3] of the emitter atom and the adsorbate atom, respectively, when the adsorbate atom is adsorbed at the bonding site. These proposals are based on a reference to "unpublished work".

When eq. (5), and eqs. (2) and (4), are expanded it is found that these expressions for ΔB(conv.) and $\Delta\epsilon$(T78) differ by a term is s^6[10]. The nature of this discrepancy can be seen more clearly as follows. The field component $F_A^{d,E}$, acting on the adsorbate atom, due to the dipole in the emitter surface atom, is given by:

$$F_A^{d,E} = 2p_E(\text{ad})/4\pi\epsilon_0 s^3. \tag{6}$$

In the context of the isolated-dipole-pair approximation this field component constitutes the difference between the local field F_A^{loc} at the bonding site and the external field; so in this context we may write:

$$\Delta\epsilon(\mathrm{T78}) = p_{\mathrm{A}}(\mathrm{ad})\{F_{\mathrm{A}}^{\mathrm{loc}} - F^{\mathrm{ext}}\}. \tag{7}$$

Eq. (2), on the other hand, can be rearranged into the form:

$$\begin{aligned}\Delta B(\mathrm{conv.}) &= \tfrac{1}{2} b_{\mathrm{A}} \{F_{\mathrm{A}}^{\mathrm{loc}} + F^{\mathrm{ext}}\}\{F_{\mathrm{A}}^{\mathrm{loc}} - F^{\mathrm{ext}}\} \\ &= \tfrac{1}{2}\{p_{\mathrm{A}}(\mathrm{ad}) + p_{\mathrm{A}}(\mathrm{ext})\}\{F_{\mathrm{A}}^{\mathrm{loc}} - F^{\mathrm{ext}}\},\end{aligned} \tag{8}$$

where $p_{\mathrm{A}}(\mathrm{ext})$ is the SI dipole moment of the adsorbate atom when in the external field.

Clearly, eqs. (7) and (8) are not equivalent, because it is the full dipole moment at the bonding site that appears in eq. (7), but a mean dipole moment that appears in eq. (8). Tsong's new method *is not physically equivalent* to the conventionally accepted method for defining the electric-field-induced component of the short-range field-adsorption binding energy.

The reason is that there are serious conceptual errors in the new method. The first results from a failure to apply classical electrostatics correctly. When two classical dipoles are in an applied field F^{ext}, their total classical electrostatic potential energy (relative to a zero of potential where they are well separated and both in remote field-free space) is given by:

$$H = -\boldsymbol{p}_{\mathrm{A}} \cdot \boldsymbol{F}^{\mathrm{ext}} - \boldsymbol{p}_{\mathrm{E}} \cdot \boldsymbol{F}^{\mathrm{ext}} + \{\boldsymbol{p}_{\mathrm{A}} \cdot \boldsymbol{p}_{\mathrm{E}} - 3(\boldsymbol{k} \cdot \boldsymbol{p}_{\mathrm{A}})(\boldsymbol{k} \cdot \boldsymbol{p}_{\mathrm{E}})\}/4\pi\epsilon_0 s^3, \tag{9}$$

where $\boldsymbol{k}$ is a unit vector directed from the E-dipole to the A-dipole. In such a situation the electrostatic energy released ($\Delta\epsilon$), when the dipoles come together from positions in which they are well separated but both in the applied field F^{ext}, is given by:

$$\Delta\epsilon = H(\mathrm{separated}) - H(\mathrm{together}). \tag{10}$$

In the context of the isolated-dipole-pair approximation, where p_{A}, p_{E} and F^{ext} are parallel, we may use scalar notation and eq. (9) reduces to:

$$H = -p_{\mathrm{A}}F^{\mathrm{ext}} - p_{\mathrm{E}}F^{\mathrm{ext}} - 2p_{\mathrm{A}}p_{\mathrm{E}}/4\pi\epsilon_0 s^3. \tag{11}$$

Substituting into eq. (10), using $p_{\mathrm{A}}(\mathrm{ext})$ and $p_{\mathrm{E}}(\mathrm{ext})$ to denote the dipole moments when the dipoles are well separated but both in the applied field, we obtain:

$$\Delta\epsilon = \{p_{\mathrm{A}}(\mathrm{ad}) - p_{\mathrm{A}}(\mathrm{ext}) + p_{\mathrm{E}}(\mathrm{ad}) - p_{\mathrm{E}}(\mathrm{ext})\}F^{\mathrm{ext}} + 2p_{\mathrm{A}}(\mathrm{ad})\,p_{\mathrm{E}}(\mathrm{ad})/4\pi\epsilon_0 s^3. \tag{12}$$

As an estimate of $\Delta\epsilon$, the expression (5) derived from Tsong's work is in error because it represents only the last term in the full expression (12).

The second major conceptual error in ref. [10] is the implicit identification of the field-induced binding energy component ΔB^{el} with the electrostatic interaction energy $\Delta\epsilon$ between classical dipoles. Whilst this would be correct for atoms with *permanent* dipole moments, it is incorrect in the case of polarisable atoms. With polarisable atoms, part of the interaction energy $\Delta\epsilon$ is converted into internal electronic energy associated with the change in the dipole moments of the atoms. If the increase in this "internal" electronic energy, as the adsorbate atom approaches its

binding position from a position in the external field, is denoted by ΔV, then the energy available for conversion into kinetic energy of the motion of the dipole as a whole (i.e. ΔB^{el}) is given by:

$$\Delta B_{el} = \Delta\epsilon - \Delta V. \tag{13}$$

In the context of the isolated-dipole-pair approximation, ΔV is given by the sum of the increases in the polarisation energies of the adsorbate and emitter atoms, ΔV_A and ΔV_E respectively. These are given by the standard elementary expressions, that for ΔV_A being:

$$\Delta V_A = \tfrac{1}{2} b_A (F_A^{loc})^2 - \tfrac{1}{2} b_A (F^{ext})^2. \tag{14}$$

Writing F_A^{loc} as $(F_A^{d,E} + F^{ext})$, and substituting into eq. (14), and then carrying out the equivalent process for ΔV_E, leads to the result:

$$\Delta V = (b_A F_A^{d,E} + b_E F_E^{d,A}) F^{ext} + \tfrac{1}{2} b_A (F_A^{d,E})^2 + \tfrac{1}{2} b_E (F_E^{d,A})^2, \tag{15}$$

where $F_E^{d,A}$ is the field component acting on the emitter atom due to the dipole in the adsorbate atom.

Expression (12) can also be simplified, by setting:

$$p_A(\text{ad}) = b_A F^{ext} + b_A F_A^{d,E}, \quad p_E(\text{ad}) = b_E F^{ext} + b_E F_E^{d,A}, \tag{16}$$

into the bracketed terms in eq. (12). This leads to:

$$\Delta\epsilon = (b_A F_A^{d,E} + b_E F_E^{d,A}) F^{ext} + 2 p_A(\text{ad})\, p_E(\text{ad}) / 4\pi\epsilon_0 s^3. \tag{17}$$

Use of eq. (6), and then eq. (16) again, to remove the dipole moments in eq. (17), gives:

$$\Delta\epsilon = 2 b_A F_A^{d,E} F^{ext} + b_A (F_A^{d,E})^2 + b_E F_E^{d,A} F^{ext}. \tag{18}$$

Hence, on substituting eq. (18) and (15) into eq. (13), we obtain an expression for the field-induced binding energy component ΔB^{el}:

$$\Delta B^{el} = b_A F_A^{d,E} F^{ext} + \tfrac{1}{2} b_A (F_A^{d,E})^2 - \tfrac{1}{2} b_E (F_E^{d,A})^2. \tag{19}$$

In the context of the isolated-dipole-pair approximation, the conventional binding-energy expression eq. (2) can be reduced by similar manipulations to the form:

$$\Delta B(\text{conv.}) = b_A F_A^{d,E} F^{ext} + \tfrac{1}{2} b_A (F_A^{d,E})^2. \tag{20}$$

In practice, when the adsorbate atom is a noble gas atom, the quantity $F_E^{d,A}$ is relatively small, and the final term in eq. (18) and (19) can be neglected in a first approximation. In these circumstances it is easily seen that:

$$\Delta B(\text{conv.}) = \Delta B^{el} = \Delta V = \tfrac{1}{2}\Delta\epsilon \quad (\text{when } F_E^{d,A} \to 0). \tag{21}$$

Half (or in the more general situation approximately half) the classical electrostatic dipole interaction energy $\Delta\epsilon$ goes into the internal electronic energy of the polar-

isable adsorbate atom. In the process of desorption this energy is released again; so only half (or, more generally, approximately half) the interaction energy $\Delta\epsilon$ needs to be provided by the hypothetical external agent. Consequently ΔB^{el} is half (or approximately half) of $\Delta\epsilon$; so failure to distinguish between ΔB^{el} and $\Delta\epsilon$ will lead to error.

The conceptual errors described above have been demonstrated in the context of the isolated-dipole-pair approximation, but equivalent errors exist in the discussion of the interaction of arrays of dipoles given in ref. [10]. In consequence, although the analysis there may represent some sort of mathematical approximation, physically it is nonsensical. It seems unfortunate that such errors should exist in a review article, when the discussion is referenced as based on "unpublished work" by the reviewer.

The present analysis, however, also reveals a conceptual error in the more conventional approach, based on eq. (2) or eq. (20). Eq. (19) shows that, when the presence of the adsorbate atom affects the sources of the electric field acting on the atom, then it is not correct to simply substitute the self-consistent local field into eq. (2) or eq. (20). One must also take into account the energy changes associated with the induced changes in the source of the local field; in the context of the isolated-dipole-pair approximation this leads to the correction term $-\frac{1}{2}b_E(F_E^{d,A})^2$ that appears in eq. (19) but not in expression (20) for ΔB(conv.).

To get some feel for the magnitude of the discrepancy, numerical values can be derived for the isolated-pair situation. Taking the helium-on-tungsten situation as a paradigm, I shall use the numerical values: $s = 0.259$ nm [10]; $b_A = 0.143$ meV V^{-2} nm^2 [2,3,11]; $b_E = 11.7$ meV V^{-2} nm^2 [3,12]; $4\pi\epsilon_0 = 0.694456$ eV V^{-2} nm^{-1}; $F^{ext} = 45$ V nm^{-1}. Taking the quantity F_A^{loc} as the self-consistent local field given by eqs. (3) and (4), we obtain:

$$\Delta B(\text{conv.}) = 1.230 \text{ eV}, \tag{22}$$

$$\Delta B^{el} = 1.166 \text{ eV}. \tag{23}$$

On the other hand, if mutual induction is completely ignored, and F_A^{loc} in the above equations is set equal to the local applied field E_A given by:

$$E_A = (1 + 2b_E/4\pi\epsilon_0 s^3)F^{ext} \quad \text{(no mutual induction)}, \tag{24}$$

then formula (2) leads to the result:

$$\Delta B \text{ (no mutual induction)} = 1.106 \text{ eV}. \tag{25}$$

The precision with which these results are stated has little physical meaning, but in the context of the isolated-dipole-pair approximation they show that the correct result lies roughly midway between the "no mutual induction" and the "conventional" results, as might be expected intuitively. The size of the discrepancy between results (22) and (23) means that it should not be neglected in detailed treatments.

The binding-energy values given above are somewhat larger than might be expected on the basis of such experimental results as do currently exist [8,13,14]. Part of the explanation could be that the proper SI polarisability of an emitter surface atom were lower than the free-space value assumed here. But a major contributory factor is the unsatisfactory nature of the isolated-dipole-pair approximation, which can lead to overestimation of binding energy by a considerable amount [15].

Array-type models, which can take into account the mutual depolarisation of emitter surface atoms, should form a better basis for calculations, particularly in the case of field adsorption on net crystal planes. To demonstrate the theoretical principle behind such models, it is best to recast the equations relating to the isolated-dipole-pair approximation into a slightly different form. The first and last terms in eq. (9), or in eq. (11), may be taken together to give the electrostatic potential energy H_A of the adsorbate atom dipole in the field F_A^{loc}, thus:

$$H_A = -p_A F_A^{loc}. \tag{26}$$

From this, a change $\Delta\epsilon_A$ in the electrostatic interaction energy associated with the adsorbate atom dipole, as it goes from a position in the external field to the adsorption site, is given by:

$$\Delta\epsilon_A = p_A(\text{ad})\, F_A^{loc}(\text{ad}) - p_A(\text{ext})\, F^{ext}. \tag{27}$$

The change in the internal electronic energy of the adsorbate atom is given by ΔV_A as defined in eq. (14). Eq. (13) can thus be rewritten in the form:

$$\Delta B^{el} = \Delta\epsilon_A - \Delta V_A - \Delta U_S, \tag{28}$$

where ΔU_S is the change in the internal energy of the source of the (self-consistent) field acting on the adsorbate atom. In the context of the isolated-dipole-pair approximation the internal energy U_S is given by:

$$U_S = -p_E F^{ext} + \tfrac{1}{2} b_E (F_E^{loc})^2, \tag{29}$$

where F_E^{loc} is the local field acting on the emitter atom; the dipole in the emitter atom and the sources of the field F^{ext} together constitute the source of the field acting on the adsorbate atom.

With an array-type model eq. (28) still holds, but U_S is given by an infinite sum of terms. These represent the internal electronic energy of polarised atoms, interaction between dipoles, interaction of the dipoles with the monopoles that are the source of the field F^{ext}, and interaction between the monopoles. A full quantum-mechanical treatment, if this were possible, might also bring in terms relating to specifically quantum effects, such as exchange energies.

In general, the calculation of the changes in the internal energy U_S induced by the movement of adsorbate atoms may become exceedingly complicated. So, in practice, it may be easier to obtain values for the force acting on the adsorbate atom, as a function of distance, and then obtain ΔB^{el} by a numerical integration. With a simple array-type model this seems to be feasible, particularly for the calcu-

lation of mean binding energies (per atom) associated with the field adsorption of a layer of noble gas atoms; so it should be possible to deduce values for ΔU_S indirectly.

References

[1] E.W. Müller, S.B. McLane and J.A. Panitz, Surface Sci. 17 (1969) 430.

[2] This paper uses the internationally-recognised four-dimensional system of electrical quantities and equations, rather than the dimensionally-inconsistent semi-Gaussian system normally used in field-ion literature. I have, however, now dropped the prime originally used to distinguish an SI quantity from its Gaussian equivalent. Where necessary, equations and quantities taken from the literature have been replaced by their international-system equivalents.

[3] R.G. Forbes, Surface Sci. 64 (1977) 367.

[4] For example, A.D. Crowell, in: The Solid–Gas Interface, Ed. E.A. Flood (Dekker, New York, 1967).

[5] R.G. Forbes, Ph.D. thesis, Cambridge Univ. (1970).

[6] R.G. Forbes, Surface Sci. 78 (1978) L504.

[7] T.T. Tsong and E.W. Müller, Phys. Rev. Letters 25 (1970) 911.

[8] T.T. Tsong and E.W. Müller, J. Chem. Phys. 55 (1971) 2884.

[9] R.G. Forbes, J. Phys. D (Appl. Phys.) 11 (1978) L161.

[10] T.T. Tsong, Surface Sci. 70 (1978) 211.

[11] 1 meV V^{-2} $nm^2 \approx 1.602\,189 \times 10^{-40}$ J V^{-2} $m^2 \approx 4\pi\epsilon_0 \times 1.439\,976$ $Å^3$.

[12] J. Thorhallsson, C. Fisk and S. Fraga, J. Chem. Phys. 49 (1968) 1987.

[13] K.D. Rendulic, Surface Sci. 28 (1971) 285.

[14] A.D. Janssen and J.P. Jones, Surface Sci. 33 (1972) 553.

[15] R.G. Forbes and M.K. Wafi, unpublished work; details available from the present author.

Surface Science 93 (1980) 192–212

AN ARRAY MODEL FOR THE FIELD ADSORPTION OF HELIUM ON TUNGSTEN (111)

Richard G. FORBES and M.K. WAFI
University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET, UK

Received 29 March 1979; accepted for publication 1 October 1979

A theory of field adsorption has been developed, using a charged-surface model comprising an infinite regular array of superimposed monopoles and dipoles, together with a distant array of monopoles of the opposite polarity. Previous theoretical treatments are shown to be oversimplified, and the effects of various corrections to the present treatment are explored. For helium on tungsten (111) in an external field of 56 V/nm, the conventionally-defined field-induced contribution to the short-range binding energy is estimated to probably lie between 40 and 75 meV, and the total differential short-range binding energy is estimated to probably lie between 30 and 50 meV. As far as is known, this is the first detailed theoretical discussion of field adsorption on a specific crystal face, and a paradigm situation has been chosen for investigation.

1. Introduction

It is now well established that there is a layer of imaging gas firmly field-adsorbed onto the surface of an operating field-ion emitter, and various attempts have been made to calculate the binding energy associated with field adsorption [1–6]. It has usually been found convenient to split the total binding energy B into two parts: a long-range part, $\frac{1}{2}b_A(F^{ext})^2$, equal to the electric-field-induced binding energy of an imaging-gas atom in the (approximately) uniform field F^{ext} assumed to exist in space somewhat above the field-adsorbed layer; and a short-range part, ΔB, that leads to localised binding at specific sites immediately above emitter surface atoms.Thus:

$$B = \tfrac{1}{2}b_A(F^{ext})^2 + \Delta B\,, \tag{1}$$

where b_A is the SI polarisability [7,8] of the adsorbate (imaging-gas) atom. By definition, ΔB is the work that needs to be done (by a hypothetical external agent) to move the adsorbate atom from its binding site to a position in the "external" field F^{ext}, under specified conditions as to the behaviour of other adsorbate atoms present. (The nature of these "specified conditions" will become apparent later.)

There are components in ΔB resulting from London dispersion forces and from repulsive "interpenetration" forces [9], but it is normally assumed that the dominant

component is the electric-field-induced part, ΔB^{el}.

Most attempts to calculate ΔB^{el} have equated it with the quantity ΔB(conv.) given by:

$$\Delta B(\text{conv.}) = \tfrac{1}{2} b_A (F_A^{loc})^2 - \tfrac{1}{2} b_A (F^{ext})^2 = \tfrac{1}{2} b_A (\beta_A^2 - 1)(F^{ext})^2 , \quad (2)$$

where F_A^{loc} is the self-consistent local field [10] acting on the adsorbate atom when at the bonding site; and β_A is the dimensionless parameter:

$$\beta_A = F_A^{loc}/F^{ext} . \quad (3)$$

Calculations of β_A, or of the equivalent parameter f_A ($=\beta_A^2$), can be performed within the context of various models and approximations. The best known approach [1,3] has been to treat field adsorption as equivalent to the interaction between a neutral emitter atom and an adsorbate gas atom in an applied field equal to F^{ext}, ignoring all other factors. This leads to the formula [8]:

$$\beta_A = [1 + 2b_E/4\pi\epsilon_0 s^3]/[1 - 4b_A b_E/(4\pi\epsilon_0)^2 s^6] \quad (4)$$

where s is the separation of the equivalent dipoles, and b_E is the *proper* SI polarisability of the surface emitter atom. (By "proper SI polarisability" we mean the parameter defined in terms of the self-consistent local field, so that in the absence of any permanent dipole moment the SI dipole moment p is given by: $p = bF^{loc}$.) This result, eq. (4), will be called the "isolated dipole pair" (or "IDP") approximation.

Forbes, however, has recently pointed out that the conventional conceptual treatment of field adsorption ignores induced changes in the internal energy of the source of the electric field [11]. A change ΔU_S is induced by the presence of the adsorbate atom(s), and consequently:

$$\Delta B^{el} = \Delta B(\text{conv.}) - \Delta U_S . \quad (5)$$

In the context of the isolated-dipole-pair (IDP) approximation it has been possible to show that ΔU_S is small relative to ΔB(conv.). Consequently, although ΔU_S must not be neglected in detailed treatments, it should be a reasonable approximation to set ΔB^{el} equal to ΔB(conv.). The major concern of the present paper is with the calculation of ΔB(conv.).

As already stated, the first published attempts to calculate ΔB(conv.) were in the context of the IDP approximation [1,3]. Concurrently, one of the present authors (R.G.F.) was attempting to develop a model which attributed localised field adsorption to the locally-high field resulting from the discrete positive charge nearest to the adsorbate atom [2]. With the appearance of ref. [1] it became obvious that both the dipole and the monopole effects associated with emitter substrate atoms ought in principle to be taken into account.

With the passage of time it has also become clear that it is unlikely to be a good approximation to disregard interaction with distant neighbours, particularly when discussing field adsorption on net crystal planes. Consequently one turns to array-type models.

Some work with finite planar arrays of monopoles has been carried out by the present authors and colleagues. The first steps were briefly reported some years ago [12], but this approach subsequently encountered difficulties arising from finite-size effects, and it became clear that the best procedure was to first explore the properties of models involving infinite regular planar arrays.

Equations relating to the interaction of two square arrays of dipoles in free space in an applied field F^{ext} have been given by Tsong [5]. He seems to conclude, though without giving the argument in full, that the accuracy of eqs. (2)–(4) is sufficient for comparing with experimental results. For the noble imaging gases, values of f_A derived by the IDP approximation are tabulated in ref. [5].

Tsong's approach, however, ignores possible monopole effects. Forbes [6,13] has thus introduced a charged-surface model in which the surface is represented by an infinite regular planar array of superimposed monopoles and dipoles, together with a distant array of monopoles of the opposite sign (needed for electrostatic self-consistency). With such a model the parameter β_A defined earlier can, in a first approximation, be split into two components β_A^m and β_A^d, resulting from monopole and dipole effects respectively. And the parameter β_A^m can be written in the form $(1 + \delta_A^m)$, so that:

$$\beta_A = \beta_A^m + \beta_A^d = 1 + \delta_A^m + \beta_A^d \,. \tag{6}$$

δ_A^m represents effects due to the discrete nature, and hence the *localisation*, of real positive charge.

The existing treatments in the literature [1,3,5] represent approximations in which $\beta_A^m = 1$ ($\delta_A^m = 0$). Hence values for β_A^d can be deduced from the values given for f_A. Comparing these derived values of β_A^d with calculated values of β_A^m, Forbes [6] inferred that localised field adsorption would (as widely believed) result chiefly from the dipole field of the emitter substrate atoms, – but that the effects of the localised nature of real positive charge should not be neglected in detailed treatments of field adsorption. Reservations were, however, expressed about the existing treatments of substrate-dipole effects.

More recently [11], it has been shown that there are conceptual errors both in the derivation of eq. (2) and in Tsong's new treatment [5] of dipole interactions in an applied field. This paper thus presents an independent calculation of β_A^d, and a more general calculation of β_A.

The structure of the paper is as follows. Section 2 contains the analysis of an array model for field adsorption. Using parameters appropriate to the helium-on-tungsten situation, section 3 then discusses numerical inadequacies involved in approximate calculations of ΔB(conv.). Sections 4 and 5 then discuss other factors affecting the estimations of first ΔB^{el}, and then ΔB. Section 6 summarises results achieved, and discusses other possibilities.

2. Analysis of the array model

The treatment here assumes a model in which the field-adsorbed gas layer is represented by a complete planar array of dipoles, one situated above every lattice point of Forbes' [6] model for a charged surface, as described above. The treatment forms an exact mathematical analysis for the quantity ΔB(conv.) in this model. Questions as to whether the model charge distribution is a good representation of the actual charge distribution at a crystallographically flat metal surface, covered with a field-adsorbed layer, are beyond the scope of the present paper, but we certainly believe the model to be satisfactory as a "first generation" treatment.

Consider a layer of atoms, of proper SI polarisability b, forming a simple regular planar lattice; and let the interatom spacing be a. If an electric field $F^{\rm imp}$, normal to the plane of the array, is impressed on each atom in the layer by charge distributions outside the layer, then the self-consistent local field $F^{\rm loc}$ acting on each atom in the layer is given by:

$$F^{\rm loc} = F^{\rm imp} - \sum_n p/(4\pi\epsilon_0\rho_n^3) = F^{\rm imp} - K_1 p/4\pi\epsilon_0 a^3 \ . \tag{7}$$

The summation term represents the depolarising field, at any atom, due to all the other atomic dipoles in the layer, each of dipole moment p; and ρ_n is the distance of the nth such dipole from the given atom. K_1 is a structure factor defined by eq. (7) and obtained by carrying out the summation. For a square lattice $K_1 \simeq 9.034$; for a hexagonal lattice $K_1 \simeq 11.034$ [14].

We shall initially assume that neither the emitter nor the adsorbate atoms have any dipole moment at zero local field. Hence, on setting $p = bF^{\rm loc}$ in eq. (7), and re-arranging slightly, we have:

$$F^{\rm loc} = M^{-1}F^{\rm imp} \ , \tag{8}$$

$$M = 1 + bK_1/4\pi\epsilon_0 a^3 \ . \tag{9}$$

The parameter M defined by eq. (9) plays the role of a relative permittivity for the layer, as has been pointed out by MacDonald and Barlow [15].

To obtain some typical values for M, we shall consider three lattice structures: first, a square lattice with a equal to the interatom spacing in the (111) face of tungsten, namely 0.4476 nm; second, a hexagonal lattice with the same value of a; thirdly, a square lattice having the same area per lattice point as the hexagonal lattice. This last has the interatom spacing equal to $(3^{1/4}/2^{1/2}) \times 0.4476$ nm, that is, 0.4165 nm.

The SI polarisability of helium, b(He), is taken as 0.143 meV V^{-2} nm^2 [7,16]. For the SI polarisability of a tungsten surface atom three values will be used.
(1) $b(W_T) = 3.19$ meV V^{-2} nm^2, which is the value used by Tsong and Müller; this will enable comparisons to be made between the present results and previous work.
(2) $b(W_{MB}) = 7$ meV V^{-2} nm^2, which is the value recommended by Miller and

Bederson in their recent review of atomic and molecular polarisabilities [17]. The Miller-Bederson value is obtained by "scaling" the value calculated by Thorhallsson et al. [18]; although the estimated accuracy of the result is stated as ±50%, we feel the "scaled" value to be a better estimate of the polarisability of a neutral tungsten atom in free space than is the theoretical value 11.7 meV V^{-2} nm^2 [18] used in previous work.

(3) $b(W_{S/F}) = 2$ meV V^{-2} nm^2, which is the value deduced by Forbes [19] from helium-ion energy measurements recently carried out by Sakurai and co-workers [20]. The Gaussian polarisabilities corresponding to the above SI quantities are, respectively: 0.206 $Å^3$, 4.6 $Å^3$, 10 $Å^3$, and 2.9 $Å^3$.

Values of M are shown in table 1. The IDP approximation has also been formally included in the theory, by setting M equal to unity. For reference purposes in calculations, parameters are shown in this and in the other tables to more places of decimals than are physically significant.

Next consider a model representing the field adsorption situation, with two such layers aligned with the atoms in one exactly over the atoms in the other, the two layers being separated by a distance s. Parameters relating to the emitter-atom layer will be subscripted "E", parameters relating to the adsorbed imaging-gas layer "A".

The impressed field F_A^{imp} acting on the adsorbed gas atoms has two components. The first (F_A^m) results from the monopole distribution, and can be written in the form:

$$F_A^m = \beta_A^m F^{ext} , \tag{10}$$

where F^{ext} and β_A^m have the same meanings as earlier. The value of β_A^m is readily calculated, for a given value of the ratio $\nu = s/a$. For helium on tungsten we shall assume that $s = 0.259$ nm [5]. The corresponding values of ν and β_A^m are shown in table 2.

The second component (F_A^{dE}) results from the dipoles in the emitter-atom

Table 1
Values of the parameter M, which plays the role of a relative permittivity for a layer of atoms, for specified lattice structures and polarisability values

Type of lattice	a (nm)	b(He) 0.143 $b^\ominus$	$b(W_{S/F})$ 2.00 $b^\ominus$	$b(W_T)$ 3.19 $b^\ominus$	$b(W_{MB})$ 7.00 $b^\ominus$
Square	0.4476	1.021	1.290	1.463	2.015
Hexagonal	0.4476	1.025	1.354	1.565	2.240
Square	0.4165	1.026	1.360	1.574	2.260
IDP aprx.		1.000	1.000	1.000	1.000

$b^\ominus = 1$ meV V^{-2} nm^2 [16].

Table 2
Values of the distance fraction ν, the monopole-induced field ratio β_A^m, and the structure factor S, for specified lattice structures

Type of lattice	a (nm)	ν	β_A^m	S
Square	0.4476	0.5786	1.067	5.934
Hexagonal	0.4476	0.5786	1.048	5.297
Square	0.4165	0.6218	1.050	4.322
IDP aprx.	0.4476	0.5786	1.000	10.323

layer (the "E-layer"), and can be written in the form:

$$F_A^{dE} = S^E p_E / 4\pi\epsilon_0 a^3 , \tag{11}$$

where S^E is a "structure factor" (corresponding to K_2 in ref. [5]), relating to the relative positions of the dipoles in the two layers, defined as follows. If ρ_n denotes the distance of the nth dipole in the E-layer from a specified atom in the adsorbate layer (the "A-layer"), then we may define dimensionless parameters r_n and ν by:

$$r_n = \rho_n / a , \qquad \nu = s/a . \tag{12}$$

In terms of these, S^E is given by:

$$S^E = \sum_n (3\nu^2/r_n^5 - 1/r_n^3) . \tag{13}$$

The nth term in this summation is a geometrical factor that determines the strength of the field component (normal to the array plane), acting on the specified A-layer atom, due to the nth dipole in the E-layer. The summation is over all dipoles in the E-layer. For reasons of symmetry there can be no lateral component in the total field acting on an A-layer atom, so summation over lateral field components need not be considered.

For a given lattice structure and given value of ν, S^E is readily evaluated by a combination of numerical and analytical means [21]. The case where all interactions are neglected except that with the nearest dipole in the E-layer is equivalent to the IDP approximation, and can be formally included in the theory by setting:

$$S^E \approx 2/|\nu|^3 \quad \text{(IDP approximation)} . \tag{14}$$

Values of S^E are shown in table 2.

We shall also need the structure factor S^A, which determines the field at a specified atom in the E-layer resulting from all the dipoles in the A-layer. Since the two layers have the same lattice structure, and eqs. (13) and (14) are even functions of ν, it follows that $S^A = S^E$. Hence the suffix is omitted in table 2.

In eq. (11), if we express p_E in the form $b_E F_E^{loc}$, we can then define a dimensionless parameter γ_A^E by:

$$\gamma_A^E = S^E b_E / 4\pi\epsilon_0 a^3 , \tag{15}$$

and hence write eq. (11) in the abbreviated form:

$$F_A^{dE} = \gamma_A^E F_E^{loc} . \tag{16}$$

The notational principle used here is that the lower suffix position represents the layer where the "effect" is, and the upper suffix position indicates the type or source of effect. γ_A^E is readily calculated for given values of the structure factor S^E and of the proper SI polarisability (b_E) of the emitter atom. Values of this and of the analogous parameter γ_E^A are shown in table 3, for the helium-on-tungsten situation.

Collecting eqs. (8), (10) and (16) together, we finally obtain an equation for the local field acting on an A-layer atom:

$$F_A^{loc} = M_A^{-1} F_A^{imp} = M_A^{-1}\beta_A^m F^{ext} + M_A^{-1}\gamma_A^E F_E^{loc} . \tag{17}$$

By direct analogy, we may obtain an expression for F_E^{loc} by interchanging "E" and "A", wherever they occur, to give:

$$F_E^{loc} = M_E^{-1} F_E^{imp} = M_E^{-1}\beta_E^m F^{ext} + M_E^{-1}\gamma_E^A F_A^{loc} . \tag{18}$$

Substituting eq. (18) into eq. (17) and solving for F_A^{loc} then leads to an expression for β_A:

$$\beta_A = \frac{F_A^{loc}}{F^{ext}} = \frac{M_A^{-1}\beta_A^m + M_A^{-1}\gamma_A^E M_E^{-1}\beta_E^m}{1 - M_A^{-1}\gamma_A^E M_E^{-1}\gamma_E^A} . \tag{19}$$

From this the enhancement factor f_A can be calculated.

A rough physical interpretation of eq. (19) can be made as follows. (1) The term $M_A^{-1}\beta_A^m$ represents a contribution to β_A resulting from the monopole charge distri-

Table 3
Values of the coefficient, γ_A^E defined by eq. (16), and of the analogous coefficient γ_E^A, for specified lattice structures and polarisability values

Type of lattice	a (nm)	γ_E^A(He)	$\gamma_A^E(W_{S/F})$	$\gamma_A^E(W_T)$	$\gamma_A^E(W_{MB})$
Square	0.4476	0.01363	0.1906	0.3040	0.6671
Hexagonal	0.4476	0.01219	0.1701	0.2713	0.5954
Square	0.4165	0.01232	0.1722	0.2747	0.6029
IDP aprx.	0.4476	0.02370	0.3315	0.5288	1.160

bution, reduced as a result of the mutual depolarising influence of the adsorbate-layer atoms. (2) The second term in the numerator represents a contribution from the emitter-atom dipoles, similarly reduced. (3) The denominator represents an enhancement effect due to the influence of the adsorbate-layer dipoles on the emitter-layer dipoles.

Owing to the symmetry of the derivation, a formula for β_E can be derived from eq. (19) by replacing "A" by "E", and vice-versa, wherever these occur.

3. The calculation of ΔB(conv.)

For an initial discussion of formula (19) it is convenient to use an approximation in which M_A^{-1} is set equal to unity and γ_E^A is set equal to zero. Physically, this corresponds to ignoring the effects of the adsorbate-layer dipoles upon themselves and upon the emitter-layer dipoles; this is a reasonable approximation, as will be shown later.

In this case eq. (19) reduces to:

$$\beta_A = \beta_A(\text{aprx.}) \equiv \beta_A^m + \gamma_A^E M_E^{-1} \beta_E^m \equiv \beta_A^m + \beta_A^d(\text{aprx.}), \tag{20}$$

where β_A(aprx.) and β_A^d(aprx.) are approximate values of β_A and β_A^d, respectively, defined by eq. (20).

Eq. (20) is the approximation employed (though not explicitly stated) in ref. [6] and in the first section of the present paper. With the help of eq. (15), β_A^d(aprx.) can be written in the form:

$$\beta_A^d(\text{aprx.}) = b_E \beta_E^m S^E M_E^{-1} / 4\pi\epsilon_0 a^3 , \tag{21}$$

which shows more clearly the parameters on which it depends.

Discussion of previous treatments. Various choices for the parameters in eq. (21) are given in table 4, together with the corresponding calculated values of β_A(aprx.) and $(f_A - 1)$, assuming a square lattice with $a = 0.4476$ nm. Also shown are the corresponding values of the binding-energy component ΔB(conv.), evaluated for the field value $F^{ext} = 56$ V/nm.

The various choices correspond to different approximations and assumptions. Set 1 is equivalent to eq. (4), but with the denominator set equal to unity rather than its calculated value of 0.987. For practical purposes, therefore, set 1 is equivalent to the IDP approximation. Physically, this choice disregards the fact that the emitter substrate atoms are the most exposed part of a highly-charged metal surface, and ignores all interactions except that between an adsorbate atom and the dipole associated with the nearest substrate atom.

In set 2 the parameters M_E^{-1} and S^E are given the values appropriate to a simple square dipole array, the other parameters remaining unchanged. Physically, this choice describes the interaction between a layer of neutral emitter atoms and a layer of adsorbate atoms, isolated in free space in an applied field F^{ext} [22]. The

Table 4
Values of specified field ratios and of the conventional short-range field-adsorption binding energy ΔB(conv.), for a square lattice with a = 0.4476 nm, for specified choices of the parameters appearing in eq. (21) and of β_A^m

	Choice					
	1	2	3	4	5	6
$b_E/b^{\ominus}$ [a]	3.19	3.19	3.19	3.19	7.00	7.00
β_E^m	1.0	1.0	0.5	0.5	0.5	0.5
S^E	10.32	5.934	5.934	5.934	5.934	5.934
M_E^{-1}	1.000	0.6836	0.6836	0.6836	0.4962	0.4962
β_A^d (aprx.)	0.529	0.208	0.104	0.104	0.165	0.165
β_A^m	1.000	1.000	1.000	1.067	1.000	1.067
β_A (aprx.)	1.529	1.208	1.104	1.171	1.165	1.232
f_A^{-1}	1.337	0.460	0.219	0.372	0.358	0.519
ΔB(conv.) [b]	0.299	0.103	0.049	0.083	0.080	0.116

[a] $b^{\ominus}$ = 1 meV V^{-2} nm^2 [16].

[b] ΔB(conv.) is evaluated for an external field of 56 V/nm, using the approximate value of the field-ratio shown, and is expressed in eV.

unapproximated array equations given by Tsong [5] correspond to this choice.

With set 3 the monopole-caused field acting on a substrate atom is set equal to $\frac{1}{2}F^{ext}$ ($\beta_E^m = 0.5$), rather than to F^{ext}. This is because, in reality, the emitter atoms form the surface of a charged conductor. Consequently they are in part the *source* of the external field. The only monopole-induced field acting on them is that due to the distant charge needed for electrostatic self-consistency [23,24]. This argument is valid whether or not the positive charge distribution is regarded as localised laterally.

Set 4 then takes into account the additional effects resulting from the localisation of positive charge and the consequent enhancement of field values at the adsorption positions above lattice sites; this is done by setting β_A^m equal to 1.067 rather than unity.

Within the overall framework of the more general approximations involved in eqs. (20) and (21), this sequence of choices represents a trend towards increasing realism in general mathematical and physical assumptions.

The numerical results in table 4 show clearly that the Tsong and Müller dipole–dipole theory (our IDP approximation) is not an adequate substitute for a full array calculation. With the data choice used here, the neglect of mutual depolarisation amongst substrate-atom dipoles causes overestimation of the binding energy by a

factor of nearly three, by itself. The degree of overestimation will, obviously, depend somewhat on the values assumed for the lattice spacing and for the surface-atom polarisability. But in many situations we would disagree with Tsong's statement, based on unpublished work cited in his review article [5], that the IDP approximation leads to numerical results sufficiently accurate for comparing with experimental results.

Monopole effects. A second point concerns the role of the field variations due to the discreteness of positive charge. The original comments on this by Forbes [6] compared the monopole contribution to field variations with the dipole contribution as calculated by the IDP approximation. However, by reference to the full-array calculation of the substrate-dipole contribution (choice 3), the monopole contribution is of relatively more significance. Including this (choice 4) increases the predicted short-range binding energy by approximately 70%, for the SI polarisability value 3.19 meV V^{-2} nm^2.

Obviously, the relative importance of monopole effects depends on the value of the surface-atom polarisability. For the somewhat higher value of 7 meV V^{-2} nm^2 used in choices 5 and 6 in table 4, the inclusion of the monopole contribution increases the predicted short-range binding energy by around 45%. (Numerical values are sensitive to the assumed lattice structure and interatom spacing, but the qualitative trends exhibited in table 4 will always be present.)

Lattice structure. The analyses in table 4 are based on a square lattice with a = 0.4476 nm, because this structure was assumed in previous work [6]. Table 5 shows how the results for this structure compare with those for the two other structure choices specified earlier, assuming the polarisability value b_E = 7 meV V^{-2} nm^2. (The binding-energy estimate derived here is based on the full formula for ΔB(conv.), namely eq. (19).)

The first three columns show that for a given value of a the use of a hexagonal rather than a square lattice leads to smaller predictions for β_A(aprx.) and β_A. Physically, this is because the lattice points are closer together: the monopole contribution to field variations is reduced, and mutual depolarisation effects are enhanced. Comparison of columns 2 and 3, however, shows that results for square and hexagonal lattices are much closer if the areas per lattice site (rather than the interatom spacings) are set equal.

We may conclude that, in modelling field adsorption on a crystal net plane, it is best to use a model with the real lattice structure of the plane, but that little error is introduced by using a square lattice with the correct area per lattice site. In other adsorption contexts this result has been known for a long time [15].

Adsorbate-layer induction effects. For clarity in the discussion of previous calculations, the analysis in table 4 used the approximate formulae, eq. (20). Table 5 shows that the true model-predicted field-ratio β_A, obtained using the full formula eq. (19), is slightly less than the approximate quantity β_A(aprx.). Physically this is the result of induction effects resulting from the presence of the adsorbate layer. The correction to the field ratio is small, about 2% in the case of each lattice struc-

Table 5
Values of the parameters appearing in eq. (23), and of the conventional short-range field-absorption binding-energy ΔB(conv.), evaluated for specified lattice structures, without approximations

Lattice type	Square	Square	Hexagonal	Hexagonal
a (nm)	0.4476	0.4165	0.4476	0.4476
b (meV V^2 nm^2)	7.00	7.00	7.00	2.00
β_A (aprx.)	1.233	1.183	1.181	1.111
M_A^{-1}	0.9800	0.9749	0.9753	0.9753
$M_A^{-1}\beta_A$ (aprx.)	1.208	1.153	1.151	1.083
D^{-1}	1.0044	1.0032	1.0032	1.0015
β_A	1.213	1.157	1.155	1.085
ΔB(conv.) [a] (eV)	0.106	0.076	0.075	0.040

[a] ΔB(conv.) is evaluated for an external field of 56 V/nm.

ture when b is taken as 7 meV V^{-2} nm^2. The correction to ΔB(conv.), however, is much greater. For the hexagonal lattice structure, for example, the full-formula result is approximately 15% lower than the approximate result.

It is easier to distinguish the two effects operating to produce this result if the symbol D is used to denote the denominator of eq. (19). Thus:

$$D \equiv 1 - M_A^{-1}\gamma_A^E M_E^{-1}\gamma_E^A . \tag{22}$$

The relationship between β_A and β_A(aprx.) can then be written in the form:

$$\beta_A = D^{-1}M_A^{-1}\beta_A(\text{aprx.}) . \tag{23}$$

The factor M_A^{-1} results from interaction within the adsorbed layer, and acts to reduce the local field acting on an adsorbed atom; the factor D^{-1} results from mutual induction between the adsorbed atoms and the emitter substrate atoms, and acts to increase the local field acting on an adsorbed atom. As shown in table 5, the factor M_A^{-1} has a somewhat greater influence than the factor D^{-1}.

In practice, the use of eq. (19) rather than eq. (20) involves negligible additional mathematical work. So we may conclude that, although eq. (20) provides a more-or-less reasonable first approximation, eq. (19) should normally be used for the calculation of ΔB(conv.) in the present model.

Choice of polarisability value. Comparison of the last two columns in table 5 shows that, after the easily avoidable approximations in the full array calculation of ΔB(conv.) have been removed, there is still an uncertainty left that is associated with the choice of proper SI polarisability for the surface atoms. Unfortunately, uncertainty over the correct value of b_E is one of the most pressing difficulties of field adsorption theory. Values mentioned in field-ion literature for the "polarisa-

bility" of a tungsten surface atom range from 0.38 meV V^{-2} nm^2 (for an atom in the (011) plane [25]) to the Thorhallsson et al. value of 11.7 meV V^{-2} nm^2, which is a computed value for an atom in free space. None of the older experimental determinations of surface-atom "polarisability" by field-ion microscope experiments constitute measurements of our parameter b_E [19]. And on the theoretical side there is, as yet, no reliable guide as to how much the proper SI polarisability of a (partially ionized?) surface atom may differ from the SI polarisability of the same atom (neutral) in free space.

As things currently stand, there seem to be two working values that have some sort of objective status. The first is the Miller–Bederson theoretical value, 7 meV V^{-2} nm^2; but there is substantial uncertainty associated with this value, and its relevance to the surface-atom situation could also be challenged. The second is the empirical working value, 2 meV V^{-2} nm^2, derived in a slightly indirect fashion by Forbes [19] from the experimental work of Culbertson et al. [20]. Although we have relatively more confidence in the latter value than the former, it is not yet clear that all the necessary corrections in the derivation of this empirical value have been properly taken into account; so there is uncertainty associated with this value too. It is lower than we intuitively expected the value of b_E to be.

For definiteness in investigating further corrections, we shall work with these two values, hoping that the true value lies somewhere between them. There seems a need for a more thorough investigation of the whole question of surface-atom polarisability, but this is beyond the scope of the present paper.

Summary. This section has largely been concerned with possible mathematical approximations in the calculation of the conventional field-induced binding-energy contribution ΔB(conv.). It has been shown that: (1) The isolated-dipole-pair approximation is not an adequate substitute for a full array calculation. (2) Mutual depolarisation, and also the field variations due to the discreteness of positive charge, must be taken into account. (3) Mutual induction effects due to the adsorbate layer can be neglected in a first approximation, but that there is no problem in including them. (4) The correct lattice structure for the crystallographic plane under discussion should be used. (5) There is a residual uncertainty by a factor of about two, due to uncertainty over the correct value of surface atom polarisability.

4. The calculation of ΔB^{el}

We now move on the consider some further corrections that must be considered when calculating the full field-induced short-range binding-energy contribution ΔB^{el}. The first two corrections concern the size of the atomic dipole moments when the field-adsorbed layer is present.

Zero-external-field dipole moment. The analyses earlier have assumed that the atomic dipole moments were directly proportional to the local field F^{loc} acting on

the atom. In principle the relationship ought to include other terms, and in particular it ought to include the SI dipole moment μ that would exist under conditions of zero local field. Ignoring second-order and higher terms, we may write:

$$p^{\mathrm{tot}} = \mu + bF^{\mathrm{loc}}. \tag{24}$$

For clarity, p^{tot} is here used to represent the total SI dipole moment when the external field is present; p^{tot} replaces p in earlier equations.

For a noble-gas adsorbate $\mu_{\mathrm{A}} = 0$, but it cannot be assumed that for an emitter substrate atom μ_{E} is zero. Within the framework of the model on which section 2 is based, this may be taken into account by substituting for b_{E} an "effective" value $b_{\mathrm{E}}(\mathrm{eff})$ given by:

$$p_{\mathrm{E}}^{\mathrm{tot}} = b_{\mathrm{E}}(\mathrm{eff})\, F_{\mathrm{E}}^{\mathrm{loc}} = b_{\mathrm{E}}(1 + \mu_{\mathrm{E}}/b_{\mathrm{E}}F_{\mathrm{E}}^{\mathrm{loc}})\, F_{\mathrm{E}}^{\mathrm{loc}} . \tag{25}$$

To obtain an estimate of the size of the correction factor, we shall assume that the electrostatic component ψ^{e} of the local work function for a given face [26] is caused by the zero-external-field dipole moment of the atoms in the array [27]. This leads to the following relationship [28]:

$$\psi^{\mathrm{e}} = -e\mu_{\mathrm{E}}/M_{\mathrm{E}}La^2\epsilon_0 , \tag{26}$$

where e is the elementary charge and La^2 is the area per lattice point ($L = 1$ for a square lattice; $L = 3^{1/2}/2$ for a hexagonal lattice).

An exact value of ψ^{e} for the (111) face of tungsten is difficult to find. For the sake of argument we shall here take it to be of order -0.5 eV. Using this value, and parameters appropriate to the hexagonal lattice considered earlier, with $b_{\mathrm{E}} = 7$ meV V^{-2} nm^2, gives:

$$\mu_{\mathrm{E}} \approx 0.01\, e\ \mathrm{nm} . \tag{27}$$

To obtain an estimate of the local field $F_{\mathrm{E}}^{\mathrm{loc}}$, it is sufficient to ignore the influence of the adsorbate atoms and put:

$$F_{\mathrm{E}}^{\mathrm{loc}} = M_{\mathrm{E}}^{-1}\beta_{\mathrm{E}}^{\mathrm{m}}F^{\mathrm{ext}} . \tag{28}$$

Further, it is sufficient in a first approximation to evaluate M_{E}^{-1} using the "uncorrected" value of b_{E}. For the hexagonal lattice structure and polarisability value under discussion, using $\beta_{\mathrm{E}}^{\mathrm{m}} = 0.5$, and setting F^{ext} equal to the "low" value 30 V/nm, we obtain $b_{\mathrm{E}}F_{\mathrm{E}}^{\mathrm{loc}}$ equal to $\sim 0.05\, e$ nm. Hence we obtain a correction factor:

$$1 + \mu_{\mathrm{E}}/b_{\mathrm{E}}F_{\mathrm{E}}^{\mathrm{loc}} \approx 1.2 . \tag{29}$$

With $b_{\mathrm{E}} = 2$ meV V^{-2} nm^2 the equivalent correction factor is approximately 1.3. The size of the correction factor will, obviously, depend on the value assumed for the electrostatic component of the local work-function, and would vary with crystallographic orientation.

The calculations made here perhaps tend to overestimate the effect of the zero-external-field dipole moment. Since the error involved in using b_{E} rather than

b_E(eff) in the analysis earlier is less than the uncertainty over the value of b_E itself, it seems satisfactory to neglect the effects of any zero-field dipole moment for the present. However, the correction is not really negligible, and the figures here suggest that a more thorough treatment will eventually be required.

Higher moments and field inhomogeneity. The influence of higher moments and field inhomogeneity on the dipole moment induced in the emitter atom could, in principle, be handled in a manner similar to that used above. That is, b_E in the earlier analysis could be replaced by an effective value that, for a given external field, takes such things into account. However, at the present time there seems no obvious way of making realistic estimations. It is felt that any error arising from this cause will probably be swamped by the general uncertainty over the value of b_E, and can thus be ignored.

Higher moments and field inhomogeneity associated with the adsorbate atom should also, in principle, be taken into account. With respect to field inhomogeneity, a similar problem arises in the adsorption of noble gases on ionic solids, and was dealt with many years ago; the early treatments are reviewed by Crowell [9] and Brunauer [29]. In the case of field adsorption the effect of field inhomogeneity has not been investigated in detail, but we have formed the impression that any binding-energy correction due to this will be relatively small.

It has sometimes been suggested, for example by Holscher [30], that a "hyperpolarisability" term, in $(F^{\text{loc}})^4$, should be included in the expression for the polarisation energy of a noble-gas atom. But we have never seen any evidence that this term would produce a significant correction.

Internal energy of the field source. According to eq. (5), the conventional estimate of the field adsorption binding energy must be reduced by an amount equal to the induced change in the internal energy of the source of the electric field. Two possibilities exist in the context of layer adsorption, depending on precisely what the mechanism of adsorption is assumed to be, and consequently some care must be exercised in defining the source of the field.

Theoretically, the simpler procedure is to regard the adsorption process as the adsorption of an already ordered planar *layer* of adsorbate atoms, this layer having been brought together in a region of space that is distant from the surface but throughout which there is a uniform field F^{ext}. These two stages, taken together, would define a *mean* (or *integral*) short-range binding energy ΔB(int.), in which an electric-field-induced component ΔB^{el}(int.) could be identified.

However, the quantity of more interest in the field-ion situation is the so-called *differential* short-range binding energy, ΔB(diff.), which is the work needed to remove a single atom from the field-adsorbed layer to a position in the external field somewhat above the layer. This quantity can also be viewed as the work needed to create a "vacancy" in the field-adsorbed layer.

This short-range binding energy can be expressed as a sum of components:

$$\Delta B(\text{diff.}) = \Delta B^{\text{el}}(\text{diff.}) + \Delta B^{\text{disp}} + \Delta B^{\text{rep}} + \Delta B^{\text{indir}} . \tag{30}$$

These represent work done against, respectively, electric-field-induced forces, dispersive forces, repulsive forces, and any forces that may exist as a result of indirect quantum-mechanical interactions between the adsorbate atoms mediated via the substrate (if these in fact exist) [31]. We are concerned here with the first of these terms; the others are discussed in section 5.

If the dipole moments in all the other atoms remained fixed whilst an adsorbate-layer vacancy was created by removing an atom, then $\Delta B^{\mathrm{el}}(\mathrm{diff.})$ would be identical with $\Delta B(\mathrm{conv.})$. However, the sizes of the dipole moments in the other atoms (and consequently the strengths of the field and force acting on the vacancy-site atom) are dependent on the instantaneous position of the vacancy-site atom and on the strength of the dipole induced in this atom. Consequently $\Delta B^{\mathrm{el}}(\mathrm{diff.})$ differs from $\Delta B(\mathrm{conv.})$, and formally we may write:

$$\Delta B^{\mathrm{el}}(\mathrm{diff.}) = \Delta B(\mathrm{conv.}) - \Delta U_{\mathrm{S}}(\mathrm{diff.}) , \tag{31}$$

where $-\Delta U_{\mathrm{S}}(\mathrm{diff.})$ is the change in the internal energy of the "source" as the vacancy-site atom is removed, the source in this case being constituted by the distant charge distribution together with all substrate and adsorbate atoms other than the vacancy-site atom.

The evaluation of $\Delta B^{\mathrm{el}}(\mathrm{diff.})$, by integration of the force acting on the vacancy-site atom along a path away from its adsorption site, is possible in principle but would certainly involve extensive, tedious, numerical calculations. Therefore one looks for an alternative approach. We think that the following argument may provide an estimate of $\Delta B^{\mathrm{el}}(\mathrm{diff.})$, though we put it forward with considerable reservations.

The effect of mutual induction within the adsorbed layer is to reduce the field acting on each adsorbate atom. If this mutual induction effect could be "turned off", whilst leaving the mutual induction effect between the substrate atoms and the adsorbate-layer atoms "turned on", then the field F^* acting on each substrate atom would be:

$$F^* = D^{-1}\beta_{\mathrm{A}}(\mathrm{aprx.})\, F^{\mathrm{ext}} , \tag{32}$$

where the symbols have the same meaning as previously. Taking F^{ext} as 56 V/nm, and using the data appropriate to the hexagonal lattice as given in table 5, with $b_{\mathrm{E}} = 7$ meV V^{-2} nm^2, we obtain a binding-energy estimate:

$$\Delta B^{\mathrm{el}}(F^*) = 0.090 \text{ eV} . \tag{33}$$

This result should be compared with the value of $\Delta B(\mathrm{conv.})$ given in table 5, namely 0.075 eV.

We now argue that during the removal process all the dipoles in the adsorbate layer (except the vacancy-site atom) would, if they remained fixed in position, have strengths intermediate between $b_{\mathrm{A}}F_{\mathrm{A}}^{\mathrm{loc}}$ and $b_{\mathrm{A}}F^*$ (where $F_{\mathrm{A}}^{\mathrm{loc}}$ is the local field acting on each adsorbate atom when the layer is complete). Consequently, the work done against electric-field-induced forces in removing the vacancy-site atom should

be intermediate between the binding-energy estimates obtained by taking the field acting on the other dipoles as F_A^{loc} and as F^*, respectively. That is, ΔB^{el}(diff.) should be intermediate between 75 and 90 meV.

For the same lattice structure, but with $b_E = 2$ meV V^{-2} nm^2, the conclusion is that ΔB^{el}(diff.) should be between 40 and 53 meV.

In reality it is likely that, as the vacancy-site atom is withdrawn, the surrounding atomic dipoles will "relax" inwards towards the vacancy centre. The consequences are difficult to estimate numerically, but we feel that the effect will be to reduce the differential binding energy slightly.

We thus think that for the field adsorption of helium on the (111) face of tungsten, the differential field-induced binding-energy contribution ΔB^{el}(diff.) could be greater than ΔB(conv.) by roughly 10 meV, when $F^{ext} = 56$ V/nm.

5. The calculation of ΔB

Finally, we must consider those contributions to the total short-range binding-energy ΔB that result from forces other than field-induced ones. Temporarily ignoring any effects due to "indirect" interaction, we may write the total potential U in which an adsorbate atom moves in the form:

$$U = U^{el} + U^{disp} + U^{rep} , \tag{34}$$

where the component potentials are due, respectively, to electric-field-induced, dispersive, and repulsive forces. The repulsive forces are those resulting from the interpenetration of electron charge clouds, and we consider their influence first.

Repulsive forces. As a simple approximation, we may take the dispersive and repulsive potentials to go inversely as the sixth and twelfth powers of the separation of the adsorbate and substrate atoms [29], and may ignore the change in the "internal energy of the source" when considering U^{el}. In these circumstances eq. (34) can be written more explicitly:

$$U(z) = -\beta_A^2(z) \cdot \tfrac{1}{2} b_A (F^{ext})^2 - C/z^6 + G/z^{12} , \tag{35}$$

where C and G are constants.

At the adsorption equilibrium separation (s) there is no resultant force acting on an adsorbate atom, so we must have:

$$\mathrm{d}U/\mathrm{d}z|_s = -\mathrm{d}(\beta_A^2)/\mathrm{d}z|_s \cdot \tfrac{1}{2} b_A (F^{ext})^2 + 6C/s^7 - 12G/s^{13} = 0 . \tag{36}$$

By definition, the dispersive and repulsive contributions (B^{disp} and B^{rep}) to the total binding energy B are given by:

$$B^{disp} = -U^{disp}(s) = C/s^6 , \tag{37a}$$

$$B^{rep} = -U^{rep}(s) = -G/s^{12} . \tag{37b}$$

Hence it follows from eq. (36) that:

$$B^{\mathrm{rep}} = (s/12) \cdot \mathrm{d}(\beta_{\mathrm{A}}^2)/\mathrm{d}z|_s \cdot \tfrac{1}{2} b_{\mathrm{A}} (F^{\mathrm{ext}})^2 - \tfrac{1}{2} B^{\mathrm{disp}} . \tag{38}$$

Since dispersion-induced forces are relatively short-range, U^{disp} is equal to zero at positions in the external field somewhat above the emitter surface; consequently, B^{disp} is equal to ΔB^{disp} as defined earlier (in connection with eq. (30)). Similarly, B^{rep} is equal to ΔB^{rep}, and these substitutions can be made in eq. (38). For purposes of discussion it is then possible to write this equation in the form:

$$\Delta B^{\mathrm{rep}} = -\eta \, \Delta B(\mathrm{conv.}) - \tfrac{1}{2} \, \Delta B^{\mathrm{disp}} , \tag{39}$$

where ΔB(conv.) is the conventional binding-energy expression given by eq. (2), and η is a parameter given by:

$$\eta = -(s/12) \cdot \mathrm{d}(\beta_{\mathrm{A}}^2)/\mathrm{d}z|_s/(\beta_{\mathrm{A}}^2 - 1) . \tag{40}$$

Eq. (30) provides an expression for the total short-range binding energy ΔB. The term $\Delta B^{\mathrm{indir}}$ is ignored here. If we also continue to neglect the change in the internal energy of the source, then ΔB^{el} in eq. (30) can be replaced by ΔB(conv.), and substitution of eq. (39) into eq. (30) gives:

$$\Delta B = (1 - \eta) \, \Delta B(\mathrm{conv.}) + \tfrac{1}{2} \, \Delta B^{\mathrm{disp}} . \tag{41}$$

Hence it may be seen that the effect of repulsive forces, in the approximation represented by eq. (35), is to reduce the dispersion-induced component of the short-range binding energy by one-half, and to reduce the electric-field-induced component by the fraction η.

The derivative in eq. (40) is readily computed [21]. With the hexagonal lattice structure assumed earlier, and $s = 0.259$ nm, we obtain $\eta = 0.51$ if b is taken as 2 meV V^{-2} nm^2, $\eta = 0.48$ if b is taken as 7 meV V^{-2} nm^2. The factor $(1 - \eta)$ in eq. (41) is thus approximately equal to one half. The effect of repulsive forces cannot be neglected.

To obtain a numerical estimate we shall assume that the factor $(1 - \eta)$ derived above may be applied to the quantity ΔB^{el}(diff.) discussed in section 4. For an external field of 56 V/nm, the interplay of field-induced and repulsive forces is thus predicted to give a binding-energy contribution lying in the range 25 to 45 meV, the lower value corresponding to our lower choice of polarisability value.

It should be noted, though, that the η-values predicted above are a direct consequence of our earlier assumption of an inverse twelfth-power law for U^{rep}. This assumption is a convenient first approximation; but a more careful treatment of the repulsive potential at a charged surface will eventually be required.

Dispersive forces. The second term in eq. (41) is the binding-energy component resulting from the interplay of dispersive and repulsive forces. This interplay also exists in the absence of the external field, and has been extensively discussed in past literature [9,29,32]. Clearly, if there were no change in the position of the adsorption site upon creation of surface charge and dipole distributions, then the $\frac{1}{2}\Delta B^{\mathrm{disp}}$ in

eq. (41) would represent the zero-external-field binding energy, at least approximately. Since the field-adsorption sites are directly above the substrate atoms, whereas the normal physisorption sites are in the positions that maximise the number of nearest neighbours, differences will exist in the details of the dispersive and repulsive interactions in the two cases. Nevertheless, it seems a reasonable first approximation to think of the second term in eq. (41) as equal to the zero-external-field binding energy.

The value of this binding energy, for helium on a metal, has normally been taken as approximately 10 meV (see ref. [9], for example). However, a recent review [32] cites a value of approximately 4 meV for helium on tungsten, derived from scattering experiments [33]. These binding-energy contributions are small, but are not negligible in comparison with the lower of the estimates made for the first term in eq. (41).

Lateral interactions. In principle there exist contributions to the total short-range binding energy that result from lateral interactions between adsorbate atoms. The effect of electrostatic dipole–dipole interactions has already been taken into account in the estimation of $\Delta B^{\mathrm{el}}(\mathrm{diff.})$. But in principle there also exist dispersive and repulsive interactions.

For two helium atoms in free space, recent calculations suggest that at a separation of 0.4476 nm there is an attractive interaction energy of around 0.1 meV [34]. As each adsorbed helium atom in a monolayer on a (111) face has six nearest neighbours, this result would predict a contribution of around 0.5 meV to the total (differential) short-range binding energy. A contribution of this order can be neglected.

Indirect lateral interaction. When noble-gas atoms are adsorbed, the presence of the surface modifies the interaction between them, even when the surface is neutral. In effect, there is an additional contribution to the binding energy, that gives rise to the term $\Delta B^{\mathrm{indir}}$ included in eq. (30). For the heavier inert gases adsorbed on graphite this contribution is known to be repulsive at large separations, and tends to reduce the magnitude of the lateral interaction energy by around 5–10% [35]. This effect will presumably exist in the adsorption of helium on a metal, but the present authors know of no relevant data. However, if its size is comparable with the percentage figure just quoted, then $\Delta B^{\mathrm{indir}}$ would be of order −0.1 meV and thus completely negligible.

Summary. We may conclude that numerical estimations of the short-range field-adsorption binding energy should include the effects of both dispersive and repulsive interactions with the substrate. Combining the estimates made earlier suggests that, for an external field of 56 V/nm, the "differential" binding energy $\Delta B(\mathrm{diff.})$ most probably lies in the range 30 to 50 meV. The limits of this range correspond to assumed surface-atom polarisability values of 2 and 7 meV V^{-2} nm^2, respectively, and are subject to some uncertainty. Fuller treatments will eventually be required.

6. General summary and discussion

The aim of this paper has been to discuss the field adsorption of helium on the (111) face of tungsten, using a particular charged-surface model involving infinite planar arrays of monopoles and dipoles, and to explore possible corrections to existing theory. Judgement is reserved on how well this model represents a real, crystallographically flat, charged metal surface (and other models should be investigated), but within this framework the following results have been demonstrated.

(1) The effects of both monopole and dipole moments in the substrate atoms, and the influence of distant neighbours, must be taken into account. (2) It is important to use the correct lattice structure for the face in question (or, failing this, a model lattice with the same area per lattice point as the real structure). (3) All forms of mutual depolarisation and mutual induction effect should be taken into account. (4) The effects of higher moments and field inhomogeneities can be neglected, but a more careful treatment of the zero-external-field dipole moment is necessary. (5) A distinction must be made between the conventional quantity ΔB(conv.), defined by eq. (2), and the field-induced component ΔB^{el}(diff.) of the differential short-range field-adsorption binding energy ΔB(diff.). (6) The effects of dispersive and repulsive interactions with the substrate must be taken into account when calculating ΔB(diff.), and more careful treatments of these will eventually be required. (7) The biggest single cause of uncertainty in the estimation of binding energy is lack of accurate knowledge of the proper SI polarisability of the tungsten surface atoms.

Numerically, our estimated value for ΔB(diff.) ("most probably between 30 and 50 meV") is much lower than the binding-energy estimates that would be given by the isolated-dipole-pair approximation. (The polarisability values used here would give predictions of around 180 meV and 900 meV, for $F^{\mathrm{ext}} = 56$ V/nm.) The main causes for this are the inclusion in our discussion of effects due to mutual depolarisation and to repulsive interactions.

For field values smaller than the 56 V/nm assumed here our estimates of ΔB(diff.) would be even lower. If such low values be physically correct for the (111) face, then there could be consequences for the theory of field-ion imaging. Current thinking tends to assume that an almost complete field-adsorbed layer is present during imaging. But if the differential short-range binding energy is very low, then the coverage could be very significantly lower than unity at temperatures near 80 K. Nonetheless the interior of the (111) plane can be resolved at these temperatures, under suitable circumstances.

For planes that are more closely packed than the (111) plane the mutual depolarisation effects would be even stronger, and the binding-energy estimates even lower. It is easy to understand that no significant degree of localised field adsorption might occur on these faces, except at extremely low temperatures. Consequently, the reported absence of field adsorption on the (110) plane of tungsten [36,37] could perhaps be rationalised.

Conversely, ΔB(diff.) would be expected to be higher for sites above planes that are more open than (111). This may be part of the reason why our numerical estimates for the (111) face are significantly lower than the experimental result of 250 meV reported for the field adsorption of helium on the tungsten (114) plane [3,5].

All these topics require rather more careful consideration than space permits here. Further, with the present model there is difficulty in discussing higher-index planes, because our model assumes that all surface charges and dipoles lie in a single plane. This is plausible for relatively-close-packed facets; but on more open faces the surface charge may be carried in part by the most exposed atoms, in part by less-exposed planes of atoms. The theory necessary to decide the relative charges carried by each layer simply does not exist. Development of a good quantum-mechanical theory of real, structured, charged surfaces seems to be required before much hard progress can be made. Regrettably, this does not appear to be imminent, so it would be helpful if field adsorption experiments could be carried out on relatively close-packed net crystal planes.

Another difficulty is that crystallographic facets on field-ion emitters are usually relatively small in size. The mutual depolarisation effects characteristic of array models will certainly be present in the real situation, but there will also be other effects operating, and infinite-array models will not be applicable unmodified. It remains to be shown how modification can best be achieved.

We conclude by stressing one fundamental point. Experimentally it is clear that localised field adsorption exists. This fact is incompatible with the use of jellium-type models: any viable model for field adsorption must therefore take the atomic nature of a real charged surface into account. However, it is clear from the present work that the isolated-dipole-pair approximation is numerically unacceptable as a substitute for an array model. A field-adsorption model must take the *crystallographic structure* of the surface into account if it is to be realistic: there can be no universally applicable calculation of field-adsorption binding energy (except as an order-of-magnitude guide). As in other areas of surface science, each crystal face and each substrate/adsorbate combination needs to be considered separately.

We believe that the present paper represents the first detailed theoretical discussion of field adsorption on a specific crystal face.

References

[1] T.T. Tsong and E.W. Müller, Phys. Rev. Letters 25 (1970) 911.
[2] R.G. Forbes, Ph.D. thesis, Cambridge Univ. (1970).
[3] T.T. Tsong and E.W. Müller, J. Chem. Phys. 55 (1971) 2884.
[4] K.D. Rendulic and E. Krautz, Acta Phys. Austriaca 42 (1975) 357.
[5] T.T. Tsong, Surface Sci. 70 (1978) 211.
[6] R.G. Forbes, Surface Sci. 78 (1978) L504.
[7] R.G. Forbes, Surface Sci. 64 (1977) 367.

[8] This paper uses the internationally-recognised four-dimensional system of electrical quantities and equations, rather than the dimensionally-inconsistent semi-Gaussian system still used in some literature. We have, however, now omitted the prime originally used to distinguish an SI quantity from its Gaussian equivalent. Where necessary, equations and quantities taken from the literature have been replaced by their international-system equivalents.
[9] For example, A.D. Crowell, in: The Solid–Gas Interface, Ed. E.A. Flood (Dekker, New York, 1967).
[10] That is, the field acting at the nucleus of the atom, in the absence of the atom itself, but in the presence of any charge re-arrangements induced by the presence of the atom.
[11] R.G. Forbes, Surface Sci. 87 (1979) L278.
[12] R.G. Forbes, 21st Intern. Field Emission Symp., Marseilles, 1974 (unpublished).
[13] R.G. Forbes, J. Phys. D (Appl. Phys.), submitted.
[14] J. Topping, Proc. Roy. Soc. (London) A114 (1927) 67.
[15] J.R. MacDonald and C.A. Barlow, J. Chem. Phys. 39 (1963) 412.
[16] $b^{\ominus} = 1$ meV V^{-2} $nm^2 \approx 1.602189 \times 10^{-40}$ J V^{-2} $m^2 \approx 4\pi\epsilon_0 \times 1.439976$ Å^3.
[17] T.M. Miller and B. Bederson, in: Advances in Atomic and Molecular Physics, Vol. 13 (Academic Press, London, 1977).
[18] J. Thorhallsson, C. Fisk and S. Fraga, J. Chem. Phys. 49 (1968) 1987.
[19] R.G. Forbes, 26th Intern. Field Emission Symp., Berlin, 1979 (unpublished).
[20] R.J. Culbertson, T. Sakurai and G.H. Robertson, Phys. Rev. B19 (1979) 4427.
[21] R.G. Forbes, R.K. Al-Dabbagh and M.K. Wafi, unpublished work; details available from first-named author.
[22] The equations correspond to this situation, but it could be argued that the polarisability value does not.
[23] For example, R.P. Feynman, R.B. Leighton and M. Sands, Lectures on Physics, Vol. II (Addison-Wesley, London, 1964).
[24] This is true for a model based on a planar array, but a much more careful analysis would be needed in any other geometry.
[25] E.W. Müller and T.T. Tsong, Field Ion Microscopy: Principles and Applications (Elsevier, Amsterdam, 1969).
[26] R. Smoluchowski, Phys. Rev. 60 (1941) 661.
[27] Within the framework of the charged surface model used in this paper this assumption is reasonable; but in a more general context it is not necessarily so, because there are other surface-charge distributions that could produce the same effect.
[28] This formula is a corrected version of one used elsewhere. Note the presence of the factor M_E in the denominator.
[29] S. Brunauer, The Adsorption of Gases and Vapours (Clarendon, Oxford, 1945).
[30] A.A. Holscher, Doctorate thesis, Leiden Univ. (1967).
[31] The inclusion of a term relating to the indirect interaction was suggested in discussion at the 26th IFES, Berlin, 1979.
[32] U. Landman and G.G. Kleiman, in: Surface and Defect Properties of Solids, Vol. 6 (Chemical Society, London, 1977).
[33] W.H. Weinberg, Ph.D. Thesis, Univ. of California, Berkeley (1971).
[34] P. Bertoncini and A.C. Wahl, Phys. Rev. Letters 25 (1970) 991.
[35] D.L. Freeman, J. Chem. Phys. 62 (1975) 4300.
[36] E.W. Müller, S.B. McLane and J.A. Panitz, Surface Sci. 17 (1969) 430.
[37] E.W. Müller, private communication.

Surface Science 108 (1981) 311–328
North-Holland Publishing Company

THE INFLUENCE OF HYPERPOLARISABILITY AND FIELD-GRADIENT POLARISABILITY ON FIELD ADSORPTION BINDING ENERGIES FOR He ON W(111)

Richard G. FORBES
University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET, UK

Received 11 August 1980; accepted for publication 23 December 1980

The effects of imaging-gas hyperpolarisability and field-gradient polarisability terms on field-adsorption binding energies have been explored. At best image fields for the noble imaging gases and molecular hydrogen, the correction to long-range binding energy is at most a few percent and may be neglected. At the tungsten evaporation field the correction is significant in the cases of Ar, Kr, H_2, and especially Xe. The system He on W(111) has been used as a paradigm in the investigation of short-range binding energy. The largest correction here is due to one of the field-gradient polarisability terms. In the field range of most relevance to the field-ion techniques the total correction is about 15–20%, and should not be neglected in detailed treatments.

1. Introduction

For some years now, following the early atom-probe experiments [1], it has been known that at the very high electric fields (~50 V/nm) necessary for the operation of a field-ion microscope there are imaging-gas atoms firmly field-adsorbed onto the surface of an operating field-ion emitter. And it has been believed that these field-adsorbed atoms have an important role in the formation of field-ion image contrast [2–5], and are present not only at kink sites but on every emitter facet from which resolved field-ion images are obtained.

The most convincing evidence of the presence of field-adsorbed atoms has perhaps been the observed existence, in field-ion energy spectra, of helium ions that must have been desorbed from a field-adsorbed position as a result of electron-stimulated excitation and/or ionization [6–8].

Although the presence of field-adsorbed imaging gas at resolved crystal facets seems adequately established, it is now less clear whether there is near *complete* coverage in the closely field-adsorbed layer under normal imaging conditions and at normal imaging temperatures (i.e. near 80 K and below). The early theoretical treatments of noble-gas field adsorption [2,3,9] predicted short-range binding energies of around 200 meV, or more, at normal imaging fields – which are large

0 039-6028/81/0000–0000/S02.50

enough to account for the presence of a nearly-complete field-adsorbed layer. And consequently it has been believed for some years that such a layer exists under normal imaging conditions. But it has also been clear, both from atom-probe experiments [9] and from observations of "hopping bright spots" [10–12], that at higher temperatures and/or lower pressure the coverage would certainly become partial. And, theoretically, coverage on specific crystal facets would be expected to be partial under normal imaging conditions if actual short-range binding energies were substantially lower than the 200 meV cited above.

Because of difficulties in estimating the effective gas pressure above the closely field-adsorbed layer (which is not identical with the pressure as measured in the microscope chamber), it is difficult to make reliable calculations concerning coverage. But the nature of the problem can be illustrated using rough calculations of the characteristic residence time τ_r of an atom in the closely adsorbed layer. I assume here that this layer can be regarded as "firmly" field adsorbed, for the purposes of discussing image contrast, if this residence time is substantially greater then the mean time interval between ionization events at a single ionization site. This latter time interval is assumed to be of the order of 10^{-4} to 10^{-3} s.

In the absence of electron-stimulated or other "non-thermodynamic" effects, τ_r is related to the activation energy Q necessary for escape from that site by:

$$\tau_r = A^{-1} \exp(Q/kT) \tag{1}$$

where T is the thermodynamic temperature of the emitter, k is the Boltzmann constant, and A is the desorption pre-exponential. For argument, let us require the residence time in the closely adsorbed layer to be 1 s or more. If A is taken as 10^{12} s^{-1}, then the minimum activation energies needed to achieve this residence time are as follows: 12 meV at 5 K; 47 meV at 20 K; 190 meV at 80 K. For an assumed residence time of 10^{-4} s the requisite activation energies would be: 8 meV at 5 K; 32 meV at 20 K; 130 meV at 80 K.

As already stated, early theoretical treatments predict short-range binding energies of around 200 meV, which are large enough to make the layer "firmly adsorbed" at normal imaging temperatures. However, a more recent theoretical treatment of field adsorption at relatively close-packed flat crystallographic surfaces [13] suggests that the earlier treatments, by neglecting mutual depolarisation effects, may have seriously overestimated short-range binding energies. Thus, Forbes and Wafi [13] estimate that, for the field adsorption of Helium on the (111) face of Tungsten, at an external field strength of 56 V/nm, the conventionally-defined field-induced contribution ΔB(conv.) to the total short-range binding energy ΔB most probably lies between 40 and 75 meV, and that ΔB itself most probably lies between 30 and 50 meV.

These estimates would be somewhat lower for fields in the imaging range (for helium the best image field is about 45 V/nm). In consequence, if it is legitimate to equate Q in eq. (1) to the short-range binding energy ΔB [14] then there must

be some doubt about the theoretical self-consistency of the assumption that a complete firmly-field-adsorbed layer of helium exists on the (111) face of tungsten during imaging.

Two main alternatives now present themselves. Either in fact *no* such complete field-adsorbed layer is present. (In which case one presumes that coverage is partial, and that the detailed mechanism of field-ion imaging is even more complicated than hitherto assumed.) Or some factor has been neglected, or wrongly discounted, in the Forbes and Wafi calculations, with the result that these calculations underestimate the true short-range binding energy.

The aim of this paper is to explore the possible influence of imaging-gas hyperpolarisability and field-gradient polarisability terms on the estimation of binding energies. The influence of the hyperpolarisability term on the long-range binding energy was explored by Southon, who adjudged it to be negligible for helium [15]. Its inclusion was again suggested by Holscher [16]. But, as far as the author knows, no evaluation of the size of such terms in the field adsorption situation has yet been published. Discussion of field-gradient polarisability terms in connection with field adsorption, or with field-ion theory generally, would seem to be a complete novelty.

For these reasons it has seemed useful, not only to present numerical results, but to set out the background theory in some detail.

As in previous discussion [13] it will be convenient to refer to the metal substrate on which field adsorption is occurring as the "emitter", and to label parameters appropriate to emitter surface atoms with the subscript "E", parameters appropriate to adsorbate-layer atoms with the subscript "A". This does not, however, necessarily imply that any emission process is taking place.

2. Hyperpolarisability terms

2.1. Basic theory and data

Following the approach of Buckingham [17,18] and of the more recent review by Bogaard and Orr [19], but changing their notation slightly to avoid clashes with previous usage in field-ion literature, we may write an expression for the dipole moment p of an electrically polarised atom or molecule as follows [20]:

$$p \triangleq \mu + bF + \tfrac{1}{2}hFF + \tfrac{1}{6}\gamma FFF\,, \qquad (2)$$

where F is the self-consistent *local* field acting on the atom or molecule; μ is the dipole moment at zero local field; b is the polarisability (α in refs. [17–19]); h is the first hyperpolarisability (β in refs. [17–19]); and γ is the second hyperpolarisability.

In general the symbols b, h and γ represent tensor quantities of the appropriate rank. However, for a spherical atom or axially-symmetric molecules the first and

third terms in eq. (2) vanish, and in the case of static polarisation the expression above simplifies to give:

$$p \triangleq bF + \tfrac{1}{6}\gamma^0 F^3 , \tag{3}$$

where b is SI polarisability as normally defined [21], and γ^0 is a scalar called the *second static hyperpolarisability*.

Strictly, eq. (2) and (3) apply only in the case of a uniform field. In a more general expression for dipole moment there are terms involving the field gradient and field-gradient polarisabilities [17,18]. These terms, and the corresponding terms in the expression for potential energy, will be disregarded here; we shall return to them in section 3.

Thus, in the approximation under discussion, the potential energy U of the polarised atom or molecule in the field F results from the field-induced dipole-moment of the atom, and is given by:

$$U \triangleq -\tfrac{1}{2}bF^2 - \tfrac{1}{24}\gamma^0 F^4 . \tag{4}$$

As made clear in the appendix, U represents the "total" potential energy (in this approximation) of the atom/field or molecule/field system, relative to a zero of energy when the neutral atom or molecule is in remote field-free space and in its electronic ground state, with U being the sum of: (a) the electrostatic potential of the induced dipole of moment p, in the field F; and (b) the polarisation-induced change in the internal electronic energy of the atom or molecule.

Experimental values for the static hyperpolarisability γ^0 are not available. However, values of the hyperpolarisability γ^K as relevant to the electro-optical Kerr effect have been deduced from experiment, by Buckingham and Dunmur [22] in the case of the noble gases, and by Buckingham and Orr [23] in the case of molecular hydrogen. These values are listed in table 1, but with the SI unit of hyperpolarisability written as J V^{-4} m^4 instead of the more conventional (equivalent) form C^4 m^4 J^{-3}. Values are also shown in the units eV V^{-4} nm^4 (1 eV V^{-4} $nm^4 \approx$

Table 1
Values of the hyperpolarisability γ^K for selected gas species, together with values of the hyperpolarisability and polarisability terms in B(ext.) evaluated at an external-field value of 57 V/nm

Species	γ^K (J V^{-4} m^4)	γ^K (eV V^{-4} nm^4)	γ^K term (meV)	b_A term (meV)	Ratio (%)
He	3.3×10^{-63}	2.06×10^{-8}	9.06	232	3.9
Ne	6.3×10^{-63}	3.93×10^{-8}	17.3	447	3.9
Ar	7.3×10^{-62}	4.56×10^{-7}	200	1850	11
Kr	1.7×10^{-61}	1.06×10^{-6}	467	2810	16
Xe	4.8×10^{-61}	3.00×10^{-6}	1320	4520	29
H_2	3.5×10^{-62}	2.18×10^{-7}	96.0	910	11

$1.602\ 1892 \times 10^{-55}$ J V^{-4} m^4); this atomic-level unit is analogous to the unit used elsewhere for SI polarisability [21], and is particularly convenient in field-ion theory.

The quantity γ^K is not identical with the static hyperpolarisability γ^0. First, the definition of γ^K involves a different combination of the components of the hyperpolarisability tensor than does γ^0. Second, hyperpolarisability components are in principal frequency-dependent, so the experimental result for γ^K depends on the wavelength of light used in the Kerr effect experiments.

In the case of atomic helium, Sitz and Yaris [24] have investigated theoretically what the difference between γ^0 and γ^K should be in the circumstances of Buckingham's experiments. They find that γ^K should be greater than γ^0 by approximately 3%. The best theoretical values for γ^0 and γ^K are 2.67×10^{-63} J V^{-4} m^4 and 2.75×10^{-63} J V^{-4} m^4, respectively [19]. This latter value may be compared with the tabulated experimental value for γ^K, namely 3.3×10^{-63} J V^{-4} m^4: a small discrepancy exists, but the difference is almost within the limits of experimental error.

We may conclude that, for our purposes, sufficient accuracy is obtained by using in eqs. (3) and (4) the tabulated values of γ^K, certainly in the case of helium, and probably in the cases of the other gases. Predicted binding-energy contributions should be slight overestimates.

2.2. Influence on long-range binding energy

The total binding-energy B(tot.) of a field-adsorbed atom is usually split into two parts, a long-range part here written B(ext.) that corresponds to the binding energy of an isolated atom in the approximately uniform field F_{ext} that exists in space somewhat above the emitter surface, and a short-range part ΔB that represents the work needed to remove a field-adsorbed atom from its bonding position to a point in the field F_{ext} [13]. Thus:

$$B(\text{tot.}) = B(\text{ext.}) + \Delta B \ . \tag{5}$$

In the present approximation B(ext.) is given from eq. (4), with the signs reversed and F set equal to F_{ext}.

Taking F_{ext} equal to the conventionally assumed evaporation field for tungsten (57 V/nm), and using the polarisability values tabulated in ref. [21], I show in table 1 the values of the polarisability and hyperpolarisability terms in eq. (4), for the noble gases and for molecular hydrogen. The last column shows the hyperpolarisability as a percentage of the polarisability term. These figures show that the hyperpolarisability term has almost negligible effect on B(ext.) in the cases of helium and neon; in the cases of the other gases, at the high field value chosen, the correction is significant, especially in the case of xenon.

At lower fields the hyperpolarisability term is relatively less important. At best imaging fields for the various gases, as estimated by Müller and Tsong [25], the

hyperpolarisability contribution to binding energy is relatively the most significant for helium, for which it contributes approximately 2.5% of B(ext.).

2.3. Influence on short-range binding energy

Imaging-gas hyperpolarisability also produces a correction to the short-range binding energy ΔB. Discussion here is restricted to the "paradigm" case of helium on the (111) face of tungsten, which is the adsorption system investigated in previous work [13].

In the light of previously published discussion [13,26], it is convenient to divide ΔB into two parts:

$$\Delta B = \Delta B(\text{main}) + \Delta B(\text{other}) . \tag{6}$$

The "main" contribution to short-range field-adsorption binding energy is obtained by using the expression for the binding energy of an atom or molecule when in an applied field F but distant from a surface, and calculating the change in this expression when the self-consistent local field F_A^{loc} (as defined below) is substituted for the external field F_{ext}. ΔB(other) represents the sum of various correcion terms, resulting from the presence of the surface and mostly not easy to evaluate in detail; the effect of ΔB(other), for the adsorption system chosen and for the field range of interest, is to reduce ΔB(main) by an estimated 25% to 40%. This splitting of ΔB into two parts is an extension of the method used previously, with ΔB(main) being the generalised equivalent to the ΔB(conv.) used earlier [13,26]. The broad physical origins of the "other" terms are discussed in ref. [13], and the introduction of hyperpolarisability and field-gradient polarisability terms into ΔB(other) is considered in the appendix.

In the approximation of this section, as inherent in eq. (4), ΔB(main) is given by:

$$\Delta B(\text{main}) \triangleq \tfrac{1}{2} b_A (\beta_A^2 - 1) F_{ext}^2 + \tfrac{1}{24} \gamma^0 (\beta_A^4 - 1) F_{ext}^4 , \tag{7}$$

where β_A is defined in terms of the self-consistent local field F_A^{loc} acting at the position of the imaging-gas-atom nucleus, when the gas atom is at its bonding site, by:

$$\beta_A = F_A^{loc}/F_{ext} . \tag{8}$$

The suffix "A" has also been added to the polarisability symbol, to make it clear that it is the adsorbate polarisability that appears in eq. (7).

Forbes and Wafi [13] calculated ΔB(main) using the first term in eq. (7) alone. Their result is denoted here by ΔB(conv.), and is shown in column 1 of table 2, for the exemplary case when the external field is taken as 57 V/nm and the proper polarisability b_E of the surface atoms of the tungsten emitter is taken as 2 meV V^{-2}

Table 2
Values of various parameters and terms involved in the calculation of ΔB(main) for helium field-adsorbed on tungsten (111), for the case when $s = 0.259$ nm, $b_E = 2$ meV V^{-2} nm^2, and $F_{ext} = 57$ V/nm; the various columns give values at different stages of approximation, the final values being in column 5

Approximation	1	2	3	4	5	6
b_A(eff.)/$b^{\ominus}$	0.143	0.143	0.1561	0.1561	0.1590	0.1576
β_A	1.0849	1.0849	1.0826	1.0826	1.0821	1.0823
β_E	0.37890	0.37890	0.37977	0.37977	0.37996	0.37987
g_A (nm^{-1})	0	0	0	−1.9558	−1.9564	−1.9561
b_A term (meV)	41.109	41.109	39.950	39.692	39.692	39.813
γ term (meV)	–	3.490	3.384	3.384	3.360	3.371
B term (meV)	–	–	–	5.583	5.579	5.581
C term (meV)	–	–	–	0.867	0.867	0.867
ΔB(main) (meV)	41.109	44.599	43.334	49.784	49.499	49.632
% increase	–	8.5	5.4	21.1	20.4	20.7

$b^{\ominus}$ = 1 meV V^{-2} nm^2 [31].

nm^2 [27]. For the inter-nuclear separation s of an emitter surface atom and its adjacent adsorbate atom the Tsong and Müller [9] value has been used, namely 0.259 nm.

As an improved approximation we may evaluate both terms in eq. (7), using the column 1 value of β_A. The result is shown in column 2 of the table. The penultimate line of the table shows the sum of the binding-energy contributions, and the last line gives the percentage difference between the new value of ΔB(main), at a given stage of approximation, and the original estimate as shown in column 1.

This result for ΔB(main) is still only an approximation, however, because the hyperpolarisability term in eq. (3) in fact leads to a small reduction in β_A (resulting physically from enhanced mutual depolarisation within the helium adsorbate layer). A more consistent calculation can be obtained, using iteration, by replacing the adsorbate-atom polarisability b_A as used in section 2 of ref. [13] by an "effective polarisability" b_A(eff.) given by:

$$b_A(\text{eff.}) = b_A + \tfrac{1}{6}\gamma^0\beta_A^2 F_{ext}^2 \qquad (9)$$

β_A is initially set equal to the column 2 value. Use of b_A(eff.) in the analysis as set out in ref. [13] then leads via their eq. (19) to a revised estimate of β_A, and hence to a new value of b_A(eff.). Three iterative steps are sufficient to give a value of β_A consistent to 1 part in 10^6; the corresponding binding energy contributions as evaluated from eq. (7) above are then consistent to 1 μeV.

The final results, as obtained in the approximation used in this section, are shown in column 3 of table 2. The inclusion of hyperpolarisability effects has led to a $5\frac{1}{2}$% increase in the predicted value of ΔB(main).

For reference purposes in connection with later discussion, I also show in table 2 (on line 2) the values of the field ratio β_E defined by: $\beta_E = F_E^{loc}/F_{ext}$, where F_E^{loc} denotes the self-consistent field acting at the nuclei of the emitter surface atoms. β_E is calculated from a formula indicated in ref. [13] – using the adsorbate polarisability value b_A to get the β_E values shown in column 1 and 2, and the finally-corrected value of b_A(eff.) to get the β_E value shown in column 3.

All numerical values in table 2 are, for reference purposes, shown to more significant figures then can be justified physically.

3. Terms involving the field gradient

3.1. Basic theory

In addition to the terms already discussed, the full expression for the potential energy U of an atom or molecule in an electric field contains terms involving the field gradient F' at the centre of mass of the atom or molecule [17,18]. In the case of an atom or axially-symmetric molecule, where also there is no permanent dipole or quadrupole moment, the full expression for U reduces to one involving scalar (rather than tensor) coefficients:

$$U = -\tfrac{1}{2}bF^2 - \tfrac{1}{24}\gamma F^4 - \tfrac{1}{4}BF^2F' - \tfrac{1}{4}C(F')^2 \ . \tag{10}$$

γ is the second static hyperpolarisability, as before, but for notational convenience the superscript has been omitted; B and C are coefficients sometimes known as "field-gradient polarisabilities".

The corresponding expressions for the induced dipole moment p and the induced quadrupole moment Θ are given by Buckingham [18] as:

$$p = bF + \tfrac{1}{6}\gamma F^3 + \tfrac{1}{2}BFF' \ , \tag{11}$$

$$\Theta = \tfrac{1}{2}BF^2 + CF' \ . \tag{12}$$

Strictly, if z denotes the direction normal to the emitter surface, in the geometrical circumstances of field adsorption, then: $F \equiv F_z$; $F' \equiv \partial F_z/\partial z$; and Θ is identical with the zz component of the quadrupole-moment tensor. Further, in the circumstances of complete-layer adsorption, the local field F_A^{loc} lies along this z-direction, so we can identify F with F_A^{loc}.

The additional terms involving F' have no effect on the long-range binding energy B(ext.), because the field F_{ext} somewhat above the emitter surface is nearly uniform. There are, however, three effects on the calculation of ΔB(main).

(1) The additional term in the expression for p will, because of the mutual depolarisation in the helium layer, lead to a small change in the estimate of F_A^{loc}, and hence to small changes in the energy terms previously calculated.

(2) There will also be additional terms in the expression for ΔB(main), which now

becomes:

$$\Delta B(\text{main}) = \tfrac{1}{2} b_A(\beta_A^2 - 1) F_{ext}^2 + \tfrac{1}{24}\gamma(\beta_A^4 - 1) F_{ext}^4 + \tfrac{1}{4} B \beta_A^2 g_A F_{ext}^3 + \tfrac{1}{4} C g_A^2 F_{ext}^2 \,. \quad (13)$$

The field gradient F'_A at the position of the nucleus of a helium atom, when at its bonding site, has been written in the form:

$$F'_A = g_A F_{ext} \,, \quad (14)$$

and is discussed further in section 3.3.

(3) The short-range field due to the induced quadrupole moment will modify the local field acting on tne emitter surface atoms, and this in turn will modify the values of β_A and g_A.

The numerical consequences of these effects are discussed in section 3.4.

3.2. Conversion of data

No experimental data about the values of B and C have been published, as far as the author is aware. However, in the case of helium, values of the equivalent Gaussian coefficients B_s and C_s have been calculated [28]. These are: $B_s = -2.271 \times 10^{-40}$ esu, $C_s = 4.018 \times 10^{-42}$ cm^5.

The conversion relationships between equivalent coefficients in the Gaussian system and the international (four electric dimensions) system are:

$$B = B_s \cdot (4\pi\epsilon_0)^{3/2} \,, \quad C = C_s \cdot 4\pi\epsilon_0 \,, \quad (15)$$

where ϵ_0 is the electric constant (permittivity of free space). These lead to the following numerical conversion factors:

$$B/\text{J} \cdot \text{V}^{-3} \cdot \text{m}^4 = 3.711\,401 \times 10^{-23} \times B_s/\text{esu} \,, \quad (16a)$$

$$B/\text{eV} \cdot \text{V}^{-3} \cdot \text{nm}^4 = 6.241\,460 \times 10^{54} \times B/\text{J} \cdot \text{V}^{-3} \cdot \text{m}^4 \,, = 2.316\,456 \times 10^{32} \times B_s/\text{esu} \,, \quad (16b)$$

$$C/\text{J} \cdot \text{V}^{-2} \cdot \text{m}^4 = 1.112\,650 \times 10^{-20} \times C_s/\text{cm}^5 \,, \quad (17a)$$

$$C/\text{eV} \cdot \text{V}^{-2} \cdot \text{nm}^4 = 6.241\,460 \times 10^{54} \times C/\text{J} \cdot \text{V}^{-2} \cdot \text{m}^4 \,, = 6.944\,561 \times 10^{34} \times C_s/\text{cm}^5 \,. \quad (17b)$$

Thus the values for B and C for helium become: $B = -5.261 \times 10^{-8}$ eV V^{-3} nm^4; $C = 2.790 \times 10^{-7}$ eV V^{-2} nm^4. As already indicated, it is convenient in field-ion theory to work with the international system of equations but to use units based on the electronvolt, volt and nanometre.

3.3. The calculation of g_A

Using an analytic approach of the type developed previously [13], we may express the local field F^{loc} acting in the vicinity of the nucleus of a field-adsorbed

helium atom in terms of a field-ratio β and the external field, and then split F^{loc} and hence β into three components, thus:

$$F^{\mathrm{loc}} = \beta F_{\mathrm{ext}} = (\beta^{\mathrm{m,E}} + \beta^{\mathrm{d,E}} + \beta^{\mathrm{d,A}})\, F_{\mathrm{ext}} \,. \tag{18}$$

The components arise, respectively, from monpoles and dipoles at the positions of the nuclei of the emitter surface atoms, and dipoles in the adsorbate atoms other than the atom in question.

Since $F' \equiv \partial F^{\mathrm{loc}}/\partial z$, we may identify g_{A} with the derivative $\partial\beta/\partial z$ taken at the position (A) of the nucleus of the field-adsorbed helium atom. The nature of the field due to the other helium atoms in the layer is such that $(\partial\beta^{\mathrm{d,A}}/\partial z)|_{\mathrm{A}}$ is equal to zero. Hence g_{A} is given by:

$$g_{\mathrm{A}} = (\partial\beta^{\mathrm{m,E}}/\partial z)|_{\mathrm{A}} + (\partial\beta^{\mathrm{d,E}}/\partial z)|_{\mathrm{A}} \tag{19}$$

Programs were previously developed for the calculation of $\beta^{\mathrm{m,E}}$ and $\beta^{\mathrm{d,E}}$ [29], so it is straightforward to evaluate the lattice sums necessary for the calculation of the derivatives. This leads to the result (for He on W(111), assuming an emitter–adsorbate separation of 0.259 nm):

$$\partial\beta^{\mathrm{m,E}}/\partial z\,|_{\mathrm{A}} = -0.81540\ \mathrm{nm}^{-1} \,, \tag{20a}$$

$$\partial\beta^{\mathrm{d,E}}/\partial z\,|_{\mathrm{A}} = \beta_{\mathrm{E}} b_{\mathrm{E}} \times -1.5015\ \mathrm{eV}^{-1}\,\mathrm{V}^2\,\mathrm{nm}^{-3} \,. \tag{20b}$$

3.4. Calculation of contributions to ΔB(main)

We now consider the numerical consequences of the effects listed earlier. As an initial approximation, g_{A} may be evaluated using the column 3 value of β_{E} in eq. (20b); and this and the column 3 value of β_{A} may then be used to evaluate the terms in eq. (13). The results are shown in column 4.

This approximation, however, neglects the effect of the third term in eq. (11) on the induced dipole moment. As in section 2.3, a more consistent calculation is obtained by defining a new "effective polarisability" by:

$$b_{\mathrm{A}}(\mathrm{eff.}) = b_{\mathrm{A}} + \tfrac{1}{6}\gamma\beta_{\mathrm{A}}^2 F_{\mathrm{ext}}^2 + \tfrac{1}{2}B\beta_{\mathrm{A}} g_{\mathrm{A}} F_{\mathrm{ext}} \tag{21}$$

Evaluating this leads, via eq. (19) in ref. [13], to new estimates of β_{A}, β_{E} and g_{A}. A second iterative step gives estimates that are then consistent to 1 part in 10^6; the resulting binding-energy contributions, consistent to 1 μeV, are shown in column 5 of the table. Physically, the field-gradient terms together make a significant binding-energy contribution, but this iterative correction has not been worthwhile.

It remains to explore the possible effect of the field due to the induced quadrupole moments in the adsorbed helium layer. The nature of the quadrupole field is such that, in a complete layer, there is no resultant field component acting on any other adsorbate atom, but the emitter surface atoms will in principle be influenced. Since quadrupole fields are very short range, I shall consider for each emitter surface atom only the quadrupole field due to the helium atom adsorbed

directly above it. Thus, at the emitter atom nucleus, the adsorbate helium quadrupole moment Θ_A gives rise to a field $F_E^{q,A}$ given by [17]:

$$F_E^{q,A} = -3\Theta_A/4\pi\epsilon_0 s^4 , \tag{22}$$

where s is the emitter–adsorbate nuclear separation. The field is normal to the array plane. As in ref. [13], the positive z-direction is taken "from the emitter to the adsorbate", and there is a convention about suffixes that the upper position labels the source and/or type of the field, etc., and the lower position labels where it acts.

Θ_A is given by eq. (12), with F equal to F_A^{loc}. Remembering that $F_A^{loc} \equiv \beta_A F_{ext}$, we may obtain a coefficient $\gamma_E^{q,A}$ given by:

$$\gamma_E^{q,A} \equiv F_E^{q,A}/F_A^{loc} = -(\tfrac{3}{2}B\beta_A F_{ext} + 3Cg_A\beta_A^{-1})/4\pi\epsilon_0 s^4 . \tag{23}$$

(This γ is not a hyperpolarisability but a field ratio.)

In the analysis that appears in ref. [13] there is a coefficient γ_E^A that, strictly, should be defined *not* by analogy with eq. (15) in ref. [13] but by:

$$\gamma_E^A = F_E^A/F_A^{loc} , \tag{24}$$

where F_E^A denotes the field component acting on the emitter atom due to multipole moments of *all* orders associated with the adsorbed atom layer. And, strictly, the coefficient γ_E^A that appears in eq. (19) of ref. [13] (and in the analogous expression for β_E) is being approximated there by the coefficient $\gamma_E^{d,A}$ that relates only to the dipole field of the adsorbate atoms. In the approximation under discussion here the "total" coefficient γ_E^A must be given by:

$$\gamma_E^A = \gamma_E^{d,A} + \gamma_E^{q,A} . \tag{25}$$

So the additional terms represented by eq. (23) above must be inserted into the analysis as described in ref. [13].

Using the column 5 values of β_A and g_A, we may derive via eq. (23) and (25) a "corrected" value of γ_E^A, and then proceed to derive new values of β_A, β_E and g_A, as before, and thence revised estimates of the contributions to ΔB(main). At the level of significance used in the table, there is no resulting change in any parameter. We conclude that, for He on W(111), the effects of quadrupole fields due to the adsorbate atoms are numerically and physically negligible.

3.5. Alternative numerical procedure

This has been the first numerical investigation of the influence of hyperpolarisability and field-gradient polarisability terms in field adsorption, and it has been thought useful to develop approximations sequentially. It is straightforward to write a program that will calculate the approximations in sequence, as described above, but a quicker numerical procedure (more suitable for hand calculations) is as follows:

(1) Assume $\beta_E = 0.5$ and calculate g_A from eqs. (19) and (20);
(2) Assume $\beta_A = 1.0$ and calculate b_A(eff.) from eq. (21);
(3) Using b_A(eff.), but ignoring the correction to γ_E^A, calculate values of β_A and β_E from the formulae indicated in ref. [13].
(4) Recalculate g_A, and then calculate binding-energy contributions from eq. (13).

This leads to the values shown in column 6 of table 2. The resulting value for ΔB(main) is within 0.2 meV of the iterated value, and is more than adequate physically. (The column 5 values are recovered after iterating steps (3) and (4) above, at most twice more.)

3.6. Precision of calculations – summary

Because of the uncertainties over the values of some of the atomic parameters (in particular b_E), and for other reasons discussed in ref. [13], it is difficult to estimate binding energies accurately: at best, ranges can be stated within which ΔB(main) and ΔB probably lie. However, for given values of the input parameters the calculation of ΔB(main) can be made as precise as required. If we take it as a working requirement that the calculation of ΔB(main) should be precise to within 1 meV, then the following conclusions can be drawn from the preceding analysis:
(1) All of the terms in eq. (13) should be evaluated, although the term involving C is of marginal significance,
(2) Effects due to the induced quadrupole moment in the helium adsorbate atoms can be completely disregarded;
(3) In the calculation of b_A(eff.), and the induced dipole moment p, the γ-term in eq. (21) needs to be included but the term in B is of marginal significance:
(4) For nearly all purposes the numerical procedure described in section 3.5 will give results of satisfactory precision, without iterating.

4. Results and discussion

For the He on W(111) system under discussion, calculations of ΔB(main) have been carried out for a range of external-field values and for the two values of emitter surface-atom polarisability used and discussed in ref. [13]. Results are shown graphically in fig. 1, and in more detail for three field values in table 3. Also shown in both cases are the corresponding values of the "uncorrected" quantity ΔB(conv.) given by the treatment of ref. [13].

In broad terms, the effect on the "main" component of short-range binding energy has been to increase this component by roughly 20%, for the field range of interest to the field-ion techniques – by somewhat less at lower field strengths. In all cases the larger part of the increase comes, somewhat surprisingly, from the field-gradient term in B, rather than from the hyperpolarisability term, although the latter also contributes significantly.

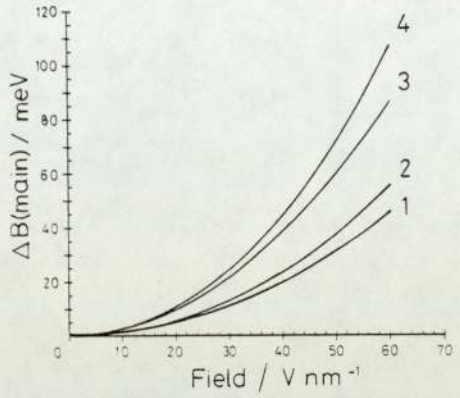


Fig. 1. The main contribution ΔB(main) to short-range binding energy, for He on W(111) assuming s = 0.259 nm, compared with the "uncorrected" quantity ΔB(conv.) appearing in ref. [13], as a function of external field strength. The curves represent: (1) ΔB(conv.), for $b_E = 2b^{\Leftrightarrow}$; (2) ΔB(main), for $b_E = 2b^{\Leftrightarrow}$; (3) ΔB(conv.), for $b_E = 7b^{\Leftrightarrow}$; ΔB(main), for $b_E = 7b^{\Leftrightarrow}$. ($b^{\Leftrightarrow}$ = 1 meV V^{-2} nm2 [31].)

It is not clearly known how accurate are the existing calculations of B. However, even an error of 20% would (in the worst case) lead only to an error of 2 meV in ΔB(main); so neither uncertainty concerning the value of B, nor the small uncertainty over the value of γ^0, is likely to have appreciably distorted the preceding results.

Table 3
Values of various terms and quantities involved in the calculation of short-range binding energy for He on W(111), as described in the text, for various choices of emitter-surface-atom polarisability b_E and external field F_{ext}

b_E	$2b^{\Leftrightarrow}$	$2b^{\Leftrightarrow}$	$2b^{\Leftrightarrow}$	$7b^{\Leftrightarrow}$	$7b^{\Leftrightarrow}$	$7b^{\Leftrightarrow}$
F_{ext} (V/nm)	30	45	57	30	45	57
b_A term (meV)	11.3	25.0	39.7	21.4	47.7	75.9
γ term (meV)	0.3	1.3	3.4	0.5	2.7	6.9
B term (meV)	0.8	2.8	5.6	1.5	5.1	10.5
C term (meV)	0.2	0.5	0.9	0.7	1.5	2.4
ΔB(main) (meV)	12.6	29.7	49.5	24.1	57.0	95.6
ΔB(conv.) (meV)	11.4	25.6	41.1	21.5	48.4	77.7
% correction	10%	16%	20%	12%	18%	23%
Gradient (nm^{-1})	−24.09	−24.25	−24.43	−22.97	−23.18	−23.38
η	0.520	0.523	0.526	0.496	0.500	0.505
$(1-\eta)\Delta B$(main) (meV)	6.0	14.1	23.4	12.1	28.5	47.4

$b^{\Leftrightarrow}$ = 1 meV V^{-2} nm^2 [31]; gradient $\equiv \partial[\Delta B(\text{main})]/\partial z \div \Delta B$(main).

4.1. The estimation of ΔB

When deriving an estimate of ΔB from the calculated value of ΔB(main), Forbes and Wafi suggested that it was important to correct for the effect of repulsive forces. The question of how to calculate reliably the repulsive (negative) contribution to binding energy, in the circumstances of field adsorption, deserves more careful exploration, but this is beyond the scope of the present paper. In what follows I use a procedure equivalent to that used in ref. [13].

The repulsive energy component corresponding to ΔB(main) can be written in the form:

$$\Delta B^{\mathrm{rep}}(\mathrm{main}) = -\eta\, \Delta B(\mathrm{main})\,, \tag{26}$$

where η is a positive coefficient. Together with ΔB(main), there is then a contribution of $(1-\eta)\Delta B$(main) to ΔB.

Elementary analysis, based on the requirement of zero resultant force on a field-adsorbed atom, shows that η is given by the equation:

$$\eta = -\frac{s}{n}\left[\frac{\mathrm{d}(\Delta B(\mathrm{main}))}{\mathrm{d}z}\,\frac{1}{\Delta B(\mathrm{main})}\right], \tag{27}$$

where s is the assumed emitter–adsorbate internuclear distance, as before, and n is the index of the assumed repulsive power law. The bracketed term has to be calculated at the bonding-point position; numerical values are shown in line 8 of table 3, and are roughly equivalent to an attractive sixth-power potential.

As in ref. [13], n is taken as 12 here, on the grounds that the emitter surface atoms are more likely to be acting in an "individual" way than in a "collective" way, which should make this choice the most appropriate of the common elementary approximations. Lines 9 and 10 in table 3 show the corresponding values of η and $(1-\eta)\Delta B$(main).

To obtain ΔB, further corrections now need to be made. The most significant are the addition of a contribution resulting from dispersive forces, estimated as probably lying in the range 5 to 10 meV, and a correction that takes into account the change in the internal energy of the source of the field $F_{\mathrm{A}}^{\mathrm{loc}}$ as an individual helium atom is moved in (from a position in the external field F_{ext}) to fill a vacancy in the field-adsorbed layer. An estimate of this second correction was made, with considerable reservations, in ref. [13]; however it now seems to the author that this correction needs investigating in greater detail. It seems unlikely to be greater than 5 meV at $F_{\mathrm{ext}} = 57$ V/nm, and would be smaller at lower field strengths. A full investigation is beyond the scope of this paper, and this correction is disregarded in what immediately follows.

To give a rough indication of plausible ranges for the value of ΔB, the following procedure is adopted. For the $b_{\mathrm{E}} = 2$ meV V^{-2} nm^2 case, 5 meV is added to $(1-\eta)\Delta B$(main); this gives an estimated lower limit at a given field strength. For the $b_{\mathrm{E}} = 7$ meV V^{-2} nm^2 case, 10 meV is added, and gives an estimated upper limit.

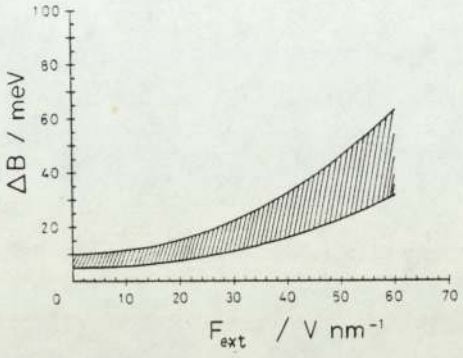


Fig. 2. Interim estimates for probable upper and lower limits for the short-range field-adsorption binding energy ΔB, for the system He on W(111) assuming $s = 0.259$ nm, as a function of external field strength. Based on assumptions stated in text.

The calculations illustrated in table 3 are carried out for the whole range of field strengths used earlier, the additions just described are made, and the resulting ranges are shown as the shaded area on fig. 2.

4.2. General implications

In broad terms, the effect of the additional terms discussed in this paper has been to raise the previous estimates [13] of the short-range field-adsorption binding energy, for the He on W(111) system, by around 20% (in the field range of most interest to the field-ion techniques).

Even in the most favourable case, the predicted values are not high enough to make it plausibe (using considerations based on eq. (1)) that a complete field-adsorbed layer would be present on a large (111) face at imaging temperatures near 80 K.

It must be concluded, either that no *complete* layer is present at such temperatures, or that – if a complete or nearly complete layer is in fact present, on which point the experimental evidence is unclear – then some other factor (such as finite-plane-size effects or bonding-point shifts) must be taken into account in the calculation of binding energies.

In passing, it may be noted that recent experiments by Ernst et al. [30] give a preliminary indication, for the system neon on tungsten, of a ΔB value of around 30 meV at $F_{ext} = 44$ V/nm. This is substantially lower than the estimate that would be derived from the Tsong and Müller treatment of field adsorption, at this field strength, namely ~240 meV – but is much closer to the values derived from a treatment similar to that of the present paper. (For the system Ne on W(111), using a treatment that includes hyperpolarisability but disregards field-gradient polarisability – because no data can be found – ΔB(main) is estimated to probably lie in the range 6 to 28 meV, when $F_{ext} = 44$ V/nm; higher estimates would be expected

for more open crystal faces.) This perhaps provides an additional reason for taking seriously the possibility that the firmly field-adsorbed layer may be incomplete, at least in some imaging circumstances and particularly at higher imaging temperatures.

Finally, a note of caution. The theory of field adsorption is still in its early stages, and it may be many years before definitive binding-energy estimates can be made theoretically, even for well-defined adsorption situations. This paper has investigated one form of correction to the basic Forbes and Wafi treatment of complete-layer field adsorption on a planar surface. There are other possible corrections that need to be explored. Thus the binding-energy values given here, although hopefully a useful indication, must be regarded only as interim estimates.

Appendix: The definition of polarisation energy

The objective of this appendix is to clarify the distinction between the "internal" and "external" components of the polarisation energy of a polarisable atom situated in an electric field in space, and to give expressions for these components.

Relative to a zero of energy that corresponds to the situation of a neutral atom or molecule in remote field-free space, the total potential energy U of the polarisable atom or molecule when in a field F is given by:

$$U = H + V , \tag{28}$$

where, as in ref. [26], H denotes the energy of interaction between the polarised atomic or molecular charge distribution and the field F, and V is the change in the internal electronic energy of the atom or molecule. (V results from the field-induced change in the electronic wave-function of the ground electronic state of the atom or molecule, and is commonly called the "internal polarisation energy".)

The quantity BE, defined by:

$$\mathrm{BE} = -U = -H - V , \tag{29}$$

is commonly called the "polarisation binding energy".

First consider the case where the field is uniform, and hence field-gradient terms and induced quadrupole moments can be neglected. I shall consider only the situation where the electric dipole is aligned with the field, so the discussion can be in terms of scalars rather than vectors.

The interaction energy H is given by the usual classical expression, so on substituting from eq. (3) for the dipole moment p (and dropping the suffix on the symbol for hyperpolarisability), we obtain:

$$H = -pF = -bF^2 - \tfrac{1}{6}\gamma F^4 . \tag{30}$$

The quantity V could be obtained by combining this expression with eq. (4), but it may be more instructive to derive an expression for V from a classical argu-

ment. Think of the polarised atom or molecule as a finite dipole consisting of positive and negative charges of magnitude q, a distance $2z$ apart when the applied field is F. These charges do not interact with each other classically, but in such a fashion that the dipole moment p is given by

$$p = 2qz = bF + \tfrac{1}{6}\gamma F^3 \,. \tag{31}$$

However, they each interact classically with the applied field, and the "internally-produced" force on each charge balances the force $\pm qF$ on it due to the field F.

The work done on the "internal forces" when the dipole length is increased by $2\mathrm{d}z$ is thus given by:

$$\mathrm{d}W = 2qF\,\mathrm{d}z = F\,\mathrm{d}p = (bF + \tfrac{1}{2}\gamma F^3)\,\mathrm{d}F \,. \tag{32}$$

And hence the total work done in creating a dipole of moment appropriate to final field F is:

$$V = W = \tfrac{1}{2}bF^2 + \tfrac{1}{8}\gamma F^4 \,. \tag{33}$$

On combining eq. (29), (30) and (33) we obtain:

$$\mathrm{BE} = \tfrac{1}{2}bF^2 + \tfrac{1}{24}\gamma F^4 \,, \tag{34}$$

which is the result that would be obtained directly from eq. (4).

It should be noted that the quantities V and BE are *not* equal. However, in the lowest approximation (which considers only the polarisability term) they do seem to be equal. In consequence, in some existing literature the name "polarisation energy" is applied indiscriminately to both V and BE. In higher approximations greater care needs to be taken with nomenclature.

In the case of a non-uniform field it is more convenient to work directly with the formulae given by Buckingham [17,18]. In a situation with the symmetry appropriate to the field-adsorption situation, the interaction energy H is given in terms of p and the quadrupole moment Θ by:

$$H = -pF - \tfrac{1}{2}\Theta F' = -bF^2 - \tfrac{1}{6}\gamma F^4 - \tfrac{3}{4}BF^2F' - \tfrac{1}{2}CF'^2 \,. \tag{35}$$

The binding energy BE is given from eq. (10) by:

$$\mathrm{BE} = \tfrac{1}{2}bF^2 + \tfrac{1}{24}\gamma F^4 + \tfrac{1}{4}BF^2F' + \tfrac{1}{4}CF'^2 \,. \tag{36}$$

Hence, from eq. (29), the internal polarisation energy is deduced to be:

$$V = \tfrac{1}{2}bF^2 + \tfrac{1}{8}\gamma F^4 + \tfrac{1}{2}BF^2F' + \tfrac{1}{4}CF'^2 \,. \tag{37}$$

In this paper we have been concerned with the higher polarisabilities of the adsorbate atoms. However, there will also be contributions to the internal polarisation energy V_E of the emitter surface atoms that result from the higher polarisabilities of these atoms. In principle, this will result in an additional component in the quantities ΔV_E and ΔU_S that appear in the discussion of ref. [26], and hence an additional component in the quantity ΔB(other) used in eq. (6) here. Calculation

of the size of this component is not possible at present, but intuition suggests that it is unlikely that the estimates of ΔB presented earlier will be substantially affected.

References

[1] E.W. Müller, S.B. McLane and J.A. Panitz, Surface Sci. 17 (1969) 430.
[2] T.T. Tsong and E.W. Müller, Phys. Rev. Letters 25 (1970) 911.
[3] R.G. Forbes, Ph.D. Thesis, Cambridge Univ. (1970).
[4] T.T. Tsong, Surface Sci. 28 (1971) 651.
[5] R.G. Forbes, Surface Sci. 27 (1971) 659.
[6] E.W. Müller and S.V. Krishnaswamy, Surface Sci. 36 (1973) 29.
[7] T. Sakurai and E.W. Müller, Surface Sci. 49 (1975) 497.
[8] R.J. Culbertson, T. Sakurai and G.H. Robertson, Phys. Rev. B19 (1979) 4427.
[9] T.T. Tsong and E.W. Müller, J. Chem. Phys. 55 (1971) 2884.
[10] W. Schmidt, Th. Reisner and E. Krautz, Surface Sci. 26 (1971) 297.
[11] A.P. Janssen and J.P. Jones, Surface Sci. 33 (1972) 553.
[12] K.D. Rendulic, Surface Sci. 28 (1971) 285.
[13] R.G. Forbes and M.K. Wafi, Surface Sci. 93 (1980) 192.
[14] Although this assumption is open to question, it has been made in all previous discussions and thus I continue to accept it here.
[15] M.J. Southon, private communication, 1963.
[16] A.A. Holscher, Doctorate Thesis, Leiden Univ. (1967).
[17] A.D. Buckingham, Quart. Rev. (Chem. Soc. London) 13 (1959) 183.
[18] A.D. Buckingham, Advan. Chem. Phys. 12 (1967) 107.
[19] M.P. Bogaard and B.J. Orr, Intern. Rev. Sci. Phys. Chem. (Ser. 2) 2 (1975) 149.
[20] This paper uses the international four-dimensional system of electrical quantities and equations. Where it is necessary to refer to them, Gaussian quantities are labelled with a subscript "s".
[21] R.G. Forbes, Surface Sci. 64 (1977) 367.
[22] A.D. Buckingham and D.A. Dunmur, Trans. Faraday Soc. 64 (1968) 1776.
[23] A.D. Buckingham and B.J. Orr, Proc. Roy. Soc. (London) A305 (1968) 259.
[24] P. Sitz and R. Yaris, J. Chem. Phys. 49 (1968) 3546.
[25] E.W. Müller and T.T. Tsong, Field-ion Microscopy: Principles and Applications (Elsevier, Amsterdam, 1969).
[26] R.G. Forbes, Surface Sci. 87 (1979) L278.
[27] R.G. Forbes, Appl. Phys. Letters 36 (1980) 739.
[28] L.L. Boyle, A.D. Buckingham, R.L. Disch and D.A. Dunmur, J. Chem. Phys. 45 (1966) 1318.
[29] R.G. Forbes and M.K. Wafi, unpublished work; details available from first-named author.
[30] N. Ernst, G. Bozdech and J.H. Block, in: Proc. 27th Intern. Field Emission Symp., Tokio, 1980.
[31] $b^{\ominus} = 1$ meV V^{-2} $nm^2 \approx 1.602189 \times 10^{-40}$ J V^{-2} $m^2 \approx 4\pi\epsilon_0 \times 1.439976$ Å^3.

Vacuum/volume 31/numbers 10–12/pages 567 to 570/1981
Printed in Great Britain

0042–207X/81/100567–04$02.00/0

Progress with the theory of noble-gas field adsorption

Richard G Forbes, *University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET, UK*

Field adsorption may provide information about the nature of highly-charged surfaces, and is thought to be essential to the formation of field-ion images. This paper briefly reviews the state of the theory.
Recent calculations, for He on W(111) at a field of 45 V nm^{-1}, suggest a short-range binding energy of order †15–30 MeV. But such binding energies would not allow the existence at 80 K of the firmly field-adsorbed layer that is currently thought to be essential to image formation; hence a discrepancy exists.
The following possibilities are considered: (1) Partial coverage in the surface-adsorbed layer; (2) Use of a more sophisticated surface model in the calculations; (3) Introducing terms related to the higher polarizabilities of the noble-gas atoms; (4) The existence of bonding-points shifts. The last of these is difficult to calculate because there seems the possibility that a 'repulsive depolarization' effect may be operating.
It is concluded that the discrepancy has not yet been resolved, and that experimental evidence as to the adsorbate-atom coverage in the interior of resolved net planes would be useful, particular for W(111) when this is resolved.

1. Introduction

The aim of this paper is to briefly describe current progress in dealing, first, with the theory of noble-gas field adsorption, second, with a consequent problem in the theory of field-ion imaging.

1.1. Field adsorption. Field adsorption is a form of physical adsorption that occurs at highly-charged surfaces (or, more generally, when the field-dependent terms in a binding-energy expression are dominant). In the context of noble-gas field adsorption, binding arises because there are electric-field-induced changes in the electronic wave-function of the gas atom or molecule, and the interaction between the resultant dipole and the electric field leads to a lowering of the potential energy of the atom/field (or molecule/field) system. Field adsorption is of interest, both because it may indirectly provide basic information about the nature of charged surfaces, and because of the presumed importance of field-adsorbed noble-gas atoms in the field-ion imaging process.

Experimentally, field adsorption was first detected by Müller[1,2], when $(W\,He)^{3+}$ and other complex metal–gas ions were discovered in mass spectra obtained with an atom-probe field-ion microscope; it was inferred that noble-gas atoms must have been locally adsorbed on the surface prior to desorption of the complex ion. But probably the most convincing evidence for the presence of field-adsorbed atoms has been the existence, in helium–ion energy spectra, of ions that must (by virtue of their observed energy) have been created at a field-adsorption site, as a result of electron-stimulated excitation and/or ionization, due to an electron shower from other gas atoms field-ionized some distance above the surface[3–6]. The existence of field adsorption can also be inferred from field-ion microscope observations of 'hopping bright spots' that are clearly due to the presence of a second (more-polarizable) gas species in an imaging-gas mixture[7–9], certainly as far as adsorption of the 'second' species in concerned.

1.2. Field-ion imaging. Normal field-ion images are formed by the ionization of imaging-gas atoms at (or slightly beyond) the so-called 'critical surface', which under imaging conditions is about 0.6 nm away from the emitter surface-atom nuclei. The Pauli exclusion principle forbids ordinary (non-stimulated) ionization inside the critical surface, as there are no vacant emitter electron levels for the imaging gas electrons to tunnel into. The field-adsorbed atoms sit between the emitter surface atoms and the critical surface, as shown in Figure 1, and are presumed to influence the generation of field-ion current in the region beyond the critical surface.

In field-ion images of, say, tungsten there are some crystal facets where every surface atom is resolved. In particular, this can be true of the (111) plane when imaged with helium, under some imaging circumstances. Adjacent atoms in a facet will be resolved if the ionization density $(dJ/dV)_D$ at a point D in the critical surface directly above an emitter surface-atom nucleus is sufficiently greater than the ionization density $(dJ/dV)_M$ at a point M in the critical surface directly above a point midway between the nuclei of adjacent surface atoms. In the spirit of the Raleigh criterion, I assume that the ratio of ionization densities has to be greater than about 1.3 for adjacent atoms to be resolved.

The ionization density $(dJ/dV)_D$ at point D is given by:

$$(dJ/dV)_D = C_D P_{eD} \qquad (1)$$

† for MeV read meV, in all cases

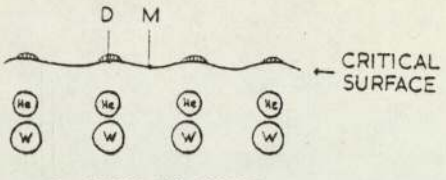


Figure 1. Schematic diagram showing the emitter surface atoms, a field-adsorbed layer of helium, the critical surface, and points D and M. The small shaded areas just outside the critical surface are the 'ionization zones' in which the ionization density is greatest.

where C_D is the concentration of imaging-gas at point D, and P_{eD} is the electron transition rate-constant for the field ionization of a gas atom at point D. C_D in this formula is, of course, being used in the sense of statistical-mechanical (time-average) probability-density for finding the nucleus of an imaging-gas atom in a small volume dV at D. Thus the ratio R of the ionization densities is given by:

$$R = (dJ/dV)_D/(dJ/dV)_M = (C_D/C_M) \cdot (P_{eD}/P_{eM}). \qquad (2)$$

Although reliable calculation of the ratios on the r.h.s. of equation (2) is difficult, there has been a feeling in the subject area for some years that, if the emitter surface is bare, then neither the gas-concentration ratio nor the rate-constant ratio would be high enough to explain the existence of atomic resolution. For example, Iwasaki and Nakamura[10] were unable to prove that the rate-constant would be higher at D than M, and calculations by Forbes[11] suggested that—at imaging temperatures of 80 K—the ratio C_D/C_M would be significantly less than 1.3. Consequently it has been felt that the presence of a complete (or nearly complete) layer of field-adsorbed atoms was essential to the formation of field-ion images, and this also has been regarded as evidence for the existence of field adsorption.

2. The calculation of field adsorption binding energies

Shortly after the experimental discovery of field adsorption, Tsong and Müller[12,13] put forward a theory that attributed field adsorption to a field-induced dipole–dipole interaction between the adsorbed noble-gas atom and the nearest emitter metal atom, and estimated binding energies of around 200 MeV. Parallel with this, the present author[14] was attempting to calculate binding energies in a model that attributed adsorption to the locally high fields above the excess positive charges in the highly-charged surface. With the appearance of ref. 12 it became clear that both the local charge and the local dipole on the emitter atom ought to be taken into account.

With the passing of time it became clear that, although the above models might be suitable for kink-site atoms, they were not suitable for adsorption on crystallographic facets, because the interactions between neighbouring surface atoms needed to be taken into account. There are difficulties with calculations using finite arrays[15], and hence one turns to the use of infinite-array models. Tsong[16] presented an analysis involving arrays of dipoles, but Forbes[17] argued that the effects of surface charges (or monopoles) ought also to be included. It was also suggested[18,19] that there were conceptual errors in existing theory. In consequence, Forbes and Wafi[19] developed a treatment of field adsorption based on the surface model shown in Figure 2, and

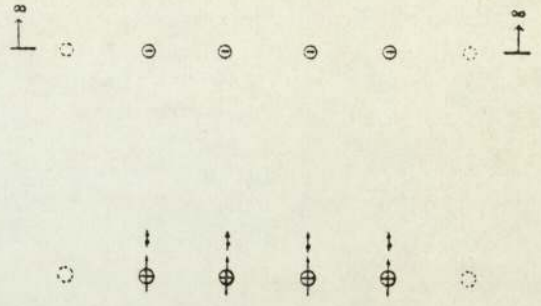


Figure 2. Charged-surface model employed by Forbes and Wafi. Each surface atom in a planar emitter is replaced by a superimposed charge and dipole, and the noble-gas adsorbate atoms are represented as a layer of dipoles. A layer of negative charge 'at infinity' is necessary for electrostatic self-consistency.

applied it to the field adsorption of helium on tungsten (111). A slightly generalized summary of their analysis is now presented.

Because of the spatial structure of the electric field above the cap of a field-ion emitter, the total binding energy B of a field adsorbed atom can be split into long-range B(ext) and short-range ΔB parts:

$$B = B(\text{ext}) + \Delta B. \qquad (3)$$

Contributions to the short-range binding energy ΔB come from field-induced forces, dispersion, repulsive forces, and indirect lateral interactions via the substrate, thus:

$$\Delta B = \Delta B^{\text{ei}} + \Delta B^{\text{disp}} + \Delta B^{\text{rep}} + \Delta B^{\text{indir}}. \qquad (4)$$

In field adsorption the first term is the dominant one, and the effect of the other terms, at fields of interest in the field-ion techniques, is between them to reduce the estimated value of ΔB^{ei} by around 30–40%.

The field-induced term is itself split into two parts, thus:

$$\Delta B^{\text{ei}} = \Delta B(\text{main}) - \Delta U_S \qquad (5)$$

Some care is needed in understanding this equation, and its origin is discussed in some detail in ref. 18. Briefly, the quantity needed in the field-ion microscope situation (or so it was believed) is the work needed to remove a single atom from the field-adsorbed layer and place it in the external field F_{ext} slightly above the surface (in a region of space where the field ripple due to the proximity of the charged surface and the adsorbed layer has become negligible, and the field has the uniform (laterally) value F_{ext}). This quantity is denoted by ΔB^{ei}: in a sense it is the work needed to create a vacancy in the field-adsorbed layer. This, however, is difficult to calculate because creating the vacancy reduces the symmetry of the situation, and causes changes in all the induced dipole moments in the vicinity of the vacancy. It is much easier to treat the situation in which the dipole moments of all the emitter and adsorbed-layer atoms (other than that of the vacancy-site atom) are taken as 'frozen' at their complete-layer values, during the removal process: the work relevant to this removal process is denoted by ΔB(main). If the frozen moments are then allowed to relax, there is a change ΔU_S in the potential energy of the system. To obtain ΔB^{ei}, this quantity ΔU_S must be subtracted from the work done to remove the vacancy-site atom whilst the other moments are frozen, and hence equation (5) results.

ΔB(main) is easily obtained from the formula for the potential energy U_{lone} of an isolated atom in a field F, as:

$$\Delta B(\text{main}) = U_{\text{lone}}(F_{\text{ext}}) - U_{\text{lone}}(F_A^{\text{loc}}) \tag{6}$$

where F_A^{loc} denotes the self-consistent local field acting on a field-adsorbed atom in the layer; this is normally expressed in terms of the ratio β_A, as:

$$F_A^{\text{loc}} = \beta_A F_{\text{ext}}. \tag{7}$$

Thus, given an expression for U_{lone}, ΔB(main) can be obtained. ΔU_S is estimated to constitute a small correction, perhaps of order 10% or so.

Up till now it has been conventional to consider in U_{lone} just the well-known F^2 ('polarizability') term, and hence equation (10) becomes:

$$\Delta B(\text{main}) \simeq \Delta B(\text{conv}) = \tfrac{1}{2} b_A (\beta_A{}^2 - 1) F_{\text{ext}}^2 \tag{8}$$

where b_A is the SI polarizability[20] of the imaging-gas atom. ΔB(conv) denotes the expression on the r.h.s. of equation (8), and is the 'conventional' expression for short-range field-adsorption binding energy.

Previous calculations of binding energy can thus be seen as attempts to calculate the field ratio β_A (or the equivalent quantity $f_A = \beta_A{}^2$ known as the enhancement factor[12,13]), using some surface model, and assumptions about gas-atom/emitter-atom 'bond-length' and the apparent polarizability of the emitter surface atoms. For tungsten this apparent polarizability is not well established, but is estimated to lie between 2 and 7 MeV V^{-2} nm^2.* With the surface model shown in Figure 2, precise formulae can be obtained for β_A, and these can readily be evaluated by computer[19,21].

For helium on tungsten (111), for an external-field value of 45 V nm^{-1} (which is the best-image-field for helium), the resulting value of ΔB(conv) is estimated to lie between 25 and 50 MeV, and the corresponding value of ΔB is estimated to lie between 15 and 30 MeV. (The limits here correspond to the two polarizability values stated above.)

3. The problem

Although these results are not unreasonable in themselves, they constitute a problem for the theory of field-ion imaging, because the predicted binding-energy values are not large enough to hold field-adsorbed atoms onto the surface for times that are long compared with the mean interval between ionization events in a given ionization zone.

This mean interval can be taken as 10^{-4} s, though in practice it might be up to about 1000 times greater (depending on instrumental considerations). Assume that the time taken for an adsorbed atom to escape from a site, where Q is the necessary activation energy, is given by the Arrhenius-type equation:

$$\tau_r = A^{-1} \exp(Q/kT) \tag{9}$$

where kT is the Boltzmann factor, with T the emitter temperature, and A is a pre-exponential, whose value I take as 10^{12} s^{-1}. A residence time τ_r of 10^{-4} s corresponds to an activation energy of 130 MeV, when $T = 80$ K.

* 1 MeV V^{-2} $nm^2 \approx 1.602189 \times 10^{-40}$ J V^{-2} $m^2 \approx 4\pi\varepsilon_0 \times 1.439976$ Å^3.

Assuming that Q can be identified with ΔB, we see that, with the ΔB values predicted above, the field-adsorbed atoms will remain adsorbed (at 80 K) only for relatively short periods of time compared with the ionization interval 10^{-4} s. But, under suitable conditions, helium-ion images of tungsten (111) are resolved at emitter temperatures near 80 K. Hence a conceptual problem exists.

Presumably, either there is something wrong with our current ideas on the process of field ion imaging, or some factor has been neglected, or wrongly discounted, in the Forbes and Wafi calculations, with the result that these calculations underestimate the true short-range binding energy.

4. Responses

4.1. Partial coverage in the adsorbed layer. A very short residence time in the surface-adsorbed layer would tend to imply either a very fluid situation, or partial coverage in the adsorbed layer (or both). Thus an initial response is to consider the effect on image contrast of a single isolated atom in a surface-adsorbed position.

As regards the rate-constant ratio, calculations by Nolan and Herman[22,23] and by Iwasaki and Nakamura[24,25] actually suggest that the rate-constant P_{eD} for helium would be *reduced* (as compared with the bare-surface case) by the adsorption of another helium atom between D and the emitter surface atom. So it seems unlikely that the ratio P_{eD}/P_{eM} would be enhanced.

To estimate the (maximum possible) effect on gas-concentration ratio, the field ratio F_D/F_M has been calculated[26] for the case of a single adsorbed atom. If thermodynamic equilibrium at temperature T were to exist, then this ratio would give rise to a gas–concentration ratio given by:

$$C_D/C_M = \exp[(\tfrac{1}{2} b_A \bar{F}^2/kT) \cdot (2\,\delta F/\bar{F})] \tag{10}$$

where $\bar{F} = \frac{1}{2}(F_D + F_M)$ and $\delta F = F_D - F_M$. In reality, complete local equilibrium will not exist, but the calculated ratio does represent an upper limit.

Results for the three situations of relevance are[21,26]:

Bare surface:	$F_D/F_M = 1.0019$	$C_D/C_M(80\text{ K}) = 1.08$
Single adatom:	$F_D/F_M = 1.0024$	$C_D/C_M(80\text{ K}) = 1.11$
Adatom layer:	$F_D/F_M = 1.13$	$C_D/C_M(80\text{ K}) = 1.7$

These figures would suggest that only the adatom layer could give rise to an atomically resolved image. Thus we look for possible defects in the binding-energy calculations.

4.2. An alternative surface model. A possible feature neglected in the surface model of Figure 2 are local dipole moments of the charge distribution, *parallel* to the emitter surface. Crudely, one might think of these as resulting from wave-function overlap between atoms. A simple method of representing effects of this sort is to insert negative charges at the midpoints of the lines joining nearest-neighbours; constancy of the mean surface charge density is ensured by also increasing the positive charges by a factor $(1+\alpha)$. α is termed the 'charge-transfer parameter'.

Calculations on such a surface model show[27] that ΔB(conv) and ΔB increase steadily as α is increased. Thus for $\alpha = 1$ the estimated limits for ΔB(conv) are that it probably lies between 40 and 70 MeV, for $F = 45$ V nm^{-1}, as compared with the 25–50 MeV estimated previously for $\alpha = 0$. Local lateral charge-transfer effects thus might in principle lead to a significant increase in the short-range binding energy, and clearly this would be of

some help in resolving the discrepancy described earlier. But even with a lateral charge transfer corresponding to $\alpha = 1$, the increase is not sufficient in itself. There is considerable difficulty in deciding what value to choose for α, especially for a highly charged surface, but intuitively I would expect an α-value of 1 to be far too high, and might look for values of 0.2 or less. Consequently this response of choosing this particular surface model is not the answer to the discrepancy.

4.3. The effect of higher polarizabilities. In the expression for U_{lone}, the F^2 term is only the first of a series of terms. If the next three terms relevant to the field adsorption situation are included (see ref 28 for detailed discussion and original references), then the expression for the potential energy of an isolated atom in a non-uniform field F becomes:

$$U_{\text{lone}} = -\tfrac{1}{2}b_A F^2 - \tfrac{1}{24}\gamma_A F^4 - \tfrac{1}{2}B_A F^2 F' - \tfrac{1}{2}C_A (F')^2 \qquad (11)$$

γ_A is the 'second hyperpolarizability', and B_A and C_A are field-gradient polarizabilities, F' denoting $\partial F/\partial z$ where z is the direction normal to the emitter surface. A fuller expression for ΔB(main) is thus obtained by combining equations (6) and (11).

In the case of He on $W(111)$, a revised calculation of field-adsorption binding energies[28] shows that, at the field strengths of interest (around 45 V nm^{-1}), the most significant of the additional terms is the one involving the field-gradient polarizability B_A. The effect of the correction is to increase the values of ΔB(main) and ΔB estimated by Forbes and Wañi[19], by about 15–20%. This increase helps, but is not by itself sufficient to resolve the problems stated earlier.

4.4. Bonding-point shifts. The existing calculations all use the tungsten–helium layer separation given by Tsong and Müller[13], namely 0.259 nm. This was obtained by adding together tabulated values of the neutral-atom radii of tungsten and helium. However, it seems possible that, under the influence of the high surface field, the helium atoms might move closer to the tungsten surface-atom nuclei. If the internuclear spacing has been overestimated by using the value 0.259 nm, then ΔB(main) will have been underestimated.

However, in attempting to estimate the magnitude of any possible bonding-point shift, an unforeseen problem arises, in that the mechanism of repulsion seems to be different from that between two neutral entities.

When two neutral atoms come into contact, the Pauli exclusion principle comes into play and the electrons tend to repel each other. However, in the field-adsorption situation it seems possible that, for the polarized helium atom, the effect of any such electron interaction might be to cause a partial depolarization of the atom—so that what keep the atoms apart is, to some extent, a reduction in the attractive forces. To give this effect a name, we might call it *repulsive depolarization*.

Currently, it is not clear (at least to the present author) how any such effects can be reliably incorporated into a calculation of bonding-point position; I would commend this problem to any theoretician interested.

It perhaps deserves note that Tsong and Müller[13] do carry out a calculation of the 'effective polarizability' of a helium atom close to a metal ('jellium') surface, and do find a reduction in the parameter they have calculated, as the helium moves closer to the surface. However their calculation is certainly not one of this 'repulsive depolarization' effect, because their perturbation expression contains no terms relating to interaction between the helium electrons and the metal electrons. In fact, their calculation seems to be non-physical, because it uses the assumption that the electric field decays inside the jellium surface. At a real surface the helium atoms are situated directly above partially-ionic surface metal atoms, and the electric field must increase as the surface atom nucleus is approached.

Conclusions

The responses made so far to the problem stated earlier have not solved the problem, although they have slightly increased our understanding of the theory of field adsorption. Some further possibilities remain to be explored, for example the consequences of the finite size of net crystallographic planes on a real emitter. However, one thing that emerges from a re-examination of the relevant experimental evidence is that we do not seem to clearly know whether field-adsorbed helium atoms are present in the interior of *resolved* net planes, under imaging conditions, nor—even if they are present—what the coverage is. Experiments designed to elucidate information about this would be useful.

References

[1] E W Müller, *Quart Rev.* **23** 177 (1969).
[2] E W Müller, S B McLane and J A Panitz, *Surf Sci.* **17**, 430 (1969).
[3] E W Müller and S V Krishnaswamy, *Surf Sci.* **36**, 29 (1973).
[4] T Sakurai and E W Müller, *Surf Sci.* **49**, 497 (1975).
[5] R J Culbertson, T Sakurai and G H Robertson, *Phys Rev.* **B19**, 4427 (1979).
[6] N Ernst, *Phys Rev Lett.* **45**, 1573 (1980).
[7] W Schmidt, Th Reisner and E Krautz, *Surf Sci.* **26**, 297 (1971).
[8] A P Janssen and J P Jones, *Surf Sci.* **33**, 553 (1972).
[9] K D Rendulic, *Surf Sci.* **28**, 285 (1971).
[10] H Iwasaki and S Nakamura, *Surf Sci.* **33**, 525 (1972).
[11] R G Forbes, 21st Int. Field Emission Symp, Marseilles, (1974). Unpublished.
[12] T T Tsong and E W Müller, *Phys Rev Lett.* **25**, 911 (1970).
[13] T T Tsong and E W Müller, *J Chem Phys.* **55**, 2884 (1971).
[14] R G Forbes, PhD Thesis, University of Cambridge (1970).
[15] R G Forbes and M K Wañi, Unpublished work.
[16] T T Tsong, *Surf Sci.* **70**, 211 (1978).
[17] R G Forbes, *Surf Sci.* **78**, L504 (1978).
[18] R G Forbes, *Surf Sci.* **87**, L278 (1979).
[19] R G Forbes and M K Wañi, *Surf Sci.* **93**, 192 (1980).
[20] R G Forbes, *Surf Sci.* **64**, 367 (1977).
[21] M K Wañi, Unpublished work.
[22] D A Nolan and R M Herman, *Phys Rev.* **38**, 4088 (1973).
[23] D A Nolan and R M Herman, *Phys Rev.* **B10**, 50 (1974).
[24] H Iwasaki and S Nakamura, *Surf Sci.* **49**, 664 (1975).
[25] H Isawaki and S Nakamura, *Surf Sci.* **52**, 588 (1975).
[26] K Chibane, MSc thesis, University of Aston in Birmingham (1980). Unpublished.
[27] S S Bains and R G Forbes, Unpublished work.
[28] R G Forbes, *Surf Sci.* **108**, 311 (1981).

V. ENERGIES OF FIELD-EMITTED PARTICLES

VA APPEARANCE-ENERGY THEORY

27 A generalised theory of field-ion appearance energies
R.G. Forbes
Surface Sci., 61 (1976) 221-40

28 On the need for new measurements of the field-ion appearance energy of helium
R.G. Forbes
Int. J. Mass Spectrom. Ion Phys., 21 (1976) 417-20

29 Comments on 'Intrinsic energy losses of field-evaporated ions'
R.G. Forbes
J. Phys. D: Appl. Phys., 9 (1976) L191-3

30 Appearance energies for tungsten ions evaporated from ionic bonding states
R.G. Forbes
J. Phys. D: Appl. Phys., 13 (1980) 1357-63

VB FIELD-ION ENERGY DISTRIBUTIONS

31 Wave-mechanical theory of field ionisation and field-ion energy distributions
R.G. Forbes
Progr. Surface Sci., 10, (1980) 249-85

32 Comments on 'Surface plasmon excitation in field-ion emission'
R.G. Forbes
Surface Sci., 60 (1976) 260-1

33 Field ionization and surface plasmons: an alternative theoretical formulation
R.G. Forbes
Nederland Tijdschrift voor Vacuumtechnick, 16 (1978) 268-9

VC ENERGY ANALYSERS

34 An advanced field electron emission spectrometer
E. Braun, R.G. Forbes, J. Pearson, J.M. Pelmore, R.V. Latham
J. Phys. E: Sci. Instr., 11 (1978) 222-8

Surface Science 61 (1976) 221–240

A GENERALISED THEORY OF STANDARD FIELD ION APPEARANCE ENERGIES

Richard G. FORBES
University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET, UK

Received 2 March 1976; manuscript received in final form 28 June 1976

By means of arguments based on a thermodynamic cycle, general formulae are derived which express standard field ion appearance energies in terms of atomic parameters and molecular term values, and in terms of thermodynamic parameters. These formulae may be applied to field ionization, field evaporation, or to the field desorption of well-behaved molecules, and apply to a very wide range of desorption mechanisms. Previous theories of energy deficits are reviewed in the light of the general result. The extent of agreement between theory and experiment is assessed, and the need for continued theoretical development is noted. Formulae are provided for converting the variety of energy-deficit parameters encountered in the literature into appearance energies; and the various terminologies are correlated.

1. Introduction

Recent developments in several field-ion emission techniques have revived interest in the measurement of field-ion energy deficits and appearance energies. On the mass spectrometry side it has been realised that the appearance energy provides useful information about the energies involved in the formation of a parent ion [1–5]. On the atom-probe side, energy deficits are now being measured for the field evaporation process, and the hope is that these might lead to information about the mechanism of field evaporation [6].

The present work was stimulated by Waugh's work on field evaporation presented at the 1974 International Field Emission Symposium [7]. At the time it was thought that the necessary theory for the interpretation of his measured energy deficits had not been articulated. The original work of Tsong and Müller [8] on the energy deficits of the noble gases was found to contain an anomaly, in that they obtained good agreement between their results and a formula whose derivation seems invalid. And the present author gave, at the 1975 International Field Emission Symposium, a corrected general formula that could be applied both to field ionization and to field evaporation [9].

At that time it was pointed out by Block (private communication) that, in the context of field-ion mass spectrometry, the meaning of a field-ion appearance energy had been considered by Goldenfeld and co-workers. A formula partly resembling

mine has indeed been stated by Goldenfeld et al. [2], but the two formulae do not coincide in all circumstances.

Other formulae for energy deficits and appearance energies have been given by Jason [10], by Anway [1], by Waugh [6], and by Röllgen and Heinen [5], for a variety of experimental circumstances.

The objective of the present work is to bring all these strands together. It will be shown that a common terminology and approach can be applied to field ionization, field evaporation, and field desorption; and that general formulae can be obtained for standard field-ion appearance energies, by means of a thermodynamic argument based on analysis of a retarding-potential energy analyser. Previous theoretical expressions, and discrepancies amongst them, are then re-examined in the light of the general result.

The theory presented here is not completely general, because discussion is confined to formulae for *standard* field-ion appearance energies as defined in section 2. (All the formulae in the literature are, in effect, expressions for some standard appearance energy.) It is hoped to deal with extensions and generalisations of the arguments given here, at a later stage.

2. Field ion appearance energy

In a retarding-potential analyser, in order to bring an ion created at a field-ion emitter (subscript e) to a halt "just outside" the collector (subscript c), a voltage difference $\Delta\zeta(=\zeta_c - \zeta_e)$ is necessary. Normally, no r-fold-charged ions reach the collector unless this voltage is more negative than some fairly well defined value $\Delta\zeta_r^{on}$, here termed the *onset voltage.*

From this, a new empirical quantity A_r^{on} is defined by:

$$-re\,\Delta\zeta_r^{on} = A_r^{on} - r\phi_c\,, \tag{1}$$

where e is the proton charge, re is the charge on the ion, and ϕ_c is a local work function associated with the collector. A_r^{on} is termed the *field-ion onset appearance energy* for the r-fold ion in question.

The concept of an appearance energy, in the present sense, was introduced into field-ion work by Goldenfield et al. [2], and a related concept has been much used by Beckey, Röllgen, and Heinen [3–5]. The merit of using A_r^{on} rather than $\Delta\zeta_r^{on}$ is that the value of the former is determined only by the process of desorption, and not by the processes involved in collection or energy analysis.

For theoretical discussion, it is better to deal with *standard* field-ion appearance energies. These are the appearance energies appropriate to an ion that has been formed (in its internal ground state) by the slow transfer of electrons from a bound atom or molecule direct to the emitter Fermi level, that in a quasi-classical approach had just sufficient energy to escape from the emitter surface, and that then

moved slowly away from the surface without suffering any anomalous deactivation effects. Standard appearance energies are denoted by the symbol A^0_{mv}: the first subscript denotes the charge state immediately prior to escape; the second subscript denotes the charge state on arrival at the collecter; and the superscript "0" indicates that the conditions stated above apply. If the surface desorption process is straightforward and well behaved, and if there are no resistive-drop effects associated with current flow along the emitter, then observed onset or peak appearance energies should correspond fairly well with the appropriate standard appearance energy.

3. Theory of standard appearance energies

This section derives formulae for A^0_{mv} in terms of atomic parameters and molecular term values near the emitter surface. The method of proof has not previously appeared in field-ion literature. Although a single general formula can be obtained, its interpretation depends on the nature of the potential structure at the emitter surface, and it is convenient to derive proofs in stages. We begin by considering the initial bonding state to be neutral, and by examining the potential structures for the more commonly discussed mechanisms of field desorption.

3.1. The location of desorption

In the image-hump model for field evaporation as originated by Müller [11], the neutral and standard ionic terms are arranged as in fig. 1a. The term U_0 represents the potential energy of the neutral atom, and the term U_r represents the potential energy of the r-fold-charged ion all of whose missing electrons have been transferred directly to the Fermi level. The pre-evaporating atom, initially neutral, makes a transition into the ionic state, but then has to surmount the Schottky hump in the ionic term before it can leave the emitter surface. In fig. 1a, and in the other potential configurations drawn in this paper, the solid line denotes the standard term of lowest value at a given point, and only the salient features of higher standard terms are shown.

In the present context it may be assumed that along any line normal to the emitter surface there is a maximum in the standard ionic term for the departing ion. The surface joining the positions of all such maxima is roughly parallel to the emitter surface and will be called the *Schottky surface* for this term. Term values taken at points in this Schottky surface are denoted by U_r^{sh}.

Because of the three-dimensional nature of the potential configuration above a real, structured field-ion emitter, the value of U_r^{sh} will vary with position in the Schottky surface. If the desorbing entity is in thermodynamic equilibrium with the emitter prior to desorption (which is assumed here to be so), then desorption will tend to occur near the point at which U_r^{sh} has a local minimum. I call this point the "local pass", call the term value there the "*pass energy*", and denote the latter by

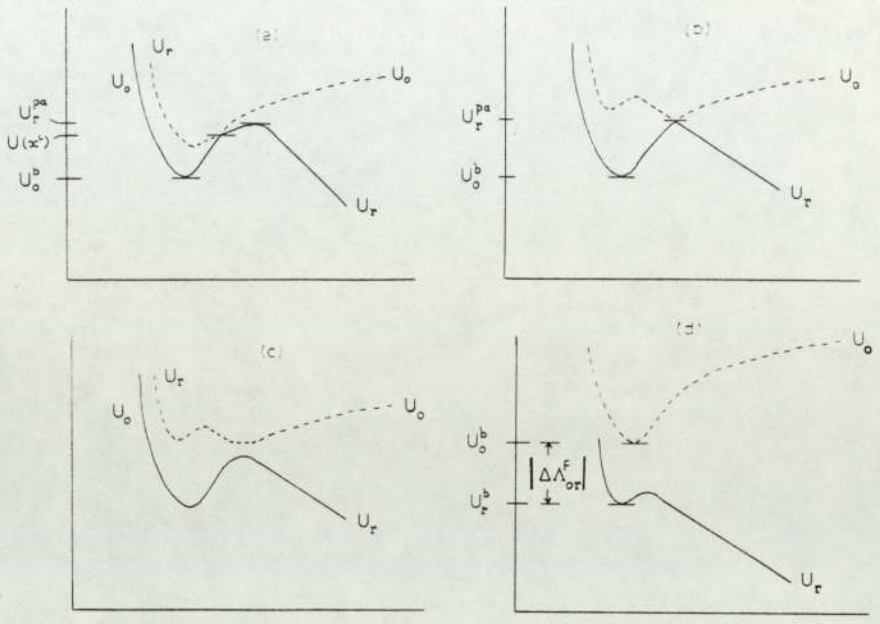


Fig. 1. Potential Configurations hypothesised in the traditional models for field evaporation. The vertical scale represents energy, and the horizontal scale distance from the emitter surface. Other annotations as defined in the text. (a) Image-hump model; (b) Simple charge-exchange model; (c) Charge-exchange model including curve repulsion; (d) Ionic bonding model.

U_r^{pa}. (The name comes from the helpful analogy between this local minimum and the crest of a mountain pass.) By definition, the standard appearance energy is that exhibited by an ion which had zero kinetic energy at the crest of the pass over which it classically escaped.

Since the pass energy may vary from desorption site to desorption site across the surface, there will be a corresponding variation in the value of the standard appearance energy. The consequences of this will be considered elsewhere.

In the simple intersection model of field evaporation originated by Gomer [12], transition and escape occur simultaneously, in the crossing surface in which the neutral and ionic Terms have common values. The relevant potential configuration is shown in fig. 1b; the crossing surface between the neutral term and the standard ionic term is called the *critical surface* for the r-fold ionization process in question.

The portion of the neutral term to the left of the critical surface and the portion of the standard ionic term to the right of the critical surface, taken together, form a kind of potential hump which (classically) the nucleus must surmount in order to escape. Thus the critical surface plays the same role in this model as does the Schottky surface in the image-hump model, and exactly analogous considerations apply: desorption tends to take place through a local "pass", at energy level U_r^{pa}.

For purposes of the derivation below, the two models are geometrically equivalent.

Potential curves for the charge-exchange model are now often drawn as in fig. 1c. The derivation below will apply to this configuration, and also to the ionic bonding model (fig. 1d), if in both cases U_r^{pa} is taken as the potential at the crest of the hump.

The derivation will also apply to the imaging-mode field ionization of the noble gases, which is a special case of configuration 1b, and where U_1^{pa} can be identified with the "characteristic energy" U'' defined by Forbes [13].

3.2. The basic appearance-energy formula

A standard appearance energy is determined empirically if ions that had just sufficient energy to escape through the local pass are brought to a halt "just outside" the collector of a retarding-potential analyser. The condition for this is:

$$U_r^c = U_r^{pa}\,, \tag{2}$$

where U_r^c is the potential energy for an r-fold ion "just outside" the collector.

Consider the cycle shown in fig. 2, which involves the following steps:

(1) Take electrons into the emitter from a neutral atom situated at the crest of the pass, placing the electrons at the Fermi level;

(2) Move resulting ion to collector;

(3) Take electrons round circuit to collector;

(4) Take electrons from collector to ion, neutralising it;

(5) Return neutral atom to its original position. (3)

If this cycle is applied in the case where $r = 1$, then the works done by a hypotheti-

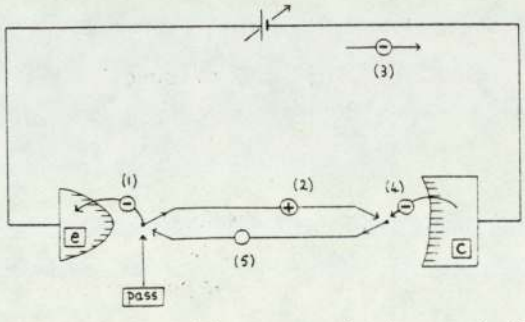


Fig. 2. Thermionic cycle used in determining an expression for a standard field ion appearance energy: (e) emitter, (c) collector; the steps are numbered as in the text.

cal external agent in each step are:

(1) $U_1^{pa}(0 \to 1) - U_0^{pa}(0 \to 1)$; (4) $\phi_c - I_1$;

(2) $U_1^c - U_1^{pa}(0 \to 1)$ (=0) ; (5) $U_0^{pa}(0 \to 1) - U_0^c$;

(3) $\mu_c - \mu_e$; (4)

where μ_c and μ_e are the local Fermi levels for electrons in the collector and emitter respectively; ϕ_c is a local work function associated with the collector; I_1 is the first ionization potential for an atom in remote field-free space; and the symbol $U_1^{pa}(0 \to 1)$ is a special case of the general symbol $U_r^{pa}(m \to r)$, which denotes the value of the r^{th} standard molecular term, taken at the particular pass where an escaping desorbate has changed from an m-fold ion to an r-fold ion in the process of escape. The expressions in steps (3) and (4) are straightforward; those in (1), (2) and (5) come from the definition of a molecular term (see Appendix C). The total work (w) done in the cycle is thus:

$$w = \mu_c - \mu_e + \phi_c - I_1 + [U_1^{pa}(0 \to 1) - U_0^c] \,. \tag{5}$$

However, the cycle is closed, and is assumed to be performed slowly, and the system is finally in the same state as it was initially; therefore w must be zero.

In a well-designed analyser $U_0^c = U_0^\infty$. The measured voltage difference $\Delta\zeta$ between the emitter and collector is given in terms of the local Fermi levels by eq. (6). Substituting into eq. (5) then gives the voltage $\Delta\zeta_{01}^0$ corresponding to condition (2):

$$e\,\Delta\zeta = e(\zeta_c - \zeta_e) = -(\mu_c - \mu_e) \tag{6}$$

$$e\,\Delta\zeta_{01}^0 = -[I_1 - \phi_c + U_0^\infty - U_1^{pa}(0 \to 1)] \tag{7}$$

The collector work-function appears in this formula because electron-ion recombination occurs at the collector.

From definition (1), it follows that the standard $0 \to 1$ field-ion appearance energy A_{01}^0 is given by:

$$A_{01}^0 = I_1 + U_0^\infty - U_1^{pa}(0 \to 1) \,. \tag{8}$$

The argument in the more general case of the direct $0 \to r$ transition is very similar, and leads to the result:

$$A_{0r}^0 = H_r^\infty + U_0^\infty - U_r^{pa}(0 \to r) \,, \tag{9}$$

where H_r^∞ is the energy needed to produce the r-fold ion from the neutral, in remote field-free space, with the electrons also being left in remote field-free space. H_r^∞ is given in terms of the free-space ionization potentials I_s by:

$$H_r^\infty = \sum_{s=1}^{r} I_s \,. \tag{10}$$

3.3. Generalisation to more complex configurations

A formula essentially similar to eq. (9) applies in the case where ionization occurs in two stages, with ionic motion occurring between two separate acts of ionization. When this situation is postulated, the second transition is customarily called post-ionization [14]. A potential configuration where this might happen is shown in fig. 3.

The derivation below provides a formula for the standard field-ion appearance energy relevant to this configuration. However, it should be realised that ions reaching point F will almost certainly not be in thermodynamic equilibrium with the emitter, and consequently it is far from certain that observed appearance energies would correspond to the standard appearance energy. This situation will be treated in more detail elsewhere.

With the potential configuration of fig. 3, steps (1) and (2) in cycle (3) have to be replaced by four sub-steps:

(a) Create an n-fold ion by taking n electrons into the emitter from a neutral atom situated at the crest of pass N, placing these electrons at the emitter Fermi level.

(b) Move the resulting n-fold ion to point F.

(c) Take an additional $(r - n)$ electrons into the emitter from the ion, placing them at the Fermi level.

(d) Move the r-fold ion to the collector.

The works done in these four sub-steps (collectively called subcycle (11)) are:

$$\text{(a)}\ U_n^{\mathrm{pa}}(0 \to n) - U_0^{\mathrm{pa}}(0 \to n), \qquad \text{(c)}\ U_r^{\mathrm{F}} - U_n^{\mathrm{F}},$$

$$\text{(b)}\ U_n^{\mathrm{F}} - U_n^{\mathrm{pa}}(0 \to n), \qquad \text{(d)}\ U_r^{\mathrm{c}} - U_r^{\mathrm{F}}, \tag{12}$$

where the reference $(0 \to n)$ indicates that the term value is in this case to be taken at the pass associated with the escape of an n-fold ion from the surface.

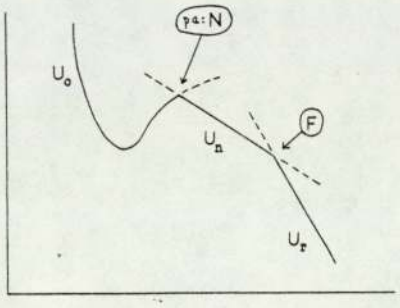


Fig. 3. A possible potential configuration in which post-ionization could follow desorption.

The works done in the remaining steps of cycle (3) are the same as in expressions (4), except that the pass appearing in step (5) has to be designated as the $0 \rightarrow n$ pass. On summing over the complete cycle, most of the terms appearing in expressions (12) cancel. Using the onset condition $U_r^c = U_n^{pa}$ $(0 \rightarrow n)$ then leads to the result:

$$A_{0r}^0 = H_r^\infty + U_0^\infty - U_n^{pa}\,(0 \rightarrow n) \tag{13}$$

In both eq. (9) and eq. (13) the first term on the right-hand-side is the heat of formation (in remote field-free space) of the ion that arives at the collector, whereas the last term is the term value at the crest of the pass through which escape occurs, for the entity that actually escapes. Thus, if one denotes this "pass energy at escape" quite generally by U^{pa}, then eqs. (9) and (13) can be combined into the common formula:

$$A_{0r}^0 = H_r^\infty + U_0^\infty - U^{pa}\,. \tag{14}$$

This formula will also apply to the potential configuration shown in fig. 4, which corresponds to a possibility suggested on experimental grounds by Moore and Spink [15], and subsequently Waugh [6], and on general theoretical grounds by Forbes [16], that field evaporation may be a multiple-stage process. This would involve: (i) escape of a neutral atom over a lateral potential hump (at C); (ii) lateral motion across the surface (though maybe only for a short distance); (iii) a field ionizing transition (at E); (iv) detachment from the surface. In this, the "neutral-diffusion" mechanism for field evaporation, the term U^{pa} is the potential at the crest of a diffusion activation-energy hump which the neutral atom must surmount in order to move across the surface.

For some purposes it may be useful to put eq. (14) into an alternative form involving binding energies. The differences in term values in eq. (14) can be expanded as in eq. (15), where U_0^b is the term value at the bonding point for the neutral, and

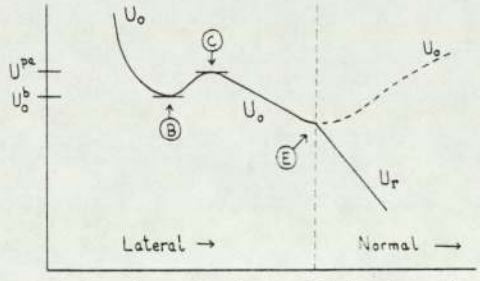


Fig. 4. The simplest potential configuration that could give rise to the neutral-diffusion mechanism. The left-hand-part of the horizontal axis denotes distance parallel to the emitter surface, the right hand-part denotes distance normal to the emitter surface. This topography may be called the "traverse to a valley" situation.

the brackets can then be substituted for as in eq. (16):

$$A^0_{0r} = H^\infty_r + (U^\infty_0 - U^b_0) + (U^b_0 - U^{pa}) , \tag{15}$$

$$A^0_{0r} = H^\infty_r + \Lambda^F_0 - Q_{0n} . \tag{16}$$

Λ^F_0 is the binding energy of the neutral, in the presence of the field, to whatever surface layer is present at the emitter; Q_{0n} is the activation energy necessary for the bound neutral to escape as an n-fold ion. If field desorption is ocurring at a resonable rate, then Q_{0n} is necessarily small, and may usually be ignored.

3.4. Generalisation to ionic bonding states

To derive a proof for the case where the desorbate is initially bound as an m-fold ion, consider the more general cycle:

(1) Take $(r - m)$ electrons into the emitter from an m-fold ion situated at the crest of the pass through which it is escaping, placing the electrons at the Fermi level;

(2) Move resulting r-fold ion to collector;

(3) Take r electrons round circuit to collector;

(4) Take r electrons from collector to ion, neutralising it; (17)

(5) Return neutral atom to emitter, and position it at the bonding point for the neutral;

(6) Convert neutral to an m-fold ion, placing the electrons at the Fermi level, and the resulting ion at the bonding-point for an m-fold ion;

(7) Move m-fold ion from bonding point to its original position at the pass.

The works done by a hypothetical external agent in each step are:

(1) $U^{pa}_r (m \to r) - U^{pa}_m (m \to r)$, (5) $U^b_0 - U^c_0$,

(2) $U^c_r - U^{pa}_r (m \to r)$ (=0) , (6) $U^b_m - U^b_0$,

(3) $r(\mu_c - \mu_e)$, (7) $U^{pa}_m (m \to r) - U^b_m$.

(4) $r\phi_c - H^\infty_r$, (18)

Here, U^b_m denotes the value of the standard m-fold ionic term at the bonding point for the m-fold ion; the bonding points for the neutral and the m-fold ion are usually not at the same position.

The total work done in the cycle is:

$$w = r(\mu_c - \mu_e) + r\phi_c - H_r^\infty + [U_r^{pa}(m \to r) - U_0^c] , \tag{19}$$

and by arguments exactly analogous to those used earlier, we obtain the result for the standard $m \to r$ field-ion appearance energy A_{mr}^0:

$$A_{mr}^0 = H_r^\infty + U_0^\infty - U_r^{pa}(m \to r) . \tag{20}$$

As before, if in fact the m-fold ion escapes as an n-fold ion and then suffers post-ionization to an r-fold ion, then the only change necessary is to designate the pass in eq. (20) as the $m \to n$ pass, and change the subscript to n.

Comparison of these results with eq. (14) shows that it is legitimate to write a general formula for A_{mr}^0:

$$A_{mr}^0 = H_r^\infty + (U_0^\infty - U^{pa}) . \tag{21}$$

This formula gives the standard field-ion appearance energy, whatever the desorption mechanism; H_r^∞ is the free space heat of formation for the ion that arrives at the collector; U^{pa} is the pass energy of the entity that actually escapes, at the pass through which it escapes.

As before, eq. (21) can be converted to a form involving binding energies:

$$A_{mr}^0 = H_r^\infty + (U_0^\infty - U_0^b) + (U_0^b - U_m^b) + (U_m^b - U^{pa}) , \tag{22}$$

$$A_{mr}^0 = H_r^\infty + \Lambda_0^F + \Delta\Lambda_{0m}^F - Q_{mn} . \tag{23}$$

Q_{mn} is the activation energy necessary for a bound m-fold ion to escape as an n-fold ion, and can usually be ignored. $\Delta\Lambda_{0m}^F$ is the energy of formation, at the surface in the conditions of the experiment, for the formation of an m-fold ion from a neutral, the electrons being placed at the Fermi level. $\Delta\Lambda_{0m}^F$ is positive if the ion is thermodynamically more stable than the neutral. The desorbate has a charge me when bound, a charge ne immediately after escape, and a charge re on arrival at the collector; the formulae hold for $r \geqslant n \geqslant m \geqslant 0$, provided $r \geqslant 1$.

4. Comparison with other theoretical results

What are in effect formulae for standard appearance energies have been stated by various previous authors, usually in the context of specific desorption situations. This section compares their results with the general formulae just derived. Where the original formula relates to an energy deficit or analogous quantity, it has been converted into a formula for appearance energy using the equations in section 2 or the Appendices. Where appropriate, the original notation has been changed to that used in the present paper.

In most cases, the cited formulae apply to the formation of a singly charged ion from a neutral. For this case, my eq. (16) reduces to eq. (24) below, if the activation energy term Q_{0n} and higher-order corrections in the term Λ_0^F are ignored:

$$A_{01}^0 = I_1 + U_0^\infty - U^{pa} = I_1 + \tfrac{1}{2}\,\alpha_0'\,(F^{pa})^2\,. \tag{24}$$

α_0' is the polarisation constant [17] for the neutral, and F^{pa} is the local field at the pass.

4.1. Jason's approximation

To interpret his measurements on hydrogen [10], Jason used the approximation:

$$A_{01}^0 = I_1\,, \tag{25}$$

which is valid because the polarisation energy for hydrogen is relatively small. Heinen et al. [3] also used this formula with organic molecules, and obtained good agreement between experiment and theory for many different species. Presumably this means that the polarisation energy (at the fields necessary for field desorption) is also small for the organic species they investigated.

4.2. The formula of Goldenfeld et al.

The formula given for A_{01}^0 by Goldenfeld et al. [2] is:

$$A_{01}^0 = I_1 + |U_0(x^i) - U_0^\infty|\,, \tag{26}$$

where $U_0(x^i)$ is the "potential energy of the neutral particle at the ionization point". Comparison of this with eq. (24) shows two major differences.

First, in my formula the term values are written with the opposite signs, and the modulus of the difference does not occur. In practice, the difference $(U_0^\infty - U^{pa})$ in my formula is positive in all normal circumstances, so both formulae agree that the appearance energy should be greater than the ionization potential.

Second, my formula contains the term value at the "point of desorption" (U^{pa}), whereas Goldenfeld et al.'s formula contains the term value at the point of ionization $(U_0(x^i))$. This difference is significant. It means that their formula is compatible with general thermodynamic arguments only for those mechanisms of field desorption in which ionization and escape occur simultaneously, namely those based on fig. 1b (and, arguably, fig. 1c). In other cases their formula is incorrect, although for configuration 1a the error is negligible. Eq. (24) is thus to be preferred.

4.3. The Tsong and Müller formula

In the case of noble-gas field ionization, investigated experimentally by Tsong and Müller [8], Goldenfeld et al.'s formula and my formula coincide in producing

eq. (24). Tsong and Müller, however, use the formula:

$$A^0_{01} = I_1 + e^2/4(4\pi\epsilon_0)x^{pa}\,, \tag{27}$$

where x^{pa} here is the critical distance of elementary field ionization theory. (Eq. (27) is a rationalised version of eq. (2) in ref. [8].)

In terms of the assumptions made in this paper, eq. (27) has been incorrectly derived, and there is no mention in ref. [8] of any special assumption that would lead to eq. (27). However, as Forbes has pointed out [18], the fact that Tsong and Müller did obtain agreement between their formula and their experimental results remains to be explained. (One possibility is that the field ionization of inert gases is not "electronically slow"; the other is systematic experimental error.)

4.4. Anway's formula

In connection with some work on the field ionization of water [1], Anway extended Jason's argument and derived a formula equivalent to eq. (28):

$$A^0_{01} = I_1 + \tfrac{1}{2}\,\alpha'_0\,(F^{pa})^2 - \tfrac{1}{2}\,\alpha'_1\,(F^{pa})^2\,. \tag{28}$$

where α'_1 is the polarisation constant for the singly-charged ion. Anway's derivation of this formula uses the (unproved) assumption that "when ionization takes place within an electric field, the ionization potential is shifted by the difference in polarisation energies between the ion and the neutral". Since this assumption leads to different conclusions from the present discussion, which is firmly based on a thermodynamic argument, one must presume that there is some fault either in the assumption itself or in another part of Anway's argument. It is now well known that there are difficulties in defining what is meant by an Ionization Potential near an emitter surface [19].

The question of whether eq. (24) is applicable to the production of H_3O^+ ions is a separate issue. The author's view is that the field desorption of molecular fragments requires a more general theory than is presented here.

4.5. Waugh's formula

Appearance energies for the field evaporation of metals have been measured by Waugh [6,20]. If it is assumed that the desorbed entity was initially bound in a neutral state, that no post-ionization occurs, and that Λ_0^F can be expanded in the form $\Lambda_0 + \frac{1}{2}\,\alpha'_0 F^2$ (Λ_0 is the zero-field value of Λ_0^F), then eq. (16) leads to the result:

$$A^0_{0r} = H_r^\infty + (\Lambda_0 + \tfrac{1}{2}\,\alpha'_0 F^2) - Q_{0r}\,. \tag{29}$$

Waugh's derivation for the image-hump model [6] also reaches this result, but his derivation for the charge-exchange model leads to the formula:

$$A^0_{0r} = H_r^\infty + \Lambda_0 + \tfrac{1}{2}\,\alpha'_0 F^2 - \Delta E - \tfrac{1}{2}\,\Gamma - Q_{0r}\,, \tag{30}$$

where ΔE is an energy shift, and $\frac{1}{2}\Gamma$ is a term attributed to broadening [21].

The simplicity and general nature of the derivation of eq. (16) strongly suggests that formula (29) is the correct result for both models of field evaporation. Waugh reaches eq. (30) by a specific integration [20]. The present author feels that the discrepancy between eq. (30) and eq. (29) results because Waugh's argument is based on the configuration of fig. 1c, and in this configuration there are certain energy-shift effects that Waugh's integration does not properly take into account. This argument will be presented in more detail elsewhere [26].

4.6. Röllgen and Heinen's formula

In connection with the field desorption of doubly-charged benzene ions from an ionic bonding state, Röllgen and Heinen derived a formula that gives the appearance parameter p_{12}^0 (cf. Appendix A) in terms of the second ionization potential and some small binding-potential contributions [5]. It may be shown (Forbes, unpublished work) that their formula can be considered as a special case of a general formula obtained by transforming eq. (15), but the proof is not straightforward. The arguments will be presented elsewhere.

5. Discussion

5.1. Assessment

The analyses in section 4 indicate that for field ionization, field evaporation, and the field desorption of simple molecular ions, all the specialised appearance-energy (or equivalent) formulae in the literature can be accounted for. Either they are special cases of the generalised formulae derived earlier, or they are approximations to a special case, or reasons can be found for supposing the specialised derivation to be in error.

Thus, from a first-principles microscopic approach based on the use of Molecular Terms and a thermionic-type cycle, general formulae have been obtained which apply to a wide range of desorption mechanisms and equally to field ionization, field evaporation and molecular field desorption.

In the case where the initial bonding state is neutral, eq. (14) can also be given a more abstract interpretation, by re-arranging it into the form:

$$H_r^\infty = -(U_0^\infty - U^{\mathrm{pa}}) + A_{0r}^0 , \tag{31a}$$

or:

$$\Delta \mathcal{U} = q + W . \tag{31b}$$

The desorbed atom or molecule can be thought of as a thermodynamic system, and the process of field desorption can be treated notionally as a device for increasing

its internal energy ($\mathcal{U}$) by an amount H_r^{∞}. Eq. (31a) is then seen as a disguised form of the Second Law of Thermodynamics: A_{0r}^{0} is the (electrical) work done on the system; $(U_0^{\infty} - U^{pa})$ is the heat given out (to the emitter) by the system.

Although this abstract mode of thinking can sometimes be useful, it should not be applied in analysing the behaviour of instruments, because the voltage generator can also do work in moving the ion around.

5.2. Limitations

The arguments presented here have obvious limitations of two main kinds. First, the treatment is confined to standard appearance energies. There are various circumstances in which the observed onset appearance energy might not correspond to a standard appearance energy; in particular, when most emitted electrons enter the metal at levels above the fermi level.

Difficulties of a different type arise if the desorbing entity is a charged molecule rather than a charged atom, because one must then consider whether it is possible and likely that the molecular ion will be formed in its internal ground state (as implied by the definition of a standard molecular term). And problems could also arise if the term value for the molecule (or molecular ion) as a whole were a strong function of its orientation relative to the emitter and/or the direction of the electric field.

In general, one might expect to classify molecular field desorption into "well-behaved" species and "others". With the "others", the formulae given earlier should apply moderately satisfactorily if the energy involved in any "non-standard" happening is small; if the energy is large, then the results may provide a probe into the nature of the actual field desorption mechanism.

The second limitation is that not all the arguments apply to situations where the arriving ion is a radical (e.g. H_3O^+) *not* derived from a parent molecule by straightforward electron removal.

It is intended to show at a later stage how the arguments here can be extended to overcome these limitations.

5.3. Comparison with experiment

Though the main objective has been to correlate theoretical expressions, it is instructive to illustrate how experiment and theory currently match up. Some representative data have been collected into table 1 from the fuller lists in the references shown, and standard appearance energies have been calculated for the noble gases and the metal species using eq. (29) but ignoring Q_{0r}. For He^+ and Ne^+, Λ_0^F has been taken equal to a polarisation binding energy of about 0.14 eV; for the metal ions, it is assumed that the species is originally bound in a neutral state, the value of the sublimation energy given by ref. [21] has been taken, and between 2 and 3 eV has been added to make some allowance for field-induced binding effects. Conceivably,

Table 1
Selected data concerning appearance energies (all energies are in eV)

Species	H_r^{∞}	A_{0r}^{0} (est)	A_r^{on} (obs)	Ref.
H_2O^+	12.60		12.65	[2]
NH_3^+	10.16		10.25	[2]
$C_3H_8^+$	11.07		11.10	[2]
$CH_3CO-OCHCH_2^+$	9.19		9.30	[2]
He^+	24.46	24.6	25.2–25.5	[8]
Ne^+	21.47	21.6	22.2	[8]
$C_6H_6^{++}$	26.0	~26.1	28.0	[22]
Rh^{2+}	25.5	~34	23.5 ± 4	[6]
Re^{3+}	60	~71	90 ± 9	[6]

somewhat higher allowance should be made. In converting Waugh's measured metal-ion deficits [20], using eq. (41), work function values as quoted in ref. [21] have been employed. For the molecular species, all the data have been taken from ref. [2]; Λ_0^F cannot be reliably estimated for these because data concerning desorption fields are not given there.

Several points stand out. For the molecular species the approximation $A_{01}^0 = I_1$ works well for both inorganic and organic molecules, even very complicated ones. The discrepancy between experiment and theory is small, and further progress in this area depends on the development of high-resolution spectrometers rather than on the development of theory.

For the noble gases $A_{01}^0 = I_1$ is still a good approximation, but there is a significant discrepancy between theory and experiment, albeit less than 1 eV. The urgent need here is for the re-performance of the basic experiments, using a carefully-calibrated high-resolution instrument [18].

For the doubly-charged benzene ion, there is a discrepancy between the observed value and the theoretical values as calculated on the assumption that the ions desorb from a neutral bonding state. Röllgen and Heinen assume that this discrepancy exists because the ions in fact desorb from a bonding state in which they are singly-charged [5,22].

For the metal species the simple approximation $A_{0r}^0 = H_r^{\infty}$ does not work, but neither does the fuller formula, eq. (29), based on the assumption that the metal atoms bind in a neutral state. Even though A_{0r}^0 cannot be estimated precisely (because the field-induced-binding term $\frac{1}{2}\alpha_0' F^2$ is not known reliably to within 2 or 3 eV), there are marked discrepancies between theory and experiment. More surprising still, the estimate is too high in one case, too low in another.

The comparison of metal-species data and theory presented here is broadly

equivalent to that presented by Waugh elsewhere [6,20]. I concur with him in thinking that in this area the theory needs to be taken further, and that there is the possibility of discovering useful things about the mechanism of field evaporation. It would also be helpful to have the experimental information confirmed and extended.

In this comparison there is no attempt to present data on associated ions, such as has been gathered by Goldenfeld et al. [2], although it seems clear that information about the energetics of surface reactions can, in practice, be deduced empirically from observed field-ion appearance energies.

In general, the agreement between experiment and theory is best when the field needed for desorption is low, or the ionic charge is low. With the higher ionic charges, there is a marked discrepancy between experiment and theory if it is assumed that the ions have desorbed from a neutral bonding state, and electron transitions have been made into states close to the Fermi level. Further theoretical work should investigate mechanisms of desorption to which these restrictions do not apply.

Appendix A: Alternative language

Use of deficits. In eq. (1) a field-ion appearance energy is defined in terms of an onset voltage. But in the literature one often finds ion energy distributions plotted against an energy deficit (D) or a voltage deficit (d) defined in terms of my $\Delta\zeta$ by:

$$-re\,\Delta\zeta = D = red\,. \tag{32}$$

The standard field-ion appearance energy is then defined in terms of a "critical energy deficit" (D_r^{on}) by:

$$A_r^{\mathrm{on}} = D_r^{\mathrm{on}} + r\phi_{\mathrm{c}}\,. \tag{33}$$

Other appearance quantities. There are two quantities closely related to an appearance energy that may sometimes be used instead of it. These are defined by:

$$A_r = rp_r = rea_r\,. \tag{34}$$

The names given here to these are: p_r is an *appearance parameter*; a_r is an *appearance voltage*. The same extra suffices (depending on context) can be added to p_r and a_r as have been used in the main text with A_r.

Appearance potential. In the literature the name "appearance potential" is used with a variety of meanings. In terms of the symbols used here, the quantities called "appearance potential" by the authors shown are to be interpreted as given in table 2. Note that Anway's quantity I is a parameter associated with the formation of

Table 2

Source	Name and symbol in source	Name and symbol used here
Goldenfeld et al. [2]	Appearance potential, $(AP_f)_n$	Appearance energy, A_r
Rollgen et al. [3–5,22]	Appearance potential, $AP(M^{n+})^{25}$	Appearance parameter, p_r
Forbes [8,18]	Appearance potential A_n	Appearance energy, A_r
Anway [1]	Appearance potential, I	No precise equivalent

an H_3O^+ ion in zero-field conditions, whereas all the other parameters refer to processes happening in the conditions of experiments.

To prevent confusion, the present paper avoids use of the name "appearance potential", although I used it in earlier papers. The subscript notation is also different here from in ref. [18].

An analogue to ionization potential. A true free-space ionization potential, I_r, can be defined by the equation:

$$I_r = H_r^{\infty} - H_{r-1}^{\infty} . \tag{35}$$

This formula suggests that the closest field-desorption analogue to a genuine ionization potential would be the quantity J_r^0 defined by:

$$J_r^0 = A_{0,r}^0 - A_{0,r-1}^0 . \tag{36}$$

This quantity has found little use in the literature so far.

Appendix B: Relationships for dispersive analysers

Dispersive ion energy analysers can be designed in two basic ways. In the scanning type, ions of a given charge are transmitted only if they have a kinetic energy close to the transmission energy K_a. If the voltage difference $(\zeta_s - \zeta_e)$ between the analyser entry slit (subscript s) and the emitter (subscript e) corresponds to onset of the energy distribution, then the field-ion appearance energy is given by:

$$A_r^{on} = -re(\zeta_s - \zeta_e)^{on} - K_a + r\phi_s , \tag{37}$$

where ϕ_s is the work function of the material associated with the entry slit. K_a is a known constant of the analyser; $(\zeta_s - \zeta_e)^{on}$ is measurable or experimentally deducible. A field-ion energy analyser of this type has been used by Jason, Hanson, and Inghram [10,27,28].

The alternative method is to fix the value of $(\zeta_s - \zeta_e)$, and allow ions of the same charge but different kinetic energies to be dispersed and focussed to different posi-

tions on some detector. The analysers described in refs. [20] and [23] operate in this way. The onset ions are characterised by a specific value (K^{on}) of kinetic energy when emerging from a defining slit. It is convenient to write this in the form of an "equivalent voltage" W^{on} defined by:

$$K^{on} = reW^{on} . \tag{38}$$

The onset appearance energy is then given by:

$$A_r^{on} = -re(\zeta_s - \zeta_e) - reW^{on} + r\phi_s . \tag{39}$$

If the appearance energies of different ions, say a metal ion and a helium ion, are measured with the same value of $(\zeta_s - \zeta_e)$, then clearly these will be related by:

$$[A_r^{on}(M^{r+})]/r - A_1^{on}(He^+) = e[W^{on}(He^+) - W^{on}(M^{r+})] . \tag{40}$$

So if the appearance energy of helium is a known quantity, then the appearance energy of the metal ion can be obtained from the "voltage difference" between the helium and metal lines.

This neat experimental procedure of performing first field evaporation of the metal, then field ionization of helium, and recording both lines on the same film frame, was devised by Waugh [6]. The voltage difference is thus directly accessible experimentally.

Waugh's analysis of his results, however, differs from that given here, because it uses the difference ($\Delta\Upsilon$) in classical electrostatic potential between points just outside the emitter and the defining slit, rather than the voltage difference between their fermi levels. In his approach eq. (39) would be replaced by:

$$A_r^{on} = re(\Delta\Upsilon - W^{on}) + r\phi_e , \tag{41}$$

where ϕ_e is a local work function associated with the emitter; but eq. (40) holds in both approaches. Waugh tabulates his results in the form of values of $(\Delta\Upsilon - W^{on})$, which he calls the "voltage deficit".

Appendix C: Surface molecular terms

In dealing with surface ionization processes in high electric fields, it is convenient to use the set of potentials introduced by Gomer and Swanson [24]. *Standard surface molecular terms* represent the potential energy of the emitter-plus-desorbate system, as a function of the position of some reference point in the desorbate; for an atom or atomic ion the reference point is its nucleus. The terms are formally defined by the equations:

$$U_0^R = U_0^\infty = 0 , \tag{42}$$

$$U_0^p - U_0^R = w_0 (R \rightarrow p) , \tag{43}$$

$$U_n^p - U_0^p = W_{0n}^p , \tag{44}$$

$$U_n^s - U_n^p = w_n (p \rightarrow s) . \tag{45}$$

With these quantities, the subscripts refer to charge states and the superscripts to positions. The superscript R labels any point in remote field-free space; eq. (42) defines the zero of potential.

The quantities on the r.h.s. of the equations denote the works done by a hypothetical external agent in the course of slow processes. w_0 (R → p) is the work done in moving a neutral from point R to point p; w_n(p → s) is the work done in moving an n-fold ion from point p to point s. W_{0n}^p is the work needed to create an n-fold ion from the neutral, at point p, the electrons being placed at the emitter Fermi level by means of slow thermodynamically-reversible processes.

It is implicit in the definition of a standard term that the desorbate shall be in its internal ground state.

Acknowledgements

I would like to thank Dr. A.R. Waugh for discussing his unpublished experimental results with me, to thank Professor J.H. Block for drawing my attention to the work of Goldenfeld and to the chemical significance of field-ion appearance energies, and to acknowledge a helpful correspondence with Professor F.W. Röllgen. I. thank the UK Science Research Council for personal financial support, and Professor E. Braun and Dr. R.V. Latham for provision of facilities.

Note added in proof

The author's attention has now been drawn to the very interesting work of Hanson on the field ionization of H_2 and HD [27], and of Hanson and Inghram on the field ionization of Helium [28].

Ref. [27] contains a thorough treatment of the complications introduced into appearance-energy theory by the vibrational and rotational behaviour of a molecular ion immediately after ionization, before subsequent dissociation. Ref. [28] reports experimental results which show that the observed onset appearance energy for Helium varies with position on the emitter surface, and discusses some possible reaons why this should differ from the standard appearance energy.

Both papers contain formulae for the critical *electrostatic* energy deficit "ΔE_c",

which is related to my A_{01}^{0} by:

$$A_{01}^{0} = \Delta E_c + \phi_e \,, \tag{46}$$

where ϕ_e is a local work function associated with the emitter. Their formulae and my eq. (24) are equivalent.

References

[1] A.R. Anway, J. Chem. Phys. 50 (1969) 2012.
[2] I.V. Goldenfeld, I.Z. Korostyshevsky and B.G. Mischanchuk, Intern. J. Mass Spectrom. Ion Phys. 13 (1974) 297.
[3] H.J. Heinen, F.W. Röllgen and H.D. Beckey, in: 21st Intern. Field Emission Symposium, Marseille, 1974.
[4] H.J. Heinen, F.W.Röllgen and H.D. Beckey, Z. Naturforsch. 29a (1974) 773.
[5] F.W. Röllgen and H.J. Heinen, Z. Naturforsch. 30a (1975) 918.
[6] A.R. Waugh, Ph.D. thesis, Cambridge Univ. (1975).
[7] A.R. Waugh and M.J. Southon, in: 21st Intern. Field Emission Symposium, Marseille, 1974.
[8] T.T. Tsong and E.W. Müller, J. Chem. Phys. 41 (1964) 3279.
[9] R.G. Forbes, in: 22nd Intern. Field Emission Symposium, Atlanta, 1975.
[10] A.J. Jason, Phys. Rev. 156 (1967) 266.
[11] E.W. Müller, Phys. Rev. 102 (1956) 618.
[12] R. Gomer, J. Chem. Phys. 31 (1959) 341.
[13] R.G. Forbes, Vacuum 22 (1972) 517.
[14] E.W. Müller and T.T. Tsong, Progr. Surface Sci. 4 (1974) 1.
[15] A.J.W. Moore and J.A. Spink, Surface Sci. 44 (1974) 198.
[16] R.G. Forbes, Surface Sci. 46 (1974) 577.
[17] The polarisation term here is considered as rationalised, in accordance with a procedure that will be described elsewhere: $\alpha_0' = (4\pi\epsilon_0)\alpha^{unr}$.
[18] R.G. Forbes, Intern. J. Mass Spectrom. Ion Phys. 21 (1976) 417.
[19] For example, H.D. Hagstrum, Surface Sci. 54 (1976) 197.
[20] A.R. Waugh and M.J. Southon, J. Phys. D (Appl. Phys.) 9 (1976) 1017.
[21] E.W. Müller and T.T. Tsong, Field ion Microscopy (American Elsevier, New York, 1969).
[22] F.W. Röllgen and H.J. Heinen, Intern. J. Mass Spectrom. Ion. Phys. 17 (1975) 92.
[23] E.W. Müller and S.V. Krishnaswamy, Surface Sci. 36 (1973) 39.
[24] R. Gomer and L.W. Swanson, J. Chem. Phys. 38 (1963) 1613.
[25] The definitions of this quantity in the papers cited contain minor errors.
[26] R.G. Forbes, J. Phys. D (Appl. Phys.) 9 (1976) 4191.
[27] G.R. Hanson, J. Chem. Phys. 62 (1975) 1161.
[28] G.R. Hanson and M.G. Inghram, Surface Sci. 55 (1976) 29.

International Journal of Mass Spectrometry and Ion Physics, 21 (1976) 417–420

Short communication

ON THE NEED FOR NEW MEASUREMENTS OF THE FIELD-ION APPEARANCE POTENTIAL FOR HELIUM

R. G. FORBES
University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET (Gt. Britain)
(Received 26 November 1975)

The objective of this note is to report the main results of a generalized theory of standard field-ion appearance potentials, and to draw attention to the need for some new, absolute, high-resolution measurements of the field-ion appearance potential for helium.

The concept of a field-ion appearance potential has been much used by Beckey, Röllgen and co-workers [1], and other researchers in field-ion mass spectrometry, in discussions of the energy deficits observed in the field desorption of molecular species. In recent papers [2, 3], Forbes has refined the concept and applied it to field ionization and field evaporation as well as to field desorption. Formulae have been given which enable experimental energy-deficit parameters encountered in the literature to be converted to "observed appearance potentials".

In ref. 3, the appearance potential theoretically manifested by an ion all of whose separated electrons have been transferred slowly and directly to the emitter fermi-level, and which in a classical approach just had sufficient energy to escape from the field-ion emitter (and then moved slowly away without suffering any anomalous thermal deactivation), is called the Standard nth Field-Ion Appearance Potential for the species in question and is denoted by A^0_{ns}. General formulae derived for A^0_{ns} can be expressed in two ways. In terms of potential-energy values (molecular term values) near the emitter surface:

$$A^0_{ns} = H_n + U^\infty_0 - U^{pa} \tag{1}$$

where H_n is the total heat of formation, in remote field-free space (the electrons also being left in remote field-free space), of the ion that arrives at the collector; U^∞_0 is the potential energy of the neutral in remote field-free space; and U^{pa} is the potential energy of the entity that desorbs, at the point of desorption. (H_n is

given by ΣI_n, where I_n is the nth ionization potential for the desorbing species, and the sum is over the first n ionization potentials.) In terms of thermodynamic parameters, for an entity bound as a neutral before desorption:

$$A^0_{ns} = H_n + \Lambda^F_0 - Q_m \tag{2}$$

whereas, for an entity bound as a k-fold ion before desorption:

$$A^0_{ns} = H_n + \Lambda^F_0 - \Delta U^*_k - Q_{km} \tag{3}$$

where Λ^F_0 is the binding energy of the entity in its neutral state, in the presence of the field; ΔU^*_k is the energy required to form a k-fold ion, from the neutral, in the conditions existing at the surface prior to desorption (ΔU^*_k is negative if ionic bonding is thermodynamically favoured); Q_m is the activation energy necessary for a neutral to field-desorb as an m-fold ion; and Q_{km} is the activation energy necessary for a k-fold ion to field-desorb as an m-fold ion. These formulae will work for $n \geqslant m \geqslant k \geqslant 0$, provided $n \geqslant 1$. If a significant desorption current is being observed, then the activation-energy term is necessarily small (probably less than 0.1 eV, at room temperature) and can usually be ignored.

These general formulae do not coincide in all cases with the formula of Goldenfeld et al. [4], nor in all cases with the specialized derivations of other workers [5, 6]. The presumption is that the formulae here are correct if it is a standard appearance potential (or equivalent) that is being calculated.

Energy deficits for field-ionized helium were measured many years ago by Tsong and Müller [5]. Their results imply an observed onset appearance potential of between 25.2 and 25.5 eV. For helium, eqn. (1) reduces to:

$$A^0_{1s} = I_1 + \tfrac{1}{2}\alpha F^2 \tag{4}$$

where I_1 is the first ionization potential for helium; α is the polarisability of helium; and F is the electric-field value at the critical distance of elementary field ionization theory. Using $I_1 = 24.46$ eV, $\alpha = 0.0002$ nm^3 and $F = 45$/V nm gives $A^0_{1s} = 24.60$ eV. There is thus a discrepancy between experiment and theory of between 0.6 and 0.9 eV, in an instrument whose resolution was estimated to be between 0.03 and 0.04 eV.

This discrepancy might, of course, be an experimental artifact. For example, one could postulate that, as a result of adsorbed contaminants, the collector work function altered in value between the time at which it was measured and the time at which the helium measurements were taken. Vacuum technique was not as good in 1963 as it is today. However the discrepancy could be real; in which case there would be significant consequences, both for the theory of field ionization and for the calibration procedure used in dispersive energy analysers such as Waugh's [6].

In general, there are many possible theoretical explanations for discrepancies between observed and standard appearance potentials. But helium field ionization

is a relatively simple process, and most can be excluded. The least unlikely explanation of a real discrepancy seems to be that the electronic transition and/or ion removal might be fast, relative to some relevant electronic relaxation time associated with the metal [2]. If this were true for helium field ionization, it might well be true for metal field evaporation, and this knowledge might aid the interpretation of field evaporation appearance potentials. To prove the "non-adiabaticity" of this electron transition would also be of wider interest in surface physics, since there are connections with arguments concerning surface plasmons [7] and with discussion concerning the validity of "Koopman's approximation" in the theory of the electron spectroscopies [8].

Knowledge of the helium appearance potential is also necessary in calibrating dispersive analysers used in the manner Waugh describes [6]. It may be shown [3] that the onset appearance potential for the metal ion M^{n+} is related to that for helium by:

$$A_n^{\circ n}(M^{n+})/n - A_1^{\circ n}(He^+) = e\Delta W \tag{5}$$

where ΔW is a quantity (the "voltage separation") directly ascertainable from measurements. An uncertainty of 0.8 eV in the appearance potential for helium is an uncertainty of ca. 2.5 eV in that for a triply charged metal ion. In the present generation of dispersive analysers the instrumental resolution is several volts [6]; if instrumental resolution were improved significantly then uncertainty about the helium field-ion appearance potential could become the biggest source of error.

Intuitively, it seems relatively improbable that the helium field-ionizing transition could be fast enough for there to be ca. 0.8 eV associated with its "non-adiabaticity", and thus one tends towards the experimental-artifact explanation of the Tsong and Müller result. But physical conditions at a highly structured, highly charged, helium-covered, field-ion emitter are so peculiar that it may be unwise to exclude from consideration things that would be improbable at flat field-free metal surfaces.

There is thus an immediate need to be more certain about the experimental value of the helium field-ion appearance potential (and it would be interesting to see if any small variations with field strength could be detected). New determinations seem to be required, with an ultra-high-vacuum, high-resolution, retarding-potential analyser in which the collector work-function can be measured by independent means (probably some variant of Holscher's technique [9]) before and after the experiments with helium. I would commend this project to any experimentalist with a suitable instrument.

NOTE ADDED IN PROOF

Following a private correspondence with Professor F. W. Röllgen, I now think that the quantities A_{ns}° and $A_n^{\circ n}$ used in this paper are best called field ion

appearance energies. These are related to the quantity $AP(M^{n+})$ used in ref. 10 by: $AP(M^{n+}) = A_n^{on}/n$. Neither quantity, it seems to the present author, is the exact analogue of a free-space ionization potential.

REFERENCES

1 For example, H. J. Heinen, F. W. Röllgen and H. D. Beckey, *Z. Naturforsch. A*, 29 (1974) 773.
2 R. G. Forbes, in *22nd International Field Emission Symposium, Atlanta*, 1975.
3 R. G. Forbes, submitted to *Surface Science*.
4 I. V. Goldenfeld, I.Z. Korostyshevsky and B. G. Mischanuk, *Int. J. Mass Spectrom. Ion Phys.*, 13 (1974) 297.
5 T. T. Tsong and E. W. Müller, *J. Chem. Phys.*, 41 (1964) 3279.
6 A. R. Waugh, Ph.D. thesis, Cambridge University, 1975.
7 A. A. Lucas and M. Šunjić, *Surface Sci.*, 32 (1972) 439.
8 For example, W. G. Richards, *Int. J. Mass Spectrom. Ion Phys.*, 2 (1969) 419.
9 A. A. Holscher, *Surface Sci.*, 4 (1966) 89.
10 F. W. Röllgen and H. J. Heinen, *Z. Naturforsch. A*, 30 (1975) 913.

J. Phys. D: Appl. Phys., Vol. 9, 1976. Printed in Great Britain.

LETTER TO THE EDITOR

Comments on 'intrinsic energy losses of field-evaporated ions'

Richard G Forbes
Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET

Received 6 July 1976

Abstract Some improvements are suggested to the theory underlying the measurement of field-evaporation energy deficits, as presented by Waugh and Southon, and their experimental results are retabulated in the form of appearance energies.

I would like to make some comments on the recent paper of Waugh and Southon (1976) chiefly about the presentation and interpretation of their interesting results and about background theory. My aims are to bring the treatment of field-evaporation energy deficits more into line with the treatment of the field-desorption energy deficits measured by Heinen *et al* (1974) and others and (I hope) to clarify the theory.

(1) Following the procedure of Jason (1966), Waugh and Southon use in their theoretical expressions a quantity V which is the difference in classical electrostatic potential between points 'just outside' the emitter (subscript 'e') and 'just outside' the analyser entry slit (subscript 's'). An alternative would be to use the voltage difference $\Delta\xi(=\xi_s-\xi_e)$ between the Fermi levels of the analyser and emitter. This is related to V by

$$-e\Delta\xi = eV - \phi_s + \phi_e \tag{1}$$

where ϕ_s and ϕ_e are local work-functions associated with the entry slit and emitter respectively.

In my opinion, it would be better to present formulae in terms of $\Delta\xi$, because this is the quantity that is observed on the meter of a voltage generator, and because $\Delta\xi$ is a parameter associated with the emitter as a whole whereas V varies from plane to plane on the emitter surface (due to patch fields). I also think it better to use the emitter, rather than the analyser entry slit, as the basic potential reference: the energy deficit on their p 1022 can then be represented as a sum of a potential energy and a kinetic energy U rather than as a difference: that is, the equation would be for $-(ne\Delta\xi+U)$ rather than for $(neV-U)$, where ne is the charge on the ion (also the equation would contain ϕ_s rather than ϕ_e).

(2) If the ionic charge is known, then in tabulating results it is probably best to use neither of the above quantities, but the so-called 'field-ion appearance energy' A_n. This is defined by

$$A_n = neV - U + n\phi_e = -ne\Delta\xi - U + n\phi_s. \tag{2}$$

The merit of using A_n is that it depends only on the processes involved in desorption, not on those involved in collection and energy analysis. (To derive a value of V from meter readings, it is necessary to know ϕ_s.)

(3) The kinetic energy U of the ion on entering the analyser can be written in terms of an 'equivalent voltage' W by means of equation (3) below; and a quantity a_n (the 'appearance voltage') can be defined by equation (4):

$$U = neW \tag{3}$$

$$A_n = nea_n. \tag{4}$$

In Waugh and Southon's experiments, deficits are recorded for two species, the metal ion and a helium ion, with the same emitter and the same value of $\Delta\xi$ (and hence of V). In this case the appearance quantities for the two ions are related by

$$[A_n(M^{n+})]/ne - [A_1(He^+)]/e = a_n(M^{n+}) - a_1(He^+) = W(He^+) - W(M^{n+}). \tag{5}$$

With their procedure of recording both the metal line and the He line on the same film frame, the difference in equivalent voltage is directly obtainable (since the voltage-to-distance scale has been determined separately); hence the appearance voltage of the metal ion is easily obtained if the appearance voltage for He is known. Clearly, appearance voltage is a useful parameter when comparing theory and experiment.

(4) This procedure of using the He line for calibration purposes is neat and convenient, particularly in a dispersive analyser. However, as I have pointed out elsewhere (Forbes 1976), there exists at present a significant discrepancy between the calculated value of $A_1(He^+)$ and its value as measured absolutely by Tsong and Müller (1964). Uncertainty about the value of $A_1(He^+)$ would be a factor limiting the precision of the results in an analyser with higher resolution than Waugh and Southon's.

(5) In a private communication to Waugh, I have expressed reservations about the presence in his energy-deficit equation of the terms relating to a configuration-interaction energy-shift (ΔE) and broadening ($\frac{1}{2}\Gamma$). These reservations arise because the terms do not appear in an appearance-energy formula I have derived on the basis of more general, thermodynamic, arguments (Forbes, submitted for publication), and I would like to comment further.

The immediate reason why the terms appear in Waugh and Southon's energy-deficit equation is as they state, but there is (I think) a more basic reason that is related to confusion in the literature over the charge-exchange models for field evaporation. This model comes in two versions: in one the potential-energy curves are drawn as crossed, and the electron transition is envisaged as a 'hopping' process; in the other the potential curves are drawn as repelled, and the electron transfer is envisaged as a continuous 'draining' process. These are two physically *different* possible mechanisms for field evaporation. In the literature there is a widespread tendency to confuse them. For example, the mathematics of the charge-hopping mechanism is often presented in conjunction with the potential diagram for the charge-draining mechanism; in some cases, the theory presented is a mixture of the mathematics for the two mechanisms.

If *either* of these two mechanisms is analysed self-consistently by means of detailed integrations, after the fashion of Waugh and Southon, then the result would be a formula for $neV - U$ that omits ΔE and $\frac{1}{2}\Gamma$. Consider the charge-hopping mechanism. In this the ionic charge is a constant once electron transfer has occurred, so Waugh and Southon's second equation on their p 1021 in fact applies to the charge-hopping mechanism. Their third equation (for Q_n) contains ΔE and $\frac{1}{2}\Gamma$: this equation applies to the charge-draining mechanism: the equation for Q_n in the charge-hopping model is their equation but with ΔE and $\frac{1}{2}\Gamma$ omitted: if this modified equation is combined with their second equation, then the result is a formula for $neV - U$ without ΔE and $\frac{1}{2}\Gamma$.

For the charge-draining model, the correct form of the kinetic-energy equation is

obtained by introducing into their first equation on p 1021 additional terms that can be represented as a variation in ionic charge with position (but may actually be rather more complicated). For consistency with the results of a general approach in terms of potential energies, these additional terms must give rise to ΔE and $\frac{1}{2}\Gamma$ in a modified version of their second equation. These then cancel the terms in their third equation as stated, leading again to a formula for $neV - U$ without the ΔE and $\frac{1}{2}\Gamma$.

Waugh and Southon's equation on p 1022 contains the additional terms because they have derived it by combining two equations, both reasonable in themselves, but relating to two different mechanisms for field evaporation.

(6) Two final points. First, the phrase 'intrinsic energy loss' is something of a misnomer, because the ion never had the energy in the first place: 'intrinsic energy deficit' is a better description. Second, although the intrinsic deficits will not significantly affect the resolution of the atom-probe, the differences in intrinsic deficit between ions of different species (and charge states) might just be detectable as very small apparent shifts in the mass-to-charge ratio, and this possibility should be remembered when very precise calibration of the atom-probe is being attempted.

Table. Empirical field-evaporation appearance voltages and energies recalculated from data in Waugh and Southon (1976).

Specimen material	Assumed work-function (eV)	Empirical appearance voltage (V)	Assumed species	Corresponding appearance energy (eV)
W	4·5	17±3	W^{3+}	51±9
Rh	4·8	12±2	Rh^{2+}	24±4
Re	5·1	30±3	Re^{3+}	90±9
Mo	4·3	10±2	Mo^{2+}	21±4
		27±2	Mo^{3+}	82±6

In summary, I agree with Waugh and Southon that their results provide interesting and useful information about the theory of field evaporation, but I suggest that it is most useful if expressed in the form of appearance voltages or appearance energies. For future reference, I retabulate their numerical results in this form, in the table above. The work-function values used are those quoted by Waugh (1975).

References

Forbes R G 1976 *Int. J. Mass Spectrom. Ion. Phys.* **21** 417–20
Heinen H J, Röllgen F W, and Beckey H D 1974 *Z. Naturforsch.* **29a** 773–81
Jason A J 1966 *Phys. Rev.* **156** 266–85
Tsong T T and Müller E W 1964 *J. Chem. Phys.* **41** 3279–84
Waugh A R 1975 *PhD Thesis* University of Cambridge
Waugh A R and Southon M J 1976 *J. Phys. D: Appl. Phys.* **9** 1017–23

J. Phys. D: Appl. Phys., 13 (1980) 1357–63. Printed in Great Britain

Appearance energies for tungsten ions field-evaporated from ionic bonding states

Richard G Forbes
Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET

Received 7 January 1980

Abstract. General formulae are derived for the standard field-ion appearance energies associated with desorption from ionic bonding states. Values for tungsten are obtained, using the image-potential approximation, and are compared with the measured appearance energy for W^{3+} ions. The measured value is compatible with field evaporation from a primarily ionic bonding state, but a better understanding of correlation-type interactions at charged surfaces is necessary if useful information about initial bonding state is to be derived from appearance-energy measurements alone.

1. Introduction

In constructing a theory of field evaporation, one of the possibilities that needs to be considered (Forbes 1974, Müller and Tsong 1974) is that evaporation takes place out of an initial bonding state that is ionic or primarily ionic in character, rather than out of a neutral bonding state as has conventionally been assumed (for example, Müller and Tsong 1969). With the (111) face of tungsten, for example, at a field strength of 57 V nm^{-1}, Gauss' theorem indicates that the excess charge per surface atom in a crystallographically flat face would be 0·55 e, where e is the elementary (proton) charge. In an emitter facet of this orientation the kink-site atoms (which are those at the highest risk of evaporation) will necessarily have a somewhat higher excess charge than this, for elementary reasons of electrostatics and geometry. It is thus difficult to avoid the conclusion that at field strengths characteristic of field evaporation the kink-site atoms will be primarily ionic in character, certainly on this tungsten (111) facet.

In the present state of surface theory it is not possible to obtain any reliable estimate of the degree of ionicity of a kink-site atom. The objective of this note is to explore, using tungsten as an example, whether it is likely that any useful information about initial bonding state could be deduced from measurements of field-ion onset appearance energies. (See Cocke and Block 1978, for a general review of the use of this technique.)

The standard field-ion appearance energy, defined by Forbes (1976a), is the appearance energy appropriate to an ion that has in a quasiclassical theory 'just managed' to escape from the emitter surface, with zero kinetic energy, and that then moved slowly away from the surface without suffering any anomalous changes in energy. This standard appearance energy is denoted by the symbol $A_{mr}{}^{0}$, where me is the charge on the desorbing entity in its original bonding state, re is the charge on the ion on arrival at the energy analyser. Standard appearance energies are useful theoretically because they are defined relative to features of the potential-energy structure at the emitter surface;

0022-3727/80/071357+07 $01.50

hence questions concerning kinetic-energy distribution and extrapolation from observed ion-energy distribution curves do not enter into their definition. The raised suffix 0 is used to distinguish this theoretical quantity from experimentally measured appearance energies, whose values are in principle influenced by such effects (though normally only to a relatively small extent).

I have shown elsewhere (Forbes 1976a) that, provided no subsidiary chemical reactions are associated with the desorption process, the standard appearance energy $A_{mr}{}^0$ is given by the generalised formula:

$$A_{mr}{}^0 = H_r + (U_0{}^\infty - U^{\mathrm{pa}}) \tag{1}$$

where: H_r is the energy of formation of an r-fold ion, from the neutral, in remote field-free space, and is given by the sum of the first r ionisation energies; $U_0{}^\infty$ is the potential energy of the neutral desorbate in remote field-free space; and U^{pa} is the standard potential energy (Forbes 1976a) of the desorbate at the top of the potential hump (or 'pass') over which it escapes from the emitter.

Formula (1) is useful when the initial bonding state is neutral, because in this case it is usually straightforward to write down a specific approximation for U^{pa}; but it is less helpful when the initial bonding state is ionic. The first objective of this note is to derive an alternative formula. We begin by developing some notation.

2. Theoretical analysis

2.1. Components of potential energy

As in previous theory, let the symbol $U_n{}^s$ denote the standard potential energy of the desorbate particle when it is in an n-fold-charged state and at position s. The work $W_n(p \to s)$ done in moving an n-fold-charged ion from position p to position s can be split into two components, done respectively against 'electric' and 'chemical' forces. Thus we may write:

$$U_n{}^s - U_n{}^p = W_n(p \to s) = n(u^s - u^p) + \eta_n{}^s - \eta_n{}^p. \tag{2}$$

u represents the electrostatic component of the potential energy of a proton, and is related to the classical electrostatic potential Υ by:

$$u^s - u^p = e(\Upsilon^s - \Upsilon^p). \tag{3}$$

η_n represents the 'purely chemical' component of the potential energy of an external atom or ion; for an external ion, η_n is often approximated by the image-potential formula. The reference zero for this potential-energy component is taken at infinity, or at some point R in remote field-free space, for all values of n. Thus we obtain:

$$\eta_n{}^R = \eta_n{}^\infty = 0 \tag{4}$$

This completes the definition of η_n.

For convenience of manipulation, consider a position x just outside the emitter surface, and let the symbol $S_n{}^s$ be defined by

$$S_n{}^s = n(u^s - u^x) + \eta_n{}^s. \tag{5}$$

For two positions s and p it is easily verified from equation (2) that:

$$U_n{}^s - U_n{}^p = S_n{}^s - S_n{}^p.$$

The quantity S_n thus represents a form of potential energy for an n-fold ion, in which the chemical component is taken as zero at infinity but the electric component as zero at the position x.

To determine an expression for the standard potential energy $U_n{}^s$, consider an electrothermodynamic cycle that starts with a neutral desorbate particle in remote field-free space, and contains the following steps.

(i) Remove n electrons from the neutral particle, one by one, leaving them well separated in remote field-free space. Each electron is then returned to the emitter fermi level as follows:
(ii) move electron from remote field-free space to position s;
(iii) move electron from position s to the position x 'just outside' the emitter surface;
(iv) place electron at emitter fermi level;
(v) finally, move the ion from remote field-free space to position s. All the above steps are considered to be done slowly, i.e. thermodynamically reversibly. The works done in each step are as follows:

$$\begin{aligned} W(1) &= H_n = \sum_{N=1}^{n} I_N \\ W(2) &= -n(u^s - u^R) + n(w^s - w^R) \\ W(3) &= -n(u^x - u^s) + n(w^x - w^s) \\ W(4) &= -n\Phi^x - nw^x \\ W(5) &= n(u^s - u^R) + \eta_n{}^s - \eta_n{}^R. \end{aligned} \tag{6}$$

In the above: w^s denotes the correlation-and-exchange component of the potential energy of an electron at position s; R labels a position in remote field-free space; I_N denotes the Nth ionisation energy of the desorbate; and H_n denotes the energy necessary to form an n-fold ion in remote field-free space. The work-function-like quantity Φ^x is given by the sum of the zero-field local work-function ϕ^E and any small correction that may arise as the result of field penetration (Tsong and Müller 1969) and/or analogous effects; Φ^x is not well defined numerically.

Taking the sum of the above works, and remembering that w^R and η^R are zero by definition, we obtain:

$$U_n{}^s - U_0{}^R = H_n - n\Phi^x + n(u^s - u^x) + \eta_n{}^s. \tag{7}$$

Hence using equation (5) we obtain:

$$U_n{}^s - U_0{}^R = H_n - n\Phi^x + S_n{}^s. \tag{8}$$

The standard potential energy $U_n{}^s$ has thus been expressed in terms of the sum of a 'configurational component' and the quantity $S_n{}^s$.

2.2. Derivation of ionic bonding-state formulae

At the crossing-point between potential energy curves corresponding to m-fold and r-fold charge states of the desorbate the standard potential energies for these states are equal. So we may write:

$$0 = (H_r - H_m) - (r - m)\Phi^x + S_r{}^c - S_m{}^c \tag{9}$$

where the superscript c denotes values taken at the crossing point.

Equation (1) may be formally expanded as follows (remembering $U_0{}^R \equiv U_0{}^\infty$):

$$A_{mr}{}^0 = H_r + (U_0{}^R - U_m{}^c) + (U_m{}^c - U^{pa}). \tag{10}$$

Hence, applying equation (8) to an m-fold-charged ion, with $s = c$, substitution into equation (10) gives:

$$A_{mr}{}^0 = (H_r - H_m) + m\Phi^x - S_m{}^c + (U_m{}^c - U^{pa}). \tag{11}$$

Now, Φ^x is not a well-defined quantity; we can eliminate it from equation (11) by making use of equation (9). After some algebraic manipulation we obtain:

$$A_{mr}{}^0 = [r(H_r - H_m) + (mS_r{}^c - rS_m{}^c)]/(r-m) + (U_m{}^c - U^{pa}). \tag{12}$$

We now expand the bracket containing the S-potentials, using definition (5). This gives:

$$mS_r{}^c - rS_m{}^c = mr(u^s - u^x) + m\eta_r{}^c - mr(u^s - u^x) - r\eta_m{}^c. \tag{13}$$

It can be seen that the electrical terms in equation (13) cancel. Hence we obtain the result:

$$A_{mr}{}^0 = [r(H_r - H_m) + m\eta_r{}^c - r\eta_m{}^c]/(r-m) + (U_m{}^c - U^{pa}) \tag{14}$$

(This formula was first presented at the 24th International Field Emission Symposium, Oxford 1977.) It can be regarded as a greatly generalised form of a result given by Röllgen and Heinen (1975) some years ago.

The merit of equation (14) is, first, that in principle it can be applied to a number of different desorption models, and second, that terms involving the electric field and

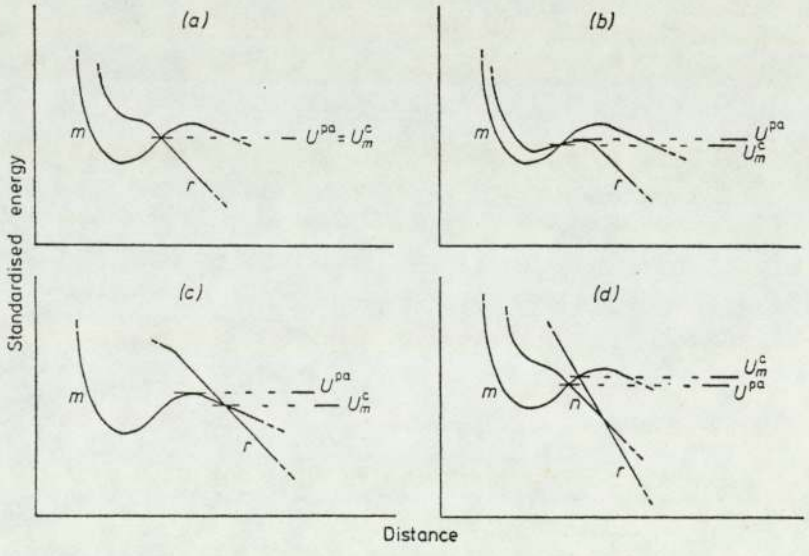


Figure 1. Some of the potential configurations to which equation (14) applies. Each plot shows distance along the horizontal axis, and standardised energy on the vertical axis. The Terms are labelled with the relevant charge numbers, as follows: m = initial (bonding) state; n = intermediate charge state (if any); r = final charge state. The standard potential energies $U_m{}^c$ and U^{pa} are defined in the text. These diagrams are schematic, being drawn to show possible relative configurations, rather than the exact shapes of potential curves.

The initial Term is drawn for the case where m is nonzero; if the initial bonding state were neutral then the potential curve would tend to flatten out as distance from the emitter increased.

work-function-like quantities have largely been eliminated. It is applicable both to neutral and to ionic bonding states, provided that $m \neq r$. Some of the various potential configurations to which the formula is applicable are shown in figure 1. The potential energy curves ('standard Terms') corresponding to different charge states are labelled with the corresponding charge numbers.

In the case of ionic bonding, a special approximation to equation (14) exists if it is legitimate to take the dominant contribution in the 'chemical' potential-energy component to be that resulting from correlation, and to use the image-potential approximation. In such a case we have:

$$m\eta_r{}^c - r\eta_m{}^c = -mr(r-m)e^2/16\pi\epsilon_0 d^c \tag{15}$$

where d^c is the distance of the crossing point from the effective imaging surface. Substitution into equation (14) gives:

$$A_{mr}{}^0 = r[(H_r - H_m)/(r-m) - me^2/16\pi\epsilon_0 d^c] + (U_m{}^c - U^{pa}). \tag{16}$$

This equation is valid for $m \neq 0$, $m \neq r$.

Equations (14) and (16) are particularly appropriate to the simple charge-hopping mechanism of field desorption (figure 1(*a*)), in which escape takes place at the crossing-point level and there is no post-ionisation process. In these circumstances $U_m{}^c$ is identically equal to U^{pa}, and the last term in equation (16) vanishes.

3. Application to tungsten

Table 1 shows how the calculated standard appearance energies for a tungsten ion would depend on the values of m and r if field evaporation were assumed to take place by means of a single-stage charge-hopping process. (This is an assumption conventionally made, although the present author has considerable reservations concerning the plausibility of multiple-electron hopping processes.) All estimates have been rounded to the nearest eV. For ionic bonding states the estimates are based on equation (16). The figures shown in brackets have been obtained by ignoring the image-potential term; the figures shown to the immediate left of each bracketed value include an image-potential contribution evaluated by arbitrarily setting d^c equal to 200 pm, which leads

Table 1. Standard field-ion appearance energies for tungsten as calculated from equations (16) and (18), shown as a function of the bonding-state charge number m and the charge number r on arrival at the analyser. (All appearance energies are rounded to the nearest eV.)

r	Standard appearance energies (eV)			
	$m=0$	$m=1$	$m=2$	$m=3$
1	17			
2	34	32 (35)		
3	58	57 (63)	61 (72)	
4	93	95 (102)	104 (118)	108 (140)

to the approximation:

$$me^2/16\pi\epsilon_0 d^c \approx m \times 1{\cdot}8 \text{ eV}. \tag{17}$$

There is considerable uncertainty both about the validity of the image-potential approximation, and about the proper value for d^c, and it is difficult to justify a more careful treatment within the framework of this approximation.

For appearance energies corresponding to neutral bonding states I have neglected polarisation effects and the activation-energy term (Forbes 1976a), and have used the formula:

$$A_{0r}{}^0 = H_r + \Lambda_0 \tag{18}$$

where Λ_0 is the tungsten sublimation energy, taken equal to 8·7 eV. The following values of ionisation energy have been used: $I_1 = 8{\cdot}0$ eV; $I_2 = 17{\cdot}7$ eV; $I_3 = 24$ eV; $I_4 = 35$ eV. These values are those used by Waugh (1975), and are attributed by him to Moore (1958); the values for I_3 and I_4 are subject to considerable uncertainty, and this has a corresponding effect on the estimates in Table 1.

Waugh and Southon (1976) measured an appearance energy of 51 ± 9 eV (see Forbes 1976b) for what they assumed to be W^{3+} ions. This result is compatible with estimates in Table 1. However, it is also clear from the table that nothing useful about the initial bonding state can be deduced: the appearance-energy estimates for the various initial charge states are so close that the uncertainties associated with the image-potential approximation render decisive discriminations impossible.

The tungsten estimates are fairly typical of those obtained for the refractory metals. It thus seems that for these metals a much more careful treatment of correlation and other 'chemical' bonding interactions at charged surfaces is a prerequisite, if any firm information about initial bonding state is to be drawn from measurements of appearance energies alone.

A more fruitful line of approach, at least in principle, might be to combine measurements of appearance energy and activation energy, as Ernst (1979) has done in the case of rhodium. This provides empirical information about the initial bonding energy and its dependence on the external field. The field-dependent part of the bonding energy consists in principle of two components: one associated with the polarisation of orbitals in the desorbate, the other with charge transfer from the desorbate to the emitter. If the first of these could be estimated independently, in a reliable manner, then the residual component could be identified as a charge-transfer induced effect and subjected to theoretical analysis.

4. Summary

The main theoretical result of the analysis here has been the derivation of an alternative general formula for the standard appearance energies of desorbates field-evaporated from ionic bonding states. An initial hope was that appearance-energy measurements could lead to hard information about the initial bonding state of the refractory metals prior to field evaporation. This hope cannot be realised at the present time, but it has been possible to establish that the measured appearance energy for tungsten is compatible with the assumption of a primarily ionic initial bonding state. This seems a necessary requirement for a self-consistent theory of tungsten field evaporation.

References

Cocke D L and Block J H 1978 *Surface Sci.* **70** 363–91
Ernst N 1979 *Surface Sci.* **87** 469–82
Forbes R G 1974 *Surface Sci.* **46** 577–601
—— 1976a *Surface Sci.* **61** 221–40
—— 1976b *J. Phys. D: Appl. Phys.* **9** L191–3
Moore C E 1958 *Atomic Energy Levels* Vol. III NBS Circular 467
Müller E W and Tsong T T 1969 *Field-Ion Microscopy: Principles and Applications* (Elsevier: Amsterdam)
—— 1974 *Prog. Surface Sci.* **4** 1–139
Röllgen F W and Heinen H J 1975 *Z. Naturf.* **30a** 918–20
Tsong T T and Müller E W 1969 *Phys. Rev.* **181** 530–4
Waugh A R 1975 *PhD Thesis* University of Cambridge
Waugh A R and Southon M J 1976 *J. Phys. D: Appl. Phys.* **9** 1017–23

Progress in Surface Science, Vol. 10, pp. 249-285, 1980
Printed in the USA.

0079-6816/80/020249-37$05.00/0

WAVE-MECHANICAL THEORY OF FIELD IONIZATION AND FIELD-ION ENERGY DISTRIBUTIONS

RICHARD G. FORBES

Department of Physics, University of Aston, Birmingham B4 7ET, England

Abstract

Although field-ion energy distributions have been measured for many years, no theoretical treatment has ever been published that is strictly compatible with the axioms of quantum mechanics. This article shows how a fully wave-mechanical theory of field ionization and field-ion energy distributions can be constructed, based on the idea that surface field ionization is the fermion analogy for Dirac's treatment of photon emission. The theory is formal and suggests, as expected, that existing quasiclassical treatments are essentially satisfactory in normal circumstances.

The discussion comprises of three main steps. The first is to clarify the various definitions of energy and potential energy used in field-ion theory, and to codify the corresponding energy diagrams. The second is to derive an expression for the rate-constant for transfer from one vibrational state to another. The third is to sum over initial states to derive a formal expression for the ion energy distribution. The article concludes by discussing the possible development of this theory.

Contents

1. Introduction

The process of surface field ionization[1] involves an atom or molecule (the "desorbate") at, or slightly, above a positively-charged emitter surface. An electron or electrons are transferred from the desorbate to the emitter, and the resulting ion is then emitted from the surface, in a process generally akin to pre-dissociation.

The theory of field-ionizing transitions can be formulated at two conceptually distinct levels. Most discussions of inert-gas field ionization and of field-ion image contrast use some form of quasi-classical formulation: in such treatments, no attempt is made to construct vibrational wave-functions, or to apply wave-mechanics to the motion of the desorbate as a whole. The alternative is a fully-wave-mechanical formulation. A particular approach of this sort, based on the Born-Oppenheimer and Zener-type approximations,[2] was developed by Gomer and Swanson[3] and applied to the calculation of field-desorption rate-constants. It is now customary to use this form of approach in the discussion of metal-atom field evaporation.

It has also become customary to employ very different types of diagrammatic representation in connection with quasi-classical and wave-mechanical descriptions of surface field ionization. In the former, emphasis is placed on the behaviour of the transferring desorbate electrons. An electron barrier diagram, such as Fig. 1, shows the barrier experienced by an electron tunnelling from an atomic desorbate to the emitter; an electron-level alignment diagram, such as Fig. 2,

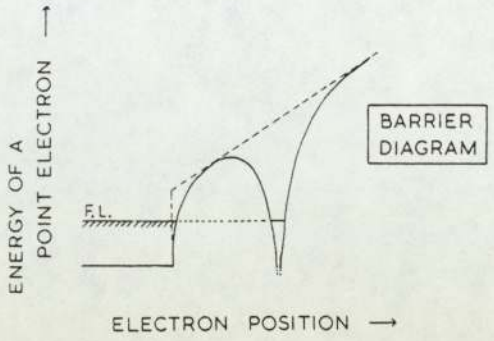


Fig. 1. Electron Barrier Diagram. Variation of the potential energy of a point electron with position as it tunnels through the barrier between an imaging-gas atom and a field-ion emitter.

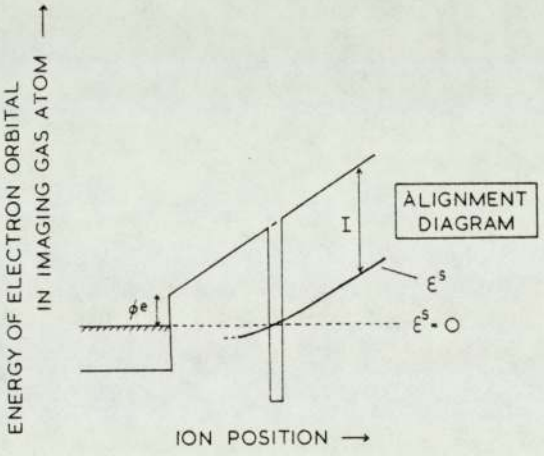


Fig. 2. Electron-level Alignment Diagram. Variation of the energy level ε^s of the topmost filled orbital in an imaging-gas atom as a function of the position s of the gas-atom nucleus, normal to the emitter surface. ε^s is defined by (15), and is zero for the nuclear position such that the electron can make a direct, energy conserving, transition to the emitter Fermi level. In the literature, confusion occurs between alignment and barrier diagrams.

shows how the energy-level of the electron orbital, from which transition occurs, varies as a function of the position of the atomic nucleus.

With a wave-mechanical description, emphasis is placed on the potential energies of the emitter-plus-desorbate system, and how these vary as a function of the emitter-to-desorbate separation, for electronic configurations corresponding to neutral and ionic states of the desorbate. A diagram of this general type is shown in Fig. 3. Such diagrams are similar, in many respects, to those used in

the theory of Molecular Spectroscopy,[4] although a different zero potential is used.

These various approaches are complementary, and are each appropriate for discussing some aspects of field-desorption theory. (In this article the term "field desorption" is used in a general sense, to cover field ionization, field evaporation and field desorption of adsorbates.)

In recent years, however, there has been an increasing interest in the measurement of field-ion appearance energies and field-ion energy distributions.[5] In these contexts, it is not helpful to make a sharp distinction between the field ionization of noble-gas atoms and the field desorption of metal substrate atoms or adsorbed species. Rather, the requirement is for a unified approach encompassing all three variants of field desorption. It has been shown elsewhere that this is possible in the theory of field-ion appearance energies.[6]

Many theoretical treatments of field-ion energy distributions have already appeared.[7-15] However, all these are quasiclassical, in the sense stated earlier. The primary objective of the present paper is to develop, formally, a wave-mechanical theory of field-ion energy distribution. This will involve the use of potential diagrams of a type formally different from those used by Gomer and Swanson. Thus, a secondary objective is to codify the various types of potential diagram used in field-ion theory, and indicate the relationships between them.

The structure of the paper is as follows. Section 2 deals with basic concepts - energies and energy diagrams. Section 3 then derives a fundamental expression for the current transferred from a single vibrational state of the desorbate when neutral to a single vibrational state when ionic. Section 4 shows formally how a summation can be carried out over initial vibrational levels, to give the total current per unit energy range entering states at a particular measured ion energy. From the resulting expression, the shape of the energy distribution is deduced, in §5, and the present treatment is contrasted with other treatments of field ionization that involve the use of vibrational wave-functions. Conclusions are briefly summarised in §6.

Two appendices deal with detailed points concerning electronic transition rate-constants, and the evaluation of the wave-function amplitudes in a "sloping one-dimensional box".

2. Basic Concepts

In a very general sense, the behaviour of the emitter and desorbate can be compared with the behaviour of the two components of a dissociating molecule. Since the theory of molecular dissociation is relatively well understood.[2,4] it can serve, in a general way, as a model for a wave-mechanical theory of the field-ionization transitions.

A. Standard desorbate potential energies ("Terms")

We begin by defining the potential energies associated with a desorbate particle. Thermodynamic arguments lead to the following mathematical definitions:[6]

$$U_0^R = U_0^\infty = 0 \tag{1}$$

$$U_0^p - U_0^R = W_0(R \to p) \tag{2}$$

$$U_n^p - U_0^p = W_{0n}(p) \tag{3}$$

$$U_n^s - U_n^p = W_n(p \to s) \tag{4}$$

In these equations, the subscripts refer to charge states of the desorbate, "0" denoting a neutral desorbate atom or molecule, and "n" denoting a desorbate ion or molecular ion carrying a charge of ne, where e is the elementary (proton) charge. The superscripts refer to the position in space of some reference point fixed in the desorbate particle. For atomic desorbates, it is convenient to take this reference point as coincident with the atomic nucleus. "s" and "p" are any arbitrary positions, and "R" labels a position in remote field-free space.

Equation (1) defines the zero of potential. The quantities "W" in the remaining equations denote the works done by a hypothetical external agent in the course of processes that are, in a certain sense, "slow" and thermodynamically reversible. $W_0(R \to p)$ is the work done in moving the desorbate, in a neutral state, from R to p; $W_{0n}(p)$ is the work done in creating an n-fold ion from the neutral desorbate, at position p, by removing n electrons one-by-one from the desorbate and placing them <u>at the emitter Fermi level</u>; and $W_n(p \to s)$ is the work done in moving the desorbate, in an n-fold-charged ionic state, from position p to position s.

The quantities U_0 and U_n, defined in this way, are sometimes called "atomic" and "ionic" potential energies. However, these names can also be applied to other, differently-defined, quantities. To single out the quantities defined by (1) to (4), we shall call them "standard desorbate potential energies", or <u>standard Terms</u>. (The name "Term" comes from the Landau and Lifshitz discussion[16] of diatomic molecules, and is discussed further in §2F.)

In the present paper, it is implicit in the definitions of U_0 and U_n that the desorbate be in the ground electronic state for the charge configuration in question.

B. Vibrational states and standardised energy diagram

Within the framework of a Born-Oppenheimer approach, as normally employed in molecular quantum mechanics,[4] vibrational states may be postulated to exist in the potential wells shown. These are vibrational states of the desorbate-as-a-whole, with respect to the emitter. Figure 3 shows one such state for the neutral Term U_0; this is marked with a blacked circle at the point where the energy level intersects the neutral Term (i.e. at the outer classical turning point). Several such states are shown for the ionic Term U_n; these are marked with open circles, where the energy levels intersect the ionic Term at the inner classical turning point. In fact, the ionic-Term states shown belong to a semi-infinite, continuous, range of vibrational states.

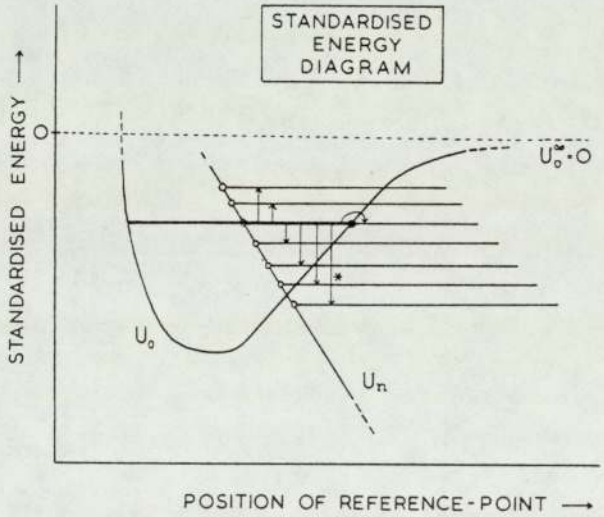


Fig. 3. Standardised Energy Diagram. Variation of the standard potential energies ("standard Terms") U_0, U_n, etc., defined by (1) to (4), with the position (normal to the emitter surface) of a reference point fixed in the desorbate. It also illustrates the changes in standardised energy E that are associated with transfers from a given initial vibrational state to various possible final vibrational states. This diagram, like the others in this paper, is schematic (see §2B) and is not drawn to scale.

If the desorbate is n-fold charged, and at position s has a total kinetic energy K^s, then we can associate with its motion a parameter E given by:

$$E = U_n^s + K^s. \tag{5}$$

For, so long as the desorbate remains in its n-fold-charged condition and suffers no collisions, the parameter E will be constant for its motion. We call E the standardised energy of the system. A definite value of standardised energy is associated with each vibrational state of the desorbate-as-a-whole.

As far as the desorbate particle is concerned, the standardised energy E plays the role of a total energy, provided that no change in charge state occurs. However, it is necessary to emphasise that E does not represent the total energy of the emitter-plus-desorbate system.

A diagram, such as Fig. 3, in which the standard Terms U_0 and/or U_n are shown as a function of the position of the reference point fixed in the desorbate particle, and in which the standardised energies of the various vibrational states are indicated by means of the vertical-axis coordinate, may be called a standardised energy diagram.

It should be realised that the diagrams in this article are schematic, being drawn to illustrate clearly the theoretical principles under discussion, rather than to depict accurately the shapes of the potential energy curves. The details of the shapes depend on the chemical natures of the particular emitter and desorbate under discussion, and on the strength of the external electric field and, in general, are not well known. We have deliberately omitted to show any detail in the ionic potential-energy curve close to the surface, as such details do not, in a first approximation, affect the theoretical discussion below, and have simply represented this potential energy by a straight line. In reality, there are short-range attractive and repulsive contributions to the ion potential energy, that cause increasing deviation from the straight-line potential as the emitter surface is approached: very close to the emitter, the short-range repulsion must dominate, but between the crossing-point and the emitter there might (depending on the external field strength) be a local potential-energy minimum. In the vicinity of the crossing-point, the ionic Term may in fact be curved, rather than straight, and this would have to be taken into account in treatments more advanced than that presented below.

It should also be noted that, in the context of noble-gas field ionization, the neutral Term is much shallower, and its gradient near the crossing-point much less, relative to the slope of the ionic Term, than is depicted in the diagrams drawn here.

C. Surface vibronic transitions

The neutral desorbate particle vibrating in its potential well can exchange energy with the phonon distribution in the emitter. Thus, the processes of ac-

commodation and activation can be represented as sequences of transitions from one vibrational state to another.

The act of field ionization can also be regarded as a surface vibronic transition, but from an initial vibrational state (of standardised energy E^i) in the neutral Term to a final vibrational state (of standardised energy E^f) in an ionic Term. Some of the infinitely-many transitions possible in principle are marked with arrows in Fig. 3. The convention has been used that a blacked circle marks an initial state, an open circle a final state.

Except in special cases, the standardised energies of the initial and final states will be different. For a particular transition, a quantity Δ^{if} can be defined by:

$$\Delta^{if} = E^i - E^f \tag{6}$$

Provided that no energy is transferred to the emitter phonon distribution during the transition, the quantity Δ^{if} has a straightforward physical interpretation. In the case of a one-electron transition, it represents the amount by which the energy level of the transferred electron exceeds the Fermi level "immediately after" this electron has entered the emitter. In a many-electron transition, Δ^{if} represents the sum of the excess energies of the two or more transferred electrons.

Since the relaxation time of the emitter electron distribution will be very short, the transferred electron(s) will very rapidly be de-excited to near the Fermi level, and the excess energy will be degraded into heat.

In a certain sense, surface field ionization can be regarded as a process whereby an electron is emitted from an atom into a "sea" of electrons; this process is thus a fermion equivalent to the emission of a photon. In the case of radiative emission, the difference between the initial-state and final-state energy levels gives the energy carried away by the photon; in the present case, the difference between the initial-state and final-state standardised energies gives the energy "carried away" by the electron(s) "emitted" from the desorbate. This analogy is quite useful, and is taken further below.

In principle, there also exists a more complicated type of field-ionizing transition, in which the transfer of the electron excites vibrations of the emitter nuclei. This type of transition can be represented on a standardised energy diagram in the same way as is the straightforward type of field-ionizing transition. However, Δ^{if} may now be split into two parts:

$$E^i - E^f = \Delta^{if} = \varepsilon_{ph} + \varepsilon \tag{7}$$

ε_{ph} represents the change in energy of the emitter photon distribution "immediately after" the transition, and ε the change in energy of the emitter electron distribution.

There also exists, at least in principle, a type of transition in which electron transfer and plasmon creation in the emitter take place simultaneously. This can only happen for final-state energies where Δ^{if} is greater than the plasmon energy. Lucas and co-workers have suggested that transitions of this type are responsible for the structure observed in field-ion energy distributions.[8-11] But this proposal has been criticised on both experimental and theoretical grounds.[17-19] Notwithstanding the reply[20] to one of these criticisms, the present author does not accept the surface-plasmon explanation of the observed energy spectra.

D. Gomer-Swanson (total-energy) diagram

When a field-ionizing transition has a non-zero Δ^{if} value, it is convenient, when discussing energy distributions, to depict the transition on a standardised energy diagram. However, for certain specialised purposes, it can be convenient to use an alternative representation.

For a particular final vibrational state in the ionic Term U_n, we may define a new desorbate potential energy U_{n*} by:

$$U_{n*} = U_n + \Delta^{if} \tag{8}$$

We may also define energy-like parameters T^i and T^f by:

$$T^i = E^i \tag{9a}$$

$$T^f = E^f + \Delta^{if} \tag{9b}$$

Equation (6) can then be rewritten in the form:

$$T^f = T^i \tag{10}$$

These quantities T represent total energy for the emitter-plus-desorbate system, referred to the zero of energy defined by (1), and (10) states the constancy of total system energy in a field-ionizing transition.

In the type of potential diagram used by Gomer and Swanson (for example, Fig. 10 in Ref.[3]), the desorbate-ion potential energy is set equal to U_{n*}, and the vertical-axis coordinate represents the total energy T. A field-ionizing transition can then be shown as a "horizontal" transition, conserving total energy. This form of diagram is illustrated in Fig. 4, where the transition marked with an asterisk in Fig. 3 is represented in this alternative manner.

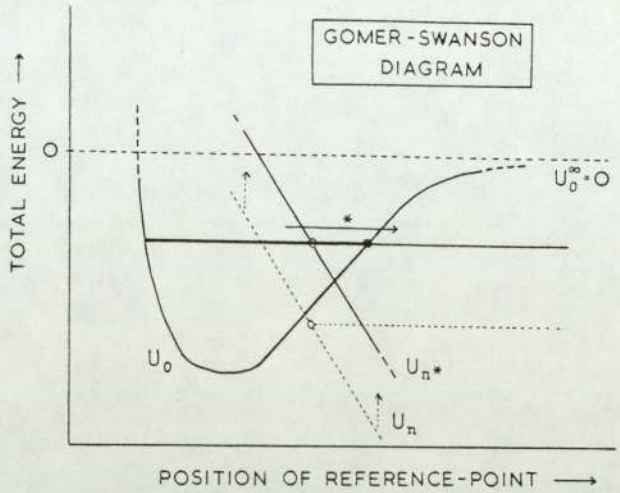


Fig. 4. Gomer-Swanson Diagram. Alternative method of depicting the vibronic transition marked with an asterisk in Fig. 3. The vertical axis here represents total energy, as defined by (9) and (10).

A total-energy diagram is useful when, as in Ref. [3], one wishes to evaluate a barrier penetration coefficient. Its disadvantage, in the context of ion-energy distributions, is that it obscures information about the relative kinetic energies of desorbate ions emitted from a single initial vibrational state into different final vibrational states. The quantity actually measured in an ion-energy analyser is the standardised energy E, not the total energy T.

Surprisingly, although the above distinctions are, of course, implicit in earlier work, there has been no previous attempt to define them explicity. In part, this may be because previous workers have often had in mind the case when the transferred electron(s) go direct to the emitter Fermi level: in this special case, the total-energy diagram coincides with the corresponding standardised-energy diagram. A simple argument showing that an energy analyser measures E rather than T is as follows.

Assume a quasiclassical viewpoint, and take a neutral desorbate particle to be initially stationary with total energy T^i. Let the particle now move, and suppose that it field ionizes, in such a fashion, that the electron(s) go to the emitter Fermi level. The resulting ion is brought to a halt in an energy analyser by applying some measured retarding voltage, V_{app}, say. This retarding voltage has a well-defined relationship with the initial energy T^i, because it is hypothesised that the electrons go directly to the Fermi level.

Now take the particle, with the same initial energy, but let it field ionize, in such a fashion, that the electron(s) go into emitter electron states above the Fermi level, the sum of their energies (relative to the Fermi level) being Δ^{if}.

This energy is rapidly converted into heat in the emitter. Owing to the constancy of the system total energy, this heat must have been extracted from the kinetic energy of the departing ion. (What happens is that the ionization event takes place further away from the emitter than in the previous case.) Consequently, a different retarding voltage is needed to bring the ion to a halt. The difference in voltage is equivalent to the energy Δ^{if}.

In the first case considered above, the measured ion energy must be equivalent to the initial system total energy T^i. In the second case, the measured ion energy is equivalent to an energy given by:

$$\text{"Energy"} = T^i - \Delta^{if} \tag{11}$$

Comparison of (11) with (9b) and (10) shows that this "energy" is, in fact, the quantity E^f defined earlier. Thus, an energy analyser measures the standardised energy of the final-ion vibrational state.

E. Standardised vibrational-level diagrams

One further type of diagram should be mentioned. If the simplifying assumption is made that the emitter is planar, then the kinetic energy K^s of the desorbate at position s may be split into components K_v^s and K_t associated with motion normal and parallel to the emitter surface, respectively. Thus,

$$K^s = K_v^s + K_t \tag{12}$$

With a planar emitter, there are no lateral forces acting on the desorbate, so the kinetic-energy component parallel to the surface is conserved, so the position suffix s may be omitted from the symbol K_t.

For a particular vibrational state, the parameter V defined by:

$$V = U_n^s + K_v^s \tag{13}$$

may be called the standardised vibrational level of the vibrational state in question. Clearly, from (5) and (12), the standardised energy E of the state is related to V by:

$$E = V + K_t \tag{14}$$

Since E and K_t are constants of the desorbate motion when in this state, it follows that so too is V. With a planar-emitter model, every vibrational state has associated with it definite values of E, V and K_t.

A diagram, such as Fig. 6, in which standard Terms are shown as a function of position, but in which the standardised vibrational levels are indicated by means

of the vertical-axis coordinate, is called a standardised vibrational-level diagram. This type of diagram is necessary in the development of detailed theories of field-ion energy distributions.

F. Quasiclassical use of standardised diagrams

The types of diagram just discussed arise naturally in the course of formulating a wave-mechanical treatment of field ionization or field desorption. However, it is also possible to use a standardised energy diagram in conjunction with a quasiclassical formulation of ionization theory; this use is illustrated in Fig. 5. The position and kinetic energy of the desorbate particle, as a neutral or as

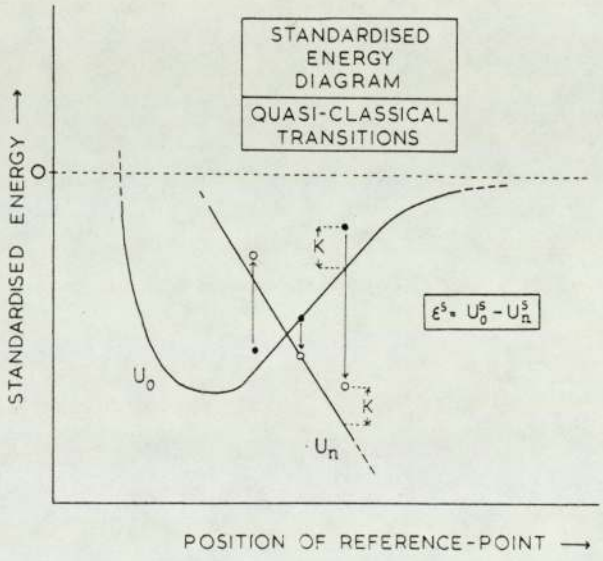


Fig. 5. Quasiclassical Transitions, shown on a standardised energy diagram. The filled and open circles denote both the position and standardised energy of the desorbate particle, before and after electron transfer. In a quasiclassical transition, there can be no change in the position nor in the kinetic energy K of the desorbate particle.

an ion, may be shown by representative points. In a quasiclassical formulation, the position and velocity of the desorbate are unchanged immediately after electron transfer: hence, the transition must be represented by a vertical arrow of length ε^s, where s is the position at which ionization occurs and

$$\varepsilon^s = U_0^s - U_n^s \tag{15}$$

ε^s being the amount of energy "carried away" by the electron(s) "emitted" into the metal during field ionization.

The existence of (15) is the basic reason why the name "Term" has been applied to the quantities U_0 and U_n. The behaviour of these quantities, as regards electron emission into the metal, is formally analogous to the behaviour of ordinary spectroscopic Terms as regards photon emission from atoms, except that we have applied the name to a quantity measured in energy units rather than in wave-numbers.

The quantity ε^s given by (15) is, by definition, the quantity plotted on the vertical axis of an electron-level alignment diagram, such as Fig. 2, so Figs. 2 and 5 are alternative methods of representing the energetics of electron transfer in field ionization. Equation (15) shows clearly, for example, that the condition $\varepsilon = 0$ and the condition $U_n = U_0$ are two equivalent methods of defining the "critical surface" of elementary field-ionization theory. In the present author's view, standardised energy diagrams are far more flexible in use than electron-level alignment diagrams. We also believe that the essential unity of field-evaporation theory and noble-gas field-ionization theory would be understood more widely, if it became normal practice to use standardised energy diagrams for both types of discussion, at least in research contexts.

3. State-to-State Transfer Process

A. Quantities and units for field-desorption theory

The next step in the formal treatment is to derive an expression for the transfer current between given initial and final vibrational states. However, before doing so, it is necessary to discuss some technicalities concerning quantities and units. These points are perhaps best understood in the context of overall emitter evaporation.

As has been pointed out elsewhere,[21,22] it is necessary to make a clear distinction between a field-evaporation rate-constant and the quantity (here, J) that represents the amount of field evaporation per unit time. The rate-constant is always expressed in the units s^{-1}; J is usually expressed in "layers/s" or "ions/s".

In 1971, changes were made to the international system of measurement. In particular, a seventh independent dimension "amount of substance" was introduced,[23,24] and there seems to be a continuing attempt to tighten up on the rules regarding the interpretation of symbols. In the spirit of these changes, it is expedient to ask: "What are the dimensions of J?"

The present author's conclusion is that normal practice and theoretical clarity both commend the convention that J have the dimensions

(amount-of-substance·time^{-1}). A quantity J with these dimensions we shall call an "amount-of-substance current", or (normally) just a "current". This convention obviously requires us to regard the "layer" as a (non-SI) unit of amount of substance.

The convention also requires us to regard the "atom", "molecule", or "ion", in J-values, such as "100 ion/s", as equivalent to units of amount of substance. Formally, it would probably be better to define an atomic-level unit of amount-of-substance, the "ordinary substance unit"* (symbol: osu), and to express J in the units "osu/s", and we shall sometimes do this. This proposal has been discussed more fully elsewhere.[22,25]

The relationship between a field-desorption current and the appropriate field-desorption rate-constant has the general form:

$$J = v \cdot (\text{rate-constant}) \tag{16}$$

In the convention used here, v represents the amount of substance at risk of ionization (or desorption), and is expressed in "layers", "atoms" (etc.), or "osu".

For general consistency, it seems best to carry this "seven-dimensional" approach over into the discussion of state-to-state transfer processes. Thus, instead of talking about the "occupation number (N^i) for state i", one talks in terms of the "population (σ^i) of state i", with $\sigma^i = N^i$ osu. And, instead of talking about the "probable number of transfers per unit time (C^{if}) from state i to state f", one talks in terms of the "transfer current (j^{if}) from state i to state f", with $j^{if} = c^{if}$ osu.

Although this new terminology will initially be unfamiliar, we feel strongly that the adoption of a seven-dimensional approach will eventually make for greater clarity in field-desorption theory.

B. State-to-state transfer current

There is no experimental evidence that excitation of emitter nuclei vibrations in the course of field ionization is a frequent event. Thus, all further discussion is restricted to the normal situation, where no energy is transferred to the emitter phonon distribution during the field-ionizing transition. In this case, we may write:

---

*The "ordinary substance unit" may be defined as "the amount of substance in a system containing one elementary entity". Just as the mole has replaced the collection of terms "gram-molecule", "gram-atom", "gram-ion", etc., the osu can replace the terms "molecule", "atom", "ion", etc., where these indicate a single unit of amount of substance, rather than the nature of the entity involved.

$$E^i - E^f = \varepsilon^{if} \tag{17}$$

where ε^{if} denotes energy transferred to the emitter electron distribution during the transfer from state i to state f, ε^{if} being measured relative to the Fermi level.

We shall also assume that the possibility of plasmon creation during field ionization may be disregarded. Plasmon-creation terms could, in principle, be included within the present treatment, but would add significantly to its complexity without producing results of any great value here.

The general expression for the transfer current j^{if} (i.e., the probable amount of substance transferred from state i to state f per unit time) is:

$$j^{if} = \sigma^i Q^{if} = \sigma^i t(\varepsilon^{if}) P^{if} \tag{18}$$

where σ^i is the population of state i (i.e., the probable amount of substance in state i), and Q^{if} is the transfer rate-constant (i.e., the effective rate-constant for transfers from state i to state f).

The rate-constant Q^{if} is given by the product of the state-to-state transition rate-constant P^{if} and the acceptance coefficient $t(\varepsilon^{if})$. The rate-constant P^{if} is the one that would be applicable if all the emitter electron states of energy appropriate to receive the transferred electron were empty. The acceptance coefficient t represents the effect of the Pauli exclusion principle on the electron transfer when these electron states are fully or partially occupied. In effect, t "inhibits" or "accepts" transitions, depending on whether transferred electrons would go into states below or above the Fermi level.

In the case of single electron transfer, if the emitter electron distribution is in thermodynamic equilibrium, $t(\varepsilon)$ is related to a centralised Fermi-Dirac distribution function $F(\varepsilon)$ by:

$$t(\varepsilon) = 1 - F(\varepsilon) \tag{19}$$

$t(\varepsilon)$ thus has the properties:

$$\begin{aligned} t(\varepsilon) &\sim 0 \quad \text{when } \varepsilon \ll 0 \\ &= \tfrac{1}{2} \quad \text{when } \varepsilon = 0 \\ &\sim 1 \quad \text{when } \varepsilon \gg 0 \end{aligned} \tag{20}$$

In the case of multiple-electron transfer, $t(\varepsilon)$ is a rather more complicated object, but fulfills the same general purpose.

The form of (18) brings out another analogy between the present theory and the theory of photon emission from atoms. In the theory of radiative emission, as developed by Einstein and Dirac,[26] the emission probability contains a factor $(1 + n')$, where n' is the number of photons in the state into which the photon is

to be emitted; the quantity n' is responsible for the phenomenon of stimulated photon emission. The corresponding factor for a fermion has the form (1 - n), where n is the number of fermions in the state into which a fermion is to be emitted;[26] the quantity n is responsible for the inhibition of fermion emission. In our case, the quantity n varies with the energy of the final electron state "immediately after transfer", and has the form $F(\varepsilon)$.

C. State-to-state transfer rate-constant

To derive a useful expression for the transition rate-constant P^{if}, it is necessary to make two assumptions. First, that it is legitimate to work in the spirit of the Born-Oppenheimer approximation and write the system wave-function Ψ in the form:

$$\Psi = \Phi_{ph} \cdot \Phi_{el} \cdot \Phi_{d} \tag{21}$$

where the wave-function components refer, respectively, to the vibrations of the emitter nuclei (Φ_{ph}), to the complete electron system (Φ_{el}), and to the motion of the desorbate relative to the emitter (Φ_{d}). Secondly, that it is permissible to use the golden rule of perturbation theory and write:

$$P^{if} = (2\pi/\hbar)\, \rho_{el}(\varepsilon^{if})\, |\langle \Psi^{f}|B|\Psi^{i}\rangle|^{2} \tag{22}$$

where Ψ^{i} and Ψ^{f} are the system wave-functions for the initial and final states in question, B is the perturbation energy, and ρ_{el} is a factor having the role of a density of final electronic states. Assumptions equivalent to these were made by Gomer and Swanson.[3]

Substitution of (21) into (22) leads to an expression of the form:

$$P^{if} = (2\pi/\hbar)\, \rho_{el}(\varepsilon^{if})\quad |\langle \Phi_{d}^{f}|\langle \Phi_{ph}^{f}|\beta(X, A)|\Phi_{ph}^{L}\rangle|\Phi_{d}^{L}\rangle|^{2} \tag{23}$$

where we have used Dirac's notation for overlap integrals, and $\beta(X, A)$ denotes the integral with respect to electron coordinates:

$$\beta(X, A) = \langle \Phi_{el}^{f}|B|\Phi_{el}^{i}\rangle \tag{24}$$

which is evaluated for a given position X of the reference point fixed in a desorbate, and given positions of the atomic nuclei in the emitter. The symbol A formally represents an assembly of these nuclear coordinates.

The actual derivation of detailed expressions for $\rho_{el}(\varepsilon^{if})$ and $\beta(X, A)$ would be far less straightforward than the above formula might suggest, even for a one-electron transition, because there are many possible final electron states

at the correct final electron energy, each with a wave-function that may extend differently into the vacuum, and these have to be summed over. However, the principles involved in such a summation are fairly well understood,[14,27,28] and it is known that the result can be written in the form used here, although there exists uncertainty over the correct form for the perturbation B.

In a standard approximation, as used in photon emission theory,[4] the overlap integral in (23) may be rewritten in the form:

$$\langle\phi_d^f|\langle\phi_{ph}^f|\beta(X, A)|\phi_{ph}^L\rangle|\phi_d^L\rangle = \beta(\bar{X}, \bar{A})\ \langle\phi_d^f|\langle\phi_{ph}^f|\phi_{ph}^L\rangle|\phi_d^L\rangle \tag{25}$$

where $\bar{X}$ and $\bar{A}$ denote appropriately chosen "average positions" for the desorbate and for the emitter nuclei. Further, since discussion here is restricted to the situation where there is no change in the emitter nuclei vibrations during field ionization, it follows that we must set:

$$\langle\phi_{ph}^f|\phi_{ph}^i\rangle = 1 \tag{26}$$

The integral on the right-hand side of (25) thus reduces to an overlap integral between the initial and final vibrational wave-functions for the desorbate particle. The position of $\bar{X}$ is in the region of space where the contribution to this integral is large.

In these circumstances, we may also assume that "on average" the electron wave-functions are probabilistically independent of the positions of the emitter-atom nuclei. Consequently, β does not "on average" depend on A. Thus, we may define a quantity $\kappa(\bar{X}, \varepsilon^{if})$, dependent on $\bar{X}$ but not on A, by:

$$\kappa(\bar{X}, \varepsilon^{if}) = (2\pi/\hbar)\ \rho_{el}(\varepsilon^{if})\ |\beta|^2. \tag{27}$$

$\kappa(\bar{X}, \varepsilon^{if})$ has the significance of an electronic transition rate-constant, although it cannot be identified exactly with the quasiclassical electron transition rate-constant P_e (see Appendix A). So, combining (23) to (27), we obtain:

$$P^{if} = \kappa(\bar{X}, \varepsilon^{if})\ |\langle\phi_d^f|\phi_d^i\rangle|^2 \tag{28}$$

If we now make the further assumption that the total vibrational wave-function for the desorbate can be separated into independent components relating to motion parallel (Ψ) and normal (χ) to the emitter surface, then the following formula is obtained for the transfer rate-constant Q^{if}:

$$Q^{if} = t(\varepsilon^{if}) \cdot \kappa(\bar{X}, \varepsilon^{if}) \cdot |\langle\Psi^f|\Psi^L\rangle|^2 \cdot |\langle\chi^f|\chi^L\rangle|^2 \tag{29}$$

D. Discussion

Formula (29) displays the four physical effects that influence the rate-constant for transfers from state i to state f. The first two terms, representing the Pauli exclusion principle and electronic factors, are relatively familiar, because these effects are also considered in quasi-classical treatments of field ionization. The remaining two terms appear only in a fully wave-mechanical treatment.

The third term has not explicitly appeared in any previous discussion. It represents a selection rule on changes in the lateral momentum of the desorbate particle during field ionization. In the context of a planar-emitter model, ψ^i and ψ^f have to be chosen from the same basic orthonormal set of functions, so $\langle\psi^f|\psi^i\rangle$ equals unity, if ψ^f and ψ^i are the same function, and zero otherwise. In such a model, only those vibronic transitions that conserve lateral momentum of the desorbate particle are allowed.

The fourth term involves an overlap integral between the normal components of the vibrational wave-functions for the initial and final states. Factors of this type have been discussed in the context of radiative transitions by Condon[29,30] and many others (for example, Herzberg[4]), and give rise to the Franck-Condon principle. The treatment here has shown formally, for the first time, that the Franck-Condon principle applies to field-ionizing transitions. As Inghram and co-workers have pointed out,[17] this result is expected.

The application of the principle, in the context of the "charge-hopping" model of field desorption, is illustrated in the standardised vibrational-level diagram (Fig. 6). As before, a black circle denotes the outer classical turning-point for

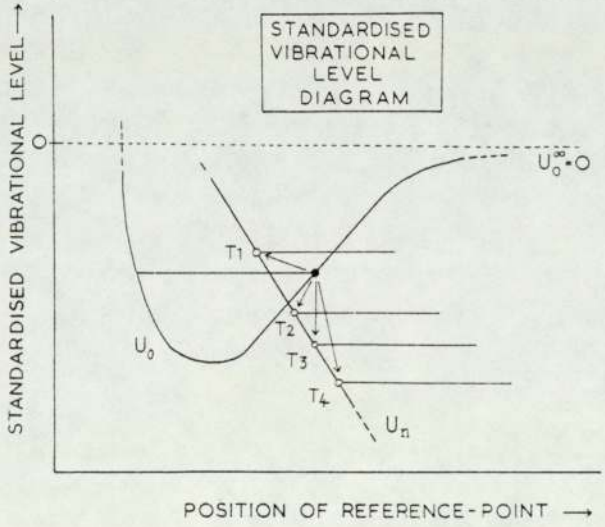


Fig. 6. Standardised Vibrational-level Diagram. This type of diagram displays on the vertical axis the quantity V defined by (13). Shown here are four possible transitions from a given initial level that have characteristically different transfer rate-constants.

an initial state, an open circle the inner classical turning-point for a possible final state. A transfer can be represented by drawing an arrow from the appropriate black circle to the appropriate open circle. The Franck-Condon factor tends to favour transfers that are "exactly vertical", such as T3 in Fig. 6. For transfers such as T4, that would be classically forbidden, the factor involves the overlap of decaying vibrational wave-function tails, and falls off rapidly as the separation of the classical turning-points increases. For transfers such as T2, that are classically allowed, the factor falls off more slowly with turning-point separation.

The Franck-Condon factor alone would not, of course, inhibit transfers that are "upwards", such as T1. Such transfers are inhibited, however, because the acceptance coefficient $t(\varepsilon^{if})$ has a low value for negative values of ε^{if}.

4. Derivation of Energy Distribution

A. Basic considerations

An ion emitted from a field-ion emitter initially travels, to a good approximation, in a central field. Consequently, a field-ion energy analyser measures standardised energy rather than the normal component of standardised energy. This principle is extremely well understood in the context of field-electron emission.[31] The objective, therefore, is to obtain a theoretical formula for dJ/dE^f, where J denotes the amount-of-substance ionized per unit time and E^f is the final-state standardised energy.

The details of the summation procedures needed to obtain a general result for dJ/dE^f are somewhat complex, so two simplifying assumptions will be made here. First, we shall consider a planar-surface emitter model: in this case, the standardised energies E^i and E^f associated with initial and final states may each be separated as in (14). Second, we shall restrict discussion to the one-dimensional population approximation. In this, it is assumed that the only initial states occupied are those for which the lateral kinetic energy K_t has some definite constant value, K_t^* say. In this case, the derivative dJ/dE^f is equal to the derivative dJ/dV^f, and the problem reduces to one of finding dJ/dV^f for the final vibrational states for which the lateral kinetic energy is K_t^*. Without significant loss of generality, we may also take K_t^* to be zero; the effect of assuming a non-zero value would merely be to produce a shift in the whole distribution by K_t^*.

Gomer and Swanson's treatment[3] also uses these approximations. They are the equivalent of the approximation in a quasiclassical treatment of "ignoring the kinetic energy of the desorbate at the instant of ionization". In the quasi-

classical treatment, it is clear, from the relative magnitudes of the likely lateral kinetic energy (a small fraction of an eV) and the scale of the structure to be discussed in field-ion energy distributions (several eV, or more), that neglecting lateral kinetic energy must physically be a reasonable approximation. This also appears to be true for a wave-mechanical treatment.

B. Total current entering final state

The total current j^f entering a particular final state f is obtained by summing (18) over all relevant initial states, thus:

$$j^f = \sum_i j^{if} = \sum_i \sigma^i Q^{if} \tag{30}$$

To proceed further, we shall assume that the initial vibrational levels are closely spaced, or can be treated as such. In the one-dimensional-population approximation, this summation may then be converted into a one-dimensional integration with respect to initial vibrational level V^i:

$$j^f = \int \frac{d\sigma}{dV^i} Q^{if}(V^i)\, dV^i \tag{31}$$

The amount of desorbate, $d\sigma$, with initial vibrational level between V^i and V^i+dV^i is given by:

$$d\sigma = \sigma^{tot}\, w(V^i)\, \rho_v^i(V^i)\, dV^i \tag{32}$$

where σ^{tot} is the total amount of desorbate in risk of ionization, in the domain over which the summation in (30) is taking place; ρ_v^i is the density of vibrational levels for the initial, neutral, Term (i.e. the number of levels per unit range of V^i); and $w(V^i)$ is the relative probability that a state of level V^i will be occupied, w being normalised such that:

$$\int w(V^i)\, \rho_v^i(V^i)\, dV^i = 1 \tag{33}$$

Thus, substituting from (29) and (32) into (31), bearing in mind that the "lateral" overlap factor is necessarily equal to unity, we obtain:

$$j^f = \sigma^{tot} \int w \cdot \rho_v^i \cdot t \cdot \kappa \cdot |\langle \chi^f | \chi^i \rangle|^2 \, dV^i \tag{34}$$

where all the terms inside the integral are functions of V^i.

We now assume that, over the range of final-state vibrational levels of interest, the ionic Term U_n can be treated as a linear function of distance. Thus, initially, the desorbate ion moves in a uniform electric field and, in the region

of the classical turning point for state f, the vibrational wave-function χ^f may be written in the form:

$$\chi^f = a(V^f) \quad \zeta(\eta^f) \tag{35}$$

where η^f is the distance measured from the classical turning point for final vibrational level V^f, $\zeta(\eta^f)$ is the solution of the Schrödinger equation in a uniform field (Appendix B) and $a(V^f)$ is a normalisation coefficient whose value depends on V^f.

Equation (34) may thus be written in the form:

$$j^f = \sigma^{tot} \, |a(V^f)|^2 \quad I(V^f) \tag{36}$$

where $I(V^f)$ denotes the following integral over initial vibrational levels:

$$I(V^f) = \int w \cdot \rho_v^i \cdot t \cdot \kappa \cdot \kappa\zeta(\eta^f)|\chi^i\rangle|^2 \quad dV^i \tag{37}$$

If the initial vibrational levels are not closely spaced, which may be the case in some circumstances,[32] then (36) still holds, but $I(V^f)$ is replaced by a summation over initial vibrational levels.

C. Conversion to differential form

The next step is to convert (36) into a differential form, i.e., to obtain an expression for the total current dJ entering final states, whose vibrational levels lie between V^f and V^f+dV^f. However, if the final states are assumed to be unbounded, then the coefficients $a(V^f)$ are infinitesimal, and the density of final vibrational levels is infinite, so the normal procedure for converting a summation does not work.

To overcome this, we introduce a procedure that has analogies in other areas of physics, but seems to be a novelty in connection with field ionization. The desorbate ion will be taken as confined within a one-dimensional box with a sloping side, as shown in Fig. 7. The vibrational levels then become discrete and ennumerable, and dJ can be written as:

$$dJ = \rho_v^f(V^f) \, dV^f \times \sigma^{tot} \quad |a(V^f)|^2 \quad I(V^f) \tag{38}$$

where ρ_v^f is the density of final vibrational levels (i.e. the number of levels per unit range of V^f).

It is shown, in Appendix B, that for a box of length L, whose base is at the standardised level V^g, the coefficient $a(V^f)$ is given by:

$$a(V^f) = 2^{\frac{1}{2}} \, L^{-\frac{1}{2}} \, (2m/\hbar^2 eF)^{\frac{1}{8}} \, n^{-\frac{1}{4}} \, (V^f - V^g)^{\frac{1}{4}} \tag{39}$$

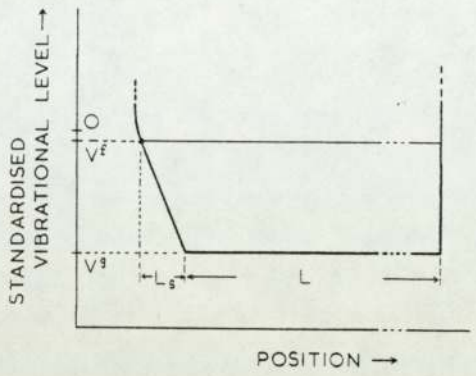


Fig. 7. The one-dimensional potential box used in calculating final-state vibrational wave-functions.

where F is the field in which the desorbate ion of mass m moves. The density of vibrational levels for the box is:

$$\rho_v^f = (L/\pi^2\hbar^2)\ (m/2)^{\frac{1}{2}}\ (V^f - V^g)^{-\frac{1}{2}} \tag{40}$$

Substitution into (38) thus gives the result:

$$dJ/dV^f = \sigma^{tot}\ D(F)\ I(V^f) \tag{41}$$

where D(F) is a field-dependent coefficient given by:

$$D(F) = \pi^{-2}\ \hbar^{-1}\ (2m/\hbar^2)^{\frac{3}{4}}\ n^{-\frac{1}{2}}\ (eF)^{-\frac{1}{4}} \tag{42}$$

The coefficient D(F) is not a function of V^f, so the structure in the ion energy distribution is directly related to the V^f-dependence of the integral $I(V^f)$, as given by (37).

D. Evaluation of $I(V^f)$

It is not possible, at present, to carry out a mathematical evaluation of the integral $I(V^f)$, because the necessary information is not available. In particular, details of the potential U_0 are not known, so the construction of wave-functions for the initial vibrational states is impeded. Nevertheless, the general form of the result can be obtained from qualitative arguments.

For simplicity, let I' denote the integrand in (37) and S the overlap factor in that equation, so we have:

$$I(V^f) = \int I'(V^i,V^f)\ dV^i \tag{43}$$

$$I'(V^i,V^f) = w \cdot \rho_v^i \cdot t \cdot \kappa \cdot S \tag{44}$$

$$S = |\langle \zeta(n^f) | \chi^i \rangle|^2 \qquad (45)$$

Consider a particular final vibrational level, and let X' be the position of the corresponding inner classical turning point, as illustrated in Fig. 8. Then

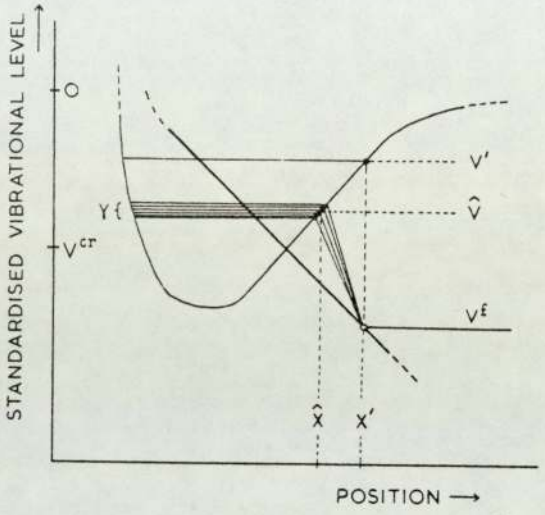


Fig. 8. Illustration of how, for a given final level V^f, the "more frequent" transfers to it may come from a band of initial levels (near $\hat{V}$) that are displaced in energy from the initial level V' corresponding to an "exactly vertical" transition.

let V' be the vibrational level of the initial state whose outer classical turning point is at the position X'. From our knowledge of the general behaviour of Franck-Condon-type overlap factors, we know that S must be an asymmetric bell-shaped function of V^i, with its maximum at, or close to, the value $V^i = V'$, as shown in Fig. 9.

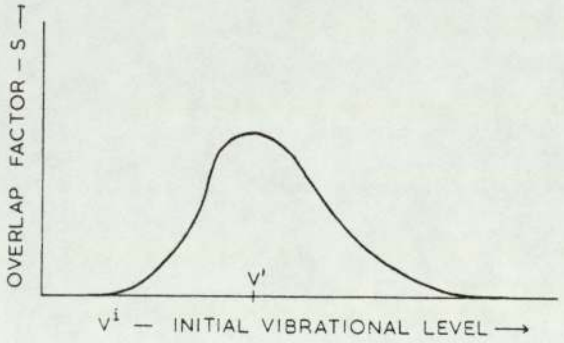


Fig. 9. Schematic illustration of how, for a given final vibrational level, the overlap factor S varies with initial vibrational level.

We now consider how the other factors in (44) affect the integrand. Under normal circumstances, we can treat the density of initial vibrational levels as a slowly varying function of V^i, so multiplication by ρ_v^i would have little effect on the function shown in Fig. 9.

As regards the occupation probability w, two situations can be envisaged. The first is when thermodynamic equilibrium (or approximate thermodynamic equilibrium) exists. In this case, w will be an exponentially-decaying function of V^i. The effect of multiplying S by w will be to skew the function shown in Fig. 9 to the left, shifting the peak to a lower value of initial vibrational level. The lower the temperature, the greater the shift of this type that will occur.

In the case of field ionization, however, where gas is supplied to the emitter from space, and ionization occurs during the process of gas accommodation to the emitter temperature, depletion of the population in the lower-lying levels could lead to a situation where, over some range of V^i, the occupation probability increases as V^i increases. The effect of multiplying S by w, in this case, would be to skew the function in Fig. 9 to the right, shifting the peak to a higher value of initial vibrational energy.

The electronic factor κ can also produce "skew" and "shift" effects of either sense, depending on the value of V'. This is because the behaviour of κ as a function of the energy ε transferred to the emitter electron distribution is closely similar to that of the quasiclassical electron transition rate-constant P_e (see Appendix A). P_e is known to go through a series of peaks as this energy increases,[9,13,15] the peaks resulting from resonance effects in the electron transfer to the emitter. Thus, there are some ranges of energy where κ is a decreasing function, others where it increases.

The acceptance coefficient t has little effect on the shape of the integrand, except when the final level V^f is close to the level of the crossing point of the potential curves U_0 and U_n. For discussion, let us denote this crossing-point level by V^{cr}. When V^f is close to V^{cr}, the coefficient t will tend to "cut off" the left-hand side of the integrand, for values of V^i lower than the value of V^f under consideration. For V^f less than V^{cr}, the effect will be to reduce the width of the integrand; for V^f greater than V^{cr} the effect will be to skew and shift the shape of the integrand to the right.

Taking all these possibilities into consideration, we conclude that the integrand I' will have an asymmetric-bell shape. The shape may vary with V^f, particularly for values of V^f close to V^{cr}. However, the result of integration can be written formally in the form:

$$I = \gamma \hat{I}' = \gamma \cdot \hat{w} \cdot \hat{\rho}_v^i \cdot \hat{t} \cdot \hat{\kappa} \cdot \hat{S} \tag{46}$$

where $\hat{I}'$ denotes the peak value of the integrand, and γ is a quantity that relates to the width of the peak. The initial vibrational level at which this peak value occurs is denoted by $\hat{V}$, and values of the factors in the integrand, taken at this level, are indicated by the circumflex sign.

The physical interpretation of this is as follows. For final states at a given vibrational level, most of the current into these states comes from a narrow band of initial levels, of width γ, centred at initial level $\hat{V}$, which is illustrated schematically in Fig. 8. Because of the other factors in the integrand (44), the "most frequent" transfer does not necessarily coincide with transfer from the level for which the Franck-Condon-type factor S is a maximum. The product $\gamma \cdot \hat{\rho}_v^i$ can be interpreted as the number of initial levels contributing significantly to the current entering final states at level V^f.

E. Special circumstances

There are some special circumstances in which the formal arguments outlined in the preceding section would not be straightforward to apply.

The above arguments assume that the width in energy-V^i of the overlap factor S is less than the width in energy-V^i of peaks in the electronic factor κ. The present author believes this to be the case in normal circumstances. However, even if the reverse were true, a result of form (46) would be obtained, but the parameter γ might basically be determined by the width of the factor κ rather than the width of S.

It was required earlier that the initial vibrational levels should be closely spaced in energy-V^i. More strictly, the requirement is that the level spacing should be comparable with or less than the quantity kT. In this case, it is legitimate to think of the levels as "smoothed out in energy" to form a continuum, and to think of S and κ as defined for all values of V^i by a process of interpolation between their values for the original discrete levels. Discussion in terms of a mathematically-continuous integrand, as above, can then take place.

This procedure, however, will tend to become unsatisfactory when the occupation probabilities of adjacent levels differ from each other by large factors.

In the limiting case of thermodynamic equilibrium at extremely low temperature, when the probability of finding the desorbate particle in any state other than its vibrational ground state is very small, most of the current into any final state must come from the lowest initial vibrational level. In this case, the product $\gamma \cdot \hat{\rho}_v^i$ in (46) is replaced by unity, and the other factors take the values appropriate to the lowest initial vibrational level.

In a less extreme case, where significant currents into a particular final state come from several vibrational levels, the quantity $I(V^f)$ would have to be determined, in practice, by a summation over these levels. However, by setting γ equal to the level spacing (or in some other appropriate, but arbitrary manner), one could formally write the result of the summation in a form similar to (46). Consequently, the discussion in the following section has a qualitative validity, even in this case.

5. General Discussion

A. Shape of energy distribution

The results achieved so far may be summarised as follows. In the present approximation, the shape of the measured energy distribution is determined by the quantity $I(V^f)$. For each value of final vibrational level V^f, there is a corresponding initial level V', from which a transition to final level V^f would be "exactly vertical", and also an initial level $\hat{V}$ from which transfers to final level V^f are "most frequent". For a given final level V^f, the factors in (46) are to be evaluated at the corresponding initial level $\hat{V}$.

However, provided that γ is approximately constant, and that the difference between $\hat{V}$ and V' is either small or approximately constant (both of which conditions should hold, except for values of V^f close to V^{cr}), the general shape of $I(V^f)$ can be deduced from a knowledge of how the relevant factors in (46) vary as a function of V'.

For values of V^f less than V^{cr}, the following features are thus predicted:

(1) The ion-energy distribution, plotted as function of decreasing V^f (i.e. as a function of increasing "relative energy deficit"), has the same general form as has the quasiclassical electron transition rate-constant P_e, plotted as a function of position X'. This is because $\hat{\kappa}$ has a behaviour similar to P_e (see Appendix A). P_e goes through a series of resonance peaks as the distance from the emitter increases.[9,13,15]

(2) If the most frequent transition to a given final level is exactly vertical, then $\hat{\kappa}$ may be identified with $P_e(X')$ (see Appendix A). In other cases, the peaks in $\hat{\kappa}$ are displaced slightly in energy from where they would be if the most frequent transition were exactly vertical. The extent of this displacement is not known, but is expected to be small.

(3) The exact positions and relative magnitudes of the peaks will also be affected by the dependence of occupation probability on initial vibrational level. Under conditions of thermodynamic equilibrium, the "high-deficit" peaks will be

diminished relative to peaks at lower deficits. Under conditions that produce an "inverted population", the high-deficit peaks will be enhanced relative to peaks at lower deficits. If $\hat{w}$ varies rapidly with level $\hat{V}$, then "skewing" and "shifting" of the peaks predicted from $\hat{\kappa}$ will also occur.

(4) The product $\hat{\rho}_v^i \cdot \hat{S}$ is expected to vary little with $\hat{V}$; it will have an effect on the total rate-constant for ionization, but will have little effect on the shape of the energy distribution. Similarly, for high values of the deficit $(V^{cr} - V^f)$, the acceptance coefficient $\hat{t}$ is approximately equal to unity, and thus has no effect on the shape of the distribution.

For values of V^f greater than V^{cr}, the acceptance coefficient $\hat{t}$ tends to approach very small values, and the effect on the ion-energy distribution is to produce a "main peak" with a relatively sharp edge on the low-energy-deficit side. Details of the shape of the peak cannot be satisfactorily derived from qualitative arguments, but should be derivable from numerical analyses within the next few years. It should then become possible to predict values for the onset Appearance Energy,[6] as measured in a retarding-potential analyser. Interesting experimental results concerning this quantity, and its variation with temperature, have recently become available,[33,34] and detailed comparisons between experiment and theory should prove instructive.

B. "Exactly vertical" form for $I(V^f)$

Although most of the current entering a final state at level V^f comes from a relatively narrow band of initial levels, the "most frequent" transfer is not necessarily exactly vertical. Nevertheless, it may sometimes be convenient to write $I(V^f)$ in a mathematical form that involves parameters appropriate to an exactly-vertical transition. Using a prime to denote factors appropriate to an exactly-vertical transition to a state of final level V^f, we may set:

$$I(V^f) = \theta \cdot \gamma \cdot w' \cdot (\rho_v^i)' \cdot t' \cdot P_e' \cdot S' \qquad (47)$$

where θ is a correction factor defined by equating expressions (46) and (47). In writing (47), we have also used the fact that for an exactly vertical transition the electronic factor κ' can be identified with the quasiclassical quantity $P_e(X')$ discussed earlier (here written P_e').

Combining (41) with (47), and rearranging slightly, leads to an alternative formal expression for dJ/dE:

$$dJ/dE = \theta \cdot \{\sigma^{tot}\ \gamma\ (\rho_v^i)'\ w'\} \cdot \{D(F)\ S'\} \cdot P_e' \qquad (48)$$

This form displays clearly four main influences that determine the shape of field-ion energy distributions. From left to right the terms represent: deviation of the "most frequent" transfer from exact verticality; the desorbate distribution prior to ionization; the wave-like nature of the motion of the desorbate as a whole; and electronic effects.

Treatments of energy-distribution theory that deal only with the last term in (48) are certainly concentrating on the factor that is most influential in normal circumstances, but formally they are incomplete.

C. Comparison with other treatments

The present approach can usefully be compared with other treatments of field ionization that, although not dealing with energy distributions, are wave-mechanical in the sense that vibrational wave-functions are taken into consideration.

By comparison with the Gomer and Swanson treatment, the present approach has three distinctive features:

(a) The definition and use of standardised energies and the corresponding diagrams. This seems an important step forward, because it facilitates the pictorial representation of the energy transfers involved in field ionization.

(b) It conducts the summation over initial states first. Formally, the argument here could be continued by summing over final states (i.e. integrating dJ/dE with respect to E) to obtain an expression for a desorption rate-constant. Gomer and Swanson's evaluation of this rate-constant carried out the summation over final states first, subsequently over initial states.

(c) It avoids the use of Zener's approach and the WKB approximation. At present, this means giving up the chance of a numerical result; but in the longer term, when numerical evaluations of the overlap factor S have been carried out, the present approach should prove superior, because it circumvents problems associated with the validity of the WKB approximation in the circumstances of field ionization. In the present author's view, these problems have not been analysed with sufficient care.

Discussions of field ionization that involve the use of vibrational wave-functions have been formulated by Schrenk and coworkers,[28,35] but these papers fail to come to grips with the problems associated with the existence of initial and final vibrational states and the need to sum over them.

In the case of helium, the effect of the gas distribution on the half-width of the main peak of the energy distribution has recently been discussed by Rendulic.[32] He has calculated energy levels and wave-functions for the vibra-

tional levels in the initial potential well U_0, and has then obtained a "dwelling probability" by taking an appropriately weighted summation of $|\chi^i|^2$ over the lowest three levels. The variation of dwelling probability with position is shown for temperatures of 20 K and 80 K, and it is noted that there is a temperature-dependent variation in the half-width of the dwelling probability. This, it is argued, is responsible for an observed temperature dependence[36] in the half-width of the measured distribution main peak.

The present author has considerable reservations about the self-consistency of Rendulic's work, since it implies an approach to the calculation of transfer rate-constants that does not involve the vibrational wave-functions for the final (ionic) states. We conclude that it does not constitute a fully wave-mechanical treatment.

D. Development

The method presented here should be seen as just the first step in the development of wave-mechanical theories of field ionization and field-ion energy distributions. Essentially, it deals with the physical situation where the act of field ionization is a one-stage process involving the "hop" of one or more electrons from the desorbate to the emitter, with no complications. It thus covers noble-gas field ionization and the "charge-hopping" model of field desorption.

There are various conceivable mechanisms of field ionization that are not dealt with within the present formulation. The most relevant of these are as follows.

(1) "Hump" type models of field desorption. With these mechanisms there is no electronic transition involved in the escape step, and the theory assumes a much simpler form.[1] However, in the present author's view there is considerable doubt as to the relevance of this mechanism to the real situation of field desorption or field evaporation, certainly in most situations.

(2) Simultaneous field ionization from a single initial charge state into two or more final charge states. In this case, the resultant distribution will be the sum of two distributions of the type discussed earlier.

(3) Two-stage mechanisms involving the post-ionization of an already desorbed ion. In principle, the second stage can be treated by methods analogous to those developed earlier, using the quantity I(V) to determine occupation probabilities for the initial levels prior to the second stage. The resultant distribution will be the sum of the "post-ionized" distribution and a modified original ion distribution.

(4) Simultaneous electron transfer and phonon excitation. As stated earlier, I know of no evidence that this occurs in practice.

(5) Simultaneous electron transfer and plasmon creation.

(6) Coupled processes of plasmon creation in the emitter and vibrational de-excitation of a departing ion.

As stated earlier, we find the theoretical and empirical arguments[17] against the likelihood of surface-plasmon creation convincing. However, it would still be of interest to formulate wave-mechanical treatments of these and other subsidiary mechanisms of field ionization, so that their a priori theoretical likelihood could be meaningfully assessed.

The most immediate task, however, is the numerical evaluation of the overlap integral S, and hence the determination of γ, for various model situations.

6. Conclusions

This article has been concerned with definitions and concepts, and to a lesser extent with formalism. By working from analogy with the long-established theory of radiative transitions in molecular spectroscopy, we have shown, in principle, how a fully wave-mechanical theory of field-ion energy distributions may be constructed. We believe this to be a significant step in the assimilation of the theory of field ionization into the main stream of physical and chemical thinking. The way now appears open for discussion of more detailed problems, taking as a guide the large body of existing theory concerning radiationless transitions,[37,38] and for the performance of well-grounded numerical calculations.

The main result of the present work has been to show that, in normal circumstances, quasiclassical treatments should give predictions that are essentially correct. In view of the good agreement between experiment and quasiclassical theories, this result was only to be expected. However, from a general point of view, it seems useful to have demonstrated that the electronic effects treated in quasiclassical formulations can be imbedded in a more complete formulation. In a minor kind of way, there is perhaps an analogy here with the work of Condon, in as much as a classical explanation of what is now known as the Franck-Condon principle existed before a justification was given for it in terms of quantum mechanics.

Appendices

A. Rate-constants for electron transition

The objective of this appendix is to clarify the distinction between the quasi-classical electron transition rate-constant P_e and the quantity κ, defined by (27), that appears in a wave-mechanical treatment of the field-ionizing transition.

In a quasiclassical formulation, the ion current (dJ) generated in a small volume of space dV well outside the critical surface is given by:

$$dJ = C\ P_e\ dV \tag{49}$$

where C is the gas concentration in the volume dV. This formula is a restatement, using the terminology of the post-1971 SI system,[22] of a formula discussed earlier,[39] and constitutes the definition of the rate-constant P_e.

The motion of the ionizing particle being ignored, application of quantum mechanics to the motion of the electron alone results in the use of one or other of the available formalisms: the "wave-packet" approach; the "barrier penetration" approach; or the "overlap integral" approach. In the last of these, the rate-constant $P_e(X)$ for ionization when the desorbate particle is at position X is given by:

$$P_e(X) = (2\pi/\hbar)\ \ \rho^*_{el}(\varepsilon^X)\ \ |\langle\phi^f_{el}|W|\phi^i_{el}\rangle|^2 \tag{50}$$

where ρ^*_{el} is a factor having the role of a density of final electronic states, and ε^X is defined by:

$$\varepsilon^X = U_0(X) - U_n(X) \tag{51}$$

W being the appropriate perturbation-energy term.

This formula, of course, ignores emitter nuclear vibrations. The definition of the quantity $\kappa(X,\varepsilon^{if})$ under equivalent conditions is obtained from (24), (25) and (27), ignoring quantities relating to emitter nuclear vibrations:

$$\kappa(X,\varepsilon^{if}) = (2\pi/\hbar)\ \ \rho_{el}(\varepsilon^{if})\ \ |\langle\phi^f_{el}|B|\phi^i_{el}\rangle|^2 \tag{52}$$

where the quantities in (52) have the meanings defined in §3C.

The possible differences between (51) and (52) lie in the perturbations, the functional dependences of the two densities-of-states on energy, the values of the energies ε^x and ε^{if} and the initial electron-state wave-functions.

In principle, the perturbation terms W and B may be different functions of position coordinates. No satisfactory consensus has ever been reached on the form of either perturbation. In practice, the same approximation eFx (where x denotes electron position) might be used for both and, in this case, the functions ρ^*_{el} and ρ_{el} would be identical (i.e. both symbols would represent the same function of energy). If different approximations are used for W and B, then ρ^*_{el} and ρ_{el} might represent slightly different functions of energy.

The complications in the preceding paragraph arise because (51) and (52), in fact, represent the results of averaging over many different final electron state wave-functions (albeit the same set of wave-functions in both cases).[14,27,28] If the perturbations are different, then the averaging process will throw up slightly different functions ρ^*_{el} and ρ_{el} in the two cases.

Apart from these differences, which are probably small in practice, there may exist a difference in the quantities ε^X and ε^{if}. This is illustrated in Fig. 10,

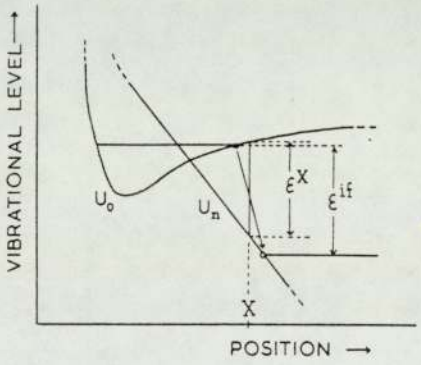


Fig. 10. Illustration of the difference between ε^X and ε^{if}, for a given position X at which the initial-state electron-orbital is centred. ε^X is the energy difference involved in the quasiclassical formulation; ε^{if} the energy difference involved in the wave-mechanical formulation.

which is drawn so as to emphasise (and probably exaggerate) the distinction. At position X, the quantity ε^X represents the vertical energy-difference between the standard potential-energy curves. However, ε^{if} represents the difference between the two vibrational levels with classical turning points on opposite sides of the position X. For a given final level, the initial level has to be so chosen that the position X is the "appropriately chosen average point" when the approximation (25) is made. (Strictly, when the equivalent approximation is made in a treatment that ignores emitter nuclei vibrations). It is clear from Fig. 10 that, if the appropriate initial and final levels were related as shown there, then ε^{if} would not be equal to ε^X. Consequently, even if the densities-of-states in (51) and (52) were the same function, they would be evaluated at different energies.

This has two consequences. The first is that the densities-of-states terms will have different numerical values in the two expressions (51) and (52); this will not be of major significance if ρ^*_{el} and ρ_{el} are slowly-varying functions of energy, which we expect to be the case, since both are "averaged" quantities. The second effect is more subtle. The final-state wave-function to be substituted into the overlap integral in (51) or (52) has to be the wave-function appropriate to an electron of energy ε^X or ε^{if} above the Fermi level. However, there may be marked differences in the extent to which final-electron-state wave-functions extend into the vacuum at different energies, and/or marked energy-dependent differences in the amplitude of extended wave-functions distant from the emitter, and these differences may produce appreciable energy-dependent differences in the values of the overlap integrals in (51) and (52), notwithstanding that the initial-electron-state wave-function is centred at the same position X in both cases.

As a result of all the above effects, it follows that $\kappa(X,\varepsilon^{if})$ and $P_e(X)$ are not identical, and will normally have different values. It should also be clear, in principle, that κ is a function of the initial and final vibrational levels involved, not simply a function of position.

For an "exactly vertical" transition, however, $\hat{\varepsilon}^{if}$ is equal to ε^X. Thus, leaving aside any differences between the assumed perturbations, and any resulting (probably small) differences between the functions ρ^*_{el} and ρ_{el}, we can identify $\kappa(X,\varepsilon^{if})$ with $P_e(X)$ in this case. This assumption is used in writing down (48).

It remains to justify the statement, made in §5A, that $\hat{\kappa}$ has behaviour similar to P_e. Clearly, this is true if the most frequent transfer to a state of given final level is exactly vertical. In other cases, it is convenient to think of $\hat{\kappa}$ and P_e as primarily functions of energy. For any given final vibrational level, there is a corresponding energy difference $\hat{\varepsilon}^{if}$ associated with the most frequent transfer to level V^f. The corresponding rate-constant $\hat{\kappa}$ is approximately equal to the rate-constant $P_e(\hat{\varepsilon}^{if})$ taken at position X^{eq}, such that:

$$\hat{\varepsilon}^{if} = U_0(X^{eq}) - U_n(X^{eq}) \qquad (53)$$

As we move to vibrational levels of higher energy deficit (i.e. V^f decreases), the corresponding energy difference $\hat{\varepsilon}^{if}$ increases, and consequently X^{eq} increases. It follows that $\hat{\kappa}$, plotted as a function of final vibrational level V^f, tends to replicate the behaviour of P_e plotted as a function of distance.

Obviously, it would be of interest to have a more detailed understanding of the above relationships, but a necessary preliminary is a better understanding of how the overlap factor S behaves numerically for transitions that are "nearly

vertical": it is precisely in this range that the standard approximations introduced by Landau and Zener[2,16] breakdown.

One further basic point deserves discussion. Implicit in this treatment is the assumption that the different formalisms for the calculation of P_e will give basically similar results, if carried out consistently. Thus, in this discussion the cause for a peak in P_e is that extended final-electron states exist for certain values of ε, and consequently the amplitude of the final-electron-state wave-function ϕ^f_{el} in the vicinity of the desorbate particle is relatively large. However, the calculations of P_e performed in the literature[7,13,15] have involved a slightly different formalism in which back-reflection of outgoing electron waves (outgoing from the desorbate) give rise to "resonances" at certain values of electron energy. These formalisms appear to be equivalent, and hence the discussion here takes over the results calculated for P_e by other workers.

B. Wave-functions for sloping box

Consider a particle moving in a "sloping one-dimensional box" of the shape shown in Fig. 7. The bottom of the box is taken at a standardised vibrational level $V^g = neV_{app}$, where V_{app} is the applied voltage difference between the emitter and its surrounding electrode, and is a negative quantity in the circumstances of field ionization. The slope of the inner end of the box is set equal to (-neF), where F is the field value near the emitter. The length L of the flat portion of the box is assumed to be much greater than the length L_s (= $- V_{app}/F$) of the sloping portion.

Consider a state f of vibrational level V^f with an inner classical turning point as shown in Fig. 7. Let η^f denote distance measured from the turning point and χ^f be the vibrational wave-function for this state. A dimensionless variable ξ^f may be defined by:

$$\xi^f = (2meF)^{\frac{1}{2}} \hbar^{-1} \eta^f \tag{54}$$

where m is the mass of the particle (desorbate ion).

It is a standard result[16] that the solution $\xi(\eta^f)$ of the Schrodinger equation in a uniform field has the form $a\phi(-\xi)$, where a is a normalisation coefficient and $\phi(\xi)$ is the Airy function defined by:

$$\phi(\xi) = \pi^{-\frac{1}{2}} \int_0^\infty \cos(\tfrac{1}{3}u^3 + u\xi)\, du \tag{55}$$

Hence, in the sloping part of the box we may write:

$$\chi^f = a(V^f)\ \zeta(\eta^f) - a(V^f)\ \phi(-\xi^f) \tag{56}$$

For large positive values of η^f and ξ^f, the asymtotic expression for $\phi(-\xi)$ is:

$$\phi(-\xi) \simeq \xi^{-\frac{1}{4}} \sin(\tfrac{2}{3}\xi^{3/2} + \tfrac{1}{4}\pi) \tag{57}$$

Consequently, within the sloping part of the box, at large positive distances η^f from the classical turning point for the particular level in question, the vibrational wave-function has the form of a rapidly oscillating sine-type wave, with a position-dependent amplitude $C(V^f,\eta^f)$ given by:

$$C(V^f,\eta^f) = a(V^f)\ (2meF)^{-\frac{1}{8}}\ \hbar^4\ (\eta^f)^{-\frac{1}{4}} \tag{58}$$

At the breakpoint (b), where the box changes from sloping to flat, η^f has the value $(\eta^f)^b$ given by:

$$(\eta^f)^b = (V^f - V^g)/neF \tag{59}$$

Hence, at this point, the sine-type wave has amplitude $C^b(V^f)$ given by:

$$C^b(V^f) = a(V^f)\ (2meF)^{-\frac{1}{8}}\ \hbar^4\ (V^f - V^g)^{-\frac{1}{4}}\ (neF)^{\frac{1}{4}} \tag{60}$$

Within the flat portion of the box, the wave-function at level V^f is a sine wave with constant amplitude and wavelength. In principle, the wave-functions in the flat and sloping parts of the box ought to be matched properly at the breakpoint. However, for present purposes the details of the matching can be disregarded. Thus, we may identify the constant amplitude of sine wave in the flat section with a coefficient $C^b(V^f)$ given by (60).

To normalise the complete wave-function one should, in principle, integrate $|\chi^f|^2$ over the whole of (one-dimensional) space. However, since the length L of the flat section is, by definition, much greater than the length of the sloping section, most of the contribution to the integral comes from the flat section of the box, and it is sufficient to apply the condition:

$$\int_0^L |C^b(V^f)|^2 \sin^2(kz)\ dz = 1 \tag{61}$$

where k is the circular wave-number appropriate to level V^f. This condition leads to the well-known result, independent of V^f:

$$C^b = (2/L)^{\frac{1}{2}} . \tag{62}$$

By comparison with (60), it follows that the coefficient $a(V^f)$ is given by:

$$a(V^f) = (2/L)^{\frac{1}{2}} (2m/\hbar^2 eF)^{\frac{1}{8}} n^{-\frac{1}{4}} (V^f - V^g)^{\frac{1}{4}}. \tag{63}$$

In this treatment, it is legitimate to ignore the fine details of the matching process, because such detail would be very sensitive to the exact shape assumed for the box near the breakpoint (for example, one could round the corner off). Since details of shape have no physical meaning, it makes no physical sense to perform a detailed matching.

We also require the density-of-states for the box at level V^f. Since L is much greater than L_s, it is sufficient to identify this density-of-states ρ_v^f with that for a particle in a one-dimensional steep-sided box of length L. This leads to (40) in §4C.

References

1. E.W. Müller and T.T. Tsong, Field Ion Microscopy: Principles and Applications (Elsevier, Amsterdam, 1969). Prog. Surf. Sci. 4, 1 (1974).
2. C. Zener, Proc. Roy. Soc. A140, 660 (1933).
3. R. Gomer and L.W. Swanson, J. Chem. Phys. 38, 1613 (1963).
4. For example: G. Herzberg, Molecular Spectra and Molecular Structure (Van Nostrand, Princeton, 1950).
5. For a recent review see: D.L. Cocke and J.H. Block, Surf. Sci. 70, 363 (1978).
6. R.G. Forbes, Surf. Sci. 61, 221 (1976).
7. A.J. Jason, Phys. Rev. 156, 266 (1967).
8. A.A. Lucas, Phys. Rev. Let. 26, 813 (1971).
9. A.A. Lucas, Phys. Rev. B4, 2939 (1971).
10. A.A. Lucas and M. Sunjic, J. Vac. Sci. Tech. 9, 725 (1972).
11. R. Brako and M. Sunjic, Surf. Sci. 54, 434 (1976).
12. G.R. Hanson, J. Chem. Phys. 62, 1161 (1975).
13. J.A. Appelbaum and E.G. McRae, Surf. Sci. 47, 445 (1975).
14. D.R. Penn, Surf. Sci. 52, 270 (1975).
15. G.R. Hanson and M.G. Inghram, Surf. Sci. 64, 305 (1977).
16. L.D. Landau and E.M. Lifshitz, Quantum Mechanics (Pergamon Press, Oxford, 1958).
17. A.J. Jason, A.C. Parr and M.G. Inghram, Phys. Rev. B7, 2883 (1973).
18. R.G. Forbes, Surf. Sci. 60, 260 (1976).
19. R.G. Forbes, Ned. Tidj. v. Vacuumtech. 16, 268 (1978).
20. R. Brako, M. Sunjic and A.A. Lucas, Surf. Sci. 60, 262 (1976).
21. R.G. Forbes, Surf. Sci. 46, 577 (1974).
22. R.G. Forbes, Proc. 7th Intern. Vac. Congr. & 3rd Intern. Conf. Solid Surfaces (Vienna 1977), p. 387.
23. International Standard ISO 31/0-1974(E): General principles concerning quantities, units and symbols.
24. International Standard ISO 31/VIII-1973(E): Quantities and units of physical chemistry and molecular physics.

25. R.G. Forbes, Physics Educ. (U.K.), 13, 5 and 269 (1978).
26. See: P.A.M. Dirac, The Principles of Quantum Mechanics (Clarendon Press, Oxford, Fourth edition, 1958).
27. S.J. Fonash and G.L. Schrenk, Phys. Rev. 180, 649 (1969).
28. S.P. Sharma, S.J. Fonash and G.L. Schrenk, Surf. Sci. 23, 30 (1970).
29. E.U. Condon, Phys. Rev. 32, 858 (1928).
30. E.U. Condon, Amer. J. Phys. 15, 365 (1947).
31. R.D. Young, Phys. Rev. 113, 110 (1959).
32. K.D. Rendulic, Surf. Sci. 70, 234 (1978).
33. N. Domke, E. Hummel and J.H. Block, Surf. Sci. 78, 307 (1978).
34. N. Ernst, Surf. Sci. 87, 469 (1979).
35. S.P. Sharma and G.L. Schrenk, Phys. Rev. B2, 598 (1970).
36. T.T. Tsong and E.W. Müller, J. Chem. Phys. 41, 3279 (1964).
37. N.F. Mott and H.S.W. Massey, Theory of Atomic Collisions (Clarendon Press, Oxford, 1949).
38. For a brief survey see: M.S. Child, Faraday Discussions 53, 18 (1972).
39. R.G. Forbes, J. Microscopy 96, 57 (1972).

Surface Science 60 (1976) 260–261

COMMENT ON "SURFACE PLASMON EXCITATION IN FIELD ION EMISSION" BY R. BRAKO AND M. ŠUNJIĆ

Received 8 March 1976

The periodic structure in field-ion energy distributions, sometimes referred to as the "Jason effect", continues to arouse interest. Jason [1] ascribed the structure to periodic variations in the electron transition rate for field ionization, with distance from the emitter surface, with the peak heights also affected by the spatial distribution of the imaging gas. But an alternative explanation was suggested by Lucas [2,3] in terms of multiple plasmon excitation by the departing ion. And in a recent note [4] Brako and Šunjić have refined this explanation.

Although Brako and Šunjić point out various deficiencies in their treatment, and stress the need for a full quantum-mechanical treatment, it seems to me that there is a major inconsistency in their argument and that this would not necessarily disappear in a fuller treatment. The inconsistency arises as follows.

From the statements on p. 435 of ref. [4], it is clear that Brako and Šunjić are assuming that "all ions have the same potential energy when created", and that they "describe the ion as a classical particle which is suddenly created at a distance d from the metal surface". Further, it is obvious from the context that this distance d is the "critical distance" of elementary field ionization theory.

Now, the experiments [1,5] show that on arrival at the collector some ions have a total energy somewhat less than that possessed by the ions making up the main peak (or "no loss line").

The ions in the main peak start from distance d with near-zero kinetic energy [6]. Therefore, if Brako and Šunjić's assumptions are correct, then either: (1) some ions are created with negative kinetic energy (whatever that may mean); or (2) some ions, created with near-zero kinetic energy, and picking up kinetic energy from the applied electric field as they move away from the emitter, then lose part of it to the surface plasmons.

Now, assumption (1) is incompatible with the authors' assertion that they are treating the ion as a classical particle. Therefore it may be rejected as inapplicable.

Assumption (2), on the other hand, implies that energy loss occurs meaningfully subsequent to ion creation, because the ion needs time to pick up energy from the field. But Brako and Šunjić state on p. 437 of ref. [4] that "it is not necessary to know the full trajectory, because the dominant contribution to the excitation strength arises from the discontinuity at time $t = 0$". In other words, their mathematics is interpreted to mean that most of the energy loss is due to the act of ioniza-

tion (at time $t = 0$) rather than to subsequent ion motion. But this interpretation directly contradicts assumption (2) above.

It therefore seems that, whichever way one looks at it, ref. [4] contains gross physical inconsistency.

There seem two possible reactions to this suggestion. First, to suppose that in a full quantum-mechanical treatment of field-ion emission the inconsistency would go away. This could be true. But it could be wishful thinking of a type peculiarly difficult to falsify.

The alternative is to suppose that the physical interpretation of the mathematics is in error, *and* that the experimental results have a different explanation.

The point I wish to stress is that, whether or not Jason's explanation of the experimental results is actually true, it is certainly a possible and logically consistent hypothesis. However, the surface-plasmon explanation as refined in ref. [4] seems *internally inconsistent.*

Richard G. FORBES

Department of Physics, University of Aston,
Costa Green, Birmingham B4 7ET, England

References

[1] A.J. Jason, Phys. Rev. 146 (1967) 266.
[2] A.A. Lucas, Phys. Rev. Letters 26 (1971) 813.
[3] A.A. Lucas, Phys. Rev. B4 (1972) 2939.
[4] R. Brako and M. Šunjić, Surface Sci. 54 (1976) 434.
[5] E.W. Müller and S.V. Krishnaswamy, Surface Sci. 36 (1973) 29.
[6] T.T. Tsong and E.W. Müller, J. Chem. Phys. 41 (1964) 3279.

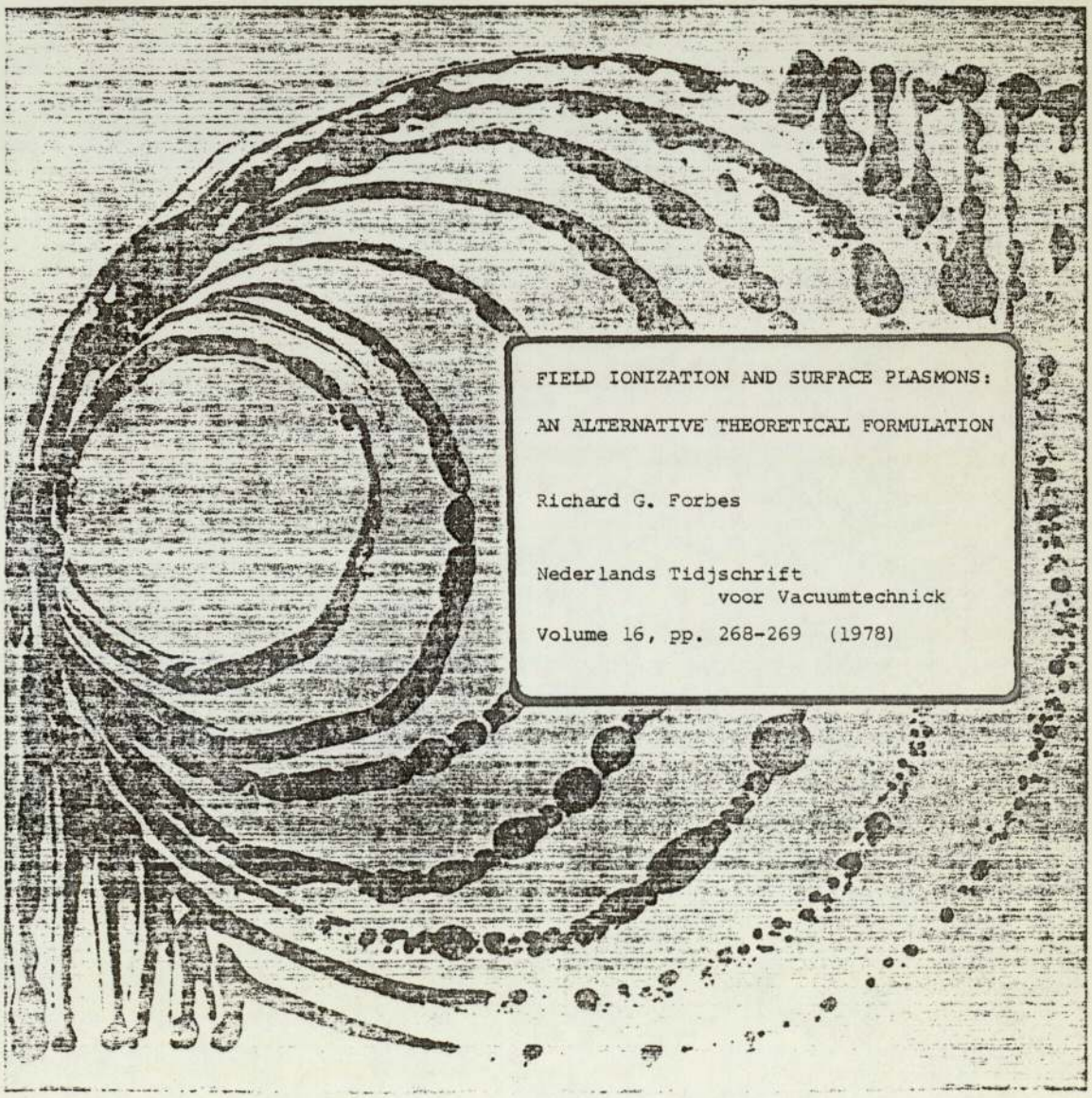
nederlands tijdschrift voor vacuumtechniek

orgaan van de NEDERLANDSE VACUUMVERENIGING

16e jaargang nr. 2/3/4 juni 5-9 1978

First European Conference on Surface Science

ECOSS I



FIELD IONIZATION AND SURFACE PLASMONS: AN ALTERNATIVE THEORETICAL FORMULATION

Richard G. Forbes

University of Aston, Department of Physics
Gosta Green, Birmingham, U.K.

INTRODUCTION

The process of surface field ionization is, for historical reasons, normally represented as a quasiclassical electron transition. It is, however, difficult to give a satisfactory a-priori theoretical justification for this procedure. Alternative approaches, based on molecular potential-energy curves (Molecular Terms), are more directly analogous to the conventional quantum-mechanical theory of radiationless molecular transitions. An approach of the latter type was used by Gomer and Swanson /1/, but has subsequently been neglected. This paper formally develops a second approach of this type, and uses it as a basis for continuing discussion of the surface-plasmon explanation of the periodic structure in field-ion energy distributions.

THE STANDARD-TERM - DIAGRAM APPROACH

The Standard Surface Molecular Term, U_n, gives the potential energy of the emitter-plus-desorbate system when the desorbate is in an n-fold charged state, it being assumed that the ion has been created by means of a thermodynamically reversible process in which the electrons for the originally-neutral desorbate have been placed at the emitter fermi-level. In the present context the desorbate is to be taken as an inert-gas atom or a Hydrogen molecule.

Field ionization of a neutral can then be represented as a non-radiative vibronic transition from a vibrational state (of the desorbate relative to the emitter) in the neutral Term U_0 to a vibrational state in the ionic Term, as illustrated in Fig. 1. The "diagonal" nature of the transition as shown on this diagram indicates that energy has been transferred to the electron gas in the metal, over and above that necessary to place the emitted electron(s) at the emitter fermi level. A diagram of this type is known as a Standard Term Diagram. The probable number of transitions per unit time from the initial state i to the final state f, j^{if}, is given by:

$$j^{if} = \sigma^i \, P^{if} \, t(\varepsilon^{if}) \qquad (1)$$

where σ^i is the probability that state i is occupied in the first place; P^{if} is an (infinitesimal) state-to-state transition rate-constant; ε^{if} is the amount by which the electronic energy-component in the final state exceeds that appropriate to the situation where all the electrons have been placed at the emitter fermi-level; and $t(\varepsilon^{if})$ is a function here called the *acceptance function* that represents the effect of the Pauli exclusion principle on the probability of electron transfer. ε^{if} is given in terms of the *standardised energies* $\mathcal{E}^i$ and $\mathcal{E}^f$ associated with the initial and final states by:

$$\varepsilon^{if} = \mathcal{E}^i - \mathcal{E}^f \qquad (2)$$

Basically, it is the distribution of emitted ions with respect to the energy $\mathcal{E}^f$ that is measured with a field-ion energy analyser. The zero of "energy deficit" can usefully be identified with the characteristic energy U'' shown on Fig.(1).

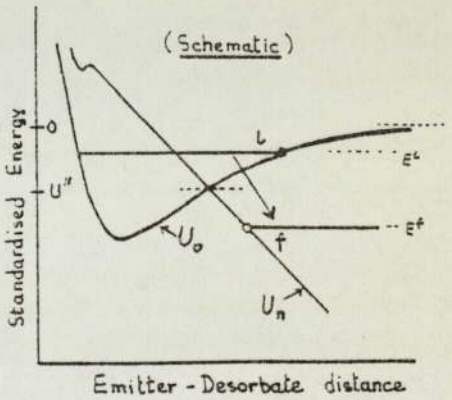


Fig. 1. Standard Term Diagram showing a radiationless vibronic transition. The blacked circle labels the initial state; the open circle labels the final state.

ALTERNATIVE EXPLANATIONS OF THE JASON EFFECT

Some years ago /2/, Jason discovered that at certain $\mathcal{E}^f$-values there were marked peaks in the ion energy distribution. The first subsidiary peak is typically some few electronvolts in energy below the main peak near U''. Jason attributed this subsidiary peak structure to the occurrence of transitions that may be represented as in Fig.(2a), assuming that the relevant ionization rate-constant had local maxima at certain emitter-desorbate distances, due to electron resonance effects. This explanation seems to provide a satisfactory explanation of the experimental observations.

Alternative explanations have been advanced by Lucas and co-workers, in terms of the creation of surface plasmons by the departing ion. Such

explanations involve one or other of the two mechanisms shown in Figs.(2b) and (2c). That shown in Fig.(2b) assumes a two-stage process: first field-ionization, and then de-energization of the emitted ion by creation of a surface plasmon. Quasiclassically, one thinks of the ion as travelling away from the surface picking up energy from the electric field; it then loses some of this energy by interaction with the surface-plasmon fields.

The mechanism of Fig.(2c) assumes simultaneous field ionization and plasmon creation, and appears to be the mechanism envisaged in a recent paper by Brako and Sunjic /3/. In their (quasiclassical) treatment Brako and Sunjic seem to describe the ion as a classical particle which is suddenly created at the "critical distance" of elementary field-ionization theory, that is at the crossing-point of the neutral and ionic Terms in Fig.(2c). Forbes /4/ has criticised the assumptions behind their treatment, on the grounds that it implies that a classical particle travels with negative kinetic energy through the region between the crossing-point and the classical turning-point for the final vibrational state (f). Brako, Sunjic and Lucas /5/, however, seem to feel that non-conservation of energy is not an important defect, provided that the ion travels across the region sufficiently quickly, because such a process is compatible with the Second Uncertainty Principle.

A corollorary to such an appeal to the Uncertainty Principle, though, is that a full wave-mechanical formulation should in principle be established. The full wave-function for the system would include vibrational wave-functions for the desorbate in the initial and final vibrational states. Clearly, in the situation of Fig.(2c) an expression for the rate-constant for simultaneous field ionization and plasmon creation should include a factor involving the overlap of decaying exponential tails in the negative-kinetic-energy region. By analogy with Gomer and Swanson's treatment /1/, this factor can probably be represented as an ion-tunnelling coefficient that is a sharp function of the final-vibrational-state energy ε^f .

The quasi-classical treatment in Ref./3/ does not throw up this factor. Numerical estimates of the factor are extremely difficult, not least because of lack of knowledge of the potential structure close to a field-ion emitter. However, it is clear in principle that neglect of this factor throws a shadow of doubt over the a-priori theoretical plausibility of surface-plasmon explanations of the Jason effect.

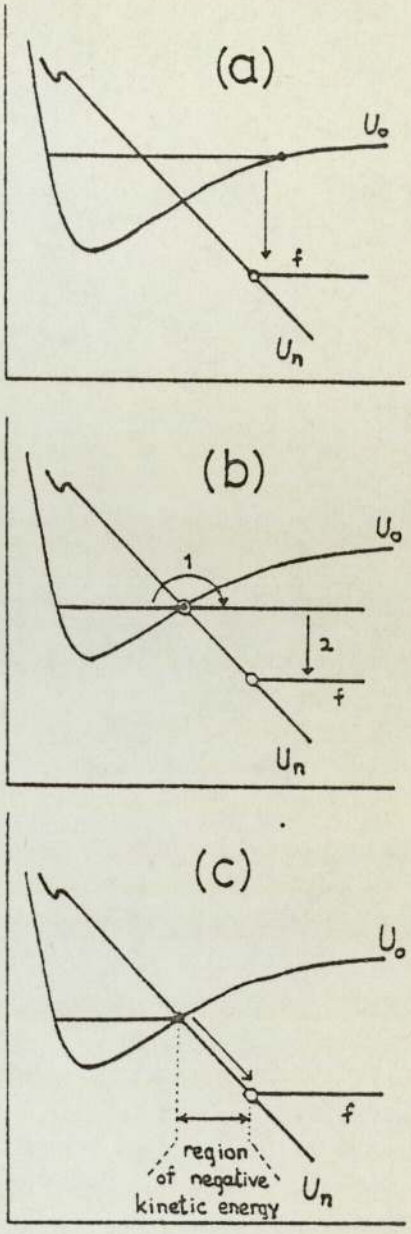


Fig. 2. Alternative mechanisms for the creation of an ion in final state (f), as observed in the Jason effect.

2a: A mechanism not involving surface-plasmon creation.
2b: Two-stage mechanism involving plasmon creation as second stage
2c: One-stage mechanism involving simultaneous field ionization and plasmon creation.

REFERENCES

/1/ Gomer, R. and Swanson, L.W., J.Chem.Phys. 38 (1963) 1613.
/2/ Jason, A.J., Phys.Rev. 146 (1967) 266.
/3/ Brako, R. and Sunjic, M., Surface Sci. 54 (1976) 434.
/4/ Forbes, R.G., Surface Sci. 60 (1976) 260.
/5/ Brako, R., Sunjic, M. and Lucas, A.A., Surface Sci. 60 (1976) 261.

J. Phys. E: Sci. Instrum., Vol. 11, 1978. Printed in Great Britain

An advanced field electron emission spectrometer

E Braun, R G Forbes, J Pearson, J M Pelmore† and R V Latham
Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET, UK

Received 21 August 1975, in final form 25 October 1977

Abstract A field electron emission spectrometer based on a hemispherical analysing element has been designed, built and tested. The instrument is versatile and convenient to operate, has a high sensitivity and a resolution of better than 30 mV. The performance of the instrument has been demonstrated in terms of the energy spectra of a tungsten emitter at 300 and 80 K, for which comparative theoretical data exist.

1 Introduction

The performance limitations of the integrating type of retarding potential electron energy analyser are well known (Plummer and Young 1970) and have been directly experienced in this laboratory during a programme concerned with the field-emitting properties of semiconducting materials. In particular, a Young–Müller spherical analyser was used for a study of cadmium sulphide (Salmon and Braun 1973) and a Van Oostrom type for a study of lead telluride (Sykes and Braun 1975).

To provide the level of experimental sophistication necessary to complement modern theoretical studies of field emission from semiconducting materials (e.g. Modinos 1974, Nicolaou and Modinos 1975), it is necessary to use a differential type of spherical deflection analyser, where it is possible to use electron multiplier techniques for greatly increasing the sensitivity of collector current detection. The output from such analysers gives the emission energy distribution directly with a nearly constant signal-to-noise ratio over the complete energy spectrum.

This approach was first used by Kuyatt and Plummer (1972), who designed an instrument based on a 135° deflection analysing element formed from concentric spheres of 25 mm mean radius. The present instrument, whilst following the general principles of the system developed by these authors, exploits both the theoretical and operational advantages of using an analysing element consisting of spheres of 50 mm radii with 180° deflection. It also uses a less complicated electron optical design for its input and output lenses, with a much simplified, though more versatile, electronic drive system. The analyser was designed by Professor D W O Heddle of The Royal Holloway College, University of London and manufactured by AEI Ltd.

To determine the resolution of the instrument and generally demonstrate its versatility, energy spectra were recorded for a clean tungsten emitter at 300 and 80 K. The halfwidths and high-energy slopes of these distributions were then compared with established theoretical results.

2 The spectrometer

2.1 *Layout of the facility*

The mechanical layout of the facility in its stainless steel vacuum vessel is shown in figure 1. The hemispherical analysing element with its associated input and output electron lens assemblies and electron multiplier is mounted horizontally on a 250 mm demountable flange, whilst the specimen stage and its electrical feedthroughs are mounted vertically from the top flange. A large viewing port assists in the specimen alignment procedure, and a smaller window at 45° to the analyser input z axis enables the field emission pattern to be viewed on the anode screen. The chamber is pumped by a standard Vacuum Generators ultrahigh-vacuum pumping system, using an oil diffusion pump, a 24 h cold trap and a titanium sublimation pump, where the ultimate pressure of the system is in the 10^{-9} Pa range.

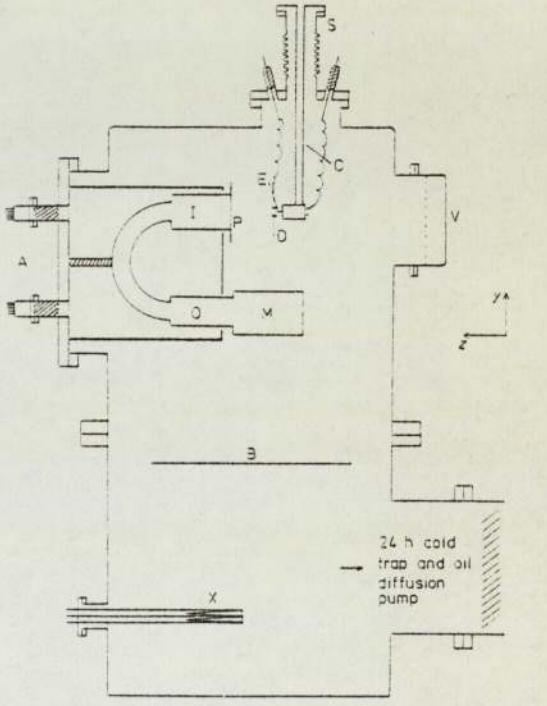


Figure 1 Layout of analyser in vacuum chamber. A, analyser; I, input lenses; O, output lenses; P, phosphor-coated anode; M, electron multiplier; E, field-emitting tip; D, deflector plates; C, cold finger; S, specimen manipulator; V, viewport; B, baffle plate; X, sublimation pump filaments.

† Present address: Department of Electronic and Electrical Engineering, University of Birmingham.

2.2 *Specimen stage and deflector plate assembly*

The specimen mount and deflector plates assembly (figure 1) is supported from a bellows-mounted manipulator stage, illustrated in figure 2. This enables the tip to move in the

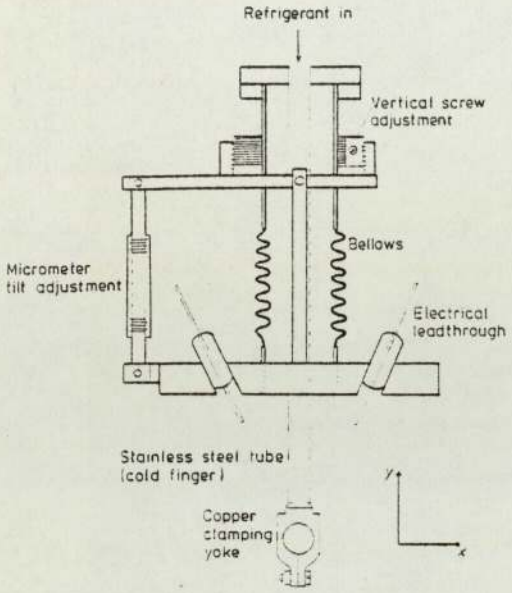


Figure 2 Specimen manipulator giving vertical (y) screw adjustment and horizontal (x) shift by tilting bellows.

x–y plane, i.e. parallel to the phosphor-coated screen anode, so that it can be axially aligned with the 1 mm diameter probe hole, which acts as the entrance aperture to the spectrometer. Subsequently, the deflector plates are used to shift the electron beam in the horizontal and vertical directions to bring selected areas of the field emission pattern over the probe hole.

Referring to figure 2, the specimen is mounted on a stainless steel tube which projects downwards from the manipulator stage into the vacuum chamber, and is vacuum-brazed to a clamping yoke machined from OFHC copper. For operation at elevated tip temperatures, or when biasing the emitter potential, a cylindrical ceramic block, carrying insulated leads appropriate to the specimen involved, is clamped into the yoke, which also supports the deflector plate assembly (figure 3(*a*)). The deflector plates are located rigidly relative to each other in V-slots accurately machined in split ceramic bushes which are clamped between stainless steel rings (figure 3(*c*)). The entire plate holder can be centred up about the position of any emitter tip, avoiding the need to control the emitter shank dimensions precisely when preparing different specimens. The use of a polar (r, θ) adjustment for this purpose keeps the overall size of the plate assembly small so that it can be removed for adjustment via the top port, along with the tip and all associated electrical connections. For operation at low temperatures, the cylindrical ceramic block is replaced by a copper mounting block, and the component parts of the deflector plate assembly are reassembled differently, so that this block can be positioned within the clamping yoke (figure 3(*b*)). The stainless steel tube is then filled with refrigerant (e.g. liquid nitrogen) to provide a 'cold finger' stage.

2.3 *Electron optical design of the spectrometer*

The choice of Kuyatt and Plummer (1972) to use a spherical deflector of mean radius 25 mm was dictated by the need for a small and compact ultrahigh-vacuum chamber; in consequence, however, it was necessary to use a 135° deflector (Gadzuk and Plummer 1973). Our choice of a deflector of mean radius 50 mm, which is based on previous experience with such hemispheres (Heddle *et al* 1973) offers several other important advantages apart from its higher theoretical resolution (Heddle 1971). The analysing energy is larger, so reducing the effects of stray magnetic fields; the apertures are larger, so reducing the effects of space charge; the voltage difference between the hemispheres is larger, so reducing the effects of patch fields (Parker and Warren 1962); machining errors are relatively less important; finally, the 180° deflector has slightly superior focusing properties.

The analysing properties of such a hemispherical deflection element have been analysed by Kuyatt and Simpson (1967),

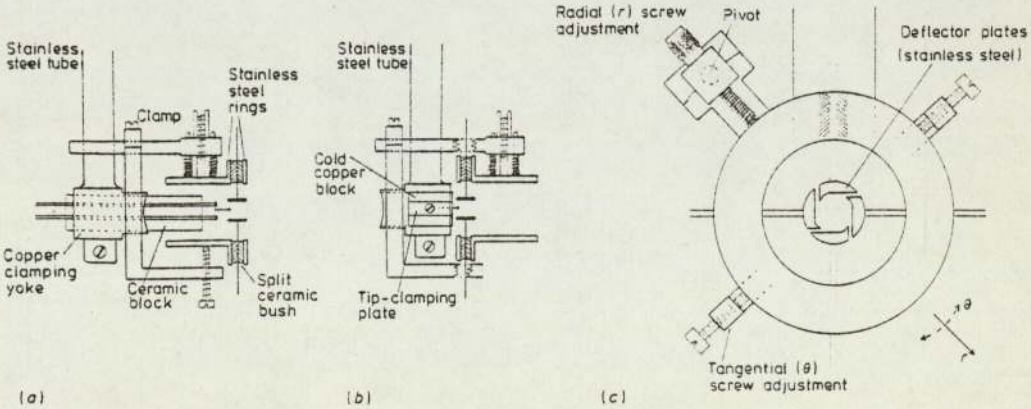


Figure 3 Deflector plate assembly. (*a*) Arrangement for elevated temperature, or for electrical biasing of emitter tip. (*b*) Arrangement for cooling of emitter. For clarity in (*a*) and (*b*), plate pins are shown rotated through 45° about axis of emitter. Correct component orientations are shown in (*c*): detail of plates.

who showed that, provided that the semi-angle x of the incident electron beam is such that $x^2 \gg w/4r$, the resolution (as measured by the width $\Delta E_{1/2}$ of the transmission function at half-height) is given by

$$\Delta E_{1/2} = E_a\, w/2r$$

where E_a is the analysing energy (that is the kinetic energy of the electrons that pass through the hemispheres along the midpath), w is the width of the input and exit apertures in the diametral plane of the hemispheres and r is the mean radius of the hemispheres. E_a is then related to the voltage difference ΔV_H applied between the hemispheres by

$$e\Delta V_H = E_a[(r_2/r_1) - (r_1/r_2)]$$

where r_1 is the radius of the inner hemisphere, and r_2 is that of the outer hemisphere.

The input lens system must now be designed to prepare a suitable beam to match the analyser requirements, i.e. to decelerate the electrons and form a real magnified image of the source at the entrance aperture of the hemispheres. Since the effective input aperture to the hemispheres was here taken as 1·2 mm diameter, the requirement is to decelerate the electrons from 2 keV to 2 eV with a magnification of 2·4. For experimental reasons, it should also be capable of operating over a range of input energies (1–4 keV) and a range of analysing energies (1–20 eV).

As the deceleration ratio is approximately 1000, a four-element design was chosen. This incorporates a three-element lens designed in accordance with the method described by Heddle (1971), followed by a fixed-ratio lens. The component lenses have different diameters so that lens elements shall be a reasonable length, varying between 9 and 20 mm, where the total length is 70 mm. The operational mode is shown by the electron trajectories of figure 4, where the

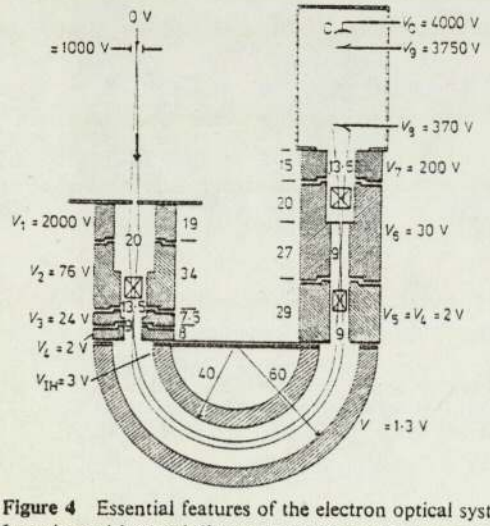


Figure 4 Essential features of the electron optical system. Lengths and internal diameters are shown to the nearest 0·5 mm. Deflector plates are contained in elements 2, 5 and 6. Also shown is a set of experimental operating voltages appropriate to an analysing energy of 2 eV; the voltages are given relative to a hypothetical stainless steel electrode outside whose surface the transmitted electrons would just come to rest. V_8 and V_9 are the voltages on the first and last dynodes of the electron multiplier and V_C is the voltage on its collector C.

focusing characteristics are adjusted by varying the potential of element 2. This element also contains a set of x–y deflector plates for fine beam-steering.

The function of the output lens system is to reaccelerate the electron beam leaving the hemispheres and at the same time to collimate it into the electron multiplier. Lens L_5/L_6 accelerates the electrons on to a rectangular aperture, where the image magnification is 0·6. (Heddle (1971) has shown that the use of a rectangular aperture markedly improves transmissions without significantly affecting resolution.) A set of deflector plates in L_5 is used for steering the beam through this aperture. The emergent beam is then accelerated to approximately 200 eV by lens L_6/L_7 so as to impinge diffusely on to the first dynode of the electron multiplier, which is a Twentieth Century ED25 model with a gain of about 10^5.

The operating conditions for our deflecting element were decided according to the criterion (D W O Heddle, private communication) that it should have a designed resolution of 20 meV for electrons entering the input lens assembly with an energy of 2 keV. Thus, assuming an entrance aperture of 0·5 mm diameter and that the emitter is effectively a point situated 50 mm away from it, the half-angle of the emergent beam will certainly not be greater than 5 mrad. Using these data in conjunction with the law of Helmholtz and Lagrange and the well established design equations for spherical deflectors of Kuyatt and Simpson (1967) summarised above leads to the following requirements: an analysing energy of 2 eV, an effective input aperture to the hemispheres of 1·2 mm diameter, and an input beam half-angle of 66 mrad. Hemispheres of radii 40 and 60 mm suffice to accommodate the beam and require a theoretical voltage difference of 1·73 V for the required resolution.

The lens elements of this electron optical system were constructed from non-magnetic stainless steel and accurately positioned by the use of sapphire spheres. A Mumetal can surrounds the low-electron-energy region of the system, and the magnetic field can be further reduced by the use of external Helmholtz coils.

3 The electrical and electronic systems

In contrast to the arrangement of Kuyatt and Plummer (1972), which used a number of separate voltage generators to supply the lens elements, our system employs a single power supply from which all the lens voltages (except V_7) are derived by means of a 2MΩ dropping resistor chain (figure 5). This has the advantage that the lens voltage ratios are independent of the magnitude of the anode voltage V_1, so that the focusing properties of the lenses are largely unaffected by change in V_1 over the operating range of the instrument from 1–4 kV. A good range of individual voltage adjustments is available for all lens elements and for the internal deflector plates. In-line jack plugs are also provided on the resistor chain for measurement of lens voltages and currents by DVM and electrometer.

The voltage applied between the hemispheres may be varied to accommodate a range of analysing energies from 2 to 18 eV. The former is required for high-resolution operation, whereas the latter offers a greater output signal at low resolution for the initial beam alignment procedure. Separate adjustments are provided for setting each hemisphere voltage relative to the hemisphere 'mid-potential' V_4.

The method adopted for energy scanning is to sweep the voltage V_4, applied to the input and output lens elements at each end of the hemispherical deflector. The output from a versatile, purpose-built ramp generator is applied to the lower end of the input lens supply resistor chain. This sweeps

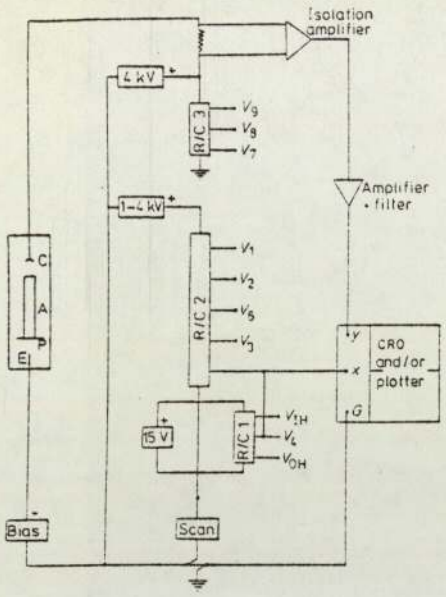


Figure 5 Voltage supply arrangements. Supplies for the beam deflector plates (not shown) are taken from a separate resistor chain.

the hemisphere mid-potential whilst leaving the voltage difference between the hemispheres (and hence the resolution) constant throughout the scan. The main anode voltage V_1 is also unaffected so that there is no interference with the emission process or beam optics in the emitter–anode space. Although the other lens voltages vary slightly, this has negligible effect on the lens focusing properties, especially so when the scan width is reduced for high-resolution work.

In terms of electron energy, this scanning method amounts to a simultaneous adjustment of the potential energy of every electron entering the hemispheres. At different points in the scan, electrons from each part of the initial kinetic energy distribution will have kinetic energies adjusted by this process, which are equal to the fixed 'analysing energy' of the spectrometer (determined solely by the dimensions and voltages of the hemispheres). A representative number of electrons from all parts of the distribution will thus pass round the hemispherical deflector, and be collected in the output signal, at the appropriate values of V_4.

The floating output signal from the electron multiplier is brought to ground potential by an isolation amplifier (Analog Devices type 273K), and fed, via a purpose-built amplifier and variable low-pass filter, to an x–y plotter or storage oscilloscope. The x input is derived directly from the ramp generator and is adjustable over a continuous range of sweep periods from 10 ms to 100 s, the faster scans permitting the observation of rapidly changing spectra.

The emission-extraction field is applied to the tip by means of the adjacent deflector plates; the emission current can therefore be controlled independently of V_1. A separate voltage generator supplies a resistor chain which incorporates parallel potentiometers and reversing switches, providing separate adjustments for x and y plate pairs. Each plate pair is controlled by ganged potentiometers, so that the plate voltages can be adjusted symmetrically about a mean value which is controlled from the power supply. The output leads to the plates contain 10 MΩ safety resistors. If the lower end of the plate supply resistor chain is isolated from earth, a common negative voltage can be applied to all four plates (up to −10 kV) for tip cleaning by field desorption.

4 Instrument performance

4.1 *Resolution*

The performance of the analyser has been evaluated using tungsten emitters, for which comparative theoretical and experimental data already exist. The procedure adopted was basically an attempt to follow the method of resolution determination described theoretically by Young and Kuyatt (1968), but hitherto not experimentally reported. (In describing their deflection analyser, Kuyatt and Plummer (1972) make only general reference to the method.) This proved less straightforward than expected. The method presupposes that a tungsten emitter can be regarded as a reliable electron source having a well calibrated and readily reproducible energy distribution. In the light of our difficulties we question here the extent to which this basic assumption is valid and describe some of the pitfalls of the method encountered in practice.

Although tungsten is the most studied field emitter, there is a measure of disagreement in the literature, even as regards the halfwidth to be expected for the electron energy distribution at 300 K. For directions other than the anomalous ⟨100⟩ and ⟨110⟩ (Swanson and Crouser 1967) experimental results reported fall broadly into two groups, having values of approximately 0·22–0·24 eV (Swanson and Crouser 1967) and approximately 0·15–0·19 eV (Young and Müller 1959, Czyzewski 1973). A division is apparent also in published theoretical curves (Young 1959, Young and Müller 1959, Young and Kuyatt 1968). For this reason, tungsten is less than ideal for instrument-testing purposes; however, owing to the general dearth of reliable energy data, no better alternative is available at present. Whilst a great deal of work has been done on the preparation and conditioning of tungsten field-emitting tips, workers who have mostly been concerned with tip lifetime and emission stability generally have not studied the associated energy distributions. Conversely, authors giving energy distribution data have not described the preparation and cleaning of their emitters. In our experience, this incomplete experimental picture is unsatisfactory because we have found that those parameters which are customarily taken as measures of analyser performance, namely, the energy distribution halfwidth and especially the slope of the high-energy edge, can both be very susceptible to the presence of tip contaminants. In particular, we have noted that the presence, under UHV conditions, of a highly stable and symmetrical emission pattern on the fluorescent screen anode of our instrument is not a reliable indication of an emitting surface which is sufficiently clean for our purpose.

Our tungsten emitters were prepared from 0·15 mm diameter wire by electronically switched DC electrolytic etching in 2N NaOH solution. After washing in water the tips were transferred directly to the vacuum chamber. Following bake-out and pump-down to pressures below 10^{-8} Pa, they emitted at first unstably, giving diffuse emission patterns and very broad energy distributions of irregular and varied shape. For room-temperature measurements, the emitter shanks were spot-welded to a lateral tungsten supporting wire that could be electrically heated to provide an *in situ* cleaning facility. Heating in this way for a few minutes at about 1170 K (measured by an optical pyrometer) invariably stabilised the emission and also changed the emission pattern, usually producing the well known display of four spots

corresponding to individual crystallographic directions. Although the energy distribution was then typically much more stable, it remained broad (approximately 0·4 eV half-width) and nearly symmetrical, i.e. without a sharp high-energy cut-off. The distribution also frequently exhibited a plateau on the low-energy side, visually similar to the well known 'Swanson hump'; however, this was evidently an unstable state since random switching occurred between this asymmetric and the more usual symmetrical distribution shapes. In all cases the plateau eventually disappeared during the subsequent procedure, so it was attributed to contamination.

Repeated heating at temperatures up to 1470 K reduced the halfwidth systematically to the region of 0·28–0·30 eV. However further heating, including flashing to temperatures in excess of 2200 K, failed to reduce the halfwidths significantly below this value. The emission remained generally unchanged apart from some reduction of emission current attributed to tip blunting, which led to the subsequent avoidance of very high temperatures. It is emphasised that the emission patterns were now symmetrical with a high degree of stability, and that the reproducibility between emitters of the energy spectra was thought to be a good indication that the emitters were clean. However the halfwidths were still typically more than 20% greater than the value expected. Figure 6 shows experimental curves from

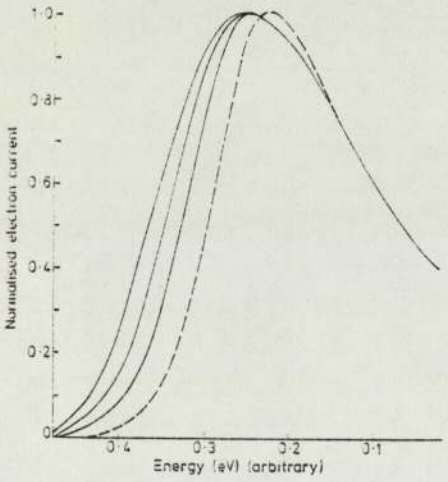


Figure 6 Electron energy distribution at 300 K from the tungsten emitter A. Full curves, experimental; broken curve, theoretical (Young 1959).

an emitter A at three consecutive stages of the cleaning process described above, with distribution halfwidths of 0·30, 0·28 and 0·26 eV respectively, the latter being the narrowest curve obtained in this way. The theoretical curve of Young (1959) is shown for comparison. The low-energy tails of the distributions, below about 70% peak height, are of identical shape and agree well with theory. Clearly, there is some distortion of the peak shape, and the curves are broadened at the high-energy side of the distribution. Reference to table 1 shows that values of halfwidth and of the width of the high-energy edge slope (between the 90% and the 10% peak height points) both fall appreciably short of theoretical expectations. The repeated appearance of this result, from several emitters, at first cast serious doubt on

Table 1 Energy distribution data.

Emitter	Halfwidth (eV)	Edge slope width (eV)†	Temperature (K)
Theory	0·18	0·04	80
	0·23	0·10	300
A	0·30	0·14	300
	0·28	0·12	300
	0·26	0·11	300
B	0·23	0·11	300
C	0·24	0·14	300
	0·22	0·08	80

† Defined as the energy interval between the 90% and 10% peak height points on the high-energy (Fermi) edge (after Young and Kuyatt 1968).

the performance of our analyser, and led to thorough checks on the requirements for lens voltage stability and the reduction of stray magnetic fields, which might have been responsible, but no such explanation was found.

The difficulty was resolved when an emitter B, which initially had followed this same behaviour pattern, was subjected to field desorption at 5 kV whilst held at 1270 K. The resulting energy distribution at 300 K is shown in figure 7, again with the theoretical curve for comparison. Table 1 shows that the agreement is now very good. If it can be assumed that there is now no contamination broadening, the edge slope width of 0·11 eV, compared to the theoretical value of 0·10 eV, would set an upper limit of about 50 meV for the analyser bandwidth (halfwidth). The result compares favourably with the experimental curve of Young and Müller (1959), whose value of 0·13 eV would correspond theoretically to a bandwidth of around 100 meV, but who from other

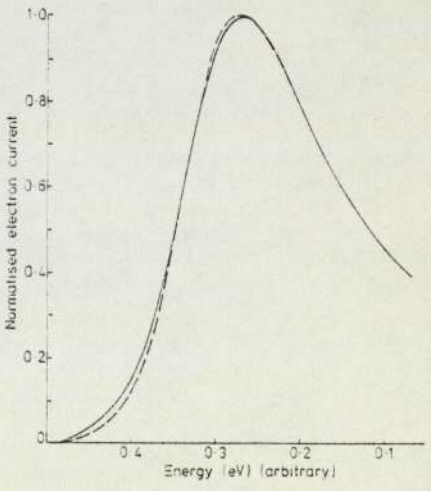


Figure 7 Electron energy distribution at 300 K from the tungsten emitter B. Full curve, experimental; broken curve, theoretical (Young 1959).

considerations claimed a resolution of 30 meV. We conclude from this that the resolution of our instrument is satisfactory for the purpose for which it was designed, despite earlier misleading evidence to the contrary, which we now attribute to the effects of a stable, long-lived contaminant.

Encouraged by the apparent superiority of field desorption as opposed to tip heating, we also attempted to clean some tungsten emitters by field desorption alone. These were positioned in our cold stage holder. The edge slope width of the low-temperature tungsten distribution should provide a more sensitive test of analyser resolution. The results obtained revealed further interesting features. Typical curves, at 300 and at 80 K, for an emitter C are shown in figure 8, again with the 300 K theoretical curve. Table 1

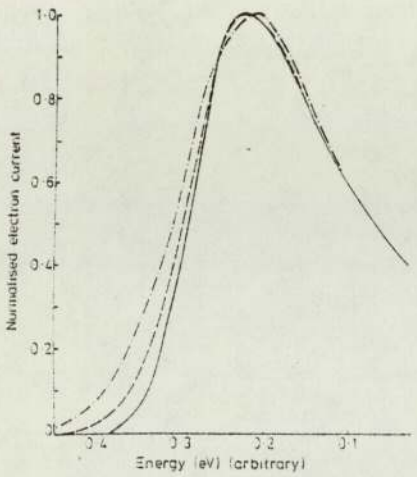


Figure 8 Electron energy distribution at 80 K and 300 K from the tungsten emitter C. Full curve, experimental at 80 K; chain curve, experimental at 300 K; broken curve, theoretical at 300 K (Young 1959).

shows that the halfwidth at 300 K was easily reduced to 0·24 eV by field desorption alone, without the difficulty previously experienced. At 80 K the halfwidth was reduced further. Although there is a marked reduction in the edge slope width upon cooling, the final value is 0·08 eV where that expected theoretically (Young and Kuyatt 1968) would be only about 0·04 eV for an analyser bandwidth of 30 meV.

This last result can be understood, however, if the distribution from this emitter at 300 K is compared with the theoretical curve. Although the halfwidth is substantially what would be expected for a clean emitter, the edge slope width is seen to be far too great (table 1). The behaviour of earlier emitters (table 1) has shown that this feature can be caused by contamination. We therefore believe that our failure here to obtain the slope width expected theoretically is attributable, not to instrumental broadening, as might at first be supposed from the 80 K curve alone, but to the fact that the energy distribution is still modified by contamination. It is important to appreciate that, at 300 K, changes in the energy distribution of this emitter C, due to contaminants, are not so readily apparent as they were in the case of the heat-treated emitter A. This is because they are now mainly confined to the region around the peak of the curve, and have very little effect on the halfwidth. Only the edge slope width is now sensitive to the presence of these contaminants. The experimental procedure must therefore include a preliminary measurement of edge slope width at 300 K, whose results should agree with theory, before measurements made on a cooled emitter can reliably be used to determine analyser resolution.

4.2 *Calibration*

This is the important procedure whereby the work function of the analysing hemispheres ϕ_H is determined from the energy distribution of a standard emitter, in this case tungsten. With this information, it is then possible to use the instrument to identify the position of the Fermi level (FL) on the energy distribution obtained from all other emitters. Such a facility will be particularly valuable for the fundamental studies of semiconductor tips for which the instrument was originally developed.

For an earthed emitting tip, the zero of potential will be at the FL of the emitter. It then follows from the earlier discussion in §2.3 that if E_a is the preset and accurately known analysing energy of the hemispheres, the condition for an electron emitted at the FL to be transmitted by the analysing element will be

$$eV_4 = E_a + \phi_H$$

where V_4 is the scan voltage and also the mid-potential of the hemispheres. Thus, if the position of the FL can be identified on an energy distribution, the corresponding value of V_4 may be used in the above equation to find ϕ_H. For a room-temperature tungsten distribution, Young (1959) has calculated that the FL occurs at a point 73% up the high-energy slope. Accordingly, ϕ_H has been measured to be $4{\cdot}23 \pm 0{\cdot}03$ eV.

For any other earthed emitter (i.e. with its FL at earth) the position of the FL on its resulting energy distribution, obtained with identical analysing conditions, will occur at the same value of V_4 used to calculate ϕ_H. This is because V_4 is a function only of the analyser parameters and not of the emitter.

5 Conclusion

The paper has described the design and operation of a high-resolution field electron emission spectrometer based on a hemispherical deflection energy analyser: it supersedes the previous generation of retarding field analysers used in our studies of field emission from semiconductors. The main advantages of the instrument arise from its being a deflection analyser. These advantages are listed below.

(i) The output provides the energy spectrum directly, rather than requiring ancillary differentiation. This adds considerably to the versatility of the instrument.

(ii) Resolution is uniformly high across the spectrum, which is particularly valuable for resolving spectral detail at low energies.

(iii) Electron multiplier detection is readily used, which increases the sensitivity and permits the use of rapid electronic display techniques and the future introduction of counting techniques.

(iv) Its performance is not critically dependent upon the electron optical positioning of the tip.

The versatility and overall performance of the instrument have been evaluated using tungsten emitters. These measurements have brought to light several practical difficulties associated with the cleaning of tungsten field-emitting tips by heating and field desorption. Results which at first sight cast doubts upon our instrument's performance, have been found to result in fact from insufficiently clean emitters. We therefore emphasise the great care which must be taken to

ensure that a tungsten emitter, used as a test specimen, really is clean. We have found that whilst the best estimates of analyser resolution must be obtained at low tip temperatures, detailed measurements of the energy distribution shape at 300 K are also necessary. We conclude that the edge slope width at 300 K is a much more important parameter in this context than has hitherto been supposed. It is apparently sensitive to a form of contamination whose presence is not revealed by more usual criteria such as the emission pattern symmetry, the stability of the analyser signal, or in some cases even by measurements of the half-width of the energy distribution. An additional use has been made of a 'clean' 300 K distribution to determine the work function of the analysing hemisphere.

Acknowledgments

We should like to thank Drs Ward Plummer and C E Kuyatt of the US National Bureau of Standards for their valuable cooperation when setting up the project; Professor D W O Heddle for designing the electron optics and subsequent general advice; and Mr F Lane for his assistance in the mechanical design of parts of the spectrometer. We also thank the Science Research Council for its financial support of this work.

References

Czyzewski J J 1973 *Surf. Sci.* **39** 1–20

Gadzuk J W and Plummer E W 1973 *Rev. Mod. Phys.* **45** 487–548

Heddle D W O 1971 *J. Phys. E: Sci. Instrum.* **4** 589–92

Heddle D W O, Keesing R G W and Kurepa J M 1973 *Proc. R. Soc.* A **334** 135–47

Kuyatt C E and Plummer E W 1972 *Rev. Sci. Instrum.* **43** 108–11

Kuyatt C E and Simpson J A 1967 *Rev. Sci. Instrum.* **38** 103–11

Modinos A 1974 *Surf. Sci.* **41** 425

Nicolaou N and Modinos A 1975 *Phys. Rev.* B **11** 3687

Parker J H and Warren R W 1962 *Rev. Sci. Instrum.* **33** 948

Plummer E W and Young R D 1970 *Phys. Rev.* B **1** 2088

Salmon L T J and Braun E 1973 *Phys. Stat. Solidi* **16** 527–32

Swanson L W and Crouser L C 1967 *Phys. Rev.* **163** 622–41

Sykes D E and Braun E 1975 *Phys. Stat. Solidi* b **69** 137

Young R D 1959 *Phys. Rev.* **113** 110–4

Young R D and Kuyatt C E 1968 *Rev. Sci. Instrum.* **39** 1477–80

Young R D and Müller E W 1959 *Phys. Rev.* **113** 115–20

VI. FIELD EVAPORATION

VIA BASIC THEORY

35 An alternative theoretical approach to field-evaporation rate sensitivities
R.G. Forbes
Surface Sci., 46 (1974) 577-601

36 Field evaporation theory: The atomic-jug formalism
R.G. Forbes
Surface Sci., 70 (1978) 239-54

37 Charge hopping and charge draining: Two mechanisms of field desorption
R.G. Forbes
Surface Sci., 102 (1981) 255-63

38 A new formula for predicting low-temperature evaporation field
R.G. Forbes
Appl. Phys. Lettrs., 40 (1982) 277-9

39 Field-evaporation theory: A re-analysis of published field-sensitivity data
R.G. Forbes, R.K. Biswas and K. Chibane
Surface Sci., 114 (1982) 498-514

40 New activation-energy formulae for charge-exchange-type mechanisms of field evaporation
R.G. Forbes
Surface Sci., 116 (1982) L195-201

41 Electrothermodynamic cycles applied to ionic potentials and to field evaporation
R.G. Forbes
J. Phys. D: Appl. Phys., 15 (1982) 1301-22

42 Theoretical arguments against the Müller-Schottky mechanism of field evaporation
R.K. Biswas and R.G. Forbes
J. Phys. D: Appl. Phys., 15 (1982) 1323-38

43 An evaporation-field formula including the repulsive ion-surface interaction
R.G. Forbes
J. Phys. D: Appl. Phys., 15 (1982) L75-7

44 Towards a criterion for the a-priori prediction of field-evaporation mechanism
R.G. Forbes
J. Phys. D: Appl. Phys., 15 (1982) L99-104

cont./

FIELD EVAPORATION - BASIC THEORY (cont.)

45 The temperature dependence of evaporation field for Gomer-type mechanisms of field evaporation
K. Chibane and R.G. Forbes
Surface Sci., 123 (1982) (in press)

46 The escape mechanism in low-temperature field evaporation
R.G. Forbes
Proc. 29th Intern. Field Emission Symp., (Göteborg, 1982) pp

VIB APPLICATIONS OF FIELD EVAPORATION THEORY

47 The derivation of surface atomic information from field-evaporation experiments
R.G. Forbes and K. Chibane
Proc. 29th Intern. Field Emission Symp., (Göteborg, 1982) pp

48 Arguments about emitter shape for a liquid-metal field-ion emission source
R.G. Forbes and G.L.R. Mair
J. Phys. D: Appl. Phys., 15 (1982) L175-8

SURFACE SCIENCE 46 (1974) 577–601 © North-Holland Publishing Co.

AN ALTERNATIVE THEORETICAL APPROACH TO FIELD EVAPORATION RATE SENSITIVITIES

RICHARD G. FORBES

Department of Physics, University of Aston in Birmingham, Gosta Green, Birmingham B4, England

Received 1 April 1974

A new theoretical approach is developed, and used to re-analyse existing field-sensitivity data. The approach is based on expansion of the rate-theory exponent as a Taylor series, about a field defined by a unity-rate-constant criterion for field-evaporation onset. Two forms of expansion are used, one of which is similar to a statistical-mechanical formula used in the theory of point-defect migration in strained metals. A procedure suggested by point defect migration theory provides a crude theoretical estimate (2.8 eV) for the first Taylor coefficient, for the field evaporation of tungsten. This is surprisingly close agreement with experimental estimates (mean 2.0 eV) derived from published data. Predictions for some other materials are listed. Advantages and implications of this alternative approach are considered. There seems a need for greater precision in the discussion of field evaporation, and for broadening of its experimental basis.

1. Introduction

By the application of a high electric field, surface atoms can be induced to desorb as positively-charged ions. This process, field evaporation, is basic to the use of the field-ion microscope and the atom-probe [1]), and offers a unique tool for investigations into the binding of surface adatoms [2]).

Unfortunately, the details of field evaporation are not well understood, as is made clear by recent reviews [3–5]). There are two competing models, the image-hump model introduced by Müller [6]), and the intersection or charge-exchange model initially proposed by Gomer [7]). Both models are capable of qualitatively predicting the relative abundances of ions of different charges [4]); neither model satisfactorily explains the effect of adsorbed image-gas atoms on rate-sensitivities [3]). Both models are partially successful in predicting the evaporation fields for metals, at low temperatures [1]). Both models are capable of predicting the values of field sensitivities measured in the manner introduced by Brandon [8]), but this is not a decisive test of their competitive merit because both models in effect contain an adjustable parameter, the effective polarisability [5]), whose value is not independently known. Further,

it is clear that current versions of the models may not properly take into account all the currently known theoretical effects[4]).

The sharpest difference between the models is that they predict different field dependences for the evaporation rate-constant[4,5]). Recent experiments[9,10]) have now produced data on the evaporation of tungsten over a reasonable range of fields, and this has been subjected to systematic statistical tests by Vesely and Ehrlich[5]). They conclude that the best representation of the results is by means of the image-hump model, taking the charge on the evaporating ion as 2+.

However, it is known from atom-probe work[1,11]) that tungsten ions usually evaporate in a triply-charged state, sometimes in a quadruply-charged state. Yet it has been shown theoretically both by Taylor[3]) and by Chambers[12]) that "post-ionization" of an evaporated ion on its way out is highly improbable. So something of an impasse now exists in field evaporation theory, from which the likeliest exits might seem the accumulation of much more experimental data, or the derivation and validation of independent means of estimating the "adjustable" parameters.

The present paper has nothing new to say about the competing merits of the two established models. Rather, I shall aim to supplement the existing analyses by exploring a novel approach to the treatment of evaporation current and rate-constant sensitivities. The approach, whose possibilities became apparent, somewhat unexpectedly, through a study[13]) of a recent paper by Page and Ralph[14]), makes use of an analogy between field evaporation theory and the theory of point-defect migration in strained metals. The outline of argument is as follows.

First, rate-sensitivity theory is developed in a way which shows that rate-constant field sensitivity may be simply related to the coefficients in a certain Taylor expansion. Changing the independent variable then puts this theory into a form whose resemblance to point-defect migration theory is clear. A standard zero-order approximation can then be used to derive a theoretical estimate of the first Taylor coefficient. Discussion ensues from the comparison of this with experimental estimates derived from published work.

The nomenclature used here is a refinement of those used by previous authors. The differences, and reasons for them, are set out in Appendix A4. The present system is felt to be clearer.

This paper is an extension of work first presented at the 20th Field Emission Symposium, Pennsylvania State University, 1973.

2. The Taylor expansion approach

Suppose that the evaporation rate-constant k_n for the atom at site n may

be written in the standard Arrhenius form:

$$k_n = A_n \exp(-Q_n/kT), \tag{1}$$

where A_n is a pre-exponential and Q_n an activation energy that depend on (amongst other things) the local electric field value F_n at site n. This equation can also be written in the "inclusive" form:

$$k_n = \exp(-\psi_n/kT), \tag{2}$$

where

$$\psi_n = Q_n - kT \ln A_n. \tag{3}$$

Though eq. (2) seems dimensionally inconsistent, this is not a real difficulty: there is an implicit convention (see Appendix A.1) governing the interpretation and transformation of such equations. It would be possible to carry out the complete analysis using dimensionally consistent equations, but only at the cost of unnecessary complications in notation and algebra. For convenience, we shall here drop the suffix n from the independent variables and from A_n, Q_n, and ψ_n, leaving it to be remembered that our equations apply to field evaporation from a specific site, that there will be variations in F_n from site to site at given applied voltage, and that there may be slight variations[5]) from site to site in the functional dependence of ψ_n on F_n and in the numerical value F_n^c of the local evaporation field as defined below.

ψ can be expanded as a Taylor series about some field F^c, to give (in first order):

$$\psi(F) = [Q(F^c) - kT \ln A(F^c)] + (F - F^c)\, \partial\psi/\partial F\,|_{F^c}. \tag{4}$$

An evaporation field is sometimes defined by the requirement that $Q = 0$. However I shall require $\psi(F^c) = 0$, i.e. that the whole of the expression in square brackets be zero. From eq. (2), this is equivalent to taking the local evaporation field F^c as the field at which the local evaporation rate-constant k_n has the numerical value unity in the system of units in use [in practice, $k_n(F^c)$ as 1 $\sec^{-1}$]. It will be seem later that, when applied to appropriate sites, this requirement corresponds reasonably with empirical definitions of evaporation field.

Using the above definition gives:

$$\psi(F) = (F - F^c)\, \partial\psi/\partial F\,|_{F^c}. \tag{5}$$

And thus, by a trivial algebraic manipulation, eq. (2) may be written in the form:

$$k_n = \exp(-\psi_n/kT) = \exp(\mu^1 f/kT), \tag{6}$$

where

$$f = (F - F^c)/F^c, \tag{7}$$

$$\mu^1 = -F^c\, \partial\psi/\partial F\,|_{F^c} = -\partial\psi/\partial f\,|_{f=0}. \tag{8}$$

f is the fraction by which the actual field at a site exceeds the evaporation field (F^e) for the site, and is termed the *fractional overfield*; it will sometimes be negative. μ^1 is a positive quantity with the dimensions of energy, and will be called the *first partial energy of field evaporation*. The mathematical transformation in eq. (8) is justified in Appendix A.2.

The usefulness of eqs. (6) to (8) is that all have very simple forms, yet the definition of μ^1 does not depend on the details of any particular model of field evaporation. [The actual numerical values of F^e, f, and μ^1 will depend slightly on the definition of evaporation field used, but to first order the exponent in eq. (6) has the same numerical value whatever definition is used.] The experimental relevance of eqs. (6) to (8) is that, in all but the most accurate experiments, the results may be interpreted as showing that $\ln k_n$ is a linear function of f: consequently, an experimental estimate of μ^1 is easily derived from the slope of the relevant rate plot.

3. Conversion to P–V form

It is instructive to put the equations of section 2 into a different form. This form is suggested by some recent work of Page and Ralph [14]) in a slightly different context, but the analysis here does not depend on the validity of their argument.

A charged surface experiences an outwards force due to the charge. If the field on the surface is F, then the "equivalent pressure" (outwards) is given by:

$$P = F^2/8\pi. \tag{9}$$

The mathematical development in section 2 can be performed in terms of P rather than F. Thus eq. (5) is replaced by:

$$\psi(P) = (P - P^e)\, \partial\psi/\partial P\,|_{P^e}. \tag{10}$$

And eq. (6) can be written in the form:

$$k_n = \exp[V^e(P - P^e)/kT], \tag{11}$$

where V^e is a positive quantity with the dimensions of volume, defined by:

$$V^e = -\,\partial\psi/\partial P\,|_{P^e}, \tag{12}$$

and called the *volume of field evaporation*. As with μ^1, the definition of V^e does not depend upon the details of any particular model of field evaporation, although its numerical value will depend on the particular definition of P^e chosen. In principle, V^e may vary in value from site to site.

The simplicity of these equations is appealing and suggestive. They bring out the point that there is an analogy between the analysis here and that normally used to discuss the diffusion of point defects in strained metals. For example, Girifalco and Welch [15]) give an analysis leading to the point-defect diffusion equation [their eq. (3.27)]:

$$\Gamma(P) = \Gamma(0) \exp(-P\bar{V}^{m}/kT), \tag{13}$$

which gives the pressure dependence of a "jump frequency", in terms of the zero-pressure jump frequency and a "partial molar volume of defect migration", $\bar{V}^{m}$. If this is put into the form:

$$\Gamma(P) = \Gamma(P^{c}) \exp[-\bar{V}^{m}(P - P^{c})/kT], \tag{14}$$

and P^{c} is chosen such that $\Gamma(P^{c}) = 1$, then the resemblance to eq. (11) becomes clearer: $\Gamma(P)$ is analogous to k_{n}; the additional minus sign is because eq. (9) defines an outwards pressure to be positive whereas the usual convention is for an inwards pressure to be positive; V^{c} is analogous to $\bar{V}^{m}$.

Use of P, P^{c}, and V^{c} is convenient for some formal discussions, but F, f, and μ^{1} are the better set for analysing experimental results. The relationships between the two sets of parameters are easily derived: the Taylor expansion leading to eq. (10) is valid when $P \simeq P^{c}$; in such circumstances:

$$\begin{aligned} 8\pi(P - P^{c}) &= F^{2} - (F^{c})^{2} = (F - F^{c})(F + F^{c}) \\ &\simeq (F - F^{c}) \cdot 2F^{c} = f \cdot 2(F^{c})^{2}, \end{aligned} \tag{15}$$

and hence f and μ^{1} are given by:

$$f = \tfrac{1}{2}(P - P^{c})/P^{c}, \tag{16}$$

$$\mu^{1} = 2P^{c}V^{c} = V^{c} \cdot (F^{c})^{2}/4\pi. \tag{17}$$

The significance of partial volumes deserves comment. In point-defect *formation* theory [15]), where one is concerned with the equilibrium concentration of defects, there appears a quantity $\bar{V}^{f}$ defined by:

$$\bar{V}^{f} = \partial \bar{G}^{f}/\partial P, \tag{18}$$

where $\bar{G}^{f}$ (the "free energy of defect formation") is the *change* in the free energy of the system when one *extra* point defect is formed. The status of $\bar{V}^{f}$ as a thermodynamic partial derivative is clear. It is possible [15]), but with less rigour (because migration is basically a thermodynamically irreversible process), to define an analogous quantity $\bar{G}^{m}$ (sometimes written $\Delta\bar{G}^{m}$), the "free energy of defect migration". However the derivative most relevant to point-defect migration is [15]):

$$\bar{V}^{m} = \partial \bar{G}^{m}/\partial P + \partial \ln \Gamma/\partial P. \tag{19}$$

This $\bar{V}^m$ contains information about the pressure dependence of the pre-exponential as well as about the dependence of the free energy: thus the $\bar{V}^m$ of eq. (19) is not strictly a thermodynamic partial derivative. Nevertheless, it plays the same mathematical and general role in migration theory as does $\bar{V}^f$ in point-defect formation theory: both parameters have conceptual existences quite independent of the details of theoretical or experimental ways of estimating them, and so have definite though perhaps unknown values. Part of the usefulness of $\bar{V}^m$ lies in this statistical-mechanical, rather than empirical, status. The V^c of eq. (12) is a parameter of the same kind as $\bar{V}^m$; and, once a convention on the definition of F^c is agreed, μ^1 is a parameter of the same kind again. The higher-order partial derivatives also have statistical-mechanical, rather than empirical, status.

4. Higher-order terms

In eq. (4) the exponent was expanded only to first order. A second-order version of eq. (6) is:

$$k_n = \exp(-\psi/kT) = \exp[(\mu^1 f + \tfrac{1}{2}\mu^2 f^2)/kT], \tag{20}$$

where μ^2, the *second partial energy of field evaporation*, is given by:

$$\mu^2 = -(F^c)^2\,(\partial^2\psi/\partial F^2)\,|_{F^c} = -\partial^2\psi/\partial f^2\,|_0. \tag{21}$$

In the P–V form the analogue of eq. (20) is:

$$k_n = \exp[V^c(P - P^c) + \tfrac{1}{2}(\partial^2\psi/\partial P^2)\,|_{P^c}(P - P^c)^2], \tag{22}$$

and the second-order equivalent to eq. (17) is:

$$\mu^2 = -4(P^c)^2\,(\partial^2\psi/\partial P^2)\,|_{P^c}. \tag{23}$$

Extension to higher orders is straightforward.

Earlier experimental work on evaporation–current field sensitivity[3,8] suggested that, within the field ranges investigated and to within the limits of experimental error, $\ln k_n$ was a linear function of f, and hence that the first-order expansion in eq. (6) is sufficient. However, Tsong's elegant rate-constant measurements[10] show the influence of higher-order terms. His fig. 9 is reproduced as fig. 1 here, but with the coordinates transformed into accordance with the conventions of the present paper. The $\ln k_n$ versus f plot is non-linear; but the $\partial(\ln k_n)/\partial f$ versus f plot is straight.

The theoretical prediction from eq. (20) is:

$$\partial(\ln k_n)/\partial f = (\mu^1/kT) + (\mu^2/kT)\cdot f. \tag{24}$$

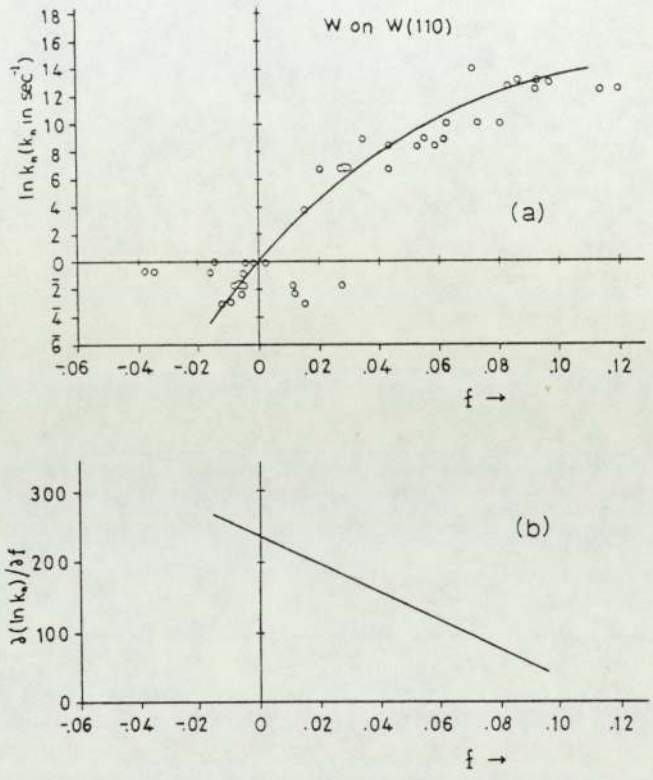


Fig. 1. (a) Evaporation rate-constant for tungsten adatoms on a tungsten (110) plane: to show the variation with the fractional overfield, f. (b) Derivative of best fit line to data in (a). Data taken from Tsong[10]), fig. 9, but plotted against transformed coordinates.

Inclusion of a third-order term in eq. (20) would lead to an f^2 term in eq. (24); what Tsong's fig. 9 shows is that, to within the errors of his data accumulation and analysis, the second-order expansion of ψ is sufficient to describe his experimental results.

Experimental estimates of μ^1 and μ^2 are, of course, easily obtained by multiplying the intercept and slope of the $\partial(\ln k_n)/\partial f$ versus f plot by kT. (Alternatively, by fitting a polynomial to the $\ln k_n$ versus f plot.) In the case illustrated, $\mu^1 = 1.7$ eV, $\mu^2 = -13.7$ eV.

5. The atomic volume approximation

A useful feature of the P–V form of evaporation theory is that it enables a quick order-of-magnitude estimate of V^c and μ^1. Data and formulae given by Girifalco and Welch[15]) suggest that volumes of formation and migration

often have numerical values equal in order of magnitude to the atomic volume. The tentative hypothesis that the same may be true of the volume of field evaporation implies, in the case of tungsten, that:

$$V^c \sim \Omega = 0.0158 \text{ nm}^3. \tag{25}$$

Hence, using eq. (17), a theoretical estimate of the first partial energy for tungsten is obtained: $\mu^1_{th} \sim 2.8$ eV. In eq. (17), F^c has been set equal to the experimental evaporation field listed by Müller and Tsong [1]), namely 57 V/nm (5.7 V/Å). The estimation method was suggested by the work of Page and Ralph [14]), but is not entirely equivalent to theirs.

In table 1 this theoretical estimate is compared with experimental estimates of μ^1 derived from the work of Brandon [8]), Taylor [3]) and Tsong [10]). Each number given corresponds to one set of data in the originals. The sets used are identified in Appendix A.3; hopefully, representative sets have been chosen. Tungsten is taken as the material for theoretical and experimental intercomparison because more experimental work has been done on it than on any other material.

The error limits shown are obtained from a simple common formula for the standard deviation of the mean. They are to be treated with the usual caution due when populations are small.

TABLE 1

Estimates of μ^1 (in eV) for tungsten

Experimental estimates	
Derived from Brandon [8])	2.0, 2.0
Taylor [3])	2.0, 2.3, 2.6
Tsong [10])	1.6, 1.7, 1.8
Mean of experimental estimates	2.0 ± 0.1
Theoretical estimate	2.8

Bearing in mind the obvious crudeness of approximation (25), and the expectation of only order-of-magnitude agreement, the closeness between the experimental estimates and that derived via the atomic volume approximation is surprising.

It is instructive to attempt an *a-priori* estimate of μ^1 with the standard models of field evaporation. For example, according to the image-hump model μ^1 is given by:

$$\mu^1 = \tfrac{1}{2}(n^3 e^3 F^c)^{\frac{1}{2}} - \alpha(F^c)^2. \tag{26}$$

The first problem is whether to take n as equal to 2 [the predicted value[1])] or to 3 or 4 (the values found from atom-probe work); corresponding values of the first term on the r.h.s. in eq. (26) would be 12.8 eV, 23.5 eV, 36 eV. The second difficulty is the derivation of an a-priori estimate of the effective polarizability α. Attempts, by Brandon and by Müller, are summarised in ref. 1: on p. 70 it is concluded that the polarization contribution $\frac{1}{2}\alpha(F^c)^2$ might be 0.59 eV on the (011) plane of tungsten, 2.72 eV on the (111) plane. Suppose we take 4 eV as typical of the value of $\alpha(F^c)^2$: combining this with the above values of the $F^{\pm}$ term gives theoretical estimates for μ^1 ranging between 8 eV and about 30 eV – way above the experimental values.

Luckily for the image-hump model, the cited estimates of α do not have to be taken too seriously, and the discrepancy is normally resolved by supposing that the method of estimation is inadequate and that α can in effect be treated as an adjustable parameter. However, this exercise does show that, in comparison with a-priori reasoning based on the image-hump model (and a-priori reasoning based on the charge-exchange model is not appreciably more successful), the atomic volume approximation gives a remarkably good answer.

This apparent success with tungsten may, of course, be a numerical coincidence. A further test would be the prediction of trends in the variation of μ^1 from metal to metal. Table 2 lists theoretical estimates (μ^1_{th}) for a number of metals for which Tsong and Müller[1]) give (though "with great reservations") experimental values of evaporation field. Experimental partial-energy estimates shown (μ^1_{exp}) are derived from Brandon's results in Appendix A.3; the first figure refers to work near 63 K, the second to work near 88 K. The

TABLE 2

Estimates of first partial energy

Material	F^c (V/nm)	Ω (nm^3)	$\mu^1{}_{th}$ (eV)	$\mu^1{}_{exp}$ (eV)	Ratio
Fe	36	0.01177	0.84		
Co	37	0.01113	0.84		
Ni	36	0.01094	0.78		
Cu	30	0.01181	0.59		
Mo	45	0.01558	1.7	1.4, 2.1	0.82, 1.2
Ru	45	0.01357	1.5		
W	57	0.01585	2.8	2.0, 2.0	0.71, 0.71
Re	48	0.01470	1.9		
Ir	50	0.01414	1.9		
Pt	47.5	0.01510	1.9	1.3, 1.3	0.68, 0.69
Au	35	0.01696	1.1		

final column gives the ratio of these to the corresponding theoretical estimate.

Inspection shows that:

(1) For platinum and molybdenum, as well as for tungsten, Brandon's experiments give experimental partial-energy values within a factor of 1.5 of the theoretical result. This provides some empirical justification for Page and Ralph's numerical approximation in the case of iridium [13, 14]).

(2) The general characteristic that μ^1 does not vary much as between different materials is true of both experimental and theoretical data.

(3) The μ^1_{exp}/μ^1_{th} ratios are very similar for tungsten and platinum, but the molybdenum ratios are discrepant. It is well known that the molybdenum field-evaporation endform is irregular near 80 °K, but tends to be less irregular at lower evaporation temperatures; there might be some connection.

However, the table's most striking feature is the paucity of experimental data: there is not really enough of it to make a test of the atomic-volume approximation sensible: judgement needs be reserved pending further experiments. I believe the numerical relationships noted in this section are sufficiently provocative to make additional experiments desirable and potentially valuable. Careful measurements are required both of evaporation fields and of field sensitivites, preferably under standardised conditions (cf. Appendix A.5).

The atomic volume approximation is, of course, no substitute for a detailed model of field evaporation, but merely a zero-order approximation. Yet some explanation of its apparent usefulness would have to emerge from the details of the correct atomistic model of field evaporation: hopefully, this is another theoretical handle on the field evaporation problem.

6. Discussion

The rest of this paper considers consequences of the basic theory now set out. The next two sections suggest that my method of data analysis has distinctive properties that will make it a useful supplement to existing methods. The final sections show that the present approach brings into greater prominence some basic questions, on which further research is needed.

6.1. The reporting of data

Earlier sections have shown that, with an appropriate definition of F^c, plotting results in the $\ln k_n$ versus f form makes the determination of μ^1 very simple and the determination of μ^2 only a little less so. The statistical-mechanical status of these parameters makes them useful to derive.

There are advantages in using f, rather than F or F/F^c, as the coordinate for the x-axis. A dimensionless parameter is better if (as is usually the case)

the field values are not precisely known. And f is to be preferred over F/F^c, firstly because differentiation with respect to f is effectively equivalent to the commonly-cited differentiation with respect to $\ln F$ (see Appendix A.2), and second because our interest is in the field sensitivity near $F=F^c$ ($f=0$) and it is convenient to have the origin of coordinates near the data points.

The particular condition on F^c chosen, which may be written in any of the various forms: $\psi(F^c)=0$, $\ln k_n(F^c)=0$, $\log k_n(F^c)=0$, $k_n(F^c)=1\ \mathrm{sec}^{-1}$, has several advantages. It makes the algebraic analysis particularly simple; it facilitates the transformation of data plotted in related forms: and it more-or-less corresponds with empirical criteria for evaporation onset (cf. Appendix A.5).

The form of this evaporation-onset criterion is itself new, in that it involves a requirement on rate-constant value, rather than on activation energy or average evaporation current (cf. Appendices A.3–A.5). This is particularly useful in rate-constant experiments: there is now a sharply-defined criterion on evaporation onset, and no uncertainty through lack of knowledge of the pre-exponential; and the position of evaporation onset can be determined from the measurements without precise knowledge of field values.

The general value of the rate-constant form of criterion is that it *compartmentalises* field-evaporation problems: this should help discussion both of field sensitivity variations and of field-evaporation endforms.

Now that field-evaporation experiments seem to be entering an era of greater precision, it would be useful to have a convention about the definition of evaporation field and the reporting of rate-sensitivity measurements. This is illustrated by table 3, which shows numbers that might be claimed to be the field sensitivity $F\,\hat{c}\,(\ln k_n)/\partial F$ at the evaporation field. Two data sets have been used, namely those shown in Tsong's[10]) Figs 9 and 10, and two different evaporation onset criteria. The left-hand column employs my "internal" criterion that the evaporation rate-constant for the adatoms be $1\ \mathrm{sec}^{-1}$; the right-hand column employs an empirical criterion, that the average evaporation current for the substrate be 0.01 layer/sec. The data set of fig. 10 refers to a larger radius specimen than does that of fig. 9.

TABLE 3

Field sensitivities

Data set	Rate-constant criterion	Empirical criterion
Ref. 10, fig. 9	241	184
Ref. 10, fig. 10	259	161
(Ratio)	(1.07)	(0.87)

The different criteria give different results; this demonstrates that they are not equivalent, which is not overly surprising. Further (though it is difficult to be sure, without analysing further data sets, that the effect is significant), the different criteria seem to behave differently as between data sets, in that set 9 has the higher field sensitivity on one criterion but set 10 on the other; the rate-constant criterion seems to give slightly the more consistent result.

The table underlines the need to have field evaporation onset criteria clearly stated, and their interrelationships clearly understood. Three different forms of criterion are examined in Appendix A.5.

Generally, if different workers use the same defining criterion, and if this criterion is an internally consistent one, then their experimental results should be strictly intercomparable; failure to use the same, consistent, criterion would lead to discrepancies both in μ^1 values and (presumably) in the values of derived parameters – particularly since, whatever the onset criterion used, further analysis of the data would at present probably have to assume that the corresponding evaporation field be 57 V/nm.

This question of onset criteria aside, there is clearly some similarity between my data-plotting method (derived from formal theoretical analysis) and one of those devised by Tsong[10]) for handling his experimental data; in fact, this part of my approach could be seen as a formalisation and generalization of his procedure. Taking this into account, it seems possible that the methods outlined in the present paper might form a suitable basis for a convention about handling field evaporation data.

6.2. The fitting of models

The actual fitting of models to reported data plots would in the Taylor expansion approach be a two-stage process. First, regression analysis gives μ values of several orders. Then these are related to model coefficients [essentially the same as the B-coefficients used by Vesely and Ehrlich[5])] via equations derived from definitions such as (8) and (21), and the equations are solved for the coefficients. Coefficient values thus derived can be compared with those predicted from the models. It is intended to describe this procedure in more detail in a later paper[13]).

This fitting procedure is more complex than Vesely and Ehrlich's straight regression fitting of a polynomial in F, and it could be more troublesome to assess goodness of match between model and data. However, there do seem possible advantages to my procedure, in that it would be possible to statistically examine internal characteristics of the data without reference to any detailed evaporation model, and without needing to know the numerical value of the relevant evaporation field. It might, for example, be possible to gain additional information about some of the points raised in Vesely and

Ehrlich's work, by testing how goodness of fit depends on the order of the polynomial in f, and by examining whether significant values of μ^3 could consistently be obtained from data plots.

If significant values of μ^3 could be obtained, these might help to discriminate between the image-hump and charge-exchange models (or, at least, between the model expressions usually quoted), because the model predictions are significantly different:

$$\text{Charge-exchange model:} \quad \mu^3 = 0, \tag{27a}$$

$$\text{Image-hump model:} \quad \mu^3 = \tfrac{3}{8}(n^3 e^3 F^c)^{\frac{1}{2}}. \tag{27b}$$

In general terms, it is felt that the fitting procedure suggested here is likely to supplement rather than replace that of ref. 5.

6.3. The determination of evaporation field

To derive experimental estimates of partial energies, the only assumption necessary is that the local field F_n is proportional to the applied voltage. However, comparison with model predictions requires a value for the local evaporation field F_n^c, since this will appear in model expressions for partial energies (or derived quantities). Ideally, one would like to determine F_n^c directly by independent means (for example, by monitoring field emission currents through a probe hole); in practice this might often be difficult to arrange, and one would need to fall back on values listed by Müller and Tsong for measured evaporation fields[1]).

Awkward questions now arise:

(i) What criterion has in practice been used to define evaporation onset in the experiments that gave rise to the list values?

(ii) Even if the criterion is consistent, how reliable is the method of field measurement?

(iii) How do the F_n^c values relate to the quantity that has been measured?

The problems raised by these questions are highly involved, and will not be considered here. The general conclusion is that precise estimates of F_n^c values are not at present available, and are not likely to be for some time. Because of this, the earlier part of the paper took the various F^c values all as equal to the list value.

Nevertheless, there is an expectation (see Appendix A.5) that the local evaporation field (F_{hr}^c) relevant to Brandon-type measurements should be somewhat higher than that (F_{ad}^c) relevant to Tsong's adatom removal experiments. If the volume of field evaporation is much the same at each site (which is necessarily true in the atomic volume approximation), then eq. (17) predicts that the first partial energy (μ_{hr}^1) relevant to measurements on end-

form evaporation should be higher than that (μ^1_{ad}) relevant to adatom removal.

Interestingly, there is some evidence for this in table 1. Reference to Appendix A.3 shows that the estimates in the table can be divided into three classes: (A) those based on the Brandon and Taylor experiments (5 members, mean 2.2 ± 0.1 eV); (B) the estimate (1.6 eV) from Tsong's endform evaporation experiment; (C) those based on Tsong's adatom removal experiments (2 members, mean 1.75 ± 0.05 eV). The estimates in class (A) are all higher than those in class (C), and the means are significantly different, as predicted. Back-applying eq. (17) would give F^c_{hr} as higher than F^c_{ad} by about 10% – which seems not unreasonable.

However, the single estimate in class (B) does not quite fit in. Expectation is that it should be consistent with the class (A) estimates. In fact, statistically it is not really significantly different from the class (A) estimates; but on a common-sense view it is strange that it should be lower than the class (C) estimates when all the class (A) estimates are higher. There is no obvious explanation. The seeming inconsistency is clearly not due to error in the formal theory: any theory which predicted a difference between μ^1_{ad} and μ^1_{hr} would have to explain the inconsistency [and any theory which predicted *no* difference would have to explain a larger inconsistency between class (A) and classes (B) and (C) taken together]. The seeming inconsistency may just be a statistical fluctuation [its inclusion with class (A) would give a revised value for μ^1_{hr} as 2.1 ± 0.1 eV, which is still significantly higher than the class (C) mean]. Or it may be due to some overlooked assumption in my method of handling the data of ref. 10, fig. 7. Or it could be due to something in the way Tsong performed his endform evaporation experiments; for example, to get the necessary range of current values, he used a combination of pulse techniques and steady evaporation. Probably further experiments will be necessary to resolve the matter.

Thus, at present it may be unwise to use variations in the partial energy estimates to deduce corresponding variations in the local evaporation field. Theoretically this should be possible; but to get significant results many data sets may need to be acquired.

The general question, of how the evaporation field used in some numerical or algebraic analysis relates to some listed value, of course occurs (though in different ways) whatever method is used to handle the data. My setting each F^c equal to 57 V/nm, Tsong's setting the F_0 in his equations equal to 57 V/nm, and Vesely and Ehrlich's plotting out of data against F (How was the F-scale decided?) are merely three different ways of ignoring some nasty problems.

Luckily, in many contexts of current interest these problems may safely

be ignored. For example, the general degree of agreement between the experimental estimates of μ^1 and the result from the atomic volume approximation would not be significantly altered by an error of 10% in the value of F^c in eq. (17). Nor would test criteria such as relationships (27) be much affected by uncertainty in F^c. But the problems will become more urgent in more precise work – for instance if discrepancies of the size exhibited in table 3 would be an embarrassment.

Clearly, there will be an increasing need for careful attention to be given to the definition, interrelation, and realization of evaporation field criteria. In particular, it might be helpful to standardise the conditions (temperature, gas nature and pressure, tip size and shape, and onset criterion used) under which, on the one hand field evaporation experiments, and on the other hand experimental measurements of evaporation field, are performed.

6.4. The field evaporation pre-exponential

A distinguishing feature of our analysis is that it is based on the "inclusive" form, eq. (2), rather than on the Arrhenius form, eq. (1). The Arrhenius form was thought inappropriate because in principle A_n represents a *temperature*-independent quantity obtainable from the intercept in an Arrhenius plot, whereas our interest is in field sensitivity and the data analysis does not involve Arrhenius plots.

In deriving expressions for μ^1, μ^2, μ^3, we have followed the practice of much of past field evaporation theory (particularly when dealing with evaporation near 80 K or above), and have concentrated on the field sensitivity of Q. However, the pre-exponential A may in principle vary significantly with F (or f), and this variation could be important when $\ln A$ is comparable with or greater than Q/kT (which is the case near F^c). It is easy to forget this with a rate-constant equation in form (1).

Physically, A contains a vibration-frequency-like component, an entropy-like component, and components which have their origins in density-of-states effects, electron-wave-behavioural effects, and nuclear wave-behavioural effects. The need to take the latter three effects into account when discussing the field sensitivity of field evaporation at low temperatures (less than about 60 K) is well known [1]). However, previous work has not pointed out that there might also be some field dependence in the vibration-frequency-like component of A.

In a general physical sense, field evaporation is a transition resembling phase transitions or mechanical transitions such as the flexed beam transition described by Pippard [16]): and there is a general physical principle that as a transition is approached the relevant vibration frequency of the system tends to slow down. In the context of field evaporation the argument is that, as

the point of evaporation is approached, the potential well in which the pre-evaporating nucleus moves gets progressively flatter and shallower, and consequently nuclear motion gets ever more sluggish.

Unfortunately, it is rather difficult to reliably estimate the magnitude of this effect, but a simple numerical example can show that the effect is not necessarily insignificant. Consider eq. (28), which can be derived from equations given elsewhere in this paper:

$$\partial(\ln k_n)/\partial f = \partial(\ln A)/\partial f - (1/kT)\,\partial Q/\partial f. \tag{28}$$

Measured values of $\partial(\ln k_n)/\partial f$ are of order 200. Suppose that in the course of a one-percent field change ($\delta f = 0.01$) near field evaporation the value of A decreased from 2.3×10^{13} to 1.0×10^{13}: this would make $\partial \ln A)/\partial f$ equal to -100. So ignoring the f-variation of A would lead to an error of 50% in the derived estimate of $\partial Q/\partial f$.

Further, in principle one expects $\partial^2(\ln A)/\partial f^2$ to be negative, and relatively large, near a transition. Such a term would make some contribution to μ^2, and thus could help to explain the incipient turning over the $\ln k_n$ versus f plot, as found experimentally by Tsong[10]).

Although the numerical change in A chosen above might seem a little on the large side, it is not easy to demonstrate it unreasonable; particularly since, according to McKinstry[4]), estimates of A derived from experimental data by Brandon's method[4, 8]) vary as between data sets by two or more orders of magnitude. It might be wise for field evaporation theory not to disregard possible variations of A with f near 80 K and above.

6.5. A SPECULATIVE HYPOTHESIS CONCERNING FIELD EVAPORATION MECHANISM

If the atomic volume approximation works, then some explanation will have to emerge from the details of the correct atomistic model of field evaporation. Since evaporation theory seems unsettled at present, some speculation may be justifiable.

Possibly the agreement between theory and experiment shown in section 5 is a numerical coincidence, or possibly there is some hidden connection between the atomic volume approximation and one or both of the established models for field evaporation.

However, it could alternatively be that the rate-controlling process does indeed have a V^c value approximated by the atomic volume, but that the rate-controlling mechanism is not that assumed by either of the existing models. For example, one could hypothesise that field evaporation is usually a several-stage process involving the following elementary steps:

(i) migration of the pre-evaporating atom, in a neutral or fractionally-charged state, to a new more field-exposed position;

(ii) draining of its electrons into the metal by means of a sequence of electronic transitions;
(iii) detachment of the ion in a multiply-charged state;
– with the first step being rate-controlling at fields near F^c. It is well known that the presence of an electric field does influence surface migration processes [17–19]).

Such a hypothesis might go some way towards explaining why an estimation method for V^c essentially similar to that used in point-defect migration theory works. And it would, for example, be compatible with the phenomenon [20]) that it is sometimes necessary to "aim-off" in atom-probe work in order to catch the atom responsible for a given image spot.

Clearly, any new hypothesis would need to be compatible with the existing achievements of field evaporation theory, in that evaporation fields can be estimated fairly well, and the relative abundances of ions in different charge states (and the variation with temperature) can be explained, at least qualitatively. Yet, on the face of it, it seems possible that relative abundances might be explained in terms of branching processes occurring in the second and third stages; and that the question of whether field evaporation occurs at all might depend on whether the second stage could occur, the theory of this being much the same as the existing models of field evaporation.

This multiple-stage hypothesis must remain extremely speculative until such time as it can be explored in adequate theoretical depth – which is no simple matter and beyond the scope of this paper.

7. Summary

It has been shown that Taylor expansion of the rate-theory exponent about a suitably defined evaporation field, and the introduction of the variable f (the fractional overfield), give rise to an empirically convenient way of analysing field sensitivity data in terms of parameters whose significance is underwritten by statistical-mechanical arguments.

Various submerged problems and possibilities of field evaporation theory have been exposed through this Taylor expansion approach; several deserve deeper investigation is work on field evaporation is to become generally more precise.

It has also been shown that in addition to the existing models of field evaporation there is, at least on the face of it, a substantially different way of accounting for the experimental data; detailed exploration of this would be an alternative general line of attack on present theoretical difficulties.

Perhaps ultimately most useful, it has been demonstrated that field evaporation may be regarded as a migration process generally similar to point-

defect migration in strained metals. The theory of such migration processes is well understood for the interior of metals, and thus should provide a fruitful source of analogies for field-ion theory, particularly when the time comes to develop a many-body theory of field evaporation.

Finally, one may concur with Tsong and Müller[9]) in urging that the experimental basis of field evaporation be broadened. I hope this paper has provided some stimuli.

Appendices

A.1. Dimensional consistency in equations involving logarithms and exponents

In handling equations such as eq. (3) an apparent difficulty arises, in that the l.h.s of the equation clearly has the dimension $\sec^{-1}$, whereas the r.h.s. seems dimensionless. Formally, this difficulty may be overcome as follows: The quantity A_n in eq. (1) is written as the product of a pure number $|A_n|$, equal in size to A_n, and a dimension-carrying constant $\hat{a}$ of numerical value unity in the system of dimensions in use. (For this paper, $\hat{a} = 1\ \sec^{-1}$.) Thus:

$$A_n = |A_n| \cdot \hat{a}. \tag{29}$$

Substituting into eq. (1) gives:

$$k_n = \hat{a} \cdot |A_n| \exp(-Q_n/kT). \tag{30}$$

Revised versions of equations (2) and (3) may then be written:

$$k_n = \hat{a} \exp(-\psi_n/kT), \tag{31}$$

$$\psi_n = Q_n - kT \ln |A_n|. \tag{32}$$

And a logarithmic version of equation (31) may be written:

$$\ln(k_n/\hat{a}) = -\psi_n/kT. \tag{33}$$

In all the above equations both sides have the same physical dimensions, and the arguments of all exponents and logarithms are dimensionless, as is logically necessary. Provided that one works consistently in a single set of units as regards time, then the constant $\hat{a}$ always has the numerical value unity, and there is an unwritten convention that it is omitted in all formulae. The only situation giving rise to difficulty is if the size of the unit of time is changed in the middle of a numerical calculation: for example, a change from seconds to minutes in the middle of a physical or numerical argument would

require $\hat{a}$ in eq. (30) to take the value 60 $\min^{-1}$, and the criterion on evaporation field to be written $k_n(F_n^e) = 60\ \min^{-1}$.

A.2. Some mathematical relationships

Field sensitivities in the literature are written in various mathematical forms. The relationships between them are demonstrated below.

From eq. (7), f is defined by:

$$f = (F - F^e)/F^e = (F/F^e) - 1,$$

where F^e is considered a constant and F a variable. Clearly:

$$\delta f = \delta(F/F^e) = \delta F/F^e = \delta V/V^e, \tag{34}$$

where V denotes voltage. Further:

$$\delta(\ln F) = \delta F/F \simeq \delta F/F^e, \tag{35}$$

the last relationship holding best for values of F near F^e. Thus:

$$\frac{\partial(\ln k_n)}{\partial f} = F^e \frac{\partial(\ln k_n)}{\partial F} \simeq \frac{\partial(\ln k_n)}{\partial(\ln F)}, \tag{36}$$

the last relationship again holding best for F near F^e. At $F = F^e$ it holds exactly:

$$\frac{\partial(\ln k_n)}{\partial f}\Big|_{f=0} = \frac{\partial(\ln k_n)}{\partial(\ln F)}\Big|_{F=F^e}. \tag{37}$$

Future field evaporation experiments will probably explore the variation of field sensitivity with field (or fractional overfield, f). Since the quantities $\partial(\ln k_n)/\partial \ln F$ and $\partial(\ln k_n)/\partial f$ vary in different ways as functions of F, and since theoretical analysis has made it clear that the latter is the more useful quantity to have, it would perhaps be useful if derivatives of the form $\partial(\ln k_n)/\partial \ln F$ ceased to be cited and experimental results were plotted on semi-logarithmic graphs.

A.3. The origins of experimental data

The experimental estimates of partial energies appearing in the tables are obtained by applying formula (38) below to the measured field sensitivities (cf. section 4):

$$\mu^1 = kT\ \partial(\ln k_n)/\partial f\Big|_{f=0}. \tag{38}$$

The data selected from the literature are identified below. For Brandon's and Taylor's results, the columns from left to right give: material; temperature of specimen holder; gas pressure and species; field sensitivity; partial energy.

(I) Brandon's results are taken from tables 2, 3, and 4 in ref. 8:

Tungsten:	62 K	2 mtorr He	378 ± 29	2.02 ± 0.15 eV,
	88 K	3 mtorr He	267 ± 26	2.02 ± 0.2 eV;
Molybdenum:	64 K	2 mtorr He	260 ± 61	1.43 ± 0.3 eV,
	88 K	2 mtorr He	275 ± 75	2.09 ± 0.6 eV;
Platinum:	63 K	2 mtorr He	236 ± 16	1.28 ± 0.09 eV,
	89 K	4 mtorr He	169 ± 25	1.29 ± 0.2 eV.

(II) Taylor's data are taken from table 3.VI on page 175 of his Doctoral thesis[3]). He gives data for four temperatures and five different gas pressures. The field sensitivities at 37 K are substantially lower (near 250) than those at the higher temperatures, and are disregarded here because field-adsorbed helium and/or nuclear-wave-behaviour may affect the evaporation process significantly at this temperature. His data also exhibit a slight pressure dependence, but all sensitivities at a given temperature lie within 10% of the figure below, which is that for the lowest gas pressure Taylor used:

Tungsten:	65 K	0.19 mtorr He	350 ± 19	1.96 ± 0.1 eV,
	80 K	0.19 mtorr He	332 ± 22	2.27 ± 0.15 eV,
	90 K	0.19 mtorr He	338 ± 17	2.62 ± 0.1 eV.

(III) Tsong's data are taken from figs. 7, 9, and 10 in ref. 10. To handle fig. 7 it is assumed that the average evaporation current at $F = F_0$ was 0.01 layer/sec: Tsong's and my criteria for evaporation onset should then nearly coincide (cf. Appendix A.5), and it then follows that:

$$\partial(\ln k_n)/\partial f\,|_{f=0} = 2.303\,[\log k_e(F)/k_e(F_0)/(F/F_0 - 1)]\,|_{F=F_0}. \quad (39)$$

The term in square brackets is written in Tsong's notation, and is plotted by him in his fig. 7. The result for μ^1 is 1.6 eV.

For figs. 9 and 10, Tsong's and my criteria for evaporation onset do not coincide, and the field sensitivity relevant to my theory has to be obtained from the transformation equation[13]):

$$\partial(\ln k_n)/\partial f\,|_{f=0} = 2.303\,\gamma\,[\partial \log\kappa/\partial(F/F_0)]\,|_{F=\gamma F_0}. \quad (40)$$

The term in square brackets is written in Tsong's notation, and is plotted in his figs. 9 and 10. γ is a number obtained from the requirement that $\log\kappa = 0$ at $F = \gamma F_0$. Values of γ measured from his figures are 0.97 for fig. 9, and 0.955 for fig. 10. Numerical values of his derivative at $F = \gamma F_0$ are obtained using his quoted slope and intercept values. The results for μ^1 are 1.7 eV and 1.8 eV. All measurements are performed at liquid nitrogen temperature, and with 1.5 mtorr He present.

A.4. Nomenclature

Tsong and Müller's discussion of field evaporation theory uses a slightly different nomenclature from Brandon's and Taylor's. Since neither seems fully satisfactory, an alternative is set out below and used in this paper.

In some situation, if at any time there are m atoms at risk of evaporation, and the evaporation of each atom is a random process with rate-constant k_n, then the number (j) of atoms evaporated per unit time is given by:

$$j = mk_n. \tag{41}$$

This paper calls k_n an *evaporation rate-constant*, m the *population at risk*, and j the *evaporation current*. k_n is measured in sec^{-1}; m in *atoms* (or sometimes in *layers* of atoms); and j in *atoms/sec* (or sometimes in *layers/sec*). Note that there is a sense in which k_n and j have different physical dimensions.

Early literature tends to ignore the distinction between k_n and j, and to call both "evaporation rate". Thus in Brandon's work the quantity k_e that appears in formulae is a rate-constant, but the measured quantity is an evaporation current. No theoretical difficulty arises there because the measurements are always reported in logarithmic form: from the arguments in Appendix A.5 it is clear that the logarithmic derivatives of j and the appropriate k_n will be almost equal, numerically. Nonetheless, in more precise work a distinction is mandatory.

Table 4 shows the names and symbols used by the different workers; the upper line gives the quantity measured in sec^{-1}, the lower line the quantity measured in atoms/sec or in layers/sec.

Table 4

Forbes	Brandon	Taylor	Tsong/Müller
Evaporation rate constant	Evaporation rate	Evaporation rate	Evaporation rate and absolute evaporation rate
k_n	k_e	k_e	κ_e
Evaporation current	Evaporation rate	Measured evaporation rate	Relative evaporation rate
j			k_e

Several points deserve notice:

(1) The use of the terms "rate-constant" and "current" for quantities measured in sec^{-1} and atoms/sec, respectively, is consistent with usage in field-

ion imaging theory, where the analogous distinction needs to be made be-between an ionization rate-constant and an ionization current[21]).
(2) Tsong and Müller use the symbol k for a current, whereas all the other authors (certainly in their theory) use k for a rate-constant. It is a common chemical convention to denote a rate-constant by k.
(3) This paper uses the subscript on k to denote the site (or type of site) to which the evaporation rate-constant is relevant; the other authors use a subscript e, to denote evaporation, and, perhaps, to distinguish the rate-constant from Boltzmann's constant.
(4) Tsong and Müller use the name "Relative Evaporation Rate" for the quantity measured in layers/second for reasons given in ref. 9. Unfortunately, their "absolute/relative" terminology could be misleading, because a result of 0.01 layer/sec could reasonably be called an absolute measurement of an evaporation current.

The advantage claimed for the present author's system is that it is clearer, and most in line with other scientific usage.

A.5. The specification of evaporation field

Experiments on the determination of evaporation fields must include two logical elements: some empirical criterion to identify the onset of evaporation, and the measurement of some field. A theory of field evaporation must include some theoretical definition of evaporation field. Those who seek to assimilate measurement and theory face two questions: (a) Does the empirical criterion for evaporation onset correspond to the theoretical criterion? (b) If so, do the experiments in fact reliably measure the field quantity that appears in the theory? This appendix deals with the first question.

We consider three methods for specifying evaporation field. It may be defined as:
(1) The field (F^0) at some point when the average evaporation current is equal to some specified value;
(2) The field (F^c) at some site when the local rate-constant for that site has some specified value;
(3) The field (F^e) at some site when the local activation energy for that site is equal to zero.
Methods (1) and (3) are in common use; method (2) is a new one, introduced in this paper.

The experimental situation in which field evaporation is just beginning to occur is here termed *evaporation onset*. In many circumstances evaporation onset would be judged qualitatively, but it is possible to use a numerical criterion in the form of "so-many layers per second". The empirical criterion is thus in terms of an average evaporation current. It can be related to

theoretical criteria derived from rate-constant theory by the following, slightly simplistic argument. For definiteness, we use terms applicable to tungsten of the normal orientation.

With a top (110) plane of average size, only a small proportion of the atoms in that plane are at any significant risk of evaporation: such atoms are said to be at *high-risk* sites. Suppose that on average (an average over all possible sizes of the top (110) plane) the *proportion* of high-risk sites is r, and the average evaporation rate-constant for these sites is k_{hr}: from eq. (41), the evaporation current j (measured in layers/sec) is given by:

$$j = rk_{hr}. \tag{42}$$

If, as the specimen evaporates and is blunted, the applied voltage is continuously increased in such a manner as to maintain k_{hr} approximately constant, then eq. (42) gives the average evaporation current.

Next, one has to guess a suitable value for r: 1% seems reasonable, but so perhaps would 5% or 0.2%. (50% would seem unreasonably high, 0.02% somewhat low.) Taking $r = 0.01$ would imply equivalence between the rate-constant criterion $k_{hr} = 1\ \text{sec}^{-1}$ and the criterion that average evaporation current equal 0.01 layer/second. By a convenient coincidence, the latter seems a reasonable quantification of the condition that most experimentalists would recognise as evaporation onset; Tsong[10]) has used a current of 0.01 layer/second to define the evaporation voltage of his substrate.

The approximate equivalence of methods (1) and (2), claimed in section 6.1, is thus established. Evaporation theory has also used method (3) to define evaporation field. Applying eq. (1), the pre-exponential lies between about 10^8 and 10^{13} per second[4]); setting Q_{hr} equal to zero gives k_{hr} the same value as the pre-exponential, so the corresponding evaporation current criterion would be between about 10^6 and 10^{11} layers/sec. This is much higher than any commonly-used empirical criterion, even allowing for our estimate of r to be too high by a factor of 50. Thus the unity rate-constant requirement is a better match to the empirical onset criterion than the $Q = 0$ requirement.

The following paragraphs deal with some subsidiary matters.

A.5.1. *The effect of tip size on the relationship between criteria*

It seems likely that the proportion (r) of high-risk sites in the top (110) plane might depend somewhat on the size and shape of the specimen. A specific requirement on rate-constant at such sites would then correspond to a range of evaporation-current criteria, the current value depending on tip size and shape. Conversely, a specified evaporation-current requirement would correspond to a range of rate-constant requirements, and hence to a range of evaporation-field values. It may be relevant that Müller and Young[22]) long

ago noted that small tip size led to a reduction in the evaporation field as defined by some empirical test of evaporation onset.

A.5.2. *Brandon's experiments*

Experiments of the type performed by Brandon and by Taylor measure the field sensitivity of average evaporation current. The relation of this to the field sensitivity of a rate-constant is easily derived. From eq. (42):

$$\ln j = \ln r + \ln k_{\mathrm{hr}},$$

whence

$$\partial(\ln j)/\partial \ln F = \partial(\ln r)/\partial \ln F + \partial(\ln k_{\mathrm{hr}})/\partial \ln F. \tag{43}$$

r is mainly determined by geometrical factors, which in a particular measurement will not depend much on the applied-voltage value, so the first term on the r.h.s. will be almost zero, and the logarithmic derivatives of k_{hr} and j will be almost equal. Thus the experiments in effect give a rate-constant field sensitivity. Note, though, that it is the rate-constant *at the high-risk sites* that appears. All past theory concerned with the evaporation of equilibrium endforms must be deemed to have been dealing with the parameters at the high-risk sites on the relevant topmost plane.

A.5.3. *The variation of evaporation field with position*

Tsong's main experiments[10]) concern adatoms on the top of the the topmost (110) plane. The adatom binding sites are in a different geometrical crystallographic environment from the kink sites in the edges of the top (110) plane. Since the adatoms will have fewer nearest neighbours, expectation is that the adatom binding energy will be less, and hence that the local evaporation field $F^{\mathrm{c}}_{\mathrm{ad}}$ for the adatoms will be less than the local evaporation field $F^{\mathrm{c}}_{\mathrm{hr}}$ for the high-risk sites. It follows that the partial energy $(\mu^{\mathrm{l}}_{\mathrm{ad}})$ for adatom removal would be less than that $(\mu^{\mathrm{l}}_{\mathrm{hr}})$ for experiments on endform evaporation.

As a first approximation, section 4 has ignored these differences, and has taken all the evaporation fields as equal to the listed value, 57 V/nm. Tsong's analysis in fact makes a similar approximation: the F_0 in his equations describing the rate-constant measurements is my $F^{\mathrm{c}}_{\mathrm{ad}}$; but the F_0 in his figs. 9 and 10 is my $F^{\mathrm{c}}_{\mathrm{hr}}$, or something very like it.

It is important to realise that we are not here discussing how the field varies across the surface at some applied voltage corresponding to evaporation onset, but how the local evaporation field F^{c}_{n} as defined by the unity rate-constant criterion varies from site to site. Different sites may reach their local evaporation field F^{c}_{n} at different applied voltages.

Note added in proof

At the 1973 Field Emission Symposium, and in section 6.5, I have speculated on general theoretical grounds that field evaporation might proceed via a multiple-stage process involving, first diffusion to a more field-exposed position, and then detachment. Experimental evidence available since submission of this paper now points to the same conclusion. Moore and Spink [23]) have analysed the relative field evaporation probabilities of atoms in various nearly-equivalent net-plane-edge sites, and conclude that in many cases a high relative evaporation probability is correlated with the presence of a diffusion path away from the site passing through a position of relatively pronounced geometrical protrusion.

References

1) E. W. Müller and T. T. Tsong, *Field Ion Microscopy* (American Elsevier New York, 1969).
2) E. W. Plummer and T. N. Rhodin, J. Chem. Phys. **49** (1968) 3479.
3) D. M. Taylor, Ph.D. Dissertation, Univ. of Cambridge (1970).
4) D. McKinstry, Surface Sci. **29** (1972) 37.
5) M. Vesely and G. Ehrlich, Surface Sci. **34** (1973) 547.
6) E. W. Müller, Phys. Rev. **102** (1956) 618.
7) R. Gomer, J. Chem. Phys. **31** (1959) 341.
8) D. G. Brandon, Phil. Mag. **14** (1966) 803.
9) T. T. Tsong and E. W. Müller, Phys. Status Solidi (a) **1** (1970) 513.
10) T. T. Tsong, J. Chem. Phys. **54** (1971) 4205.
11) S. S. Brenner and J. T. McKinney, Appl. Phys. Letters **13** (1968) 29.
12) R. S. Chambers, G. Ehrlich and M. Vesely, in: *17th Field Emission Symp.*, Yale University, 1970.
13) R. G. Forbes, manuscripts in preparation.
14) T. F. Page and B. Ralph, Surface Sci. **36** (1973) 9.
15) L. A. Girifalco and D. O. Welch, *Point Defects and Diffusion in Strained Metals* (Gordon and Breach, London, 1967).
16) A. B. Pippard, *Reconciling Physics with Reality* (Cambridge Univ. Press, Cambridge, 1972).
17) J. A. Becker, Bell System Tech. J. **30** (1951) 907.
18) M. Drechsler, Z. Electrochem. **61** (1957) 48.
19) J. P. Barbour, F. M. Charbonnier, W. W. Dolan, W. P. Dyke, E. E. Martin and J. K. Trolan, Phys. Rev. **117** (1960) 1452.
20) S. S. Brenner and J. T. McKinney, Surface Sci. **23** (1970) 88.
21) R. G. Forbes, J. Microscopy **96** (1972) 57.
22) E. W. Müller and R. D. Young, J. Appl. Phys., **32** (1961) 2425.
23) A. J. W. Moore and J. A. Spink, Surface Sci. **44** (1974) 198.

Surface Science 70 (1978) 239–254

FIELD EVAPORATION THEORY: THE ATOMIC-JUG FORMALISM

Richard G. FORBES
University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET, UK

Received 19 April 1977

Discrepancies are demonstrated to exist in the conventional mathematical treatment of the charge-exchange model for field evaporation, and a new treatment is presented that concentrates on the behaviour of the atomic nucleus prior to evaporation. The new approach seems physically more correct, produces markedly better agreement with experimental data concerning rate-constant field-sensitivity, and resolves several outstanding puzzles in field evaporation theory. The main numerical achievement is to predict a value of −6 for the ratio of partial energies μ^2/μ^1.

1. Introduction

The object of this paper is to present a new mathematical treatment of the charge-exchange model for field evaporation, that enables several existing difficulties with field evaporation theory to be resolved.

Traditionally, there have been two main contending models for field evaporation. Müller, using an image-hump model, was the first to propose that field evaporation was a thermally activated process [1], and a later variant of this model was able to predict evaporation fields with a moderate degree of success [2]. But subsequent work [3] showed that an equivalent or greater degree of success could be achieved with the charge-exchange model originated by Gomer [4].

The merit of the image-hump model, in its simplest version, is that the associated mathematics is extremely simple. However, as has been pointed out by various workers (in particular Tsong [5] and McKinstry [6]), considerable difficulties arise if more sophisticated versions of the model are examined in detail. the evaporating ion is assumed to be three-fold or four-fold charged, then the predicted distance of the hump from the metal's surface is unrealistically small, and the validity of the classical image-force potential would in any case be in doubt. Further, if short-range repulsive interactions of the type proposed by Brandon [7] are taken into account, then it is far from obvious that any hump actually exists in the ionic potential curve. For these and other reasons I shall follow Müller and Tsong [3] in thinking that, although the image-hump model may provide a convenient mathematical formalism, it is unwise to regard it as physically realistic.

The charge-exchange model is, prima facie, a much more likely mechanism. But, unfortunately, there currently exist certain discrepancies between the experimental results and theoretical analyses based on this model, particularly in the prediction of rate-constant field sensitivities. It will be shown below that these discrepancies are due to a mathematical oversight in the conventional analyses, and disappear when this oversight is corrected.

Much of the argument will be based around the rate-constant measurements made by Tsong [5], and the method of analysing rate-constant data developed by Forbes [8]; section 2 summarises these. Section 3 displays the difficulties arising with the conventional treatments, and the remaining sections then present and discuss the new approach.

Several points of notation deserve comment. In conformity with current trends, the equations in this paper are presented in rationalised form. In particular, the symbol α' is used to denote the quantity "*SI (absolute) polarisability*" recently discussed elsewhere [9]. α' is related to the *Gaussian polarisability* α_s by:

$$\alpha' = (4\pi\epsilon_0)\,\alpha_s\,, \tag{1}$$

where ϵ_0 is the electric constant (permittivity of free space). "Gaussian polarisability" is the official international name [10] for the quantity represented by the symbol α in most existing field-ion literature – most of which uses dimensionally-inconsistent expressions for polarisation energy [9].

I shall also employ a tighter convention about the arguments of logarithms and exponentials than has previously been thought necessary in field-ion literature, by making use of the special symbol "$\langle\,\rangle$". By definition, $\langle k_d \rangle$ in this paper stands for "the numerical value of the quantity k_d when k_d is expressed in s^{-1}"; and a similar definition holds for $\langle A \rangle$. Strictly, the entity $\ln k_d$ is mathematically improper, because k_d is not a pure number; the entity $\ln\langle k_d \rangle$, however, is not improper.

The work described here was first presented at the 23rd International Field Emission Symposium, at the Pennsylvania State University, in 1976. A more wide-ranging account of existing field-evaporation theory can be found in the review by Müller and Tsong [11].

2. The description of rate-sensitivity data

Data concerning the variation of the evaporation rate-constant k_d typically comes in the form of a plot of $\ln\langle k_d \rangle$ against applied voltage V. This may be converted into a field-dependent plot by the following procedure. An evaporation field F^c may be defined by the "unity rate-constant criterion" [8]:

$$k_d(F^c) = 1\ s^{-1}\,. \tag{2}$$

A quantity f, the *fractional overfield*, is then defined by:

$$f = (F - F^c)/F^c\,. \tag{3}$$

The *first-order field sensitivity* $S^c(f, T)$ is defined by:

$$S^c(f, T) = \mathrm{d}\ln\langle k_d\rangle/\mathrm{d}f = F^c\,\mathrm{d}\ln\langle k_d\rangle/\mathrm{d}F\,, \tag{4}$$

and higher-order field sensitivities by:

$$S^{nc}(f, T) = \mathrm{d}^n\ln\langle k_d\rangle/\mathrm{d}f^n = (F^c)^n\,\mathrm{d}^n\ln\langle k_d\rangle/\mathrm{d}F^n\,. \tag{5}$$

(The derivatives with respect to f and F are here represented as total derivatives, for reasons that will become clear later. However, since the relevant experiments are conducted at a fixed temperature T, there is also a sense in which S^c can be represented as a partial derivative taken with temperature held constant: field sensitivities are not used in this way here.)

The convert the V-dependent data plot into an f-dependent plot, the former is used to obtain the applied-voltage value V^c such that $k_d(V^c) = 1\ \mathrm{s}^{-1}$. On the assumption that $F \propto V$, it follows that:

$$f = (V/V^c) - 1\,. \tag{6}$$

Hence the V-axis can be redesignated, and values of field sensitivities obtained. This approach has the advantage that the quantity F^c is well-defined theoretically, and that data corresponding to $F = F^c$ can be identified empirically.

If data are presented in the form of a plot against some field F, then a similar redesignation procedure can be applied, using eqs. (2) and (3).

Fig. 1, reproduced from fig. 10 in Tsong's paper [5], but with transformed coordinates, shows how $\ln\langle k_d\rangle$ varies with f, for a tungsten adatom on a tungsten (110) plane, at 80 K. The curve drawn by Tsong has been used in the identification of F^c.

Though field-sensitivity values are often used in discussions of field evaporation

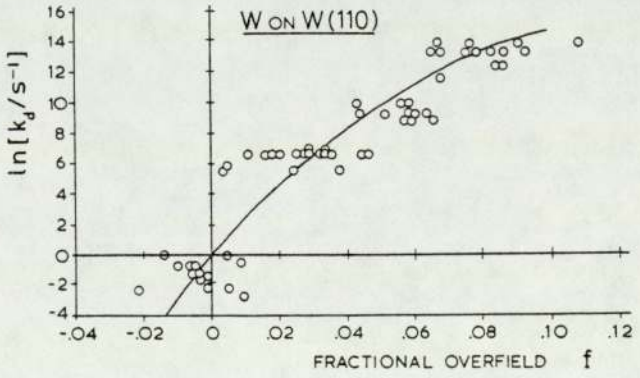


Fig. 1. The variation of evaporation rate-constant with fractional overfield, for tungsten adatoms on a tungsten (11) plane. Data taken from Tsong [5], fig. 10, but shown plotted against transformed coordinates.

theory, it is sometimes more convenient to represent the data in terms of the coefficients μ^n defined by:

$$\mu^n = kT\, S^{nc}(0, T) \tag{7}$$

where k is the Boltzmann constant, and $S^{nc}(0,T)$ is the n^{th} order field-sensitivity value, taken at $f = 0$ (that is, at the evaporation field F^c). The coefficients μ^n can be simply related to the coefficients that appear in a regression polynomial fitted to fig. 1 [12,13].

These coefficients also have a direct theoretical interpretation. At sufficiently high temperatures, field evaporation obeys an equation that may be written in the Arrhenius form:

$$k_d = A \exp(-Q/kT)\,, \tag{8}$$

where Q is a classical activation energy, and A is a pre-exponential [3]. Following ref. [8], this equation can also be written in the form:

$$k_d = t_s^{-1} \exp(M/kT)\,, \tag{9}$$

where

$$M = kT \ln\langle A\rangle - Q\,, \tag{10}$$

and $t_s \equiv 1$ sec. (The dimensional constant t_s must be included to keep eq. (9) dimensionally consistent.) From eqs. (7), (5) and (9), it follows that the coefficients μ^n are given theoretically by:

$$\mu^n = (\mathrm{d}^n M/\mathrm{d}f^n)|_{f=0} = (F^c)^n\, (\mathrm{d}^n M/\mathrm{d}F^n)|_{F=F^c}\,. \tag{11}$$

That is, the μ^n are the values of the various-order derivatives of M, taken at the evaporation field F^c.

The quantity M has the dimensions of energy, but does not correspond exactly to any existing thermodynamic quantity. Forbes [14] has proposed that it be given a distinctive name, and has tentatively suggested "*Puissance*". The coefficients μ^n also have the dimensions of energy, and have been termed "partial energies" or "*partial puissances*".

Puissance is a function of the site field F, and may be Taylor-expanded about the evaporation field F^c, as follows:

$$M(F) = M(F^c) + (F - F^c)\left.\frac{\mathrm{d}M}{\mathrm{d}F}\right|_{F^c} + \tfrac{1}{2}(F - F^c)^2 \left.\frac{\mathrm{d}^2M}{\mathrm{d}F^2}\right|_{F^c} + \ldots \tag{12}$$

$M(F^c)$ is zero by definition, as may be seen by substituting $M = 0$ into eq. (9) and comparing the result with definition (2). Hence, on introducing the variable f defined by eq. (3), and using relationships (11), one obtains:

$$M(f) = \mu^1 f + \tfrac{1}{2}\mu^2 f^2 + \ldots\,. \tag{13}$$

Table 1
Experimental estimates of partial puissances

Tungsten:	μ^1	2.0 eV
	μ^2	−10 eV
Molybdenum:	μ^1	1.4−2.1 eV
Platinum:	μ^1	1.3 eV

This equation provides an alternative way of defining the partial puissances μ^n.

If an expression for Q as a function of F (or f) is available, and if it may be assumed that any field dependence in the pre-exponential A is small and hence

$$\mathrm{d}M^n/\mathrm{d}F^n \triangleq -\mathrm{d}^nQ/\mathrm{d}F^n\,, \tag{14}$$

then theoretical expressions for the μ^n in terms of atomic parameters can readily be obtained. Consequently, comparisons of experimental and theoretical estimates of the various μ^n are a useful way of comparing experiment and theory.

Experimental estimates of μ^1 can be obtained from the work of Brandon [15], Taylor [16] and Tsong [5], and are discussed in ref. [8]. Tungsten is the only material that has been studied by more than one worker: the corresponding estimates of μ^1 are remarkably consistent, and we can take the "typical" value of μ^1 for tungsten as 2 eV. Table 1 above also shows estimates of μ^1 for molybdenum and for platinum, derived from Brandon's work.

Tsong's experiments are the only source for a quantitative estimate of μ^2, and this only for tungsten. Unfortunately, there is some uncertainty about the statistical treatment of his data [12,13]. The "typical" estimate of μ^2 as −10 eV is based on a re-analysis by Patel [12] of Tsong's data, and is the rounded average of the estimates derived from the two sets of adatom data published in ref. [5].

I shall employ the data in table 1 when discussing the results of the new treatment presented later. However, for discussion of the conventional treatment, in the following section, it seems more appropriate to use data derived directly from the curve drawn by Tsong and shown in fig. 1. The data are given in table 2. These are derived from the numbers given by Tsong in the caption to fig. 10 in ref. [5], using the transformation formula given in ref. [8] and assuming that $\ln\langle k_\mathrm{d}\rangle$ equals zero

Table 2
Rate-constant field-sensitivity data derived from fig. 1

$f = 0$	$S^\mathrm{c}(0, 80\ \mathrm{K})$	~260
$f = 0.08$	$S^\mathrm{c}(0.08, 80\ \mathrm{K})$	~80
$f = 0$	$S^{2\mathrm{c}}(0, 80\ \mathrm{K})$	~−2250

when Tsong's quantity F/F_0 equals 0.955. The qualitative conclusions of section 3 are the same whether one uses the data in table 1, the data in table 2, or data derived from fig. 9 in ref. [5].

3. The conventional mathematical treatment

This section now displays a major difficulty associated with the conventional mathematical treatment of rate-sensitivities. Quantitative discussion requires a definite activation-energy expression. For present purposes, I shall disregard the broadening, energy-shift, and short-range terms sometimes included [11], and suppose that in the charge-exchange model the evaporation activation-energy $Q(F)$ is given by:

$$Q(F) = [\Lambda_0 + H_n^{\infty} - n\phi^{e} + C(x^{cr})] - neFx^{cr} + \tfrac{1}{2}\alpha'' F^2 , \tag{15}$$

where e is the elementary (proton) charge; ne is the charge on the ion after field evaporation; Λ_0 is the zero-field binding-energy of the neutral atom to the surface; H_n^{∞} is the energy of formation, in remote field-free space, for the ion from the neutral; ϕ^{e} is the appropriate local work-function; and x^{cr} is the effective distance of the crossing point ("escape point") from the metal's electrical surface. (x^{cr} may be identified with the quantity "$x_c + \lambda^{-1}$" appearing in Müller and Tsong's discussion [11]).

$C(x)$ represents the correlation interaction between the ion and the surface, and is usually approximated by the image potential:

$$C(x) \triangleq -n^2e^2/16\pi\epsilon_0 x , \tag{16}$$

where ϵ_0 is the electric constant, and x is the distance of the ion from the metal's electrical surface. The exact applicability of this approximation is, however, open to some doubt [11].

α'' is the "effective SI polarisability", and is given by:

$$\alpha'' = \alpha'_0 - \alpha'_n \tag{17}$$

where α'_0 and α'_n are polarisability-like quantities, having the dimensions of SI absolute polarisability, associated with the neutral and with the evaporating ion, respectively.

The conventional mathematical treatment supposes that, in evaluating dQ/dF, any variation with field in the position of the crossing point may be disregarded. This assumption will also be made here.

If, initially, the polarisation term in eq. (15) is disregarded, then:

$$dQ/dF \triangleq -neFx^{cr} , \tag{18}$$

and hence from earlier equations we obtain:

$$S^{c} \triangleq neF^{c}x^{cr}/kT , \tag{19}$$

$$S^{2c} \triangleq 0 . \tag{20}$$

There are considerable difficulties in making exact a-priori estimates of F^c [8], or of x^{cr} [11]. For the evaporation field F^c the conventional value for tungsten of 57 V/nm will be used. As regards x^{cr}, the author's view is that to substitute the relevant atomic diameter is a crude but more-or-less sufficient approximation, that takes some cognisance of field penetration [11] and is roughly equivalent to supposing that the metal's electrical surface is effectively in the vicinity of the substrate nuclei. For tungsten the value $x^{cr} = 0.27$ nm will be assumed: this is the nearest-neighbour distance in the tungsten lattice [17]. Taking $n = 3$ (which is the commonest tungsten evaporation charge-state found in atom-probe work [18]), one obtains $neF^c x^{cr} \triangleq 46$ eV. Since $k \cdot 80$ K equals 0.0069 eV, the predicted first-order field sensitivity S^c is approximately 6700.

Clearly, eqs. (19) and (20) are unable to predict the curvature observed in fig. 1, but it also deserves note that the discrepancy between the experimental (260) and theoretical (6700) estimates of the field sensitivity at the field F^c is far too large to be attributed to uncertainties in n, x^{cr}, or F^c.

If the polarisation term is included, then the following expressions (in terms of f) can be obtained:

$$S^c(f, T) = [neF^c x^{cr} - \alpha''(F^c)^2 (1 + f)]/kT , \tag{21}$$

$$S^{2c}(f, T) = -\alpha''(F^c)^2/kT . \tag{22}$$

Clearly, equations of this form can qualitatively explain the curvature shown in fig. 1.

The observed value of S^{2c}, namely -2250, implies a value for $\alpha''(F^c)^2$ of approximately 16 eV. This implies a value for α'' equivalent to a Gaussian polarisability of about 7 Å^3. Such a value appears to be compatible with the Gaussian-polarisability value 9.2 Å^3 quoted [11] in connection with experiments on surface diffusion in high electric fields.

A semi-theoretical estimate of $S^c(0, T)$ can now be obtained by setting $f = 0$ in eq. (21) and substituting back the empirical value for $\alpha''(F^c)^2$. Taking $neF^c x^{cr}$ as 46 eV, as before, leads to the prediction $S^c(0, 80\text{ K}) \triangleq 4300$.

This estimate is still much larger than the experimental result (260). To get the estimate of $S^c(0, 80\text{ K})$ as low as is found experimentally, the square bracket in eq. (21) must be equal to approximately 2 eV, which requires that $neF^c x^{cr}$ have a value of around 18 eV. But, if $n = 3$, then this in turn implies a value for x^{cr} of just over 0.1 nm.

Such a value is not inconceivable, but would tend to suggest that intuitive considerations concerning the position of the metal's electrical surface were faulty. The author's opinion, shared with McKinstry [6], is that a value for x^{cr} of 0.1 nm or so is too low to be plausible. Thus, even when the polarisation term is included, the fit between theory and experiment is unconvincing.

There is also a more general difficulty associated with the form of eq. (21). Both

the terms in the square bracket are large; but experiment requires that they be roughly equal; which will be the case only if:

$$nex^{cr} \triangleq \alpha'' F^{c} . \tag{23}$$

But there is no obvious reason why this near-equality should hold for each of a wide range of materials; so one might expect to observe wide variations in field sensitivity S^{c} (or, equivalently, μ^{1}) from material to material. But if such an effect operated then it should be well known, because field evaporation is basic to endform preparation. No such effect exists, and for the three materials for which μ^{1} has been measured the values of μ^{1} lie within a factor of 1.6 (cf. table 1). There is an unhappiness of form about eq. (21), and hence about theories that explain curvature in the $\ln\langle k_{d}\rangle$ versus f plot in terms of the influence of polarisation effects.

Further evidence of some inadequacy in the existing treatment comes from the work of Vesely and Ehrlich [19]. They took the series expansions, for Q in terms of F, conventionally associated with the image-hump and charge-exchange mechanisms and fitted them to the rate-constant data published by Tsong [5], using regression techniques. Their conclusion, albeit qualified in various ways, was that statistically the best fit to the data was given by the image-hump series expansion, with $n = 2$. This creates an anomaly, because the charge-state of tungsten on arrival at the detector is normally 3+, and the image-hump mechanism can be discounted on general grounds, as explained earlier.

A further well-known difficulty with both the existing models is that, if one uses the criterion of lowest predicted evaporation field to predict the evaporation charge state [2,3], then the predicted charge-state is usually lower than that found experimentally in atom-probe work.

4. The atomic-jug formalism

One path forward from the difficulties just described might be to include additional terms in the expansion for Q in terms of F. An alternative is to re-examine the mathematical assumptions behind the customary treatment.

As was pointed out by McKinstry [6], the correct mathematical expression for the total derivative dQ/dF is:

$$\frac{dQ}{dF} = \left(\frac{\partial Q}{\partial F}\right)_{x^{cr}} + \left(\frac{\partial Q}{\partial x^{cr}}\right)_{F} \frac{dx^{cr}}{dF} . \tag{24}$$

The conventional treatment is equivalent to ignoring the second term in eq. (24), on the grounds that $dx^{cr}/dF \triangleq 0$. It would be possible to directly investigate the effect of including the second term: instead, it seems better to approach the problem from an entirely different angle, that concentrates on the behaviour of the vibrating nucleus prior to field evaporation.

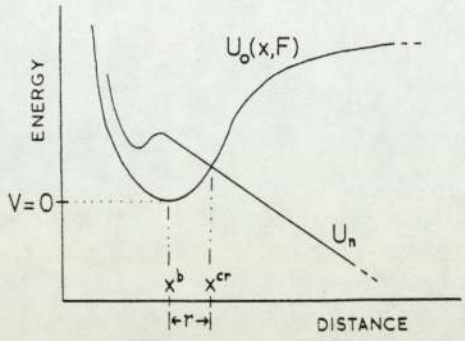


Fig. 2. Potential configuration associated with the simple charge-exchange mechanism of field evaporation.

As shown in fig. 2, the pre-evaporating nucleus moves in a potential well. Let $V(x, F)$ denote the binding-well potential, measured relative to a zero of potential at the bottom of the well, at x^b. In terms of the standard neutral Molecular Term $U_0(x, F)$ defined in ref. [20] (i.e. in terms of the atomic potential energy), $V(x, F)$ is given by:

$$V(x, F) = U_0(x, F) - U_0(x^b, F) . \tag{25}$$

In terms of $V(x, F)$, the evaporation activation-energy $Q(F)$ is given by:

$$Q(F) = V(x^{cr}, F) . \tag{26}$$

One thus has a picture of the pre-evaporating nucleus as moving in a jug-like potential: if the nucleus gets out as far as x^{cr}, then evaporation takes place. For the nucleus, the process of field evaporation is an atomic-level analogy to the pouring of water from a jug. As the field in increased, the "lip" of the jug is lowered, and the probability of evaporation increases.

4.1. The parabolic-jug approximation

As a first approximation, it may be hypothesised that V has the form appropriate to a simple harmonic oscillator:

$$V = \tfrac{1}{2}\kappa(x - x^b)^2 \tag{27}$$

where κ is a force-constant. This is the *parabolic-jug approximation*, introduced by Forbes [21]. It follows that the activation energy $Q(F)$ is given by:

$$Q(F) = \tfrac{1}{2}\kappa(x^{cr} - x^b)^2 = \tfrac{1}{2}\kappa r^2 \tag{28}$$

where r is a new variable representing the distance between the bonding point and the crossing point.

For notational clarity, it is useful to here introduce a new symbol (R) to denote

the distance of the crossing point from the metal's electrical surface, and a new symbol (a) to represent the distance of the bonding point from the electrical surface. That is:

$$a \equiv x^{\mathrm{b}} , \tag{29}$$

$$R \equiv x^{\mathrm{cr}} = a + r . \tag{30}$$

From eq. (28), using this new notation, it follows that the total derivative of Q with respect to F is given by:

$$\frac{\mathrm{d}Q}{\mathrm{d}F} = \left(\frac{\partial Q}{\partial \kappa}\right)_r \frac{\mathrm{d}\kappa}{\mathrm{d}F} + \left(\frac{\partial Q}{\partial r}\right)_\kappa \frac{\mathrm{d}r}{\mathrm{d}F} \tag{31}$$

$$= \tfrac{1}{2} r^2 \frac{\mathrm{d}\kappa}{\mathrm{d}F} + \kappa r \left(\frac{\mathrm{d}R}{\mathrm{d}F} - \frac{\mathrm{d}a}{\mathrm{d}F}\right). \tag{32}$$

If the initial bonding state is assumed to be neutral, as in the conventional treatment of field evaporation, then it is probably a good approximation to treat the derivatives $\mathrm{d}\kappa/\mathrm{d}F$ and $\mathrm{d}a/\mathrm{d}F$ as negligibly small, and eq. (32) reduces to:

$$\mathrm{d}Q/\mathrm{d}F = \kappa r \, (\mathrm{d}R/\mathrm{d}F) . \tag{33}$$

An expression for $\mathrm{d}R/\mathrm{d}F$ may be obtained as follows. Expressions (15) and (28) for activation energy must be equivalent, so:

$$\tfrac{1}{2}\kappa r^2 = (\Lambda_0 + H_n^{\infty} - n\phi^{\mathrm{e}} + C) - neFR + \tfrac{1}{2}\alpha'' F^2 . \tag{34}$$

Taking differentials and rearranging leads to:

$$(\kappa r - C' + neF)\, \mathrm{d}R = -(neR - \alpha'' F)\, \mathrm{d}F . \tag{35}$$

In doing this it has been assumed that $\mathrm{d}r = \mathrm{d}R$ (because $\mathrm{d}a \doteq 0$), and that any variation of α'' with position may be ignored; C' denotes $\mathrm{d}C/\mathrm{d}R$. Substitution into eq. (33) now leads to:

$$\frac{\mathrm{d}Q}{\mathrm{d}F} = -\frac{\kappa r}{\kappa r - C' + neF} (neR - \alpha'' F) , \tag{36}$$

which is the basic result.

We shall also require, later, explicit expressions for r and Q in terms of F. Approximate expressions can be obtained as follows. In eq. (34): the term $\frac{1}{2}\kappa r^2$ is ignored, being small; the term C is evaluated at some specific value of R, and is then combined with the other terms inside the brackets to give a constant W^0; $neFR$ is expanded as $neF(a + r)$; and the result is rearranged, to give:

$$r \doteq \frac{W^0}{neF} - a + \frac{\alpha'' F}{2ne} . \tag{37}$$

This approximation holds for r not too large, and is adequate for fields near F^{c}. At

an equivalent level of approximation, $Q(F)$ is given by:

$$Q \triangleq \tfrac{1}{2}\kappa \left(\frac{W^0}{neF} - a + \frac{\alpha'' F}{2ne} \right)^2 . \tag{38}$$

This expression should be compared with the conventional expression, eq. (15), in which the dependence of x^{cr} on F is not shown explicitly.

5. Application to the charge-exchange model

5.1. Estimation of μ^1

If, as before, any field dependence in the Arrhenius-form pre-exponential A is ignored, then from earlier equations:

$$S^c(F, T) = -F^c(\mathrm{d}Q/\mathrm{d}F)/kT . \tag{39}$$

Substituting from eq. (36) and re-arranging leads to:

$$S^c(F, T) = \frac{\kappa r R}{kT} \frac{F^c}{F} \left(\frac{neF}{\kappa r - C' + neF} \right) \left(1 - \frac{\alpha'' F}{neR} \right). \tag{40}$$

In the interests of algebraic simplicity, two inessential approximations are now introduced. Consider the final bracket in eq. (40), and compare the quantities $neFR$ and $\alpha'' F^2$. From earlier estimation, $neFR$ is approximately 46 eV. For an a-priori estimate of $\alpha'' F^2$, one might take twice the largest estimate of polarisation energy given by Müller and Tsong (ref. [3], p. 70): this would give $\alpha'' F^2$ as about 5 eV. In this case $\alpha'' F/neR \ll 1$, and the final bracket can be approximated to unity, which we shall do here. (If the actual value of $\alpha'' F^2$ is somewhat larger, as might seem the case from Tsong's experiments on adatom diffusion [22], then a correction factor would need to be introduced.)

In general the second bracket can be set equal to a quantity γ that has a value somewhat less than one and is a slowly-varying function of field. However, in a first approximation this bracket also may be set equal to unity. With these approximations one obtains:

$$S^c(F, T) = (\kappa r R/kT)\,(F^c/F) . \tag{41}$$

Hence, the first partial puissance μ^1 is given by:

$$\mu^1 = kT\, S^c(F^c, T) = \kappa r_c R_c , \tag{42}$$

where the suffix "c" on r_c and R_c indicates that the values are those appropriate to the evaporation field F^c defined by the unity-rate-constant criterion.

Expression (42) may be re-written as:

$$\mu^1 = 2(\tfrac{1}{2}\kappa r_c^2)\,(R_c/r_c) . \tag{43}$$

The first bracket may be identified with the activation energy $Q(F^c)$, which may be

taken as about 0.2 eV. An a-priori estimate of R_c/r_c, made intuitively, might be around 10. These values lead to a semi-theoretical estimate for μ^1 of around 4 eV.

This estimate is about twice the experimental estimates listed in table 1, and is much closer than the a-priori estimates that would be derived in the conventional treatment. The remaining discrepancy by a factor of two may be due to the approximations made earlier, or it may be due to inaccuracy in my intuitive estimate of R_c/r_c. For consistency with the experimental estimate, it will be supposed here that in the case of tungsten the value of R_c/r_c is "actually" 5. Since $R_c = a + r_c$, the value $a = 0.27$ nm leads to an estimate for r_c of about 70 pm (0.07 nm). Although perhaps a little higher than one might intuitively expect, such a value is certainly not unreasonable.

5.2. Field-sensitivity formulae

By combining eq. (37) with eq. (40), expressions may be obtained that explicitly exhibit the dependence of $S^c(F, T)$ on the site field F. The general result is algebraically complicated, and not very informative. It seems better to use different approximations to it in different field regimes: three such regimes have been identified [21]:

The low-field regime. At low fields the distance r becomes large, and the approximation of setting the second bracket in eq. (40) equal to unity fails. Rather, it has to be set equal to $neF/\kappa r$, which leads to:

$$S^c(F, T) \triangleq neF^cR/kT\,. \tag{44}$$

This result is identical with eq. (19).

The regime near F^c. In eq. (41) the main field dependence not explicitly exhibited is that of r. Substituting from eq. (37) gives:

$$S^c(F, T) \triangleq \frac{\kappa RF^c}{kT}\left(\frac{W^0}{neF^2} - \frac{a}{F} + \frac{\alpha''}{2ne}\right). \tag{45}$$

This is the approximation given in ref. [21], and neglects the relatively weak field-dependence of R. This can be included explicitly by using eq. (37) a second time, to give:

$$S^c(F, T) \triangleq \frac{\kappa F^c}{kT}\left(\frac{W^0}{neF^2} - \frac{a}{F} + \frac{\alpha''}{2ne}\right)\left(\frac{W^0}{neF} + \frac{\alpha''F}{2ne}\right). \tag{46}$$

The high-evaporation-field regime. At higher fields, when the distance r becomes very small, $S^c(F, T)$ becomes a gently-varying function of field F; and at fields very close to the field F^e at which the activation energy becomes zero it follows from eq. (41) that:

$$S^c(F, T) \triangleq 0\,. \tag{47}$$

5.3. Estimation of μ^2

A general expression for μ^2 is tedious to derive, so a simplified approach will be used here. From earlier definitions:

$$\mu^2 = kT\,F^c(\mathrm{d}S^c/\mathrm{d}F)|_{F^c}\,. \tag{48}$$

An approximate expression for $\mathrm{d}S^c/\mathrm{d}F$ may be obtained by differentiating eq. (45), ignoring the field-dependence of R. Substitution into eq. (48) then yields:

$$\mu^2 \triangleq \kappa R_c[-(2W^0/neF^c) + a]\,. \tag{49}$$

An analogous expression for μ^1 may be obtained by substituting eq. (45) into eq. (42), to give:

$$\mu^1 = \kappa R_c[(W^0/neF^c) - a + (\alpha'' F^c/2ne)]\,. \tag{50}$$

By combining these we obtain:

$$2\mu^1 + \mu^2 \triangleq \kappa R_c[-a + (\alpha'' F^c/ne)]\,. \tag{51}$$

Using eqs. (30) and (42), some algebraic rearrangement gives:

$$\mu^1 + \mu^2 \triangleq -\mu^1(R_c/r_c)\,[1 - (\alpha'' F^c/neR_c)]\,. \tag{52}$$

In consistency with previous assumptions, we may take the term in square brackets as unity, and R_c/r_c as 5. Hence one obtains the semi-theoretical prediction:

$$\mu^2/\mu^1 \triangleq -6\,. \tag{53}$$

More exact approximations tend to produce values lying within ±2 of this figure, with some uncertainty arising from our lack of exact values for parameters such as R and α''.

From the data in table 1, a typical experimental estimate of μ^2/μ^1 is −5. Considering the imprecise nature of both experiment and theory at the present time, the agreement between them is entirely satisfactory.

5.4. The force-constant

An empirical estimate of the force-constant κ can be obtained by substituting the values $Q(F^c) = 0.2$ eV, $r_c = 0.068$ nm, into the formula $Q(F^c) = \frac{1}{2}\kappa r_c^2$. This gives κ as approximately 90 eV/nm^2.

If the assumption of a parabolic well is taken to have some physical reality, then a vibration frequency ν may be obtained from the formula for a linear harmonic oscillator:

$$2\pi\nu = (\kappa/m)^{1/2} \tag{54}$$

where m is the mass of the vibrating atom. Using the mass appropriate to a tungsten atom leads to the result: $\nu \triangleq 1 \times 10^{12}$ Hz.

This value is well in accord with prior expectation [23]. In fact, the agreement is better than we have any right to expect, and may be partly fortuitous.

6. Discussion

In the quantitative interrelation of experimental data, the present version of the parabolic-jug approximation seems significantly more effective than the conventional treatment of the charge-exchange mechanism. However, the treatment here involves many approximations, and some of these deserve detailed exploration. For the time being, the numerical agreements achieved are best regarded as indicating future potential rather than hard performance. The real merits of the new approach seem to lie in some more-general characteristics.

(A) The predictions made here concerning the values of μ^1 and μ^2/μ^1 do not depend significantly on the value assumed for the ionic charge-state after escape. The present treatment is a "partial theory", compatible with the atom-probe results. Hopefully, it should be possible to integrate it with some future theory concerning final charge-state.

(B) All the predicted parameter values are readily plausible. Further, the quantities κ, r_c and R_c should not be sensitively dependent on the nature of the material involved, so one might reasonably expect μ^1 to be much the same for different materials, as the experimental results tend to suggest.

(C) Prediction of a high-field regime where rate-sensitivities are relatively small appears to be compatible with the experience of atom-probe operators that such a regime exists experimentally (Waugh, private communication).

(D) The explicit series expansion for Q in terms of F derived here, namely eq. (38), contains terms in F^{-2} and F^{-1}. No such terms appear in the series expansion used in the conventional treatment of the charge-exchange mechanism. In their curve-fitting exercise with Tsong's data, mentioned earlier, Vesely and Ehrlich concluded that the image-hump series expansion produced a statistically-better fit than the conventional charge-exchange series expansion [19]. The present treatment suggests that at fields near the evaporation field F^c the conventional expansion is mathematically faulty. Consequently, no physical significance need be attached to the Vesely and Ehrlich result. In particular, it constitutes no valid reason for prefering the image-hump mechanism over the charge-exchange mechanism.

(E) In the new treatment the effective polarisability α'' does not have a sensitive influence on the theoretical estimates of μ^1 or μ^2. This characteristic helps resolve some past controversy. Tsong [5] has claimed, first, that variations in effective polarisability ought on theoretical grounds to exist, dependent on the crystallographic environment of the bonding site; second, that such variations have been detected in his experiments on rate-constant sensitivity, as analysed using a conventional charge-exchange series expansion for Q. Vesely and Ehrlich [19], and Patel and Forbes [12,13], however, assert that Tsong's analysis of his data is statistically

inadequate, and that regression procedures assuming this conventional series expansion lead to α'' values that the are significantly different from Tsong's. Vesely and Ehrlich further assert that Tsong's published data, when analysed by regression techniques, contain no statistically significant evidence of crystallographic variation in α''.

With the present treatment, the absence of significant differences in μ^2, as between crystallographically different data sets, is entirely compatible with the existence of significant variations in α'', as between crystallographically different bonding sites. In effect, another potential anomaly has been eliminated.

(F) In the present treatment there are no assumptions that are highly sensitive to the orientation assumed for the axis of the potential barrier. It is possible to consider r_c, R_c and a as measured along a path of escape that is partly parallel to the emitter surface. The numerical agreements demonstrated here are thus compatible with variants of the charge-exchange mechanism that assume that nuclear motion during evaporation is partly parallel to the surface. There seems an increasing amount of experimental evidence to suggest that lateral motion may be important (for example, refs. [24] and [25]), but up till now theoretical speculations [8,25] have been in terms of a mechanism involving activation over a diffusion hump. A lateral charge-exchange mechanism now seems a viable alternative.

The essential difference between the old and the new treatments is as follows. The conventional treatment evaluates the evaporation activation energy as the difference between the ion energy at a fixed point and the bonding-point energy; the atomic-jug formalism evaluates it as the difference between the ion energy at the crossing point (wherever that may happen to be) and the bonding-point energy. With the conventional treatment the reason why the rate-constant versus field curve flattens out is the mathematical fact that a squared (in F) term increases more rapidly than a linear (in F) one: with the atomic-jug formalism the reason is the physical fact that a binding-potential curve is flattest near to the bonding point.

Although several approximations made deserve systematic exploration, the present simplified treatment's main fault is that it appears to assume that the pre-evaporating atom is bound (prior to evaporation) in an essentially neutral condition. Tsong and Kellogg's theoretical interpretation of the polarisabilities derived from their experiments on adatom diffusion [22] seem to support the present author's intuition [8] that the bonding state may be partly ionic. In fact, the theory here can be extended to cover the ionic bonding situation [26], and in the case of bonding as a singly-charged ion the numerical results are fairly similar to those for neutral bonding.

Overall, it seems clear that, as compared with the conventional treatment of the charge-exchange mechanism, the atomic-jug formalism is more correct physically, is more sophisticated mathematically, and leads (even in the basic version of the parabolic-jug approximation) to superior scientific performance. We may conclude that the conventional treatment is mathematically inadequate, and that the variation of the crossing-point position with field is of major physical significance.

By a strange chance, the effect of this is to make the charge-exchange model closer in spirit to Erwin Müller's original image-hump model, for in the latter the variation with field in the position and energy-level of the top of the hump is the factor of significance. Plus ça change, plus c'est la même chose, even in field evaporation theory.

Acknowledgements

I thank the UK Science Research Council for personal financial support, and Professor E. Braun and Dr. R.V. Latham for provision of facilities.

References

[1] E.W. Müller, Phys. Rev. 102 (1956) 618.
[2] D.G. Brandon, Surface Sci. 3 (1964) 1.
[3] E.W. Müller and T.T. Tsong, Field Ion Microscopy: Principles and Applications (Elsevier, Amsterdam, 1969).
[4] R. Gomer, J. Chem. Phys. 31 (1959) 341.
[5] T.T. Tsong, J. Chem. Phys. 54 (1971) 4205.
[6] D. McKinstry, Surface Sci. 29 (1972) 37.
[7] D.G. Brandon, Brit. J. Appl. Phys. 14 (1963) 474.
[8] R.G. Forbes, Surface Sci. 46 (1974) 577.
[9] R.G. Forbes, Surface Sci. 64 (1977) 367.
[10] Draft International Standard ISO/DIS 31/V, entitled "Quantities and units of electricity and magnetism".
[11] E.W. Müller and T.T. Tsong, Progr. Surface Sci. 4 (1973) 1.
[12] C. Patel, M. Sc. thesis, Univ. of Aston in Birmingham (1974).
[13] C. Patel and R.G. Forbes, in: 22nd Intern. Field Emission Symp., Atlanta, 1975; manuscript in preparation.
[14] R.G. Forbes, in: 21st Intern. Field Emission Symp., Marseille, 1974.
[15] D.G. Brandon, Phil. Mag. 14 (1966) 803.
[16] D.M. Taylor, Ph.D. thesis, University of Cambridge (1970).
[17] K.M. Bowkett and D.A. Smith, Field Ion Microscopy (North-Holland, Amsterdam, 1970).
[18] For example, S.S. Brenner and J.T. McKinney, Appl. Phys. Letters 13 (1968) 29.
[19] M. Vesely and G. Ehrlich, Surface Sci. 34 (1973) 547.
[20] R.G. Forbes, Surface Sci. 61 (1976) 221.
[21] R.G. Forbes, in: 23rd Intern. Field Emission Symp., Pennsylvania State Univ., 1976.
[22] T.T. Tsong and G. Kellogg, Phys. Rev. B12 (1975) 1343.
[23] Note that ν here represents a genuine vibration frequency, and not an Arrhenius pre-exponential.
[24] A.J.W. Moore and J.A. Spink, Surface Sci. 44 (1974) 198.
[25] A.R. Waugh, E.D. Boyes and M.J. Southon, Surface Sci. 61 (1976) 109.
[26] R.G. Forbes, unpublished work.

Surface Science 102 (1981) 255–263

CHARGE HOPPING AND CHARGE DRAINING: TWO MECHANISMS OF FIELD DESORPTION

Richard G. FORBES
Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET, UK

Received 12 February 1980; accepted for publication 17 June 1980

It is suggested that the so-called charge-exchange model for field desorption in fact comprises two alternative mechanisms between which we may discriminate theoretically, and possibly experimentally. This note also corrects some conceptual errors that have arisen in past literature through failure to make this discrimination.

1. Introduction

The theoretical discussion of field evaporation is often presented in terms of two "models": the "image-hump" model [1–3] and the "charge-exchange" model [4–6]. I suggest in this note that the so-called charge-exchange model in fact comprises two different (albeit related) mechanisms of field desorption. Failure to distinguish properly between these mechanisms seems to have led to misunderstandings in the literature concerning the theory of transition rate-constants, and – in connection with the theory of appearance energies – to the statement of formulae that are incompatible with the principle of conservation of energy. The objective of this note is to clarify the theoretical situation, and to suggest that it might be possible to distinguish experimentally between the mechanisms.

2. Theoretical principles

Charge exchange is a phenomenon that occurs in the context of the crossing or near-crossing of potential-energy curves. The principles involved have been extensively discussed in connection with the theory of atomic collisions and the theory of internal conversions in molecules (see, for example, refs. [7,8]), and are basically very well understood. I begin by setting down how these might be expected to apply in the circumstances of field desorption.

The curves shown in figs. 1a and 1b represent the *standard* potential energy ("standard Term") for a desorbate particle, as defined in ref. [9]. The emitter, with

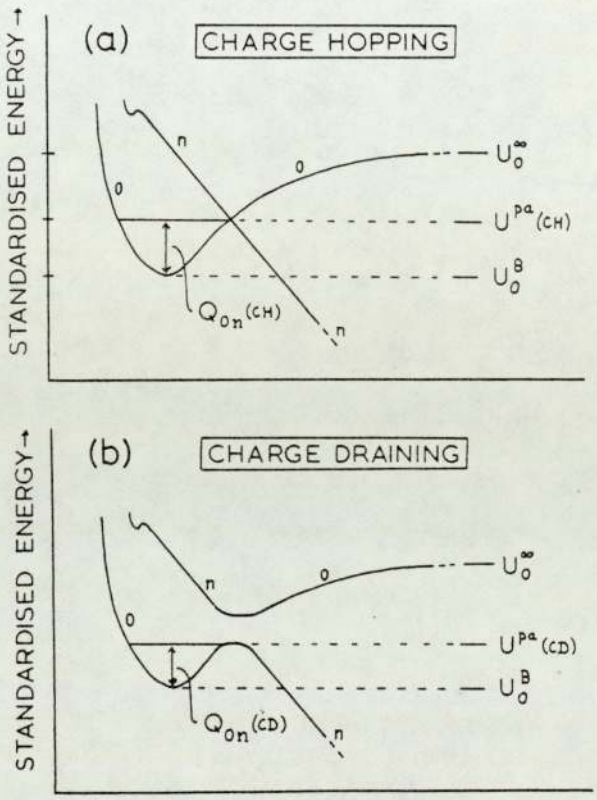


Fig. 1. Potential configurations for the two alternative mechanisms of field desorption involving charge exchange, in the special case where the initial bonding state is neutral ($m = 0$). (a) Charge-hopping mechanism, involving curve intersection. (b) Charge-draining mechanism, involving curve repulsion. (The potential curves shown are schematic.)

its reservoir of electrons, is to the left of the crossing region, and vacuum is to the right. As the desorbate particle moves from left to right, one or more electrons are transferred from the desorbate particle to the emitter. Away from the crossing region the potential-energy curves correspond to well-defined charge stages of the desorbate particle [10], and these can be labelled by the corresponding charge-numbers, m and n, me being the charge on the desorbate in its initial state (where e is the elementary charge), ne the charge in its final state. For simplicity, fig. 1 depicts the situation where the initial bonding state is neutral ($m = 0$), but the equations below and also fig. 2 are valid for non-zero values of m.

Depending on the circumstances of the electron transfer, the potential-energy curves in the crossing region are drawn in the alternative ways shown in fig. 1, either as "intersecting" or as "repelled". (These configurations are sometimes

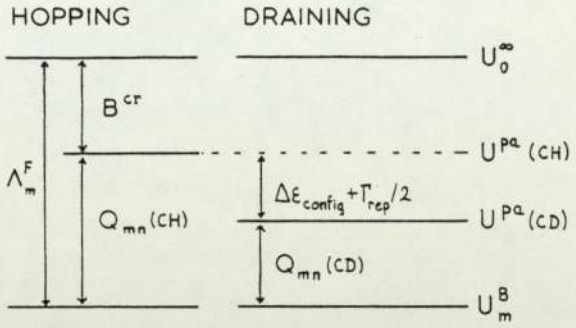


Fig. 2. General relationships between the energy terms involved in the charge-hopping and charge-draining mechanisms, shown schematically.

referred to as the "non-adiabatic" (fig. 1a) and "adiabatic" (fig. 1b) cases, but the present author finds this terminology unilluminating.)

The choice between the two configurations basically depends on the relative speeds associated with the electron behaviour and with the motion of the desorbate particle as a whole. If the desorbate passes quickly through the crossing region, and transfer of an electron is unlikely on a single passage through the region, then the curves are drawn as intersecting. Transfer of an electron (or electrons), if it occurs, is best envisaged as a sharp "hop". Hence I call this the "*charge-hopping*" mechanism of field desorption.

At the opposite extreme, if the desorbate passes very slowly through the crossing region, and transfer of an electron (or electrons) is almost inevitable on a single passage, then the curves must be drawn as "repelled". In this case the electron transfer is best envisaged as a slow draining of electron charge out of the desorbate into the emitter, and I call this the "*charge-draining*" mechanism of field desorption.

The intersecting curves are (in principle) obtained by requiring the desorbate electrons to be in orbitals centred on the desorbate, with the emitter regarded as causing a perturbation. In contrast, the repelled curves are obtained by treating the emitter and desorbate as a single potential configuration: when the desorbate is in the crossing region the wave-function for the transferring electron has significant amplitude both in the emitter and in the desorbate.

It can be shown (for a simple treatment, see ref. [11]), that the standardised energies of the intersection-point in the charge-hopping model, U^{pa}(CH), and of the top of the potential hump in the charge-draining model, U^{pa}(CD), are related by an expression of the form:

$$U^{pa}(CD) = U^{pa}(CH) - \Delta\epsilon_{config} - \Gamma_{rep}/2 \qquad (1)$$

where $\Gamma_{rep}/2$ is the half-width associated with the curve repulsion shown in fig. 1b (i.e. the minimum separation of the curves is Γ_{rep}). $\Delta\epsilon_{config}$ is an energy shift result-

ing from "configuration interaction" and is defined here in such a fashion that a downwards shift would be positive. (In practice this $\Delta\epsilon_{config}$ would be a positive quantity.)

Field-desorption experiments inevitably involve a number of individual desorption events, and we may assume that the initial vibrational energy of the desorbate particle is statistically distributed. If the number of events is large enough (and also assuming that a charge-exchange process is rate-limiting), then there will be some events that proceed by a charge-hopping mechanism, and some that proceed by a charge-draining mechanism. The question thus arises: What selection of events does the experiment in question measure? An onset appearance energy, for example, might be determined by the relatively few events proceeding by a charge-draining mechanism, even though the vast majority of the events proceeded by the charge-hopping mechanism.

At the present time, the lack of reliable numerical information about the values of $\Delta\epsilon_{config}$ and Γ_{rep} (to quote but one difficulty) makes it impossible to give decisive answers to all such questions. This note thus adopts a more modest position. For each of the two mechanisms discussed here I shall assume that effectively all field desorption proceeds by the mechanism in question; certain clearly-defined theoretical consequences follow; these consequences can then be contrasted.

3. Differences between the two mechanisms

3.1. Rate-constant theory

With the charge-hopping mechanism, the transition rate-constant P^{if} associated with transfer of the desorbate particle from a given initial vibrational state i in the m-fold-charged Term to a given final vibrational state f in the n-fold-charged Term involves both an electronic transition probability and a factor relating to the vibrational motion of the desorbate particle as a whole.

With the charge-draining mechanism, the electronic changes are envisaged as happening so fast that the rate-constant P^{if} is determined *solely* by the factor relating to vibrational desorbate motion.

The existence of these two different cases was first pointed out by Gomer and Swanson [4] in their discussion of field desorption (see remarks following their eq. (49)). It was also made clear that their theoretical treatment applied to the charge-hopping situation.

However, in a subsequent textbook account (see p. 65 in ref. [12]), Müller and Tsong show in sequence: the diagram associated with the charge-draining mechanism; an equation that clearly relates to the charge-hopping mechanism (their eq. (3.28)); a definition of activation energy that appears to relate to the charge-draining mechanism (their eq. (3.29)); and further equations (their eq. (3.31) and following) that clearly relate to the charge-hopping mechanism. This somewhat unsatis-

factory summary of Gomer and Swanson's work has, in my view, greatly contributed to the confusion that persists around the charge-exchange mechanisms of field desorption.

Although there will exist differences between the detailed expressions for desorption rate-constant for the two mechanisms, the general formulation of field-desorption theory at the present time is not completely satisfactory, and in practice differences in rate-constant expressions could probably not be reliably tested.

3.2. Appearance energies

For a desorbate particle initially bound in an m-fold-charged state, that arrives at the detector in an r-fold-charged state, the *standard* appearance energy A^0_{mr} is given by [9]:

$$A^0_{mr} = H^\infty_r + (U^\infty_0 - U^{\mathrm{pa}}) , \tag{2}$$

where H^∞_r is the energy needed to form the desorbate ion, from the neutral, in remote field-free space; U^∞_0 is the energy of the neutral desorbate in remote field-free space; and U^{pa} is the standard potential energy at the top of the potential hump (or "pass") over which the desorbate escapes from the emitter. If the desorbate particles are initially in thermodynamic equilibrium with the emitter, then measured onset appearance energies should correspond fairly closely with the relevant standard appearance energy.

For the charge-hopping mechanism, U^{pa} represents the intersection-point level, as shown in fig. 1a, and hence eq. (2) gives rise to the alternative formulae:

$$A^0_{mr} = H^\infty_r + \Lambda^{\mathrm{F}}_m - Q_{mn}(\mathrm{CH}) = H^\infty_r + B^{\mathrm{cr}} , \tag{3}$$

where Λ^{F}_m represents the *total* bonding energy of the desorbate (including polarisation energy), in its m-fold-charged state, relative to the reference zero defined by the neutral desorbate in remote field-free space; $Q_{mn}(\mathrm{CH})$ is the activation energy shown in fig. 1a; and B^{cr} is the energy by which the intersection level is below the zero-level.

For the charge-draining mechanism, eq. (2) gives rise to the formula:

$$A^0_{mr} = H^\infty_r + \Lambda^{\mathrm{F}}_m - Q_{mn}(\mathrm{CD}) , \tag{4}$$

where $Q_{mn}(\mathrm{CD})$ is the activation energy shown in fig. 1b. When relationship (1) is taken into account, eq. (2) leads to the alternative formula for the charge-draining model:

$$A^0_{mr} = H^\infty_r + B^{\mathrm{cr}} + \Delta\epsilon_{\mathrm{config}} + r_{\mathrm{rep}}/2 . \tag{5}$$

All these formulae are derived directly from eq. (2), which is itself derived by means of a reversible electrothermodynamic cycle [9] that involves only basic considerations of conservation of energy. The energy relationships involved in the various formulae are shown diagrammatically in fig. 2.

In the literature, however, (most recently in ref. [13]) a formula has been given in connection with the "charge-exchange model" that has the form:

$$A^0_{mr} = H^\infty_r + \Lambda^F_m - Q_{mn} - \Delta\epsilon - \Gamma/2 \,, \tag{6}$$

where Q_{mn} is the relevant activation energy (for transition from an m-fold to an n-fold desorbate charge state). The interpretation of the symbols $\Delta\epsilon$ and Γ in this formula is discussed further below. Whatever their interpretation, though, this equation does not correspond with any of the formulae derived earlier, for either charge-hopping or charge-draining. An equation of form (6) is incompatible with conservation-of-energy arguments as associated with an electrothermodynamic cycle.

Eq. (6) was first given by Waugh [14,15]. In this work the symbols $\Delta\epsilon$ and Γ actually refer to a configuration-interaction energy shift and to curve separation, as in the earlier part of this note, but his derivation is faulty because it combines two equations one of which relates to the charge-hopping mechanism, the other to the charge-draining mechanism (see ref. [16]).

A partial derivation of a formula of form (6) is also attributed to Tsong et al. [17]. Their derivation, however, involves the concept of an effective ionization energy (for a metal atom near the emitter surface), which is said to be given by the formula (see eq. (18) in ref. [17]):

$$I_{\text{eff}} = I - \Delta\epsilon_G - \Gamma_G/2 \,. \tag{7}$$

I have added the suffix G to the symbols used here, because they refer to energy-shift and "atomic-level broadening" effects discussed by Gadzuk [18]. (These effects are physically different from the shift effects discussed earlier.) I_{eff} is the effective ionization energy, and I is the corresponding ionization energy in remote field-free space.

Now it is questionable whether eq. (7) is a correct formula for effective ionization energy, since the normal convention of theoretical physics is to represent atomic-level broadening in terms of an *imaginary* energy (and it is also questionable whether "effective ionization energy" is a useful and well-defined concept). However, these points are unimportant here, because it is clear that I_{eff} (written $I(x_c)$ in ref. [17]) *should not appear* in Tsong et al.'s eq. (17). Eq. (17) is an expression for the standard potential energy of an ion, and in the reversible thermionic cycle used to define an ionic potential-energy curve the ionization process is carried out formally in remote field-free space. Hence it is the *free-space* ionization energy I that should appear in their eq. (17). As a consequence the Gadzuk terms would not appear in their eq. (19), and the eventual result would be equivalent to eq. (3) and (4) here. The statement in ref. [17], that the Gadzuk terms appear in their eq. (19) as a result of some "quantum effect", is based on spurious reasoning.

It follows from eqs. (3) and (4) that for both of the charge-exchange mechanisms it is true that:

$$A^0_{mr} + Q_{mn} = H^\infty_r + \Lambda^F_m \,, \tag{8}$$

where Q_{mn} is the relevant activation energy. In fact this is a *thermodynamic* relationship, determined by the energetics of the situation, and it is also true for the Müller–Schottky ("image-hump") and other mechanisms of field desorption.

This means that knowledge of H_r^{∞} can be combined with measurements of appearance energy and activation energy to give information about the size and field dependence of bonding energy, as has been done by Ernst [13], without detailed knowledge as to the mechanism of field evaporation. (Other than the fact that it is thermodynamically determined, as opposed, say, to electron-impact determined.) Ernst, working with eq. (6), was obliged to assume that the $\Delta\epsilon$ and Γ terms were negligible. His work is strengthened by the proof, above, that they should not be there at all.

Small difficulties remain, since measured appearance energies and activation energies do not exactly correspond with the theoretical quantities appearing in eq. (8), but any necessary correction should be small, possibly of the order of a few kT.

Returning to the question of discriminating between the charge-hopping and charge-draining desorption mechanisms, it might seem from comparison of eqs. (3) and (5) that this might be done on the basis of appearance energy measurements. In practice, however, the a-priori calculations of the quantity B^{cr} and of the configuration-interaction terms are not sufficiently good at present. Discrimination is not possible by this means.

3.3. Ion energy distribution half-width

The most promising approach for discriminating between the charge-hopping and charge-draining mechanisms may lie in the measurement of the half-width of the ion energy distribution.

A formal wave-mechanical theory of field-ion energy distributions has recently been given by Forbes [19], in the context of the charge-hopping potential-energy diagram. This suggests that in normal circumstances a quasi-classical treatment would constitute an adequate approximation. There have been many of these (for example refs. [20–22]), mainly applied to the field ionization of imaging gas. In broad terms they suggest that, at the field strengths characteristic of the field-ion techniques and at sufficiently high temperatures (near 80 K, say), the half-width of the main peak in the ion energy distribution is basically determined by electronic considerations, and is of order 1 eV or somewhat less. For metal field evaporation the equivalent calculations do not exist, but I would expect a result similar in order of magnitude.

By contrast, with the charge-draining mechanism the distribution half-width is, at sufficiently high temperature, essentially determined by the statistical distribution of the desorbate particle initial vibrational energy (to a lesser extent by the three-dimensional geometry of the potential-energy surfaces). If the desorbate-particle assembly can be considered to be initially in thermodynamical equilibrium,

then the distribution half-width would be of order kT, which is less than 0.01 eV at 80 K.

In practice, in connection with field evaporation, the measurement of very small half-widths might be impeded both by instrumental limitations and by site-to-site variations in the level U^{Pa}(CD) of the top of the potential hump. The latter would make it difficult to sample a sufficient number of desorption events from essentially identical initial bonding sites (except, perhaps, for liquid metals, if replenishment by diffusion could occur). Nevertheless, because the predicted half-widths for the two mechanisms are substantially different, and would have different temperature dependences, this type of experiment might give rise to useful information.

3.4. Conclusions

We may conclude that, from a conceptual standpoint, the "intersecting" and "repelled" curve configurations have different desorption characteristics associated with them. Experimental discrimination may eventually be possible. There is thus good justification for considering the two curve configurations to constitute different mechanisms of field desorption, between which a distinction should be made in theoretical discussions.

References

[1] E.W. Müller, Naturwissenschaften 29 (1941) 533.
[2] E.W. Müller, Phys. Rev. 102 (1956) 518.
[3] The image-hump model provides a useful elementary mathematical formalism, but it is unwise to regard it as physically realistic.
[4] R. Gomer and L.W. Swanson, J. Chem. Phys. 38 (1963) 1613.
[5] T.T. Tsong and E.W. Müller, Phys. Status Solidi (a) 1 (1970) 513.
[6] R.G. Forbes, Surface Sci. 70 (1978) 239.
[7] N.F. Mott and H.S.W. Massey, The Theory of Atomic Collisions (Clarendon, Oxford, 1949).
[8] M.S. Child, Faraday Disc. 53 (1972) 18.
[9] R.G. Forbes, Surface Sci. 61 (1976) 221.
[10] For simplicity of discussion I make this assumption here, but it is not regarded as absolutely essential. Much of what follows would apply if the initial desorbate bonding state were well-defined electronically, but "polar" (i.e. not corresponding to any definite integral charge-number).
[11] L.D. Landau and E.M. Lifshitz, Quantum Mechanics (Pergamon, New York, 1958).
[12] E.W. Müller and T.T. Tsong, Field Ion Microscopy: Principles and Applications (Elsevier, Amsterdam, 1969).
[13] N. Ernst, Surface Sci. 87 (1979) 469.
[14] A.R. Waugh, Ph.D. thesis, Cambridge University (1975).
[15] A.R. Waugh and M.J. Southon, J. Phys. D (Appl. Phys.) 9 (1976) 1017.
[16] R.G. Forbes, J. Phys. D (Appl. Phys.) 9 (1976) L191.

[17] T.T. Tsong, W.A. Schmidt and O. Frank, Surface Sci. 65 (1977) 109.
[18] J.W. Gadzuk, Surface Sci. 6 (1967) 133.
[19] R.G. Forbes, 25th Intern. Field Emission Symp., Albuquerque, 1979;
R.G. Forbes, Progr. Surface Sci., to be published.
[20] A.J. Jason, Phys. Rev. 156 (1967) 266.
[21] J.A. Appelbaum and E.G. McRae, Surface Sci. 47 (1975) 445.
[22] G.R. Hanson and M.G. Inghram, Surface Sci. 64 (1977) 305.

A new formula for predicting low-temperature evaporation field

Richard G. Forbes
University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET, United Kingdom

(Received 3 September 1981; accepted for publication 5 November 1981)

An elementary formula is derived that enables the approximate prediction of the field strength for low-temperature field evaporation, using chemical thermodynamic data. The new formula closely resembles the familiar Müller–Schottky formula, but its derivation relies only on a general argument concerning energetics, and does not employ the physically dubious arguments concerning potential curves that are implicit in image-hump formalisms. It is suggested that the new formula could usefully supplant the Müller–Schottky formula in elementary discussions of field evaporation.

PACS numbers: 79.70. + q, 68.45.Da, 73.90. + f

There has for many years been an incompletely resolved question as to the nature of the escape mechanism in low-temperature field evaporation. Two types of mechanism have been proposed: The Müller–Schottky (or "image-hump") mechanism[1,2] in which ionization precedes escape, and charge-exchange type mechanisms (first proposed by Gomer[3,4]) in which ionization and escape occur simultaneously.

Consensus seems to exist that charge exchange is the more plausible mechanism physically (see, for example, Ref. 5), and this is supported by recent work by Forbes *et al.*[6] which shows that measurements[7] of the field sensitivity of

 0003-6951/82/030277–03$01.00

evaporation flux are not compatible with the simple formalisms normally used to analyse the Müller–Schottky mechanism.

Nonetheless, it is well known that both the charge-exchange and image-hump formalisms can produce predictions[8,9] of evaporation field strength in adequate agreement with observed values at cryogenic temperatures.

Thus the present author has long suspected that in the prediction of evaporation fields the evaporation kinetics (i.e. details of mechanism) are subsidiary, and that approximate prediction can be achieved on the basis of a more general "thermodynamic" argument using energetics alone. I show in this letter how a suitable formula can be derived.

Suppose that the field-evaporating atom is initially bound in a neutral state, with bonding energy Λ_0^F (in the presence of the field), and subsequently escapes as an n-fold ion over an activation energy barrier that is at a distance x^p from the metal's electrical surface. An expression for the work done in creating the ion can be obtained from discussion of an electrothermodynamic cycle (see, for example, Ref. 10), and leads to the expression

$$\text{work} = (\Lambda_0^F + H_n - n\phi^F) - neFx^p + \eta_n, \tag{1}$$

where H_n is the energy needed to form a n-fold ion from the neutral in remote fieldfree space, being given by the sum of the first n free-space ionization energies; ϕ^F is the local work function of the surface at which the ion is formed; F is the external field; e is the elementary (proton) charge; and η_n is the "purely chemical" component of the ion-surface interaction potential.

This work is of course equal to the activation energy for field evaporation Q. If, as a first-order approximation in Eq. (1), we replace Λ_0^F by its zero-field value Λ_0^0, and assume that the largest component in η_n is the correlation energy and that this can be represented by a classical image potential, then we obtain

$$\text{work} = Q = K_n - neFx^p - n^2e^2/16\pi\epsilon_0 x^p, \tag{2}$$

where to simplify subsequent algebra we have put

$$K_n \equiv (\Lambda_0^0 + H_n - n\phi^F). \tag{3}$$

The simplest formalisms use a "zero activation energy" definition of the evaporation field F_n^{ev} for an n-fold ion. This is satisfactory in the context of low-temperature field evaporation, and here results in the formula

$$neF_n^{ev}x_n^p = K_n - n^2e^2/16\pi\epsilon_0 x_n^p, \tag{4}$$

where x_n^p gives the point of escape for the n-fold ion. This formula is essentially the same as that derived in the context of a charge-exchange formalism (see, for example, Ref. 9).

Now the problem with formula (4) is that the value of x_n^p is difficult to decide precisely, largely because the position of the metal's electrical surface (relative to the positions of the nuclei of the metal surface atoms and the evaporating atom) is not well known. Let us therefore—and this is the new element in the discussion—circumvent the problem by merely specifying that when $Q = 0$ the correlation energy is some significant fraction α_n of the quantity K_n. That is, we put

$$n^2e^2/16\pi\epsilon_0 x_n^p = \alpha_n K_n. \tag{5}$$

It is assumed that the fraction α_n may vary somewhat with n.

It follows that

$$x_n^p = (16\pi\epsilon_0/n^2e^2)\alpha_n^{-1}K_n^{-1}. \tag{6}$$

Substituting back into Eq. (4) then gives

$$F_n^{ev} = \sigma_n(16\pi\epsilon_0/n^3e^3)K_n^2, \tag{7}$$

where the coefficient σ_n is defined by

$$\sigma_n = \alpha_n(1 - \alpha_n). \tag{8}$$

For a given value of n the only unknown in Eq. (7) is σ_n.

The convenient thing about formula (7) is that if $\alpha_n \sim \frac{1}{2}$ then σ_n is a relatively slowly varying function of α_n. So, if all that is required is approximate (say within a factor of 1.5) prediction of F_n^{ev}, then the lack of precise knowledge of α_n (or equivalently x_n^p) would be relatively unimportant. This point is illustrated in Fig. 1: a value of σ_n lying within $\pm 25\%$ of $\sigma_n = 0.2$ is given by α_n values lying in the relatively wide range $0.183 < \alpha_n < 0.817$.

I now roughly estimate σ_n values for the low-temperature field evaporation of metals. As already stated, values of x_n^p are not well known, but we can reasonably assume that x_n^p is not less than the neutral-atom radius, because the metal's electrical surface must be inside the electron charge clouds of the substrate. Most metal atoms of interest to field-ion emission have radii of around 120 pm or slightly greater, so I take this as a "low" value for x_n^p. On the other hand, Culbertson *et al.*[11] found values as high as 180 pm for the distance from the metal's electrical surface of the bonding point of a field-adsorbed helium atom; I thus take 200 pm as a "high" value for x_n^p. On further assuming that typical values of K_n for metals are $K_1 \sim 9$ eV and $K_2 \sim 20$ eV, we reach the estimates that σ_1 lies between 0.16 and 0.22, while σ_2 lies between 0.23 and 0.25.

I thus conclude that for a rough prediction of evaporation fields we may employ formula (7), using $\sigma_1 = 0.2$ and $\sigma_2 = 0.24$.

A trial of this formula is shown in Table I, for five metals selected because they cover a good range in values of both K_1 and observed evaporation field strength (F^{obs}). The values of K_1 and K_2 are derived from the data tabulation in Ref. 5, and (with the exception of Ag) the information about observed evaporation field and lowest observed evaporation charge state is also taken from Ref. 5. The Ag evaporation field comes from Ref. 12.

If we assume that the metal in question escapes in the lowest observed charge state (and any higher observed

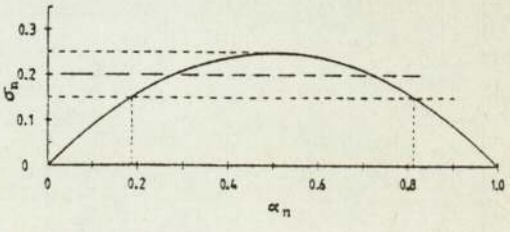


FIG. 1. Illustration of how σ_n varies as a function of α_n. Note that the range $\sigma_n = 0.2 \pm 25\%$ covers all values of α_n lying between 0.183 and 0.817.

TABLE I. Evaporation fields predicted using Eq. (7), for five selected metal species, compared with observed evaporation fields. σ_1 has been taken as 0.2 and σ_2 as 0.24.

Species	K_1 (eV)	K_2 (eV)	F_1^{ev} (V/nm)	F_2^{ev} (V/nm)	Lowest observed n value	F^{obs} (V/nm)
W	12.14	25.64	82	55	2^+	57
Ir	10.73	22.43	64	42	2^+	53
Pt	9.55	22.81	51	43	2^+	48
Fe	7.79	19.55	34	32	1^+	35
Ag	5.94	22.83	20	43	1^+	23

charge states result from postionization,[13,14] then it is clear that formula (7) succeeds in its objective of approximately predicting evaporation field strengths. (In fact there seems something of a tendency to underpredict; this could be explained by our neglect of "polarization terms," which would in effect make K_n in Eq. (7) slightly higher.)

The simple Müller–Schottky formula for F_1^{ev} can be derived from Eq. (7) by setting $\sigma_n = \frac{1}{4}$. The closeness of this value to the values of 0.2 and 0.24 used above explains the well-known success of the simple image-hump formalism in predicting evaporation fields approximately. Conceptually, however, formula (7) is superior to the Müller–Schottky formula, because the derivation of the former relies only on energetic considerations, and does not involve the highly dubious[9,15] assumptions about the shape and relative configuration of potential energy curves that are implicit in the simple image-hump formalism. Despite its known shortcomings, Müller–Schottky theory is often used in elementary discussions because of its mathematical simplicity. I submit that the derivation of Eq. (7) here is of equivalent simplicity, and that there are pedagogical advantages in preferring the new treatment.

Finally, I would stress that the new formula applies only to low-temperature field evaporation, when Q is negligibly small at the onset of significant evaporation current. (A modified treatment is necessary for high-temperature field evaporation.) And also that the new formula is intended only as an elementary approximation. For the precise estimation of evaporation fields, and/or discussion of the field sensitivity of evaporation rates, a detailed discussion based on specific potential energy curves is necessary.

[1] E. W. Müller, Phys. Rev. 102, 618 (1956).
[2] E. W. Müller, Adv. Electr. Electron. Phys. 13, 83 (1960).
[3] R. Gomer, J. Chem. Phys. 31, 341 (1959).
[4] R. Gomer, *Field Emission and Field Ionization* (Oxford University, London, 1961).
[5] T. T. Tsong, Surf. Sci. 70, 211 (1978).
[6] R. G. Forbes, R. K. Biswas, and K. Chibane, Surf. Sci. 114, 498 (1982).
[7] T. T. Tsong, J. Phys. F 8, 1349 (1978).
[8] D. G. Brandon, Surf. Sci. 3, 1 (1964).
[9] E. W. Müller and T. T. Tsong, *Field Ion Microscopy: Principles and Applications* (Elsevier, New York, 1969).
[10] R. G. Forbes, J. Phys. D 13, 1357 (1980).
[11] R. J. Culbertson, T. Sakurai, and G. H. Robertson, Phys. Rev. B 19, 4427 (1979).
[12] S. Nakamura, J. Electron. Micros. 15, 279 (1966).
[13] R. Haydock and D. R. Kingham, Phys. Rev. Lett. 44, 1520 (1980).
[14] D. R. Kingham, Surf. Sci. (to be published).
[15] T. T. Tsong, J. Chem. Phys. 54, 813 (1971).

498

Surface Science 114 (1982) 498–514
North-Holland Publishing Company

FIELD EVAPORATION THEORY: A RE-ANALYSIS OF PUBLISHED FIELD SENSITIVITY DATA

Richard G. FORBES, R.K. BISWAS and K. CHIBANE

Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET, UK

Received 2 July 1981; accepted for publication 29 September 1981

Data concerning the field sensitivity of field evaporation flux were published for six refractory metals by Tsong in 1978. These data are converted into the form of partial energies, and used in a re-examination of the consistency of published field evaporation formalisms describing the escape step. The general conclusion is that the data are not compatible with the conventional simple image-hump formalisms, even when experimental error is taken into account, but are compatible with an analysis of charge exchange based on a simple parabolic-jug formalism.

1. Introduction

Field evaporation has an established significance as the emission process in the atom-probe field-ion microscope [1,2] and related techniques, and a new importance as a possible emission mechanism in connection with liquid-metal field-ion sources [3,4].

Following recent work by Ernst [5,6] and by Kingham [7,8], it now seems virtually certain that for most metals field evaporation is a two-stage process. First a metal atom *escapes* as a singly-charged or possibly doubly-charged ion; subsequently, it may be *post-ionized* into higher charge states. Kingham's demonstration, that theoretically post-ionisation is a likely process, provides a most plausible explanation of the high charge states sometimes observed amongst field-evaporation products [9], and serves to redirect theoretical attention towards the nature of the escape process.

As is well known [10], there are two commonly discussed alternatives for the escape process; the Müller–Schottky (or "image-hump") mechanism [11,12] in which ionization precedes escape; or some form of charge-exchange process [13] in which ionization and escape occur together. These proposed mechanisms are incompatible alternatives, that lead to quantitatively different physical interpretations of basic field evaporation data [14], and it would be useful to decide firmly between them.

Detailed comparison of the mechanisms has, however, been handicapped by two things. First, the Müller–Schottky mechanism is often discussed in terms of simple image-hump formalisms that are of dubious validity [15–17] (but do

0039-6028/82/0000–0000/$02.75

apparently lead to reasonable predictions of evaporation field—see, for example, ref. [9]). Second, most mathematical analyses of the charge-exchange mechanism prior to 1978 contain a disabling mathematical flaw [14,18].

Recently, Tsong [19] has reported some measurements on the field dependence of evaporation flux ("evaporation current"). The objectives of this note are to retabulate these in a form analogous to that used by Forbes [18,20], and to explore whether the resulting data are compatible with the existing field-evaporation formalisms.

The structure of the paper is as follows. In section 2 we set down some basic theory relating to the concept of "partial energies". Section 3 converts the experimental data into the form of partial energies. Sections 4, 5 and 6 then discuss this data in the contexts of the published formalisms. Section 7 draws our conclusions together.

2. Background theory

It is first necessary to deal with some matters of notation and terminology. The amount of material evaporated per unit time from a field-ion emitter may be expressed in the form [21]:

$$J = n_{\mathrm{hr}} k_{\mathrm{hr}}, \tag{1}$$

where n_{hr} is the amount of material (or "count of atoms") at high risk of evaporation, and k_{hr} is the *field-evaporation rate-constant* for the atoms at "high risk" sites (which in practice are the kink sites). The rate-constant k_{hr} is measured in s^{-1}. The quantity n_{hr} is regarded as having the dimension [amount-of-substance], but is measured in the units "atoms" or "layers" (these being regarded as non-SI units of amount of substance). The quantity J is regarded as having the dimensions [amount-of-substance/time], is measured in layers/s (or atoms/s), and is here called the *field-evaporation flux*.

J has previously been called "evaporation current" [20], this notionally being an abbreviation for "amount-of-substance current". However, it now seems to us that the term "current" is better used to refer to the *electric* current associated with field evaporation, particularly in the context of liquid-metal field-ion sources. Thus we have adopted the term "flux" as a name for rate-like quantities having dimensions [amount-of-substance/time], and the name "field -evaporation flux" for J. In the literature, J is also called "evaporation rate", "measured evaporation rate", and "relative evaporation rate", but we prefer not to use these names.

Our system has been adopted in order to avoid ambiguity, to be consistent with the revised SI system introduced in 1971, and to keep k as having the traditional chemical meaning of "rate-constant". Note that Tsong's k is our J, and Tsong's κ is our k_{hr}.

The three parameters in eq. (1) are all functions of the field strength F at the

high-risk sites. Consider some reference situation, in which F has the value F^0, and let the corresponding values of these parameters be J^0, n_{hr}^0 and k_{hr}^0. It follows that:

$$\ln(J/J^0) = \ln(n_{hr}/n_{hr}^0) + \ln(k_{hr}/k_{hr}^0). \tag{2}$$

Each of the terms in eq. (2) has value zero at $F = F^0$, and it is convenient to Taylor-expand them about F^0. If we introduce a variable g by:

$$g = (F - F^0)/F^0, \tag{3}$$

then the expansion of the rate-constant term (for example) can be put in the form:

$$\ln(k_{hr}/k_{hr}^0) = (\mu_1/kT)g + \tfrac{1}{2}(\mu_2/kT)g^2 + \ldots. \tag{4}$$

The (μ_1/kT), (μ_2/kT), etc., are the Taylor coefficients, which are written in this way because the quantities μ_1, μ_2 are expected to be temperature independent if field evaporation is a thermally-activated process obeying the Arrhenius equation.

By analogy with eq. (4), we may expand the other terms in a similar way, thus:

$$\ln(J/J^0) = (\lambda_1/kT)g + \tfrac{1}{2}(\lambda_2/kT)g^2 + \ldots, \tag{5}$$

$$\ln(n_{hr}/n_{hr}^0) = (\nu_1/kT)g + \tfrac{1}{2}(\nu_2/kT)g^2 + \ldots. \tag{6}$$

Clearly we have:

$$\lambda_1 = \mu_1 + \nu_1, \quad \lambda_2 = \mu_2 + \nu_2. \tag{7}$$

Existing theories of field evaporation are, of course, formalisms that deal with rate-constant field sensitivity rather than evaporation-flux field sensitivity. Each formalism provides an expression for the field-evaporation activation energy $Q(F)$; expressions for the "partial energies" μ_1 and μ_2 can then be obtained as follows [18,20]. The Arrhenius equation may be written in the form:

$$k_{hr} = \exp(M/kT), \tag{8}$$

where

$$M = kT \ln\{A\} - Q, \tag{9}$$

where $\{A\}$ is the numerical value of the field evaporation pre-exponential A, expressed in s^{-1}. It follows that:

$$\ln(k_{hr}/k_{hr}^0) = [M - M(F^0)]/kT. \tag{10}$$

Taylor-expanding M in terms of F, about $F = F^0$, we obtain:

$$M = M(F^0) + (F - F^0)\left.\frac{dM}{dF}\right|_{F=F^0} + \tfrac{1}{2}(F - F^0)^2\left.\frac{d^2M}{dF^2}\right|_{F=F^0} + \ldots. \tag{11}$$

Substituting from eq. (3) and then substituting into eq. (10) gives:

$$\ln\left(k_{\mathrm{hr}}/k_{\mathrm{hr}}^{0}\right)=\left(F^{0}g/kT\right)\left.\frac{\mathrm{d}M}{\mathrm{d}F}\right|_{F^{0}}+\tfrac{1}{2}\left[\left(F^{0}\right)^{2}g^{2}/kT\right]\left.\frac{\mathrm{d}^{2}M}{\mathrm{d}F^{2}}\right|_{F^{0}}+\ldots. \tag{12}$$

Hence, by comparison with eq. (4):

$$\mu_{1}=F^{0}\left.\frac{\mathrm{d}M}{\mathrm{d}F}\right|_{F^{0}}\simeq-F^{0}\left.\frac{\mathrm{d}Q}{\mathrm{d}F}\right|_{F^{0}}, \tag{13a}$$

$$\mu_{2}=\left(F^{0}\right)^{2}\left.\frac{\mathrm{d}^{2}M}{\mathrm{d}F^{2}}\right|_{F^{0}}\simeq-\left(F^{0}\right)^{2}\left.\frac{\mathrm{d}^{2}Q}{\mathrm{d}F^{2}}\right|_{F^{0}}. \tag{13b}$$

The second step in each of these relationships assumes that the field dependence of the pre-exponential A may be neglected. Previous analyses, for example refs. [14,17,18], have taken this to be an adequate approximation, particularly at temperatures near 80 K and above, and we continue with this assumption. The arguments principally involved are as follows. First, the pre-exponential A involves a factor relating to the vibrational frequency of the bound neutral. This frequency depends on the shape of the neutral potential well, which (in comparison with the activation energy) is unlikely to be sensitively affected by a small field variation. Second, the pre-exponential involves a factor that relates to the possibility of "ion tunneling" [16] and is in principle field-dependent for this reason. It has always been supposed that, at temperatures near 80 K and above, any field dependence due to this cause is insignificant.

Neglect of posssible field dependence in the pre-exponential A is, in practice, an integral part of the formalism under discussion in later sections. The question of whether this neglect is really justified is far from straightforward and would require detailed mathematical analysis that regretfully is beyond the scope of the present paper.

3. Experimental data

3.1. Conversion of data

Tsong's experimental results [19] are in the form of a plot of $\lg(J/J^{0})$ versus F/F^{0}, where F^{0} is the field at which the evaporation flux is 10^{-2} layers/s. He seems to have treated this data by making the (perfectly reasonable [22]) assumption that $\lg(n_{\mathrm{hr}}/n_{\mathrm{hr}}^{0})$ is effectively zero, and hence that:

$$\lg\left(J/J^{0}\right)=\lg\left(k_{\mathrm{hr}}/k_{\mathrm{hr}}^{0}\right)=a_{0}+a_{1}\left(F/F^{0}\right)+a_{2}\left(F/F^{0}\right)^{2}. \tag{14}$$

He does not explicitly mention any assumption equivalent to taking $\lg(n_{\mathrm{hr}}/n_{\mathrm{hr}}^{0})$ zero, but he does state that a_{1} and a_{2} are best-fit coefficients to the experimental data, so we shall assume that the correct logical identification between his

parameters and ours is:

$$\lambda_1 = \ln(10) \times kT \cdot (a_1 + 2a_2), \tag{15a}$$

$$\lambda_2 = 2\ln(10) \times kT \cdot a_2. \tag{15b}$$

For convenience in dealing with errors, it is useful to define working parameters α_1, α_2 by:

$$\alpha_1 = \ln(10) \times kT \cdot a_1, \tag{16a}$$

$$\alpha_2 = \ln(10) \times kT \cdot a_2. \tag{16b}$$

Tsong's values of a_1 and a_2, together with corresponding values of α_1 and α_2 are shown in table 1; T has been put equal to 77.3 K.

Existing theories of field evaporation deal with rate-constant field sensitivity rather than evaporation-flux field sensitivity, so we require μ_1 and μ_2 rather than λ_1 and λ_2. We shall thus follow Tsong and other workers (for example, Brandon [15]) in assuming that there is no significant field dependence in the amount of material at high risk of evaporation. This implies that ν_1 and ν_2 can be neglected, and hence that experimental estimates of partial energies are given by:

$$\mu_1(\text{exp}) \simeq \lambda_1(\text{exp}) = \alpha_1 + 2\alpha_2, \tag{17a}$$

$$\mu_2(\text{exp}) \simeq \lambda_2(\text{exp}) = 2\alpha_2. \tag{17b}$$

Partial energies derived via eq. (17) are shown in table 2, together with corresponding values of their ratio. For simplicity in later argument, we label the columns here and later in terms of μ_1, μ_2 and derived expressions. However, it should be remembered that strictly (except in the case of table 3) the columns show experimental estimates of λ_1, λ_2 and derived expressions.

For comparison with the tungsten estimates in table 2, we show in table 3 values of μ_1^c, μ_2^c and μ_2^c/μ_1^c, derived by Forbes and Patel [23,24] from Tsong's measurements [17] on the field sensitivity of evaporation rate-constant, in the

Table 1
Values of F^0, a_1 and a_2, as given in ref. [19], together with working parameters α_1 and α_2 derived via eq. (16); the error in a_1 is not explicitly stated in ref. [19], but is assumed here to be ± 5 in each case; for working purposes, the values and errors for α_1 and α_2 are given to more places than are physically significant

Species	F^0 (V/nm)	a_1	a_2	α_1 (eV)	α_2 (eV)
Mo(110)	46	715±5	−310±10	10.967±0.077	−4.754±0.153
Ru(11$\bar{2}$1)	42	200±5	−75±20	3.068±0.077	−1.150±0.307
Hf(10$\bar{1}$0)	40	335±5	−135±20	5.138±0.077	−2.071±0.307
W(110)	55	805±5	−350±10	12.347±0.077	−5.368±0.153
Ir(100)	52	205±5	−80±20	3.144±0.077	−1.227±0.307
Pt(100)	48	200±5	−75±20	3.068±0.077	−1.150±0.307

Table 2
Experimental estimates of the partial energies μ_1 and μ_2, and their ratio, derived from the data in table 1, using eq. (17); it has been assumed that the quantities ν_1 and ν_2 appearing in eq. (7) are negligibly small

Species	μ_1 (eV)	μ_2 (eV)	μ_2/μ_1
Mo(110)	1.46 ± 0.4	-9.51 ± 0.3	-6.5 ± 2
Ru(11$\bar{2}$1)	0.77 ± 0.7	-2.30 ± 0.6	-3.0 ± 3
Hf(10$\bar{1}$0)	1.00 ± 0.7	-4.14 ± 0.6	-4.2 ± 3.5
W(110)	1.61 ± 0.4	-10.74 ± 0.3	-6.7 ± 2
Ir(100)	0.69 ± 0.7	-2.45 ± 0.6	-3.6 ± 3.5
Pt(100)	0.77 ± 0.7	-2.30 ± 0.6	-3.0 ± 3

case of tungsten. (The suffix c on the partial energies indicates that these relate to an evaporation field defined by a unity-rate-constant criterion [20] rather than the evaporation-flux criterion used in table 2.)

Bearing in mind that different criteria may lead to different field values being identified as the "evaporation field" [20], and that the value of μ_1 will depend on the choice of "evaporation field", we feel that the comparison of table 3 (based on rate-constant field sensitivity) with the tungsten estimates in table 2 (based on flux field sensitivity) is entirely satisfactory, and goes some way towards confirming that ν_1 and ν_2 can be neglected in the context of low-temperature field evaporation.

3.2. Treatment of errors

Some comment on our treatment of errors is necessary. In deriving the errors on the μ's from those on the α's we have used simple linear formulae. For example, the standard deviation $\sigma(\mu_1)$ is obtained from the formula, based on eq. (17a).

$$\sigma(\mu_1) = \sigma(\alpha_1) + 2\,\sigma(\alpha_2). \tag{18}$$

Table 3
Experimental estimates of the partial energies μ_1^c and μ_2^c, and their ratio, derived by Patel [23] from original data concerning rate-constant field sensitivity taken by Tsong [17]; in the analysis of the fig. 9 data, certain deviant points have been excluded from the analysis on statistical grounds; the original data were obtained from pulse-type experiments on the field evaporation of tungsten adatoms from a tungsten substrate

Data source	μ_1^c (eV)	μ_2^c (eV)	μ_2^c/μ_1^c
W, ref. [17], fig. 9	1.43 ± 0.07	-12 ± 2	-8.5 ± 1
W, ref. [17], fig. 10	1.31 ± 0.08	-9 ± 3	-7 ± 3

This simple procedure will usually give an error estimate larger than any other reasonable procedure, and — as will become clear later — is adequate for the purposes of this paper.

But in reality this procedure is not statistically correct, because α_1 and α_2 are not statistically independent variables. A better estimate of error would be obtained from:

$$\sigma^2(\mu_1) = \sigma^2(\alpha_1) + 4\,\sigma^2(\alpha_2) + 4\rho_{12}\,\sigma(\alpha_1)\,\sigma(\alpha_2), \tag{19}$$

where ρ_{12} is the correlation coefficient between α_1 and α_2.

Information about the covariance of a_1 and a_2 is not available in ref. [19], and to obtain such information it would be necessary to carry out new regression calculations on the original data sets.

New regression calculations should result in estimates of error lower than those given in table 2. However, because our objective is to test formalisms, it has not been found necessary to perform these calculations. A formalism that fails when errors are overestimated would fail when the errors were correctly estimated; on the other hand, if derived parameter values are "reasonable", then the error estimate is of lesser significance.

In deriving the error estimates in tables 4 to 6 simple methods have been used, and in consequence the errors are probably overestimated. For this reason, in some places parameter values are given to more significant figures than the stated error limits would seem to justify.

Finally, we would point out that if the temperature of the emitter tip is in fact higher than the boiling point of liquid nitrogen (77.3 K), then there will be a small systematic error in some of the parameter values stated.

4. Comparison with theory: I. Tsong's formula

In his paper, Tsong [19] has interpreted his coefficients a_1 and a_2 in terms of an analysis of the charge-exchange mechanism that he gave some years earlier [17]. There is, however, a mathematical flaw in this earlier analysis. This arises in the differentiation of the "pseudolinear" term in F, in the expression for activation energy. The correct expression [14,18] is:

$$\frac{\mathrm{d}}{\mathrm{d}F}(neFx^{\mathrm{p}}) = nex^{\mathrm{p}} + neF\frac{\mathrm{d}x^{\mathrm{p}}}{\mathrm{d}F}, \tag{20}$$

where e is the elementary (proton) charge, ne is the charge on the ion, and x^{p} is the distance of the "crossing point" (of neutral and ionic standard potential curves) from the emitter's electrical surface.

Tsong has neglected the second term in eq. (20), apparently thinking it to be relatively small [17]. But in reality the second term is almost equal to the first, but opposite in sign. A corrected analysis [18], assuming the bound neutral to be moving in a parabolic potential well, with force-constant κ, leads to the

result (see appendix):

$$\frac{\mathrm{d}}{\mathrm{d}F}(neFx^{\mathrm{p}}) \simeq \kappa r x^{\mathrm{p}}/F, \tag{21}$$

where r is the distance of the crossing-point from the bottom of the well. This term is smaller than nex^{p} by about an order of magnitude.

The physical implication of this flaw is that Tsong's derivation of values of "polarisability" and of "$x_c + \lambda$" from his tabulated regression data is meaningless. His values may be disregarded.

5. Comparison with theory: II. Simple image-hump formalisms

5.1. Data manipulation

With the simple image-hump formalism, it may be shown by elementary analysis that, to second order in F, the activation energy $Q(F)$ for the field evaporation of an atom bound as a neutral is given by:

$$Q(F) = \left(\Lambda^0 + H_n - n\phi^E\right) - \left(n^3e^3F/4\pi\epsilon_0\right)^{1/2} + \tfrac{1}{2}(c_0 - c_n)F^2, \tag{22}$$

where H_n is the energy needed to form a n-fold ion, from the neutral, in remote field-free space, being given by the sum of the first n ionization energies; Λ^0 is the binding energy of the neutral to the emitter surface, in the absence of the field; ϕ^E is the relevant local work-function of the emitter; and ϵ_0 is the electric constant. c_0 is a parameter associated with polarisation of the neutral in an external field F (or, possibly, with partial charge transfer from evaporating atom to emitter [25]), and c_n is the corresponding parameter for an n-fold ion. Note that, although these coefficients are sometimes called "polarisability", neither coefficient is a polarisability in the sense of the ordinary text-book definition [26].

Using eq. (20) in eqs. (13a) and(13b), we obtain:

$$\mu_1 = \tfrac{1}{2}\left(n^3e^3F^0/4\pi\epsilon_0\right)^{1/2} - (c_0 - c_n)F^{02}, \tag{23}$$

$$\mu_2 = -\tfrac{1}{4}\left(n^3e^3F^0/4\pi\epsilon_0\right)^{1/2} - (c_0 - c_n)F^{02}, \tag{24}$$

$$\mu_2/\mu_1 = -\tfrac{1}{2}(1+\theta), \tag{25}$$

where θ is a "polarisation correction factor" given by:

$$\theta = 6(c_0 - c_n)F^{02}/\left[\left(n^3e^3F^0/4\pi\epsilon_0\right)^{1/2} - 2(c_0 - c_n)F^{02}\right]. \tag{26}$$

By further manipulation of eqs. (23) and (24) we obtain:

$$\left(n^3e^3F^0/4\pi\epsilon_0\right)^{1/2} = \tfrac{4}{3}(\mu_1 - \mu_2) \simeq \tfrac{4}{3}\alpha_1, \tag{27}$$

$$\tfrac{1}{2}(c_0 - c_n)F^{02} = -\tfrac{1}{6}(\mu_1 + 2\mu_2) \simeq -(\alpha_2 + \alpha_1/6). \tag{28}$$

Finally, by substituting eqs. (27) and (28) back into eq. (22), we obtain the consistency relationship, at field F^0:

$$K_n - Q(F^0) = \tfrac{3}{2}\mu_1 - \mu_2 \simeq \tfrac{3}{2}\alpha_1 + \alpha_2, \tag{29a}$$

where, for notational convenience, we have defined a new symbol K_n by:

$$K_n = \Lambda^0 + H_n - n\phi^E. \tag{29b}$$

5.2. Detailed comparisons

The simplest version of the image-hump formalism ignores the F^2 ("polarisation") terms in eqs. (22) to (25). If this were a reasonable approximation, then we should expect $\mu_2/\mu_1 \simeq -\frac{1}{2}$. Inspection of table 2, however, shows that experimental estimates (ignoring error limits) lie between −3 and −7.

Because of the relatively wide error limits stated for μ_2/μ_1, it is only in the cases of Mo and W that it is possible to say decisively that the individual experimental results are incompatible with the simple no-polarisation formalism. This incompatibility can also be illustrated graphically. From earlier equations it follows that:

$$\lg(J/J^0) \simeq \left[Q(F^0) - Q(F)\right]/kT\ln(10). \tag{30}$$

Hence, on using eq. (22), but ignoring the F^2 term, we obtain:

$$\lg(J/J^0) \simeq B\left[\left(F/F^0\right)^{1/2} - 1\right], \tag{31a}$$

where B is a constant defined by

$$B = \left(n^3e^3F^0/4\pi\epsilon_0\right)^{1/2}/kT\ln(10). \tag{31b}$$

Fig. 1 shows Tsong's results [19] for tungsten, replotted with a superimposed theoretical curve obtained by adjusting B to fit the observed slope at $F/F^0 = 1$. The discrepancy is obvious and it also deserves note that the F^0 values derived from eq. (31b), using the observed B, are much less than the observed field. (The derived values are: for $n = 1$, $F^0 = 7$ V/nm; for $n = 2$, $F^0 = 0.9$ V/nm: the observed value is 55 V/nm.) When the observed field is used to plot a theoretical curve (assuming $n = 2$), the discrepancy is even more obvious.

If all the six ratio estimates are taken together as a group a clear result is also achieved. Because the simple formalism, without F^2 terms, makes a prediction that is independent of species and of charge-state at escape, the six ratio values in table 2 can be taken as six independent estimates of a quantity whose value is known. From these estimates (ignoring the a-priori errors) we may deduce that $\mu_2/\mu_1(\text{exp}) = -4.48 \pm 0.69$. Using a Student t test, it is found that even at the 0.5% level there is a significant difference between this "experimental" mean and the "true" mean of −0.5.

Thus there is at least a 99.5% probability that these experimental results are not compatible with the simple formalism that neglects the F^2 terms.

The next simplest image-hump formalism, in which the F^2 terms are included, as in eqs. (22) to (25), has been widely used in the literature. A test of its self-consistency may be based on eq. (29). Values of the l.h.s. of this

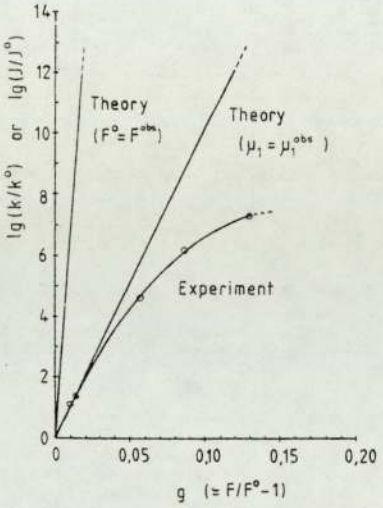


Fig. 1. A comparison of Tsong's (1978) experimental field sensitivity data [19] for tungsten with the predictions of the simple image-hump (no F^2 terms) formalism. Two theoretical comparisons are shown: one in which the parameter F^0 in eq. (31b) is taken as equal to the observed field, with $n=2$ assumed; the other in which the parameter F^0 is adjusted so that the theoretical curve has the same slope at the origin as does the experimental curve. (In making these comparisons it has been assumed that the quantity ν is effectively constant.)

equation, shown in table 4 for $n = 1$ to 3, are obtained by: (a) setting $Q(F^0) = 0.2$ eV; (b) using the ϕ^E values shown, which have been taken from refs. [27] and [28]; (c) using values of Λ^0 and ionization energy I_n taken from ref. [9]. Values of the r.h.s. are obtained from the data in table 1.

Inspection of table 4 shows that in all cases, in the framework of the formalism in use, the hypothesis that $n > 1$ is wildly improbable. Even in the $n = 1$ case the difference between the l.h. and r.h. sides of the equation is at least 7 times the standard derivation of the experimental estimate; so, statistically, we have negligible probability that the two sides of the equation are equal within experimental error. That is, the experimentally-derived data are not compatible with the image-hump formalism under discussion.

A further test of the simple image-hump formalism (including F^2 terms) can be based on eq. (27). Values of the evaporation field F^0 predicted from the field-sensitivity data, for $n = 1$ to 3, are shown in table 5. These can be compared with observed values of evaporation field as noted in ref. [19].

Even allowing for 30% error in the "observed" field values, there is only one

Table 4
Self-consistency test of the simple image-hump formalism (including F^2 terms), based on eq. (29); values of the l.h.s. of this equation have been derived as explained in the text, and are shown for values of $n=1$ to 3; values of the r.h.s. of the equation are derived from the data in table 1; the last column expresses the difference between the l.h.s. (for $n=1$) and the r.h.s., in terms of the standard deviation σ for the r.h.s., in each case

Species	ϕ^E (eV)	$K_n - Q(F^0)$ (eV)			$\frac{1}{2}\alpha_1 + \alpha_2$ (eV)	Deviation
		$n=1$	$n=2$	$n=3$		
Mo(110)	5.12	8.59	19.6	41.7	11.7±0.3	12σ
Ru(11$\bar{2}$1)	4.86	8.93	20.8	44.4	3.5±0.4	13σ
Hf(10$\bar{1}$0)	3.65	9.50	20.8	40.4	5.6±0.4	9σ
W(110)	5.14	11.30	24.2	43.0	13.2±0.3	7σ
Ir(100)	5.27	10.56	22.3	44.0	3.5±0.4	7σ
Pt(100)	5.84	8.81	21.5	43.7	3.5±0.4	13σ

case (Hf) in which the observed value is compatible with any value predicted from the field-sensitivity data using an image-hump formalism.

5.3. Conclusion

From these comparisons we may conclude with considerable confidence that, if Tsong's data [19] concerning evaporation-flux field sensitivity may be taken as good estimates of the corresponding data concerning evaporation rate-constant field sensitivity, then these data are not compatible with the conventional simple image-hump formalisms used to analyse the Müller–Schottky mechanism of field evaporation.

Table 5
Values of the evaporation field F^0, as derived from eq. (27), using the data in table 1, for values of $n=1$ to 3; the last column gives the direct experimental estimate of F^0, at liquid-nitrogen temperature, made by Tsong [19]

Species	Derived field (V nm^{-1})			Observed field (V nm^{-1})
	$n=1$	$n=2$	$n=3$	
Mo(110)	148 ±2	18.6 ±0.3	5.49±0.08	46
Ru(11$\bar{2}$1)	11.6±0.6	1.45±0.07	0.43±0.02	42
Hf(10$\bar{1}$0)	32.6±1	4.1 ±0.1	1.21±0.04	40
W(110)	188 ±3	23.5 ±0.3	6.97±0.08	55
Ir(100)	12.2±0.6	1.53±0.07	0.45±0.02	52
Pt(100)	11.6±0.6	1.45±0.07	0.43±0.02	48

Obviously, this result has been proved only in the six cases for which data have been taken. However, there is nothing particularly "special" about these six cases, and the presumption must be that these image-hump formalisms are not generally appropriate for metal field evaporation.

The inability of simple image-hump formalisms to explain field-sensitivity data has a further implication. It suggests that the use of such formalisms to predict evaporation fields (by setting Q in eq. (22) equal to zero) is without proper scientific legitimacy — even though this procedure leads to reasonable agreement with experiment. Closer investigation [29] in fact shows that equally good agreement can be achieved on the basis of a simple general argument concerning the energetics of field evaporation; it seems that ability of a formalism to approximately predict evaporation field is not a useful test of the corresponding mechanism of field evaporation.

6. Comparison with theory: III. Parabolic jug formalism

6.1. Derivation of formulae

By working with a formalism in which the neutral atom is assumed to be bound in a parabolic potential well, of force-constant κ, Forbes [18] has been able to derive expressions for partial energies that can be put in the following form:

$$\mu_1 = \kappa r^0(a + r^0) = 2Q(1 + a/r^0), \tag{32}$$

$$\mu_2 = -\mu_1(2 + a/r^0), \tag{33}$$

where a is the distance of the neutral-atom bonding point from the metal's electrical surface, and r^0 is the distance of the "point of escape" (at the crossing-point of neutral and ionic potential-energy curves) from the neutral-atom bonding point, at the evaporation field F^0 in question. These expressions are derived from eqs. (43) and (52) in ref. [18]. Certain inessential mathematical approximations have been made in the derivation of eqs. (32) and (33), and hence these equations constitute a basic version of the formalism.

Ref. [18] in fact works in a context where the evaporation field is defined by a unity-rate-constant criterion, but the same expressions are relevant to the present context (where the evaporation field is defined by an evaporation-flux criterion), although obviously μ_1, μ_2 and Q have slightly different meanings.

By algebraic manipulation of eqs. (32) and (33) we may derive expressions for Q and for r^0/a as follows:

$$Q = -\tfrac{1}{2}\mu_1^2/(\mu_1 + \mu_2), \tag{34}$$

$$r^0/a = -\mu_1/(2\mu_1 + \mu_2). \tag{35}$$

Note that all the expressions derived in this section are independent of the charge state of the escaping ion.

6.2. Detailed comparisons

The experimental estimates of μ_1 and μ_2, listed in tables 2 and 3, can be used in various ways. In the original treatment in ref. [18], the Arrhenius equation was used to make an a-priori estimate of Q as 0.2 eV, and this was then combined with the experimental estimate of μ_1 to make an estimate of a/r^0 and hence a semi-theoretical estimate of μ_2/μ_1. This semi-theoretical estimate compared favourably with the experimental estimate derived from the data in table 3.

An alternative approach is used here. Table 6 shows values of Q and r^0/a derived from the data in table 2, using eqs. (34) and (35), and the question is whether the derived values are "physically reasonable".

An a-priori estimate of activation energy may be made by combining eq. (1) with the Arrhenius equation:

$$k_{\mathrm{hr}} = A \exp(-Q/kT), \tag{36}$$

and taking: $J = 0.01$ layer/s, $n_{\mathrm{hr}} = 0.01$ layers (which implies $k_{\mathrm{hr}} = 1\ \mathrm{s}^{-1}$); $T = 77.3$ K. This leads to the result: $Q = 0.18$ eV.

If the field-evaporation pre-exponential A were "anomalously" low, which has been suggested for some materials in the past [30,31], then slightly lower values of Q would be expected. For example, taking $A = 10^8\ \mathrm{s}^{-1}$ would lead to the result $Q = 0.12$ eV.

The experimental estimates of Q are clearly compatible with these theoretical estimates, to within the limits of experimental error, and would still be compatible with the theoretical estimates if the limits of error were substantially reduced. Thus the activation-energy estimates are certainly "reasonable".

It remains to consider the experimental estimates of r^0/a, which range from 0.2 to 1.0. The values for Mo, W, and Hf are less than 0.5 and we feel these to be entirely reasonable. The remaining values, particularly the Pt and Ru values (both 1.0), are higher than one would perhaps expect; however, the error limits on these particular estimates are very wide, and in consequence the stated

Table 6
Values of the activation energy $Q(F^0)$ and the ratio r^0/a, as derived from eq. (34) and eq. (35), using the data in table 2

Species	Q (eV)	r^0/a
Mo(110)	0.13 ± 0.06	0.22 ± 0.02
Ru($11\bar{2}1$)	0.20 ± 0.2	1.0 ± 1
Hf($10\bar{1}0$)	0.16 ± 0.15	0.46 ± 0.1
W(110)	0.15 ± 0.06	0.21 ± 0.02
Ir(100)	0.14 ± 0.15	0.64 ± 0.6
Pt(100)	0.20 ± 0.2	1.0 ± 1

values of r^0/a could well be somewhat higher than the true values; so we do not regard these high estimated values as casting doubt on the formalism.

In general, we conclude that a charge-exchange mechanism, in the form of a basic version of the parabolic-jug formalism, is apparently well able to provide a physically satisfactory explanation of partial energies derived from flux field-sensitivity data.

At the same time, it should be realized that — because this paper employs only the basic version of the formalism — the tabulated values of Q and r^0/a are not necessarily the best estimates that could be derived from the data.

7. General discussion

The conclusions of this paper need to be approached with care. Our immediate objective has been, not to decide the mechanism of metal field evaporation, nor to derive reliable atomic-level information from published experimental data, but the more modest one of testing the consistency of experimental data with the existing published theoretical formalisms.

What we have found is that the evaporation-flux field-sensitivity data are not compatible with existing image-hump formalisms, but are compatible with a charge-exchange formalism.

Because the existing image-hump formalisms are manifestly incomplete — for example, past numerical analyses do not include any term relating to the repulsive ion-metal interaction that must physically be present — it is logically inappropriate to conclude from the present discussion that the field-sensitivity data are necessarily incompatible with the Müller–Schottky mechanism. For example, the proposition might be advanced that the demonstrated inconsistencies could be removed by including a repulsive term, or by taking a more sophisticated approach to the calculation of the correlation interaction. Or it might be argued that the inconsistencies result from the neglect of field dependence in the pre-exponential. But, intuitively, these propositions look unpromising, particularly when set against the demonstrated compatibility of the field-sensitivity data with a simple charge-exchange formalism.

Charge-exchange thus looks much more likely than the Müller–Schottky mechanism. In view of often-expressed past doubts about the self-consistency of the simple image-hump formalisms, this conclusion is unsurprising. The new feature of the present discussion is that the incompatibilty of experimental data with these image-hump formalisms has been established by means of statistical tests.

Appendix. Differentiation of the term pseudo-linear in F

There appears in the expression for field-evaporation activation energy Q, an "electrostatic" term (denoted here by v) that has the form:

$$v = neFx^p, \tag{37}$$

where x^{p} is the distance of the "point of escape" from the metal's electrical surface, and the other symbols have their usual meanings. In the context of a charge-exchange formalism, x^{p} corresponds to the crossing-point of neutral and ionic potential energy curves.

The expression for the total differential of Q with respect to F has $(-\mathrm{d}v/\mathrm{d}F)$ as one of its terms. Because x^{p} is itself a function of F, the total differential $\mathrm{d}v/\mathrm{d}F$ is given by:

$$\frac{\mathrm{d}v}{\mathrm{d}F} = \left(\frac{\partial v}{\partial F}\right)_{x^{\mathrm{p}}} + \left(\frac{\partial v}{\partial x^{\mathrm{p}}}\right)_{F} \frac{\mathrm{d}x^{\mathrm{p}}}{\mathrm{d}F}. \tag{38}$$

That is

$$\frac{\mathrm{d}v}{\mathrm{d}F} = nex^{\mathrm{p}} + neF\frac{\mathrm{d}x^{\mathrm{p}}}{\mathrm{d}F}. \tag{39}$$

As stated in section 4, the second term in this expression has often been neglected. In the context of a model in which the neutral atom moves in a parabolic potential-energy well, an expression for this second term, and hence for $\mathrm{d}v/\mathrm{d}F$, can be derived as follows. (The treatment is based on that in ref. [18].)

The field-evaporation activation energy can be written in the alternative forms:

$$Q = \tfrac{1}{2}\kappa r^2 = \left(\Lambda^0 + H_n - n\phi^{\mathrm{E}}\right) - neFx^{\mathrm{p}} + P(F) + C(x^{\mathrm{p}}), \tag{40}$$

where r is the distance of the point-of-escape (the crossing point) from the bottom of the potential well, κ is the force constant for the well, P(F) is a (primarily) field-dependent correction term relating to the polarisation of the neutral and the ion, or equivalent effects, and $C(x)$ is a (primarily) distance-dependent function describing an interaction between the ion and the surface due to correlation and (in principle) repulsive forces. $C(x)$ is usually approximated by the image-potential. The other symbols have their usual meanings, as listed in section 5.1.

Taking differentials we obtain:

$$\kappa r\,\mathrm{d}r = nex^{\mathrm{p}}\,\mathrm{d}F + (\mathrm{d}P/\mathrm{d}F)\,\mathrm{d}F - neF\,\mathrm{d}x^{\mathrm{p}} + (\mathrm{d}C/\mathrm{d}x^{\mathrm{p}})\,\mathrm{d}x^{\mathrm{p}}. \tag{41}$$

The relationship of r and x^{p} is:

$$x^{\mathrm{p}} = a + r. \tag{42}$$

It is a good approximation to assume that a, the distance of the well bottom from the metal's electrical surface, is independent of field. Hence $\mathrm{d}r = \mathrm{d}x^{\mathrm{p}}$ and we obtain:

$$\mathrm{d}x^{\mathrm{p}}/\mathrm{d}F = -(nex^{\mathrm{p}} - \mathrm{d}P/\mathrm{d}F)/(neF + \kappa r - \mathrm{d}C/\mathrm{d}x^{\mathrm{p}}). \tag{43}$$

As a first approximation we may neglect the polarisation and correlation + repulsion terms. Thus

$$\mathrm{d}x^{\mathrm{p}}/\mathrm{d}F \simeq -nex^{\mathrm{p}}/(neF + \kappa r). \tag{44}$$

Substituting into eq. (39) then gives

$$dv/dF \simeq nex^{p}\left[1 - neF/(neF + \kappa r)\right] = nex^{p}\kappa r/(neF + \kappa r). \quad (45)$$

At this stage we can neglect κr in comparison with neF; the basic result stated in section 4 follows:

$$dv/dF \simeq \kappa r x^{p}/F. \quad (46)$$

Obviously, higher-order approximations can be obtained by including the omitted terms, but these do not dramatically affect the result, numerically.

Acknowledgement

This work forms part of a research project funded by the UK Science and Engineering Research Council. In addition, one of us (K.C.) wishes to thank the Ministry of Higher Education and Scientific Research of the Republic of Algeria for personal financial support.

References

[1] E.W. Müller, J.A. Panitz and S.B. McLane, Rev. Sci. Instr. 39 (1968) 83.
[2] J.A. Panitz, CRC Critical Rev. Solid State Sci. 5 (1975) 153.
[3] R. Clampitt, Nucl. Instr. Methods 189 (1981) 111.
[4] L.W. Swanson, in: Microcircuit Engineering '80, Amsterdam, 1980.
[5] N. Ernst, Surface Sci. 87 (1979) 469.
[6] N. Ernst and G. Bozdech, in: Proc. 27th Intern. Field Emission Symp., Tokyo, 1980.
[7] R. Haydock and D.R. Kingham, Phys. Rev. Letters 44 (1980) 1520.
[8] D.R. Kingham, Surface Sci. 108 (1981) L460.
[9] T.T. Tsong, Surface Sci. 70 (1978) 211.
[10] E.W. Müller and T.T. Tsong, Progr. Surface Sci. 4 (1974) 1.
[11] E.W. Müller, Phys. Rev. 102 (1956) 618.
[12] E.W. Müller, Advan. Electron. Electron Phys. 13 (1960) 83.
[13] R. Gomer and L.W. Swanson, J. Chem. Phys. 38 (1963) 1613.
[14] D. McKinstry, Surface Sci. 39 (1972) 37.
[15] D. Brandon, Brit. J. Appl. Phys. 14 (1963) 474.
[16] E.W. Müller and T.T. Tsong, Field Ion Microscopy: Principles and Applications (Elsevier, Amsterdam, 1969).
[17] T.T. Tsong, J. Chem. Phys. 54 (1971) 813.
[18] R.G. Forbes, Surface Sci. 70 (1978) 239.
[19] T.T. Tsong, J. Phys. F (Metal Phys.) 8 (1978) 1349.
[20] R.G. Forbes, Surface Sci. 46 (1974) 577.
[21] R.G. Forbes, in: Proc. 7th Intern Vacuum Congr. and 3rd Intern. Conf. on Solid Surfaces, Vienna, 1977, p. 387.
[22] This assumption is reasonable at temperatures near 80 K, but would not be reasonable at temperatures at which metal atom diffusion could occur.
[23] C. Patel, MSc Thesis, University of Aston in Birmingham (1974).
[24] R.G. Forbes and C. Patel, 22nd Intern. Field Emission Symp., Atlanta, 1975 (unpublished).

[25] T.T. Tsong and G. Kellogg, Phys. Rev. 12B (1975) 1343.
[26] The field that appears in the conventional definition of polarisability is the self-consistent local field acting on the atom. At a metal surface this field may be significantly different from the external field.
[27] J.C. Rivière, in: Solid State Science, Vol. 1, Ed. M. Green (Dekker, New York, 1969) p. 179.
[28] Z. Knor, in: Surface and Defect Properties of Solids, Vol. 6 (1977) p. 139.
[29] R.G. Forbes, Appl. Phys. Letters, to be published.
[30] D.G. Brandon, Phil. Mag. 14 (1966) 803.
[31] In our view, Brandon's deductions concerning the value of the field evaporation pre-exponential should now be treated with reservations.

Surface Science 116 (1982) L195–L201
North-Holland Publishing Company

SURFACE SCIENCE LETTERS

NEW ACTIVATION-ENERGY FORMULAE FOR CHARGE-EXCHANGE TYPE MECHANISMS OF FIELD EVAPORATION

Richard G. FORBES

University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET, UK

Received 29 September 1981; accepted for publication 29 January 1982

A simple proof is given of a new approximate formula for the dependence of field-evaporation activation energy on external field, F, in the context of the charge-hopping desorption mechanism. The formula shows the explicit dependence on field, and is found to contain terms in F^{-1} and F^{-2}. It is also shown how corrections can be made for the effects of field on the interaction between the desorbing entity and the surface, and for the effects of forces due to correlation and repulsion interactions between the departing ion and the surface.

Field evaporation has significance in the field-ion techniques [1], and may be an emission process in liquid-metal field-ion sources [2,3]. Conventional field-evaporation theory assumes the escape of an ion from the emitter surface to be a thermally-activated process, and central to many aspects of the theory is some formula for the activation energy Q as a function of the external field F at the typical evaporation site. The nature of the formula depends on the mechanism of escape.

As is well known, two mechanisms have been widely discussed: The Müller–Schottky mechanism [4,5], and Gomer-type mechanisms involving a surface charge-exchange process [6,7]. With the Müller–Schottky mechanism, explicit formulae for Q in terms of F have been available since 1956, but there are difficulties in principle with the mechanism, particularly when applied to the metals conventionally used in low-temperature field-ion emission [8–11].

A Gomer-type mechanism is usually more plausible, but it is customary to discuss this in terms of formulae equivalent to eq. (3) below, in which Q is not given explicitly as a function of F. This has created errors and misunderstandings in the past, particularly in discussion of field-sensitivity data.

An explicit (though approximate) formula for $Q(F)$, for the charge-hopping mechanism (one of the variants of charge exchange [12]) was given by the author some years ago [13]. I have now realised that a simpler proof can be given, and that the result can be generalised, and these are the main aims of this note. It is convenient to begin with a careful basic discussion.

The energetics of field evaporation. A pre-evaporating atom is initially bound, at the emitter surface, with binding energy $\Lambda(F)$ [14]. The work

0039-6028/82/0000–0000/$02.75

$W_n(x, F)$ necessary to formally create, from this initial bonding state, a n-fold-charged ion at a distance x from the emitter's electrical surface, when the external field is F, can be derived from an electrothermodynamic ("Born–Haber type") cycle. Using the usual elementary approximations we obtain:

$$W_n(x, F) = \left[\Lambda(F) + H_n - n\phi^{\mathrm{E}}\right] - neFx - n^2e^2/16\pi\epsilon_0 x + G/x^t - \tfrac{1}{2}c_n F^2, \quad (1)$$

where e is the elementary charge; ne the charge on the ion; H_n the energy needed to form the ion from the corresponding neutral atom, in remote field-free space, being given by the sum of the first n free-space ionization energies; ϕ^{E} is the relevant emitter work-function; ϵ_0 the electric constant; t the repulsive exponent; G a constant; and c_n the ionic polarization-energy coefficient [15]. This equation can also be put, with greater generality, in the form:

$$W_n(x, F) = \left[\Lambda(F) + H_n - n\phi^{\mathrm{E}}\right] - neFx + \eta_n(x, F), \quad (2)$$

where $\eta_n(x, F)$ denotes the "purely chemical" component of the ion–surface interaction, and is approximated in eq. (1) by image-potential, repulsion, and polarization terms. In these "*work-formulae*" the x and F are independent variables.

To obtain the field-evaporation activation energy Q_n for escape as an n-fold ion, we need to formally create the ion at the top of the relevant barrier in the potential structure. Denoting the distance of the top of the barrier from the emitter's electrical surface by x^{p}, we obtain from eq. (2):

$$Q_n = W_n(x^{\mathrm{p}}, F) = \left[\Lambda(F) + H_n - n\phi^{\mathrm{E}}\right] - neFx^{\mathrm{p}} + \eta_n(x^{\mathrm{p}}, F). \quad (3)$$

In this formula (and related approximate formulae) the distance x^{p} is *not* an independent variable but is related to the external field F, in some as-yet unspecified way. An equation of this type is an "*implicit Q-formula*", is based on energetics alone, and applies to all thermally-activated mechanisms of field evaporation.

To eliminate x^{p} from eq. (3), a relationship (the "*subsidiary condition*") between x^{p} and F has to be postulated. With the Müller–Schottky mechanism, the condition in physical terms is that the position x^{p} be at the top of a Schottky hump; mathematically, that:

$$\left(\partial W_n/\partial x\right)_F|_{x=x^{\mathrm{p}}} = 0 \qquad \text{(IH)}, \quad (4)$$

where (IH) labels this "image-hump" criterion.

In the well-known basic image-hump (BIH) formalism this gives:

$$Q_n(F) = \left[\Lambda^0 + H_n - n\phi^{\mathrm{E}}\right] - \left(n^3e^3F/4\pi\epsilon_0\right)^{1/2} \qquad \text{(BIH)}, \quad (5)$$

where Λ^0 is the zero-field bonding energy. The variable x^{p} does not appear in eq. (5), and we may call such an equation a *Q-formula explicit in F.*

Q-formulae for Gomer-type mechanisms. An explicit Q-formula for the charge-hopping mechanism is readily obtained when the initial bonding state is primarily neutral [16]. To get a realistic subsidiary condition, we must assume a

shape for the initial-state binding well; relative to the bottom of the well, let the potential energy be $W_0(x)$. It follows that:

$$Q_n = W_0(x^p) \qquad \text{(CH)}, \qquad (6)$$

where (CH) labels charge-hopping.

For example, if we use the Einstein [17] simple approximation of a parabolic potential well (of force-constant κ), then:

$$Q_n = \tfrac{1}{2}\kappa(x^p - a)^2 = \tfrac{1}{2}\kappa r^2 \qquad \text{(CH)}, \qquad (7)$$

where a is the distance of the potential minimum from the emitter's electrical surface, and r is the distance of the "point of escape" (at x^p) from this minimum.

With the charge-hopping mechanism, x^p is defined by the intersection of the initial and final potential energy curves, i.e. by the requirement that:

$$W_n(x^p, F) = W_0(x^p). \qquad (8)$$

This is the subsidiary condition in its general form. The simplest explicit Q-formula is obtained by using eq. (3) to derive an expression for x^p in terms of F, and then substituting into eq. (7).

We may arrange eq. (3) into the form:

$$neFx^p = \left[\Lambda(F) + H_n - n\phi^E\right] + \eta_n(x^p, F) - Q_n \equiv h(x^p, F) - Q_n, \qquad (9)$$

where $h(x^p, F)$ is a relatively slowly varying parameter I call the "*pivot height*", that has the dimensions of energy and is defined by the right-hand side of eq. (9). When the intersection point corresponds with the potential minimum (so $Q_n = 0$), we have:

$$neF^e a = \left[\Lambda(F^e) + H_n - n\phi^E\right] + \eta_n(a, F^e) = h^e, \qquad (10)$$

where F^e is the evaporation field as defined by a zero-activation-energy criterion; and h^e a "standard pivot height", being a constant given by eq. (10). The relationship between h^e, a and F^e can be illustrated graphically, as in fig. 1: the x-axis is placed at the level of the bottom of the potential well, and the y-axis is placed to coincide with the metal's electrical surface: the third side of the triangle is drawn through a "pivot" on the y-axis at an energy level h^e.

For a value of F slightly less than F^e, x^p is somewhat greater than a: this is illustrated in fig. 2. As a first approximation we may neglect Q_n in eq. (9) and assume that the variation in h is also small enough to be neglected: that is, we put:

$$neFx^p = h(x^p, F) - Q_n \simeq h^e. \qquad (11)$$

In fig. 1 this is equivalent to swinging the third side of the triangle about the pivot on the y-axis and taking x^p as the intercept on the x-axis. We obtain:

$$x^p \simeq h^e/neF = aF^e/F. \qquad (12)$$

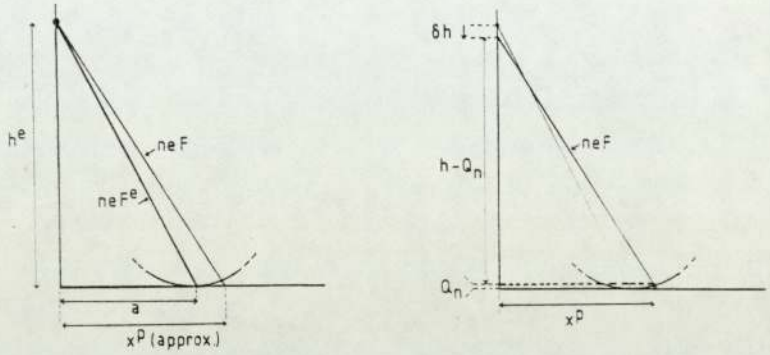


Fig. 1. A schematic diagram to illustrate, first, the relationship between the limiting evaporation field value F^e, the standard pivot height h^e, and the distance a: second, an approximate relationship between field F, pivot height h, and distance x^p, when in eq. (9) the activation energy Q_n is neglected and h is set equal to h^e. In the diagram the x-axis coincides with the energy level of the bottom of the well, and the y-axis with the metal's electrical surface: note that this diagram is not a potential diagram.

Fig. 2. Schematic diagram to illustrate the correct relationship between field F, pivot height h, activation energy Q_n, and distance x^p.

The activation energy Q_n can then be obtained by substituting into eq. (6). In the case of a parabolic well we obtain:

$$Q_n \simeq \tfrac{1}{2}\kappa\left(\frac{h^e}{neF} - a\right)^2 = \tfrac{1}{2}\kappa a^2\left(\frac{F^e}{F} - 1\right)^2 \qquad \text{(CH)}. \tag{13}$$

This is a Q-formula explicit in F, for the charge-hopping mechanism.

Eq. (13) is probably a "reasonable", rather than a good, approximation; but it serves to demonstrate several points. First, the alternative forms given enable either h^e to be estimated theoretically, or F^e to be estimated empirically. Second, the formula contains $1/F^2$ and $1/F$ terms: the dependence of Q_n upon F is manifestly *not* linear, as used to be suggested in field-ion literature. Third, the formula contains the type of physical parameter that one might intuitively expect to be relevant, in particular the force-constant for the vibration of the bound particle, and does not in the first approximation involve any "polarization" terms.

In past literature, it has sometimes been assumed that x^p can be treated as a constant, a recent example being ref. [18]. Fig. 1 shows clearly the erroneous nature of this assumption.

An improved approximation can be obtained by Taylor-expanding $h(x^p, F)$ in eq. (11), while continuing to neglect Q_n. From eq. (9) we obtain:

$$h(x^p, F) \simeq h^e + \left[(\mathrm{d}\Lambda/\mathrm{d}F) + (\partial\eta_n/\partial F)\right](F - F^e) + (\partial\eta_n/\partial x^p)\, r, \tag{14}$$

where the derivatives in eq. (14) are all evaluated for $F = F^e$ and $x^p = a$. The expression takes a simpler-looking form if we use η_n' to denote $(\partial \eta_n / \partial x^p)|_a$ and define a parameter π^e by:

$$\pi^e = [(d\Lambda/dF) + (\partial \eta_n/\partial F)]|_{a,F^e}. \tag{15}$$

The final result, after some manipulation, is:

$$Q_n = \tfrac{1}{2}\kappa\left(a - \frac{\pi^e}{ne}\right)^2\left(1 - \frac{\eta_n'}{neF}\right)^{-2}\left(\frac{F^e}{F} - 1\right)^2 \qquad \text{(CH)}. \tag{16}$$

The parameter π^e has the dimensions of SI dipole moment, but is not in any meaningful sense a physical dipole moment.

Eq. (16) is a better approximation than eq. (13), but does not have the attractive simplicity that the latter has. In eq. (16) the first bracket is a correction due to the effects of field on the purely chemical interactions between the desorbing entity and the surface; the second bracket is a correction due mainly to the correlation and repulsion interactions between the ion and the surface. Neither correction is really negligible, but it deserves note that in the evaluation of η_n' the contributions due to correlation and to repulsion have opposite signs and tend to cancel (physically, this is because the ionic bonding point in the absence of the field is in the vicinity of the neutral bonding point).

In obtaining eq. (16), the term Q_n in eqs. (9) and (11) has been neglected; at cryogenic temperatures this term is very small in comparison with h (at 80 K, 0.2 eV compared with an h-value of order 10 eV), so the neglect is perhaps justified. Nevertheless, its neglect means that there will be an error in evaluating dQ_n/dF from eq. (13) or (16). So if a good approximation is required, it is better to use the formula:

$$dQ_n/dF = (dQ_n/dx^p)(dx^p/dF), \tag{17}$$

where dQ_n/dx^p is obtained from eq. (6), and dx^p/dF from eq. (8) – or, rather, from whatever approximate versions of these relationships are in use. The result for the simple parabolic-well approximation has been discussed elsewhere [13].

Strictly, eqs. (13) and (16) apply to the charge-hopping variant of charge exchange. In reality, as recently remarked by Kingham [19], field evaporation may occur by a mechanism more akin to the charge-draining variant of charge-exchange. In this case the calculation of x^p as a function of F might still have approximate validity, but the activation-energy expression would need correction. Formally we might write [12]:

$$Q_n(CD) = Q_n(CH) + \Delta U^p, \tag{18}$$

where (CD) and (CH) label charge-draining and charge-hopping quantities, and ΔU^p is a correction term, negative in value, that represents a lowering [20] of system potential energy as compared with what this would be if the

transferring electron were confined to orbitals localized in the desorbing entity or in the emitter. But the correction term would be a function of x^P and certainly field-dependent. At present there are no reliable calculations of ΔU^P, so there is little option but to use the charge-hopping formulae when discussing Gomer-type mechanisms of field evaporation.

Three final points may be worth making, the first about the Müller–Schottky mechanism. Despite its known shortcomings, this mechanism is often described in elementary discussions because the associated formulae are simple. I suggest that the analysis based on fig. 1 is equally simple; so the objections to the use of charge-exchange-type mechanisms in elementary discussions have been considerably diminished.

Second, I feel that a prime cause of confusion in past discussions of escape mechanism has been the custom of discussing the Müller–Schottky mechanism in terms of its *explicit* Q-formula, but Gomer-type mechanisms in terms of *implicit* Q-formulae. Few accounts have made clear that the formulae given in connection with the two escape processes were of differing logical status.

Finally, I would point out that the approach here is readily generalised to deal with other shapes of potential well. It is straightforward, for example, to use or add in higher powers of r in eq. (7), or to use a Morse potential for $W_0(x)$. The approach to charge-exchange presented here thus has considerable empirical flexibility.

This work forms part of a research project funded by the UK Science and Engineering Research Council.

References

[1] E.W. Müller and T.T. Tsong, Progr. Surface Sci. 4 (1973) 1.
[2] R. Clampitt, Nucl. Instr. Methods 189 (1981) 111.
[3] G.L.R. Mair, Nucl. Instr. Methods 172 (1980) 567.
[4] E.W. Müller, Phys. Rev. 102 (1956) 618.
[5] E.W. Müller, Advan. Electron. Electron Phys. 13 (1960) 83.
[6] R. Gomer, J. Chem. Phys. 31 (1959) 341;
R. Gomer, Field Emission and Field Ionization (Harvard Univ. Press, Cambridge, MA, 1961).
[7] R. Gomer and L.W. Swanson, J. Chem. Phys. 38 (1963) 1613.
[8] E.W. Müller and T.T. Tsong, Field Ion Microscopy: Principles and Applications (Elsevier, Amsterdam, 1969).
[9] T.T. Tsong, J. Chem. Phys. 54 (1971) 4205.
[10] R.K. Biswas and R.G. Forbes, J. Phys. D (Appl. Phys.), to be published.
[11] R.G. Forbes, unpublished work.
[12] R.G. Forbes, Surface Sci. 102 (1981) 255.
[13] R.G. Forbes, Surface Sci. 70 (1978) 239.
[14] Note that, as used here, this symbol denotes the total binding energy, including the polarization term.
[15] Strictly, c_n is not a polarizability, because it is associated with the external field F rather than the local field.

[16] The present approach should also be applicable if the entity is bound in a partially ionic state; but if the position of the initial-state potential energy minimum is a significant function of field, then modification would become necessary.
[17] A. Einstein, Ann. Physik 22 (1907) 180.
[18] M. Konishi, M. Wada and O. Nishikawa, Surface Sci. 107 (1981) 63.
[19] D.R. Kingham, Surface Sci. 108 (1981) L460.
[20] L.D. Landau and E.F. Lifschitz, Quantum Mechanics (Pergamon, New York, 1958).

J. Phys. D: Appl. Phys., 15 (1982) 1301–1322. Printed in Great Britain

Electrothermodynamic cycles applied to ionic potentials and to field evaporation

Richard G Forbes
Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET, UK

Received 4 January 1982

Abstract. A formal treatment of electrothermodynamic ('Born–Haber-type') cycles is presented, and is applied to construct general expressions for the standard potential energy of a particle external to a conducting emitter surface; a tabulation is given of effects and terms suggested over the years for inclusion in approximate representations for this potential energy. It is shown qualitatively how effects due to non-integral charge states, to 'fast' processes, and to correlation energy in two-electron hopping processes, can be included within the discussion.

General expressions are given for electron orbital level in field-evaporating ions, and for the activation energy for field desorption from ionic bonding states, and simple approximate formulae are obtained for the corresponding evaporation fields. Application of these formulae to tungsten suggests that tungsten atoms field evaporate from primarily neutral bonding states, but also contributes doubt as to the validity and/or usefulness of elementary approximations.

1. Introduction

An electrothermodynamic (or 'Born–Haber type') cycle is used in field desorption theory to calculate the work done in creating an ion in the presence of a field, in the vicinity of the emitter surface. From the resulting expression the activation energy for desorption and the desorption field can be calculated.

Simplified expressions obtained by carrying out such a cycle are often stated in the literature (for example, Gomer 1959), but to the best of the present author's knowledge there has been no clear and complete discussion of the cycle itself.

The intention of this paper is thus twofold: to present a general treatment of electrothermodynamic cycles; and to give the results for the creation of ions from lower ionic states, rather than from the neutral bonding state normally assumed. The general approach used can be seen as a development and clarification of that used elsewhere (Forbes 1976, 1980a).

The structure of the paper is as follows: § 2 deals with some basic matters, including the definition of standard potential energies, and the idea of 'purely electric' and 'purely chemical' potential-energy components; § 3 derives a formal expression for standard potential energy; § 4 shows how this expression can be used to obtain formulae for electron orbital level and desorption activation energy, and goes on to derive approximate expressions for evaporation fields; § 5 applies these simple formulae to tungsten field evaporation; and § 6 deals with corrections to the basic arguments needed in some specific circumstances; § 7 then summarises the whole analysis.

0022-3727/82/071301 + 22 $02.00

The physical quantities used in this paper are based on the International ('four electric dimensions') System of Measurement, and a glossary of symbols used is given in Appendix 1.

2. Basic matters

2.1. The system

A particle, which may be atomic or molecular, and initially neutral or ionic, is situated close to the charged surface of a conducting 'emitter'. The particle is 'external', and has its location defined by some internally-fixed reference point. (For an atom the nucleus is convenient.) The particle is initially assumed to have an integral charge-number; 'partially ionic' situations are considered later.

Potential energies are defined below for various particle charge-states. These are conveniently treated as potential energies for the particle; but in reality they refer to the emitter + particle system.

2.2. Suffix notation

With a quantity that depends both on charge state and on spatial position, lowered suffices indicate the charge state(s) involved and raised suffices the position(s) involved. For example, W_{mn}^{qs} denotes the standard work needed to create an n-fold ion at position s from the corresponding m-fold ion initially at position q. Left-to-right suffix ordering corresponds to before–after ordering.

Labels, rather than coordinates, are used to indicate positions in space. This is because of the varying conventions about locating the origin of coordinates at a charged surface, and resulting literature inconsistency. The general results below do not depend on the choice of coordinate system; the elementary approximations do. Labels used are listed in Appendix 1.

Note that in a consistently-analysed one-dimensional model the conventional 'point at infinity' is not in field-free space; this concept is thus replaced by the 'point in remote field-free space'.

2.3. Standard potential energies (Standard Terms)

If the external particle is in its internal ground electronic state, before and after ionisation, and the ionisation process is 'slow' in a thermodynamic sense discussed later, potential energies can be formally defined by

$$U_0^{\mathrm{R}} = 0 \tag{1}$$

$$U_n^s - U_0^{\mathrm{R}} = W_{0n}^{\mathrm{R}s} \tag{2}$$

U_0^{R} denotes the potential energy of the relevant neutral entity in remote field-free space; so equation (1) defines the zero of potential energy. Relative to this zero, equation (2) formally defines the potential energy of the relevant entity when n-fold-charged and at position s. $W_{0n}^{\mathrm{R}s}$ denotes the standard work done in creating the n-fold ion at position s from the corresponding neutral in remote field-free space. (The phrase '*standard work*' refers to a 'slow' process in which the electrons are removed one by one and returned

separately to the emitter fermi level, with each of the intermediate ions being in its internal ground state).

The quantities U_n defined in this way are widely used in field-ion literature and I call them *standard potential energies* or (following Landau and Lifshitz 1958) *Standard Terms*. The name 'Term' is used because these quantities play a role with respect to the emission of electrons that is directly analogous to the role played by ordinary spectroscopic terms with respect to the emission of photons (Forbes 1980b).

2.4. *Chemical and electrical potential-energy components*

The potential-energy difference $(U_n^s - U_n^q)$ is equal to the work W_{nn}^{qs} done by a hypothetical external agent in moving the relevant n-fold-charged entity from q to s. As when discussing the motion of electrons near charged surfaces, it is convenient to split W_{nn}^{qs} into two components, performed against 'purely electric' and 'purely chemical' forces.

The 'purely electric' component represents work done against the electric fields that would exist in the absence of the external particle (and in the absence of any effects that would be induced by its presence). This component can be written in terms of electric potentials, as

$$W_{nn}^{qs}\,(\text{electric}) = ne(\Upsilon^s - \Upsilon^q) = n(u^s - u^q) \tag{3}$$

where e is the elementary (proton) charge; ne the charge on the ion; $(\Upsilon^s - \Upsilon^q)$ the difference in classical electrostatic potential between points q and s (as conventionally defined by a vanishingly-small 'test charge'); and $(u^s - u^q)$ the corresponding difference in the electrostatic component of the potential energy of a proton. In practice, u is more convenient to work with than Υ. The question of a reference zero for u is considered later.

Note that the quantity u as introduced here refers to the *complete* electrostatic component of proton potential energy, not merely the component arising from the presence of an applied voltage.

The 'purely chemical' component of W_{nn}^{qs} represents work done against all forces other than the 'purely electric' ones. (The name 'purely chemical' is not ideal, but it is difficult to find a better one.) Introducing a symbol η_n to represent the 'purely chemical' component of standard potential energy U_n, we formally have

$$W_{nn}^{qs}\,(\text{chemical}) = \eta_n^s - \eta_n^q. \tag{4}$$

The reference zero for η_n is taken in remote field-free space, so

$$\eta_n^{\mathrm{R}} = 0. \tag{5}$$

Conventional discussion has various effects contributing to η_n. For an external ion with integral charge-number the main effects are: correlation (i.e. effects due to the rearrangement of substrate electrons so as to screen the field of the external ion); polarisation of the external particle; London-type dispersion interactions; and short-range repulsion. (The conventional representations of these effects are shown in table 1.) When the external particle is neutral and physisorbs (e.g. a noble-gas atom), the main interactions are polarisation, dispersion and repulsion; when the particle is neutral and chemisorbs, the interaction is often represented by a bonding energy, with an additional term to represent the effects of the external field.

In elementary discussions, therefore, η_n is represented as a sum of individual terms, with the number of terms included depending on the nature of the elementary approx-

imation. However, in quantum-mechanical treatments it seems likely that the classically recognised interactions will (at least to some extent) lose their individual identities, and that the treatments will lead directly to values of η_n (or possibly U_n) as a function of position and field strength.

2.5. Conclusions

Combining equations (4) and (5) gives the basic relationship:

$$U_n^s - U_n^q = n(u^s - u^q) + \eta_n^s - \eta_n^q. \tag{6}$$

The standard-potential-energy difference has been split into components. The merit in this is that an electron experiences the same electric effects as a proton (but with reversed sign) but different chemical effects. When an electrothermodynamic cycle involves the motion of both ions and electrons, the electric terms cancel to some extent, but the chemical terms do not.

The above discussion applies formally both to ions and to the neutral ($n = 0$) case, but obviously the electric terms vanish in the latter case.

The above theory assumes the external particle to have integral charge-number. But, physically, the possibility exists that either in the bound state or during the escape process, or both, the desorbing entity does not have a well-defined integral charge (for example see Tsong 1978, Forbes 1981a). In such circumstances the splitting of standard-potential-energy differences into electrical and chemical components is problematical or impossible; the only really satisfactory approach may be the direct quantum-mechanical evaluation of values of U_n.

Nevertheless, in the case of field ionisation, and in the simplest models of field desorption (the 'charge hopping', 'ionic bonding' and 'Müller-Schottky' mechanisms), it is assumed that the desorbing entity has effectively integral charge in the initial bonding state and well-defined integral charge immediately after escape. To these cases the above theory certainly applies.

3. The standard potential energy U_n^s

We next require an expression for the standard potential energy U_n^s at some position close to the emitter surface. As is well known, this is obtained by starting with a neutral entity in remote field-free space and summing the works done in the following individual stages of an electrothermodynamic cycle.

(1) Remove n electrons one by one from the neutral entity at position R in remote field-free space.

(2) Place the n electrons one by one at the emitter fermi level.

(3) Move the resulting ion from position R to position s.

It is expedient to analyse step (2) in greater detail. Because the 'chemical' (i.e. correlation-and-exchange) and 'electrical' interactions of an electron with the emitter + field system are essentially independent of each other, we may deal separately with them. Hence we consider the following notional thermodynamic sub-steps.

(2*a*) Move 'chemically-interacting electron' from position R to the emitter fermi level.

(2*b*) Move 'electrically-interacting electron' from position R to position s.

(2*c*) Move 'electrically-interacting electron' from position s to the emitter interior i.

Bearing in mind that there are n electrons in step (2), we have that the works done in the above steps are

$$\left.\begin{aligned} \text{work (1)} &= H_n\left(=\sum_{k=1}^{n} I_k\right) \\ \text{work (2}a\text{)} &= -n\chi^{\mathrm{E}} \\ \text{work (2}b\text{)} &= -n(u^s - u^{\mathrm{R}}) \\ \text{work (2}c\text{)} &= -n(u^{\mathrm{i}} - u^s) \\ \text{work (3)} &= n(u^s - u^{\mathrm{R}}) + (\eta_n^s - \eta_n^{\mathrm{R}}). \end{aligned}\right\} \tag{7}$$

In the above expressions H_n is the energy needed to create an n-fold ion, from the neutral, in remote field-free space, being given by the sum of the first n free-space ionisation energies I_k; χ^{E} is the chemical component of the emitter work function; and the labels i and R denote points in the interior of the emitter, and in remote field-free space, respectively. The minus sign appears in steps (2b) and (2c) because of the negative charge on the electron.

Bearing in mind that U_0^{R} and η_n^{R} are zero by convention, and that the electric term in step (2b) cancels the corresponding term in step (3), we find that the total work done and hence the potential energy U_n^s are given by

$$U_n^s = W_{0n}^{\mathrm{R}s} = H_n - n\chi^{\mathrm{E}} + n(u^s - u^{\mathrm{i}}) + \eta_n^s. \tag{8}$$

The third term in equation (8) now requires closer attention. In sophisticated treatments it would be necessary to calculate it as a whole, but for most elementary purposes this term is usefully split into two parts: a 'zero-voltage' part that exists in the absence of any applied voltage between the emitter and its surroundings; and an 'applied' part that exists as a consequence of an applied voltage. Formally:

$$(u^s - u^{\mathrm{i}}) = (u^s - u^{\mathrm{i}})_{\mathrm{zv}} + (u^s - u^{\mathrm{i}})_{\mathrm{app}}. \tag{9}$$

The zero-voltage part is itself usefully split into two parts. If we use the suffix E to label a point 'somewhat'† outside the emitter surface, then

$$(u^s - u^{\mathrm{i}})_{z-v} = (u^{\mathrm{E}} - u^{\mathrm{i}})_{\mathrm{zv}} + (u^s - u^{\mathrm{E}})_{\mathrm{zv}} \tag{10a}$$

$$= -\psi^{\mathrm{E}} + \delta^s \tag{10b}$$

By definition, the first bracket on the RHS of equation (10a) relates to the electrostatic component ψ^{E} of the local work-function for the face in question; the second bracket is a small correction that depends somewhat on the position s, and could be described as due to 'patch fields and/or surface fields'. I denote this small correction term by δ^s. (The minus sign in front of ψ^{E} results from the negative charge of the electron.)

Substituting equations (10b) and (9) back into equation (8), and remembering that the complete local work-function φ^{E} is given by the sum of its chemical (χ^{E}) and electrostatic (ψ^{E}) components, we obtain

$$U_n^s = H_n - n\varphi^{\mathrm{E}} + n\delta^s + n(u^s - u^{\mathrm{i}})_{\mathrm{app}} + \eta_n^s. \tag{11}$$

† The requirement that position E be 'somewhat' outside the emitter is to ensure that this position is, for practical purposes, outside the range of both correlation-and-exchange forces and the short-range electrostatic fields due to the detailed charge distribution in the electrical double layer at the emitter surface (as it exists in the absence of any applied voltage). The local work-function for the face in question can then be defined as the work done in taking an electron, initially at the fermi level, from the interior of the emitter to position E.

This expression can be further simplified if we define the electrostatic potential-energy component u to be zero in the interior of the emitter, i.e. we set

$$u^{\mathrm{i}} = 0. \tag{12}$$

Provided that (if necessary) position i is so chosen that the first bracket in equation (10*a*) gives the electrostatic component of local work-function, this convention does not disturb the arguments above. We then obtain:

$$U_n^s = H_n - n\varphi^{\mathrm{E}} + n\delta^s + n(u^s)_{\mathrm{app}} + \eta_n^s. \tag{13}$$

Equation (13) is a formal expression for the standard potential energy U_n^s; it involves no significant approximations.

3.1. The S-component and its representations

Expression (13) falls naturally into two parts: a 'fixed' or 'configurational' part, and a 'variable' (i.e. position and field dependent) part that I denote by S_n^s. Formally:

$$U_n^s = (H_n - n\varphi^{\mathrm{E}}) + S_n^s \tag{14a}$$

$$S_n^s = n(u^s)_{\mathrm{app}} + n\delta^s + \eta_n^s. \tag{14b}$$

I refer to this variable part as the *S-component* of standard potential energy. ('*S*' stands for 'split-referenced': the three terms in equation (14*b*) are zero-referenced at different points: the emitter interior, the point E, and remote field-free space, respectively.) Note that this *S*-component is somewhat different from the '*S*-potential' defined in an earlier paper (Forbes 1980a). The present quantity is better defined, and seems the more useful of the two.

The quantity S_n finds use, as such, in some manipulations, but numerical applications require specific algebraic approximations. For the applied electrical term in equation (14*b*) all one-dimensional charged-surface models (e.g. Tsong and Müller 1969, Theophilou and Modinos 1972, Lang and Kohn 1973) give much the same physical result: that, at sufficient distance from the emitter surface, this term has a linear form:

$$(u^s)_{\mathrm{app}} \simeq -eFx^s \tag{15}$$

where F is the external field (i.e. the field component, somewhat above the surface, resulting from the applied voltage); and x is distance measured from the emitter's electrical surface.

(The different charged-surface models differ in where to locate this electrical surface relative to definable features of the emitter substrate, and over whether there is a field-dependent shift in the position of the electrical surface. The result can also be expressed in alternative mathematical forms involving a small distance parameter: but there is no advantage in the more complicated algebraic forms.)

This linear form is also found in analysis of a structured surface model (R G Forbes and M K Wafi unpublished work, Wafi 1981).

The zero-voltage electrical term has not appeared explicitly in previous discussions. As will be seen, it has a significant formal role in the analysis of one-dimensional models and the definition of zero-field activation energy, but in other circumstances is probably small enough to be completely neglected.

The nature of the 'purely chemical' term in equation (14*b*) was discussed earlier. The elementary approximations used to represent the various contributory effects are

well known, and are listed in table 1. Alternatively, η_n can be treated as a single entity and modelled by one of the standard approximations (parabolic potential well, Morse potential, etc.), this procedure being particularly useful in the vicinity of the initial bonding point.

Table 1 brings together nearly all the algebraic expressions that have been proposed for inclusion in the S-component over the last 25 years. Thus something of the history of field-desorption theory is manifested in this table.

Table 1. Effects and terms proposed for inclusion in the S-component. Note that all notation has been adjusted to the present author's preferred form. (A glossary of symbols is given in the Appendix.)

Effect	Term	Originator
1[a] Electrostatic potential energy	$-neFx$	Müller (1956)
2[a] Correlation (image-potential)	$-n^2e^2/16\pi\varepsilon_0 x$	Müller (1956)
3 Polarisation	$-\frac{1}{2}c_nF^2$	Gomer (1959)
4 Short-range repulsion	$+G/x^t$	Brandon (1963)
5[c] Interaction of dipole with field	$-pF$	Gomer (1959)
6[d] Configuration interaction	$+\Delta U$	Gomer (1959)
7[b] Bonding of neutral to surface (also dispersion)	$-D/x^t$	Brandon (1963)
8[b] Field-induced reduction in metal-atom binding energy	$+\theta F$	Tsong (1971)
9 Sum over all relevant ion-core repulsions	$+\Sigma K_j/x_j^t$	Tsong (1971)
10 Hyperpolarisability	$-(1/24)\gamma_n F_{\text{loc}}^4$	Tsong (1971)
11 Field-gradient polarisability	$-\frac{1}{4}B_nF_{\text{loc}}^2F'_{\text{loc}}$	Forbes (1981b)
12 Field-gradient polarisability	$-\frac{1}{4}C_n(F'_{\text{loc}})^2$	Forbes (1981b)
13[a] Patch and/or surface fields	$+n\delta$	Present paper
More-general forms		
14[a] Potential due to applied field	$+nu_{\text{app}}$	Present paper
15 Purely-chemical interactions	$+\eta_n$	Forbes (1980a)
16[b] Potential near bonding point (Parabolic-well approximation)	$+\frac{1}{2}\kappa(x-x^{\text{B}})^2$	Forbes (1978)

Notes: (a) Effects 1, 2, 13 and 14 apply only to ions. (b) Effects 7, 8 and 16 were envisaged as applying only to neutrals. (c) Effect 5 applies only if adsorption produces an effective dipole moment. (d) Effect 6 applies in the case of charge-draining mechanisms. (e) Additional terms applying in the case of an adsorbed layer are given by Swanson and Gomer (1963).

4. Applications

Given an expression for U_n^s, it is trivial to derive an expression for the work of ionisation W_{mn}^{qs}, from the relationship

$$W_{mn}^{qs} = U_n^s - U_m^q. \tag{16}$$

The resulting formulae have two main applications—in determining electron orbital level and in deriving activation-energy expressions—which we look at in turn.

4.1. Electron orbital level

Consider the formation of a $(m+1)$-fold ion from a m-fold ion, at some position s where $U_m^s > U_{m+1}^s$. The work done by the hypothetical external agent in creating this ion, if the

transferred electron is placed at the emitter fermi level, is by definition

$$W^{ss}_{m,m+1} = U^s_{m+1} - U^s_m \tag{17}$$

This work is a negative quantity, which we write in the form $-\varepsilon^s_m$.

In reality no energy is given to any external agent. Rather, the energy ε^s_m is transferred (with the electron) to the emitter electron energy distribution, where it ultimately degrades to heat. In a formal cycle the energy ε^s_m can be used to promote the transferred electron to a level above the fermi level. Thus, in the context of the approach to field ionisation that regards it as an energy-conserving one-electron transition, we deduce that at position s the orbital energy level of the topmost filled orbital in the m-fold ion was at a level ε^s_m above the fermi level, where:

$$\varepsilon^s_m = U^s_m - U^s_{m+1} = -I_{m+1} + \varphi^E - \delta^s - (u^s)_{app} + \eta^s_m - \eta^s_{m+1}. \tag{18}$$

This relationship is illustrated in figure 1. (Note that the energy carried away by the

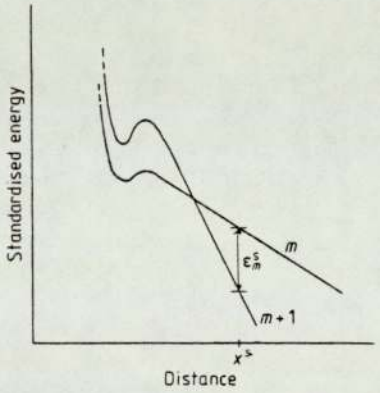


Figure 1. To illustrate the relationship between the electronic level ε^s_m of the topmost filled orbital in an m-fold-charged ion and the standard potential energies for the m-fold and $(m + 1)$-fold charged ions.

This diagram, like the others in this paper, is *schematic*. In each case the diagrams are drawn to illustrate the point under discussion; they are not intended either to be realistic representations of actual potential curves, or to be totally consistent with each other. With the exception of figure 7, each diagram shows standardised energy (or standard potential energy) on the vertical axis, and position on the horizontal axis. The curves shown represent standard potential energies (standard Terms), and are labelled with the charge number of the corresponding charge state.

electron is equal to the energy difference between the standard potential energies; as already noted, this analogy with optical transitions is the reason why the description 'Term' is used as an alternative name for the quantities U_n.)

In the elementary approximation using the basic electrostatic, correlation, and polarisation terms shown in table 1 we have

$$\varepsilon^s_m = \varphi^E + eFx^s - I_{m+1} + (2m + 1)e^2/16\pi\varepsilon_0 x^s - \tfrac{1}{2}(c_m - c_{m+1})F^2. \tag{19}$$

For $m = 0$ this equation reverts to the well-known simple formula, widely used in the field ionisation of inert gases; the 'ionic' ($m \geq 1$) form may be of interest in connection with post-ionisation (for example, Haydock and Kingham 1980). Note that correlation and polarisation terms corresponding to both 'before' and 'after' ionic states appear in the expression for orbital energy level if $m \geq 1$.

4.2. Field desorption and field evaporation

4.2.1. Introduction. At sufficiently high temperatures (near 80 K and above) field desorption is a classical thermally-activated process, with a field dependent activation energy. Simple theories are possible when the desorbing particle effectively has integral charge-number both in its initial bound state and immediately after escape. A conceptual description of the escape process and associated theory is as follows.

Prior to field desorption let the desorbing particle be bound in an m-fold-charged state. To escape from the emitter the particle has to climb (energy-wise) to the crest of a 'pass' in the potential structure, at position p; let it be in an n-fold-charged state immediately after escape. The escape process requires an activation energy $Q_{mn}(F)$ that is a function of the external field F at the desorption site. After escape the desorbing particle may experience one or more post-ionisation processes (Kingham 1981). This implies that the observed charge-state is in most circumstances not to be identified with the charge state immediately after escape, but does not otherwise affect the calculation of activation energies.

Full theoretical discussion of escape requires three inputs. First—an expression for Q_{mn} in terms of x_m^{B}, x_n^{p} and F. (x_m^{B} is the distance of the initial bonding point from the emitter's electrical surface, and x_n^{p} the distance of the point of escape.) Since Q_{mn} is the work necessary to create an n-fold ion at p from an m-fold ion at B, it can be found by using appropriate approximations in the relationship:

$$Q_{mn} = W_{mn}^{\mathrm{Bp}} = U_n^{\mathrm{p}} - U_m^{\mathrm{B}} \tag{20}$$

The resulting expression I call an *implicit Q-formula*, 'implicit' because it contains the variables x_m^{B} and x_n^{p} which may themselves be functions of F.

Second, we require *subsidiary conditions* that give the values and/or field-dependence (depending on circumstances) of x_m^{B} and x_n^{p}. This enables, at least in principle, the derivation of a Q-formula that is explicit in F. (Different models may employ different subsidiary conditions, and hence have the same implicit Q-formula but different explicit Q-formulae.)

Third, we need a *desorption-field criterion* that states how 'desorption field' is defined. The criterion may set the numerical value of either the activation energy or the desorption rate-constant at a given or 'typical high-risk' emission site, or may set the desorption flux from the whole or a portion of the emitter. Criteria of all these types are in use. The simplest approach is to require that Q_{mn} be zero, but other criteria can be more realistic, and with these the desorption field as evaluated theoretically is a function of temperature.

It is not the intention of this paper to become involved in detailed algebraic analyses concerning the temperature dependence of desorption field, as this will be presented elsewhere. Rather, I shall use the zero-activation-energy criterion to obtain first some general formulae, and then some simple basic approximations, thereby exposing a number of physical points about existing discussions of field evaporation.

4.2.2. Activation-energy expressions. Three general forms may be of some use. The first comes by using equation (14*a*) twice in equation (20), to give

$$Q_{mn} = (H_n - H_m) - (n - m)\varphi^{\mathrm{E}} + S_n^{\mathrm{p}} - S_m^{\mathrm{B}} \tag{21}$$

The second comes by using the symbol Λ_m^F to represent the bonding energy of the m-fold ion to the surface, in the presence of the field, this being defined relative to the same

zero as standard potential energy. It follows that

$$\Lambda_m^F \equiv -U_m^B \tag{22}$$

and hence, from equations (14*a*) and (20)

$$Q_{mn} = (\Lambda_m^F + H_n - n\varphi^E) + S_n^p. \tag{23}$$

Alternatively, Λ_m^F may be split into a zero-field part Λ_m^0 and a field-dependent part $\Delta\Lambda_m^F$, thus

$$\Lambda_m^F \equiv \Lambda_m^0 + \Delta\Lambda_m^F \tag{24}$$

which on substitution into equation (23) gives

$$Q_{mn} = (\Lambda_m^0 + H_n - n\varphi^E) + \Delta\Lambda_m^F + S_n^p. \tag{25}$$

In practice, form (21) is most useful when $m \neq 0$, and forms (23) and (25) more useful when $m = 0$, but all the above forms are general. Obviously, numerous approximate versions of the Q-formulae can be obtained by expanding the S-components in terms of the expressions listed in table 1.

In the neutral bonding case ($m = 0$) it is convenient to define a parameter K_{0n} by

$$K_{0n} = (\Lambda_0^0 + H_n - n\varphi^E) \tag{26}$$

so that equation (25) takes the form

$$Q_{0n} = K_{0n} + \Delta\Lambda_0^F + S_n^p. \tag{27}$$

If, further, we ignore the field-dependent term $\Delta\Lambda_0^F$, and take only the basic terms in S_n^p due to electrostatic potential energy, patch/surface fields, and correlation, then we obtain:

$$Q_{0n} = K_{0n} + n\delta^p - neFx_n^p - n^2e^2/16\pi\varepsilon_0 x_n^p. \tag{28}$$

The usual elementary formula has been recovered, but with the addition of the patch/surface field term.

Some basic points deserve making about K_{0n}. First, note that it contains the *local* work-function of the surface from which emission is occurring. There has been disagreement over this in the past, with Southon (1963) and Brandon (1964) suggesting that the so-called 'total' work-function of the emitter ought to appear, but Müller and Tsong (1974) insisting that the inclusion of the local work-function was correct. The formal arguments here confirm that the Müller and Tsong view is correct in the circumstances of field desorption, when $n\delta^p$ is negligible.

Second, note that strictly K_{0n} *cannot* be identified with the formal work necessary in the absence of any external field to remove an n-fold ion to a remote point in space, as is often suggested in elementary discussions (e.g. Müller and Tsong 1974). This work is obtained from equation (28) by taking F zero and position p to be in remote space: but in these circumstances the $n\delta^p$ term is not small: it constitutes a patch-field correction: so equation (28) becomes

$$Q_{0n}(F = 0) = K_{0n} + n\delta^R = \Lambda_0^0 + H_n - n\varphi^{tot} \tag{29}$$

where R labels the remote point and φ^{tot} ($= \varphi^E - \delta^R$) is by definition the 'total' work-function. Thus, to be strictly correct, K_{0n} should be described as 'the work necessary, in the absence of any external field, to desorb the n-fold ion *locally* from the emitter surface

facet on which the corresponding neutral entity is adsorbed'. (The ion then interacts with the patch-field system surrounding the emitter, and—depending on the face where it was originally adsorbed—either reaches a remote point with finite kinetic energy or needs to find a small additional amount of activation energy before being able to escape completely from the emitter.)

An exception to the above argument occurs in the case of the consistently-analysed one-dimensional model, in which it is postulated that the emitter is planar and of infinite extent. In this artificial case there can be no patch-field correction, and the activation energy needed to remove an n-fold ion to a remote point, in the absence of any external field, can be identified with K_{0n}. Perhaps misunderstandings have arisen through the use of one-dimensional models in field-desorption theory.

4.2.3. Basic predictions of evaporation field F^{ev}. We now move on to the prediction of evaporation field, for neutral and ionic bonding states, using the simple criterion that Q_{mn} be zero. For given charge numbers m and n, this defines a 'zero-Q evaporation field', F^{ev}_{mn}: actual evaporation fields are expected to be less than this.

With a neutral initial bonding state the simplest approximation is to use equation (28), but neglect the patch/surface-field term. The question then arises as to how x^p_n is to be chosen. The Müller–Schottky basic-image-hump formalism (e.g. Müller 1960, Brandon 1964) takes x^p_n to be at the top of a simple Schottky hump; but it is now clear that in conventional low-temperature field evaporation no Schottky hump exists (Biswas and Forbes 1982); so this approach can be discarded. Other formalisms (e.g. Gomer 1961, Müller and Tsong 1969) assume that the ionic potential-energy curve intersects the neutral curve at the bottom of the potential well in the latter, and hence that $x^p_n = x^B_0$. This results in the formula

$$F^{ev}_{0n} x^B_0 = K_{0n}/ne - ne/16\pi\varepsilon_0 x^B_0. \tag{30}$$

An estimate of F^{ev}_{0n} can then be made either by estimating the value of x^B_0, or by requiring that the second term on the RHS of equation (30) be some definite fraction α_{0n} of the first RHS term. This second approach forms the basis of a new method (Forbes 1982) of estimating evaporation field, and leads to the formula

$$F^{ev}_{0n} = \sigma_{0n}(16\pi\varepsilon_0/n^3e^3)K^2_{0n} \tag{31}$$

where $\sigma_{0n} = \alpha_{0n}(1 - \alpha_{0n})$, and in practice is a numerical factor with a value in the vicinity of 0.2, for the metals conventionally discussed.

The value of F^{ev}_{0n} predicted from equation (30) depends on the value chosen for x^B_0. Straightforward analysis shows that the maximum zero-Q evaporation field predictable from equation (31) is

$$F^{ev}_{0n}\,(\text{maximum}) = (4\pi\varepsilon_0/n^3e^3)K^2_{0n} \tag{32}$$

which is equivalent to taking $\sigma_{0n} = \frac{1}{4}$ in equation (31). (By coincidence, this value is identical with that predicted using the basic image-hump formalism.)

With an ionic initial bonding state various evaporation mechanisms would in principle be possible. We first consider a straightforward charge-hopping mechanism. In this case it is convenient to construct a basic theoretical approach directly analogous to that just presented for a neutral initial state.

A constant K_{mn}, characteristic of the species and charge states concerned, can be defined by

$$K_{mn} = (H_n - H_m) - (n - m)\varphi^E \qquad (m \neq 0). \tag{33}$$

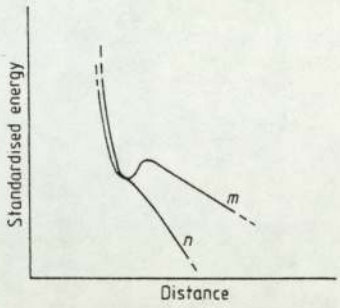


Figure 2. Term diagram representing the charge-hopping model applied in the case of ionic bonding.

This constant is chosen for algebraic convenience and has no straightforward physical interpretation. By using just the basic electrostatic and correlation terms in the S-components in equation (21), by requiring that Q_{mn} be zero, and by assuming that if field evaporation from an m-fold ionic state is actually occurring then the potential curve for the n-fold ion must intersect the curve for the m-fold ion at the bottom of a bonding well, and hence that $x_n^p = x_m^B$ (as in figure 2), we obtain

$$F_{mn}^{ev} x_m^B = K_{mn}/(n-m)e - (n+m)e/16\pi\varepsilon_0 x_m^B. \tag{34}$$

This equation is the generalised equivalent of equation (30), and reduces to it when $m = 0$. (But note that equation (33) is not valid for $m = 0$.)

The generalised equivalent to equation (32) is

$$F_{mn}^{ev}\,(\text{maximum}) = [4\pi\varepsilon_0/(n-m)(n^2-m^2)e^3]K_{mn}^2. \tag{35}$$

A generalised equivalent to equation (31) also exists, but it is not perceived to be useful.

It deserves comment at this stage that there are inconsistencies in the simple approximations leading to equation (34). In particular, we have assumed that there is a potential well in the m-fold term, but have omitted from equation (34) the repulsive-interaction term that would produce the well. Partial justification may be offered for this, as follows: first, repulsive terms should be included in both S-components at x_m^B, and these tend to cancel; second, the repulsive component is in any case much less in magnitude than the attractive image-potential energy.

Schottky-type desorption. A second mechanism of evaporation from an ionic bonding state is genuine Schottky-type desorption, in which the ion escapes over a Schottky hump in the same charge state as that in which it was originally bound, and no ionisation

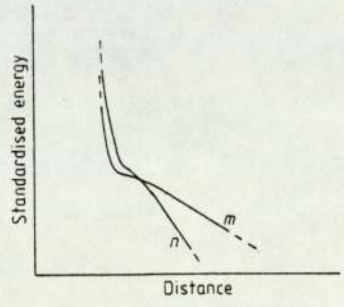


Figure 3. Ionic Term diagram for the case when no ionic bonding well exists.

event is involved in the escape process. Equation (21) is still valid, but reduces to the triviality

$$Q_{mm} = S_m^{\mathrm{P}} - S_m^{\mathrm{B}}. \tag{36}$$

The simplest useful Q-formula is then obtained by using the basic electrostatic, correlation, and repulsion terms in table 1, which together give rise to the well-known ionic potential containing a Schottky hump.

An expression for the field F_m^{HD} at which this hump 'just disappears' can be obtained analytically (Biswas and Forbes 1982), and has the form

$$F_m^{\mathrm{HD}} = \Gamma \cdot me/16\pi\varepsilon_0 a_m^2 \tag{37}$$

where a_m is the distance from the emitter's electrical surface of the bonding point for the m-fold ion, *in the absence of the field*; and Γ is a numerical factor whose value depends on the assumed form of the repulsive power law. For an inverse 9th-power repulsive potential, $\Gamma = 0.535$; for an inverse 12th-power law $\Gamma = 0.602$.

Use of a zero-activation-energy criterion for defining evaporation field leads us to identify the evaporation field with this hump-disappearance field F_m^{HD}.

Clearly, for fields above F_m^{HD} there is no possibility (within the framework of the simple formalisms discussed above) that an m-fold ion can be bound to the surface. Hence, if the quantity F_{mn}^{ev} predicted from equation (34) is greater than F_m^{HD} then the potential configuration depicted in figure 2 may not obtain, and F_{mn}^{ev} as given from equation (34) cannot be interpreted as an evaporation field.

In these circumstances the potential-curve configuration may be as in figure 3, and it is better to write equation (34) in the revised notation:

$$Fx_{mn}^{\mathrm{cr}} = K_{mn}/(n-m)e - (n+m)e/16\pi\varepsilon_0 x_{mn}^{\mathrm{cr}} \qquad (m \neq 0) \tag{38}$$

and to think of equation (38) as an equation relating the field F and the position x_{mn}^{cr} of the crossing-point of the m-fold and n-fold potential curves. (Though whether it is reasonable to ignore the repulsive terms in such a case is another matter.)

5. Application to tungsten field evaporation

For illustration I apply these formulae to tungsten. The possibility was raised elsewhere (Forbes 1980a) that some tungsten surface atoms may be essentially ionic prior to field evaporation, and it was shown that observed onset appearance energies were compatible with neutral or with ionic bonding.

Table 2 shows evaporation fields obtained from equations (30) and (34) for postulated single-step evaporation from 0^+, 1^+ and 2^+ initial bonding states into 1^+, 2^+ and 3^+ escape states†. The following thermodynamic data (Tsong 1978) were used: $\Lambda_0^0 = 8.66$ eV; $I_1 = 7.98$ eV; $I_2 = 18$ eV; $I_3 = 24$ eV; $\varphi^{\mathrm{E}} = 4.5$ eV; and the tungsten-atom radius ρ was taken as 137 pm. Bonding distances are not well known (particularly for ionic bonding), so calculations used three postulated values for x_m^{B}, namely ρ, $3\rho/2$, and 2ρ.

† In reality, single-step evaporation into a 3^+ state seems an implausible process; the data are included simply for illustration. The observed 3^+ and higher charge states of field-evaporated tungsten ions are almost certainly the result of post-ionisation after escape (see Kingham 1981). D R Kingham (private communication; also *Vacuum*, to be published) has also queried whether 2^+ ions can be formed from 0^+ in a single-step process, although the present author currently believes this to be possible.

Table 2 also shows values of the maximum zero-Q evaporation field as calculated from equations (32) and (35).

Table 2. Evaporation fields for tungsten atom of radius ρ, based on the zero-activation-energy criterion, predicted using the simple approximations inherent in equations (30), (32), (34) and (35). Predicted values are rounded to the nearest $\frac{1}{2}$ V nm^{-1}. (The conventionally given value for the observed evaporation field of tungsten is 57 V nm^{-1}.)

Charge state			F^{ev}_{mn} (V nm^{-1}) for bonding distances			
m	n	$K_{mn}/(n-m)$ (eV)	ρ	$3\rho/2$	2ρ	F^{ev}_{mn} (maximum) (V nm^{-1})
0	1	12.14	69.5	50.5	39.5	102.5
0	2	12.82	55	45.5	37	57
0	3	15.05	52.5	47.5	40.5	52.5
1	2	13.5	41	40	35	42
1	3	16.5	43.5	46	41	47.5
2	3	19.5	46.5	52.5	47	53

Table 3 shows values of the hump-disappearance field F^{HD}_m, for 1^+, 2^+ and 3^+ charge states, calculated for postulated zero-field ionic bonding distances of ρ, $3\rho/2$, and 2ρ, taking Γ to be 0.602 (i.e. 12th-power repulsion).

It deserves emphasis that the equations used in these calculations are the simplest possible approximations, and the tabulated results are valid only to the extent that these approximations are adequate.

Table 3. Values of the field F^{HD}_m at which the Schottky hump disappears, as predicted from equation (37) using $\Gamma = 0.602$ (12th-power repulsion). (ρ denotes the tungsten atomic radius.)

Charge state	F^{HD}_m (V nm^{-1}) for assumed zero-field bonding distances		
m	ρ	$3\rho/2$	2ρ
1	11.5	5.1	2.9
2	23.1	10.3	5.8
3	34.6	15.4	6.7

Three features of the tabulated results deserve comment. First, for both 1^+ and 2^+ initial states the hump-disappearance field F^{HD}_m is markedly lower than any of the calculated fields F^{ev}_{mn} for evaporation from these bonding states, and is also significantly lower than the accepted value (57 V/nm) for the low-temperature (near 80 K) evaporation field of tungsten. Within the framework of the equations in use, this suggests that it is inconsistent to assume that tungsten can be bound in an ionic state. For consistency we must assume that tungsten is bound in a neutral or 'primarily neutral'† state. One must also assume that at the evaporation field the configuration of the 1^+ and 2^+ potential

† When a surface is electrically charged the surface atoms must physically be partially ionic, to some extent. By describing a surface atom as 'primarily neutral' I mean that we treat the situation conceptually as if the atom were neutral, and the theory contains a field-dependent binding-energy correction that is presumed to behave much more like a polarisation term than like an electrostatic or image-potential term.

curves is basically as depicted in figure 3, and that the field values listed in table 2 for $m > 0$ are to be interpreted as (estimates of) the fields at which a crossing-point between the m-fold and n-fold curves is at the distance stated (as per the discussion of equation (38)).

Second, in each case one or other of the fields F_{02}^{ev} and F_{03}^{ev} is less than the field F_{01}^{ev} (as has been found by other authors, for example Müller and Tsong 1969). This suggests that the initial escape step involves transfer into a 2^+ or 3^+ state, rather than into a 1^+ state. Whether 2^+ or 3^+ is the preferred state depends (in this formalism) on the value assumed for x_0^B; my current view is that the larger values are more plausible, and hence that 2^+ is the expected escape state. However, the closeness of F_{02}^{ev} and F_{03}^{ev} shows that the intersection point of the 2^+ and 3^+ curves must be only slightly outside the bonding point, so the probability of post-ionisation of 2^+ to 3^+ is expected to be high.

Third, because F_{02}^{ev} is less than F_{01}^{ev} at each of the postulated bonding distances, it follows that the intersection of 0^+ and 2^+ curves is closer to the emitter's electrical surface than is the intersection of 0^+ and 1^+ curves. Similarly, since F_{12}^{ev} is less than F_{02}^{ev}, it follows that the intersection of 1^+ and 2^+ curves is closer in than the intersection of 0^+ and 2^+ curves. Since this last intersection is supposed to occur at the bonding distance, it follows that the $1^+/2^+$ intersection is inside the neutral-atom bonding point. An arrangement of potential curves schematically consistent with these results is shown in figure 4.

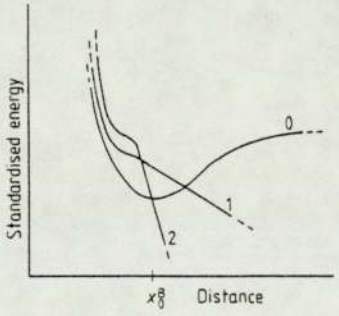


Figure 4. Term diagram for the field evaporation of tungsten, as indicated by the simple theory described in the text.

Oversimplifications. Points about the adequacy of simple approximations arise from the above analysis. First, if the intersection of 1^+ and 2^+ curves is found to be inside the neutral-atom bonding point, as it is above, then it was probably not an adequately self-consistent procedure to ignore the repulsion terms in formulating equations (34) and (35).

A more subtle point is that, even when repulsion terms are included, it is doubtful whether a curve configuration as shown in figure 2 could arise. The form of equation (37) shows that, if a_1 and a_2 are approximately equal, and if roughly the same repulsive law (and hence roughly the same Γ-value) applies to both singly- and doubly-charged ions, then the field at which the hump in the 2^+ curve vanishes will be greater than the field at which the hump in the 1^+ curve vanishes. This implies that it is not possible to have a situation in which there is a Schottky hump in the 1^+ curve but not in the 2^+ curve. Within the framework of the simple approximation involving correlation, repulsion and linear-electrostatic terms, the configuration shown in figure 2 is almost certainly not possible.

More generally, the closeness of the predicted evaporation fields for various parameter combinations suggests that (even if one postulates the validity of the elementary expressions listed in table 1) more effects and more terms ought to be included in the calculations. (Or possibly tungsten is a bad material to try out simple approximations on, its extensive employment in experiments notwithstanding.)

Summarising the results of the simple approximations, these suggest that tungsten surface atoms cannot be bound in an ionic state, and that the escape step involves a transition from a primarily neutral to a twofold-charged state.

These results have to be set in the context of experimental observations, and in particular the field evaporation of highly field-exposed atoms, such as those at the kink-sites of a tungsten (111) facet. Gauss' theorem indicates that, at a field strength of 57 V/nm, the excess charge per surface atom in a crystallographically flat (111) face would be about 0.55 e, and the actual effective charge on the kink-site atoms might be slightly higher than this. So it is by no means obvious that it is reasonable to think of the field-exposed atoms as 'primarily neutral', in the sense defined earlier.

Clearly there is an urgent need for a better theoretical understanding of the nature of charged surfaces on an atomic scale, and for improved methods of calculating the potential in which a desorbing ion moves. The hope that better approximations can be found for the S-component, particularly for 'partially ionic' initial bonding states, was one motive for presenting a generalised treatment of electrothermodynamic cycles.

6. Modified cycles

6.1. Non-integral charge states

It remains to deal with some matters left over from earlier sections. The first is the problem of non-integral charge-states. Either at the bonding point, or at the top of the potential barrier over which escape occurs, the desorbing entity may be in a charge state that is non-integral or ill-defined. This can occur if it is energetically favourable for the topmost electron (or electrons) originally associated with the desorbate to be in electron states whose wave-functions have some amplitude in the emitter and some amplitude in the desorbate, or if these electron(s) are in an electronic configuration resembling a chemical bond between the desorbate and the emitter.

Non-integral charge-state at the point of escape arises in the 'charge-draining' model of field desorption (Forbes 1981a), shown in figure 5. Let the potential energy at the top of the charge-draining hump (at position x_n^{p}) be $U^{\mathrm{pa}}(\mathrm{CD})$. Formally, this is related to the potential energy U_n^{p} for a 'pure' n-fold ion at position x_n^{p} by:

$$U^{\mathrm{pa}}(\mathrm{CD}) = U_n^{\mathrm{p}} + \Delta U^{\mathrm{p}}. \tag{39}$$

ΔU^{p} is a correction†, negative in value, that is it is sometimes described as due to 'configuration interaction'. Construction of energetic arguments similar to those used earlier now leads to revised equations incorporating the ΔU^{p} term.

The situation of a partially ionic or 'polar' (Gomer and Swanson 1963) bonding state is easy to treat formally. We use the subscript α to designate a polar bonding state and Λ_α^F to represent its bonding energy. (The reference zero is the same as before.) Thus, when both forms of correction exist, equation (23) becomes:

$$Q_{\alpha n} = (\Lambda_\alpha^F + H_n - n\varphi^{\mathrm{E}}) + S_n^{\mathrm{p}} + \Delta U^{\mathrm{p}}. \tag{40}$$

† In the literature ΔU^{p} is often written $-\Delta E - \frac{1}{2}\Gamma$, but there seems no advantage in this latter form, and possibly some disadvantage—see Forbes (1981a).

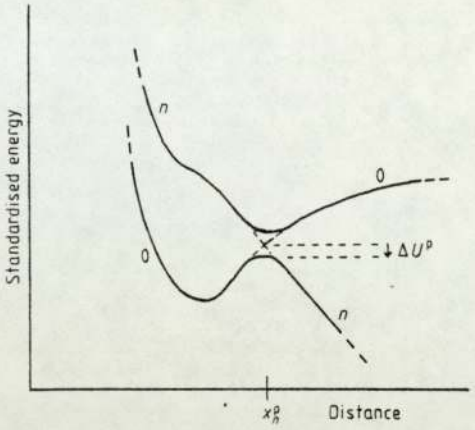


Figure 5. Term configuration assumed for a charge-draining mechanism of field evaporation from a neutral bonding state, showing the shift ΔU^P due to configuration interaction.

This equation is formally satisfactory, but of little help in numerical analysis of results. Currently, no satisfactory calculations of Λ_α^F or ΔU^P exist. Unfortunately, it now seems likely that for refractory metals the desorbate charge-state will be non-integral and/or ill-defined, both at the bonding point and at the escape point. So I feel that the absence of detailed treatments of Λ_α^F and ΔU^P, and their variation with field—and, more generally, the lack of a detailed theory of charged surfaces—will increasingly be seen as key problems in field desorption theory.

6.2. *Fast processes*

The treatments of electrothermodynamic cycles given here and elsewhere require that the individual cycle steps be 'slow' and 'thermodynamically reversible'. This ensures that the emitter charge distribution is always in electrical equilibrium under the influence of the field of the external particle, and avoids the complications associated with (for instance) dynamic image potentials.

If the requirement is not met (i.e. the external particle motion is 'fast') then heat is generated in the emitter, and the particle kinetic energy turns out correspondingly less than expected from the standard potential energy. This happens whether the motion is towards or away from the emitter: heat is generated in both cases: hence fast motion is in a certain sense *not* thermodynamically reversible.

We need to enquire whether surface field ionisation is actually 'slow' or actually 'fast', and if 'fast' then how to represent it diagrammatically. The relevant comparison seems to be between, on the one hand, a characteristic time associated with the relaxation of the emitter electron distribution, and on the other hand characteristic times associated with: (*a*) the electron transfer process in field ionisation and (*b*) the ion departure speed.

Because field ionisation involves electron tunnelling, and because the ion has only thermal energy on escape, it has always seemed to the author that field desorption must be 'slow' in the above sense. (Detailed discussion of characteristic times is not entirely straightforward, and is beyond the scope of this paper.) However, it is clear how fast processes could be represented. If ion departure were fast, then in effect the ion moves in a potential slightly lower than the standard potential energy, so a 'vertical' (i.e. Franck–Condon) transition results in the ion having lower standardised energy (see figure 6). If electron transfer were fast, then in effect heat is transferred to the emitter,

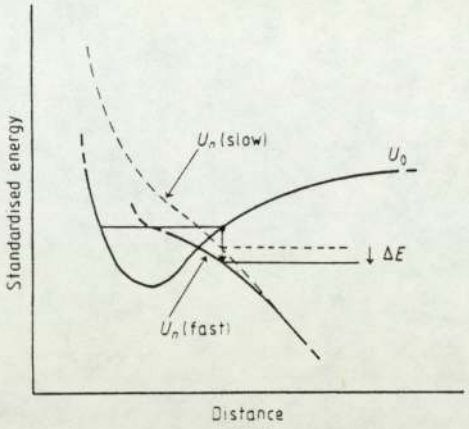


Figure 6. To show that an ion formed by a Franck–Condon transition has lower final standardised energy if the ion departure is 'fast'. Away from the emitter, the ion kinetic energy is different by an amount ΔE. (ΔE is a negative quantity.)

and on an alignment diagram a one-electron transition must be represented as 'diagonally downwards', as in figure 7; this implies that the critical surface inside which auto-ionisation is significantly inhibited must (for fast transfer) be shifted slightly outwards from its standard 'slow' position.

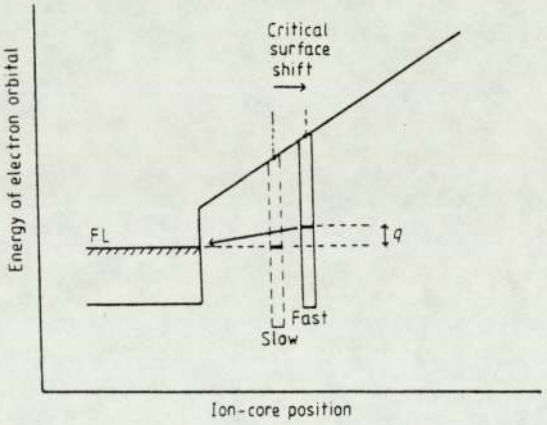


Figure 7. To illustrate how a 'fast' electron transition is represented on an electron-level alignment diagram; also, that there is an outwards shift in the critical surface. In a fast transition an electron tunnelling to the fermi level is presumed to create an amount of heat q in the emitter electron distribution. The 'dashed' well shows the critical orbital-level alignment if the transition is presumed to be slow, the 'full' well the alignment if the transition is presumed to be fast. (For clarity the effect is exaggerated.)

The above analysis is believed to be essentially correct at the classical level of debate normally employed in field-ion theory. But a fuller, quantum-mechanical, analysis may lead to a somewhat more complicated story.

6.3. Correlation energy in two-electron hopping processes

Another effect not previously discussed arises with multiple-electron-hopping processes (if these exist). The conventional electrothermodynamic cycle specifies that, in creating

the ion, the electrons be removed one by one, in remote field-free space. The total work needed (H_n) is then the sum of the relevant free-space ionisation energies. However, if two or more electrons were removed together and simultaneously then additional work would be needed. This follows because, once well away from the resulting ion the removed electrons would have electrostatic interaction energy. This could be re-adsorbed by the hypothetical agent as the electrons then separated. But the net work done by the external agent in the whole removal + separation process must still be H_n, so it follows that the initial removal needed extra work, this work being identified with a correlation energy between the removed electrons.

It seems that broadly analogous considerations could apply during the actual field desorption process. There would be an electrostatic interaction between the 'hopping' electrons, if two or more hop simultaneously, and the energy would presumably manifest itself eventually as heat in the emitter. The energy obviously has to come from the applied field, so it seems that the critical surface (inside which no significant auto-ionisation occurs) for a multiple electron transition must be slightly outside the crossing surface defined by the standard 0^+ and n^+ potential-energy curves, as depicted in figure 8.

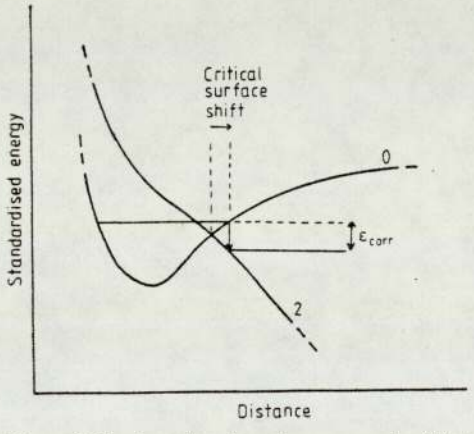


Figure 8. To show that there is an outwards shift in the critical surface if, in a two-electron hopping process, a correlation interaction energy ε_{corr} is transferred with the electrons.

In practice, it is doubtful if transitions involving more than two electrons occur. (There is also the question as to whether two-electron transitions in fact occur, raised by Kingham.)

7. Summary

The achievements of this paper have been as follows.

(1) General expressions have been put forward for the standard potential energy of an ion near a charged surface, and the elementary approximations used to represent contributions to this potential energy have been tabulated.

(2) It has been shown how these expressions are used to provide formulae for field-desorption activation energy and for electron orbital level.

(3) It has been shown qualitatively how effects associated with non-integral charge states, with 'fast' atomic processes, and with correlation energy in two-electron transitions, can be accommodated within the discussion.

(4) Simple algebraic expressions, based on the common elementary approximations, have been derived for predicting the field strength for desorption from ionic bonding states. Applying these formulae suggests that tungsten atoms are bound, prior to evaporation, in a primarily neutral state; but doubts also exist about the validity of elementary approximations and simple formulae, certainly in this case.

More generally, I have no doubt that the full description of the field desorption process will require relatively accurate quantum-mechanical calculations of the various potentials involved. (And probably sooner rather than later.) However, my perception is that such calculations may be more profitable if performed against the background of properly formulated classical arguments. Thus this paper has attempted to set out formally, in more detail than hitherto, a classical theory of electrothermodynamic cycles. But it must be seen only as an intermediate stage, in a developing scientific analysis.

Acknowledgment

This work forms part of a research programme funded by the UK Science and Engineering Research Council.

Appendix 1. Glossary of symbols used

The following table identifies the meaning of the symbols used in this paper. Where possible, an attempt has been made to ensure consistency with other publications, but the very large number of standard symbols needed in field-ion theory makes the coherent choice of symbols a continuing problem, and some duplication and inconsistency (as between papers) is inevitable.

Quantities	
a_m	Distance of bonding-point for m-fold ion from metal's electrical surface, in the absence of any applied field
b_n	Proper SI polarisability of n-fold ion
B_n	First field-gradient polarisability for n-fold ion
c_n	Polarisation-energy coefficient for n-fold ion [$= b_n \times (F_{\text{loc}}/F)^2$]
C_n	Second field-gradient polarisability for n-fold ion
D	Constant in rth power attractive law
e	Elementary (proton) charge
E	Standardised energy (i.e. the sum of the ion kinetic energy and the standard potential energy for the charge state in question)
ΔE	Change in standardised energy of departing ion, relative to the 'slow' case, if the ion departure is 'fast'
F	External field (often called 'applied field' in literature)
F^{ev}_{mn}	Evaporation field (i.e. value of F defined by zero-activation-energy criterion), for entity bound in m-fold-charged state to escape in n-fold-charged state
F^{HD}_m	Value of F at which the Schottky hump in the potential U_m (or S_m) 'just disappears'
F_{loc}	Self-consistent local field
F'_{loc}	Gradient of self-consistent local field

G	Constant in tth power repulsive law (between atom or ion and surface)
H_n	Energy needed to form n-fold ion, from corresponding neutral, in remote field-free space, with the electrons being left well-separated in remote field-free space
I_k	kth free-space ionisation energy
K_{mn}	Thermodynamic quantity defined by equation (26) if $m = 0$, and by equation (33) if $m \neq 0$
K_j	Constant in tth power repulsive law (between ion cores)
m	Charge-number of entity in bonding state or in initial state
n	Charge-number of entity after escape and before any subsequent post-ionisation
p	Effective dipole moment of desorbate, in initial bonding state
q	Heat created in emitter electron distribution as result of 'fast' electron transition
Q_{mn}	Activation energy, for entity bound in m-fold-charged state to escape in n-fold-charged state
r	Exponent in attractive power law
S	Split-referenced potential-energy component
t	Exponent in repulsive power law
u	(Total) electrostatic potential energy of proton
u_{app}	Component of electrostatic potential energy of proton due to applied voltage
U_n	Standard potential energy (standard Term) for n-fold ion
ΔU^{p}	Change in potential energy, due to configuration interaction
W_{mn}^{qs}	Standard work done in creating n-fold ion at position s from m-fold ion at position q
x	Distance of reference point in desorbate entity from emitter's electrical surface
γ_n	Second hyperpolarisability for n-fold ion
Γ	Constant appearing in equation (37) for F_m^{HD}
δ^s	Value of potential correction due to patch and/or surface fields, at position s
ε_m^s	Level of topmost filled electron orbital in m-fold-charged entity, relative to emitter fermi level, when the reference point in this entity is at position s
ε_0	Electric constant (sometimes called 'permittivity of free space')
$\varepsilon_{\text{corr}}$	Inter-electron correlation energy transferred to emitter in multiple-electron-hopping process
η_n	Purely-chemical component of standard potential energy, for n-fold-charged entity
θ	Constant in Tsong's (1971) expression for the field-induced reduction in the binding energy of a metal atom to a surface
κ	Force-constant in parabolic-well approximation
Λ_m^0	Binding energy of m-fold-charged entity to surface, in absence of any applied voltage or field
Λ_m^F	Binding energy of m-fold-charged entity to surface, when applied field is present
$\Delta\Lambda_m^F$	Field-induced increase in binding energy of m-fold-charged entity to surface
ρ	Atomic radius (of tungsten, here)
σ_{0n}	Constant appearing in equation (31) for F_{0n}^{ev}
Υ	Classical electrostatic potential
φ^{E}	Local work-function of emitter
φ^{tot}	So-called 'total' work-function of emitter (i.e. work needed to remove electron to large distance from emitter)
χ^{E}	Chemical component of work-function of emitter
ψ^{E}	Electrostatic component of local work-function of emitter

Labels (excluding quantities used as labels)

app	'Due to applied field'
B	Value appropriate to a bonding point (whether field present or not)
cr	Value appropriate to a crossing point (or surface) between potential curves
CD	Quantity appropriate to charge-draining mechanism of field desorption
E	Indicates *local* emitter work-function; can also be seen as label for a position 'somewhat outside the emitter surface' used in the definition of local work-function
ev	Field value defined by zero-activation-energy criterion for evaporation or desorption field
HD	Field value appropriate to disappearance of Schottky hump
i	Value appropriate to interior of emitter
I	Value appropriate to point of inflexion in ionic potential-energy curve
j	Labels individual surface atoms
p	Value appropriate to 'point of escape'
pa	Potential energy value at 'top of pass' over which escape occurs
q, s	Labels for arbitrary positions above emitter surface
R	Value appropriate to point in remote field-free space
zv	Value for zero applied voltage
α	Value appropriate to entity in bonding charge-state that is 'polar'
0	As subscript—value appropriate to neutral entity
0	As superscript—value appropriate to zero external field

References

Biswas R K and Forbes R G 1982 *J. Phys. D: Appl. Phys.* **15** 1323–38
Brandon D G 1963 *Br. J. Appl. Phys.* **14** 474–84
—— 1964 *Surface Sci.* **3** 1–18
Forbes R G 1976 *Surface Sci.* **61** 221–40
—— 1978 *Surface Sci.* **70** 239–54
—— 1980a *J. Phys. D: Appl. Phys.* **13** 1357–63
—— 1980b *Prog. Surface Sci.* **10** 249–85
—— 1981a *Surface Sci.* **102** 255–63
—— 1981b *Surface Sci.* **108** 311–28
—— 1982 *Appl. Phys. Lett.* **40** 277–9
Gomer R 1959 *J. Chem. Phys.* **31** 341–5
—— 1961 *Field Emission and Field Ionization* (Cambridge, Mass.: Harvard University Press)
Gomer R and Swanson L W 1963 *J. Chem. Phys.* **38** 1613–29
Haydock R and Kingham D R 1980 *Phys. Rev. Lett.* **44** 1520–3
Kingham D R 1981 *PhD thesis* University of Cambridge
Landau L D and Lifschitz E M 1958 *Quantum Mechanics* (New York: Pergamon)
Lang N D and Kohn W 1973 *Phys. Rev.* B **7** 3541–50
Müller E W 1956 *Phys. Rev.* **102** 618–24
—— 1960 *Adv. Electron. Electron Phys.* **13** 83–179
Müller E W and Tsong T T 1969 *Field-Ion Microscopy: Principles and Applications* (Amsterdam: Elsevier)
—— 1974 *Prog. Surface Sci.* **4** 1–139
Southon M J 1963 *PhD thesis* University of Cambridge
Swanson L W and Gomer R 1963 *J. Chem. Phys.* **39** 2813–36
Theophilou A K and Modinos A 1972 *Phys. Rev.* B **6** 801–12
Tsong T T 1971 *J. Chem. Phys.* **34** 4205–16
—— 1978 *Surface Sci.* **70** 211–33
Tsong T T and Müller E W 1969 *Phys. Rev.* **181** 530–4
Wafi M K 1981 *PhD thesis* University of Aston in Birmingham

J. Phys. D: Appl. Phys., 15 (1982) 1323–1338. Printed in Great Britain

Theoretical arguments against the Müller–Schottky mechanism of field evaporation

R K Biswas and Richard G Forbes
Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET, UK

Received 13 January 1982

Abstract. In a theoretical investigation of the Müller–Schottky mechanism of field evaporation, an 'extended' image-hump formalism has been developed that includes a term relating to the repulsive ion–surface interaction. This formalism is used to calculate the field at which the Schottky hump disappears. It is then shown that: (i) the conventional simple image-hump formalisms are invalid because they predict evaporation fields higher than the hump-disappearance field; (ii) the ability of the simple formalisms to roughly predict observed evaporation fields is a numerical coincidence rather than a meaningful test of the formalisms; (iii) observed evaporation fields are in most cases higher than the fields at which the Schottky hump disappears. It is concluded that normal metal field evaporation at low temperatures cannot be described by any image-hump formalism yet considered, and that such evaporation almost certainly takes place via a Gomer-type surface charge-exchange mechanism.

1. Introduction

The application of a sufficiently high electric field to a metal surface will cause the surface atoms to desorb as ions. This field evaporation process is of significance in atom-probe field-ion microscopy and related techniques, and may be an emission mechanism in liquid–metal field–ion sources (see, for example, Aitken and Mair (1980)). The sources have considerable technological potential in the areas of lithography and microcircuit fabrication (Clampitt 1981), and it would be valuable to have a better understanding of how they worked. A better understanding of conventional low-temperature (near 80 K) refractory-metal field evaporation might be a useful preliminary, and also is of interest in its own right.

Following recent work by Ernst (1979), Ernst and Bozdech (1980), Haydock and Kingham (1980) and Kingham (1982), it now seems virtually certain that for most metals field evaporation is a two-stage process. First a metal atom *escapes* as a singly-charged or doubly-charged ion; subsequently it may be *post-ionised* into higher charge states.

Many refractory metals can be induced to evaporate into 3^+ or higher charge states (see, for example, Müller and Krishnaswamy (1976)), and it is difficult to see how these ions could be formed in a single-step process. The demonstration that post-ionisation is likely has removed this difficulty, and allows us to accept as a paradigm the proposition that escape occurs into a singly or (possibly) doubly-charged state. It also serves to redirect theoretical attention towards the nature of the escape step.

0022-3727/82/071323 + 16 $02.00

As is well known, two types of escape process are commonly discussed: the Müller–Schottky ('image-hump') mechanism, in which ionisation precedes escape; and Gomer-type mechanisms, in which ionisation and escape occur simultaneously in a surface charge-exchange process. Corresponding potential diagrams are shown as figure 1.

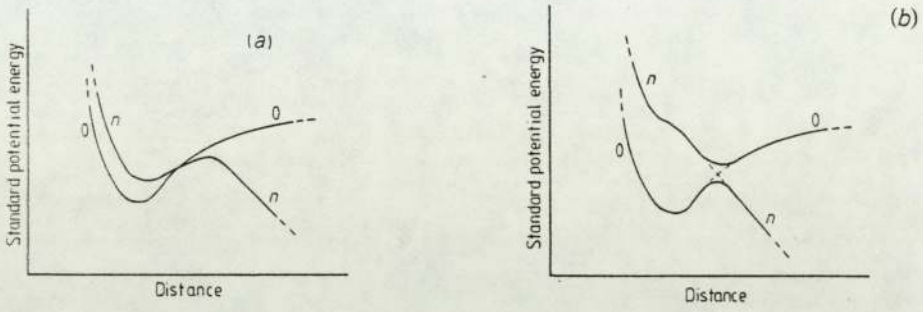


Figure 1. Potential energy diagrams corresponding to the (*a*) Müller–Schottky and (*b*) charge-exchange-type mechanisms of field evaporation. The curves are labelled with the charge numbers of the associated charge states.

At any particular field strength these mechanisms are incompatible alternatives. To get reliable atomic-level information from field-evaporation experiments one must know what the evaporation mechanism is, because the relevant formulae are mechanism dependent (as stressed by McKinstry (1972)). There is some consensus that a Gomer-type mechanism is more likely (see Tsong (1978a) for example), but the Müller–Schottky model is often used, probably because there is a simple associated mathematical formalism that happens to give roughly correct predictions of evaporation field. However, approximately correct field predictions can also be obtained from a charge-exchange based analysis (e.g. Müller and Tsong (1974)). So ability to predict evaporation field is not a test of existing formalisms.

This situation, and our recent work on field-sensitivity data (Forbes *et al* 1982), have stimulated a new look at the internal self-consistency of image-hump formalisms. Physical reservations about simple image-hump formalisms have been expressed on numerous occasions (for example: Brandon 1963, Müller and Tsong 1969, Tsong 1971, McKinstry 1972, Forbes 1978, Tsong 1978a). The gist of the objections is: (i) the distance of the Schottky hump from the metal's electrical surface, as calculated by the simplest formalism, using observed evaporation fields, is too small to be plausible (because it is less than an atomic radius); (ii) if repulsive forces between the ion core electrons and the metal substrate electrons were taken into account, then it is doubtful whether any Schottky hump actually exists in the ionic potential-energy curve.

The present authors fully share these reservations. The purpose of this paper is to put actual numbers onto the second of the above objections. By choosing a specific mathematical form for the repulsive potential, it is possible to make numerical estimates of the field F^{HD} at which the Schottky hump disappears. Values of this field can then be compared with: (a) evaporation fields as predicted by the simpler image-hump formalisms (which tests their theoretical self-consistency); and (b) observed evaporation fields (which tests whether evaporation occurs by the Müller–Schottky mech-

anism). The idea of doing this is hardly new, but detailed results of such calculations have not previously been published, as far as we are aware.

The structure of the paper is as follows. In § 2 we present a basic analysis of the energetics of field evaporation and of the Müller–Schottky mechanism, followed by accounts of the conventional 'simple' image-hump formalisms and of new 'extended' formalisms that take the repulsive term into account. In § 3 we investigate the self-consistency of the simple formalisms; § 4 compares the results of the extended formalisms with observed evaporation fields. Section 5 explains away the apparent success of the simplest image-hump formalism in predicting evaporation fields. We summarise our conclusions in § 6.

The notation and conventions in this paper generally follow those in Forbes 1982b, but with minor simplifications in use of suffices. The International (four electric dimensions) System of Measurement is employed.

Finally, here, we would draw attention to the logical distinction that now has to be made between a *mechanism* and a *formalism*. A mechanism is a physical thing, corresponding to an arrangement of potential curves relative to each other. A formalism is a mathematical thing, determined by the precise choice of algebraic expressions used in representing the potential curves. In particular, this paper discusses four different image-hump formalisms. These are different approaches or approximations in formalising (mathematically) the Müller–Schottky mechanism. It follows that the phrase 'image-hump model', much used in past discussion, is now logically ambiguous; hence we now avoid it, and prefer to call the physical mechanisms by the names of their originators.

2. Analysis of the Müller–Schottky mechanism

2.1. Basic energetics

Consider the interaction of a neutral or 'primarily neutral†' atom with a metal surface as represented by a 'jellium' type model (see, for example: Müller and Tsong (1974), Lang and Kohn (1973)), and take the zero of energy to be the potential energy of the neutral atom in remote field-free space.

The potential energy U_0^B of the neutral atom at its bonding point is, in the conventional simple approximation, written in the form:

$$U_0^B = -\Lambda^F = -\Lambda^0 - \Delta\Lambda = -\Lambda^0 - \tfrac{1}{2}c_0F^2 \tag{1}$$

where Λ^0 is the neutral-atom binding energy in the absence of any external field; Λ^F is the neutral-atom binding energy in the presence of an external field F; $\Delta\Lambda$ is the field-induced increase in binding energy; and c_0 is a parameter, with dimensions of SI polarisability, that is associated with a field-induced modification of the electronic structure of the bound 'primarily neutral' atom.

c_0 is sometimes called 'polarisability' in field-ion literature, but it cannot be identified with the parameter called 'polarisability' in textbooks, because it is not defined in terms of the local field acting on the atom in question. We thus call it a 'polarisation energy

† When a surface is electrically charged the surface atoms must physically be partially ionic, to some extent. By describing a surface atom as 'primarily neutral' we mean that the situation is treated conceptually as if the atom were neutral, and the theory is deemed to contain a field-dependent binding-energy correction term that behaves more like a polarisation term than like an electrostatic or image-potential term. Where appropriate in this paper the phrase 'neutral atom' means surface atoms that are 'primarily neutral' in the above sense.

coefficient'. It also deserves comment that there could be doubt about the suitability of an F^2 form to represent the binding energy increase $\Delta\Lambda$: it might be that c_0 as *defined* by equation (1) is not constant, particularly if the physical effects operating include things other than polarisation of the orbitals of the neutral atom. However, in the present context these reservations are of little significance, the important thing being that c_0 should be positive and greater than the ionic polarisation-energy coefficient c_n: this hypothesis is physically plausible and has always been accepted.

Now consider the interaction of the corresponding atomic ion with the metal surface. The standard potential energy $U_n(x, F)$ of an n-fold-charged ion, at position x outside a surface of local work-function φ^E, above which there is an external field F, can be written in terms of a 'configurational' part and a 'variable' component $S_n(x, F)$ (see Forbes 1982b). Thus:

$$U_n(x, F) = H_n - n\varphi^E + S_n(x, F) \tag{2}$$

where n is the ion charge-number, and H_n the energy needed to form an n-fold ion from the neutral, in remote field-free space, being given by the sum of the first n free-space ionisation energies.

Differences of opinion exist as to the best place to locate the origin of the distance coordinate normal to a charged surface. Our convention is based partly on intuition, partly on Lang and Kohn's (1973) work. Using a jellium model, they showed that within the electron charge distribution there is a single plane with the property that, at sufficiently large distances from the plane, the electrostatic component of ion potential energy is proportional to, and the correlation component of potential energy inversely proportional to, distance from the plane. We assume that such a plane also exists for a real (structured) charged surface: we call the plane the *electrical surface*, and measure distance x from it: this simplifies the form of algebraic expressions. (Note that our x is thus different from Lang and Kohn's, and our origin different from that of Müller and Tsong (1974).)

$S_n(x, F)$, as defined in equation (2), can then be represented in the form:

$$S_n(x, F) = -neFx - n^2e^2/16\pi\varepsilon_0 x - \tfrac{1}{2}c_nF^2 + G/x^t \tag{3}$$

where e is the elementary (proton) charge, ε_0 the electric constant, c_n the ionic polarisation-energy coefficient, and G a constant associated with a tth power repulsive law. The terms in equation (3) represent, respectively, electrostatic potential energy, correlation potential energy, ionic polarisation energy, and a repulsive interaction energy.

The assumption, involved in equation (3), that the repulsive-interaction term is correct (or at any rate adequate) when distance is measured from the electrical surface, might be subject to challenge, but it constitutes a useful first approximation.

It should be realised that, by defining distance x as measured from the metal's electrical surface, we are avoiding awkward questions as to where this electrical surface is located, relative to the positions of the nuclei of the metal surface atoms. As will be seen later, this question of the location of the electrical surface is in fact important, but no general consensus exists in the literature, and its detailed discussion must regretfully be put beyond the scope of the present paper.

Returning to the energetics of ion behaviour, we may deduce from equations (1) and (2) that the ion potential energy, *relative to the level of the neutral bonding state*, is given by:

$$W_n(x, F) = U_n(x, F) - U_0^B = (\Lambda^F + H_n - n\varphi^E) + S_n(x, F). \tag{4}$$

Thus, if we use x_n^{Sh} to denote, for the n-fold ion, the distance of the Schottky hump from the metal's electrical surface, then the activation energy $Q_n(F)$ necessary for an atom, initially bound as a neutral, to escape over this Schottky hump as an n-fold ion is:

$$Q_n(F) = W_n(x_n^{Sh}, F) = (\Lambda^F + H_n - n\varphi^E) + S_n(x_n^{Sh}, F). \tag{5}$$

The condition for an ion to be at the top of a Schottky hump can be expressed in the form:

$$(\partial W_n/\partial x)|_{x_n^{Sh}} = (\partial S_n/\partial x)|_{x_n^{Sh}} = 0. \tag{6}$$

This condition provides a relationship between x_n^{Sh} and F that, at least in principle, enables x_n^{Sh} to be eliminated from equation (5), leaving this as a function of F.

2.2. The conventional image-hump formalisms

In the simple image-hump formalisms only the first two or three terms in the expression (3) for $S_n(x, F)$ are considered. Thus we take:

$$S_n(x, F) = -neFx - n^2e^2/16\pi\varepsilon_0 x - \tfrac{1}{2}c_n F^2. \tag{7}$$

Application of condition (6) then gives:

$$neF = n^2e^2/16\pi\varepsilon_0(x_n^{Sh})^2 \tag{8}$$

or

$$x_n^{Sh} = \tfrac{1}{2}(ne/4\pi\varepsilon_0)^{1/2}F^{-1/2}. \tag{9}$$

Hence on substituting into equation (7), and then into equation (5), and also using equation (1) to expand Λ^F, we obtain the familiar formula:

$$Q_n(F) = (\Lambda^0 + H_n - n\varphi^E) - (n^3e^3F/4\pi\varepsilon_0)^{1/2} + \tfrac{1}{2}(c_0 - c_n)F^2. \tag{10}$$

The next step is to define, within the framework of the Müller–Schottky mechanism, a 'zero-Q' evaporation field F_n^{ev} by the requirement that:

$$Q_n(F_n^{ev}) = 0. \tag{11}$$

If the polarisation terms are ignored, then we obtain the following formula for evaporation field, in terms of thermodynamic parameters:

$$F_n^{ev} = (4\pi\varepsilon_0/n^3e^3)(\Lambda^0 + H_n - n\varphi^E)^2. \tag{12a}$$

Equation (12*a*) is the *'basic'* image-hump formalism. Alternatively, if the polarisation terms are included, a self-consistent solution must be found to the equation:

$$(n^3e^3F_n^{ev}/4\pi\varepsilon_0)^{1/2} - \tfrac{1}{2}(c_0 - c_n)(F_n^{ev})^2 = (\Lambda^0 + H_n - n\varphi^E). \tag{12b}$$

As long as the polarisation term is sufficiently small, this equation has a solution that leads to a value of F_n^{ev} slightly or somewhat greater than would be obtained from equation (12*a*).

Equations (12*a*) and (12*b*) we refer to collectively as the *'simple* image-hump' (SIH) formalisms; specifically, equation (12*b*) is the SIH formalism with F^2 terms.

Tabulations of F_n^{ev} as given by one or other of the simple formalisms have been presented by various authors, including Müller (1960), Brandon (1963) and Tsong (1978a). If, using Brandon's criterion, it is assumed that the evaporation field is the smallest of those calculated for different values of n, then it is found that calculated fields

agree with observed fields to within a factor of 1.5 or better. This agreement has sometimes been held to justify the simple formalisms.

2.3. Extended formalisms including the repulsive term

We now develop '*extended*' formalisms that include the repulsive interaction term. It is convenient to initially disregard the polarisation terms, because in our surface model they just produce a uniform energy shift, independent of position, in the relative levels of the ionic and neutral potential curves.

The shape of $S_n(x, F)$, and hence $U_n(x, F)$, is dependent on field, as shown in figure 2. At zero field (figure 2(a)) the shape is that of an ionic bonding potential. At low fields (figure 2(b)) the Schottky hump develops. As the field increases there comes a critical situation, at some field F_n^{HD} (figure 2(c)), at which the hump has just disappeared, and the curve is flat at the point of inflexion. Then, above this field (figure 2(d)), there is no hump at all in this curve.

The objection to the simple formalisms is as follows. If, at the field F_n^{ev} as calculated in the preceding section from equation (12), there is in fact no Schottky hump in the potential curve as given by the full expression (3), then it was *not* mathematically legitimate to apply condition (6) to the abbreviated expression (7) in which the repulsive term is ignored. At high fields, above F_n^{HD}, equation (7) is non-physical, and the maximum in it is a procedural artefact. Consequently it is an exercise of doubtful validity to equate the energy level of this artefact maximum to the bonding level of the neutral, as the simple formalisms do.

The value of F_n^{HD} is thus of some interest, and we now show how it can be calculated analytically in the framework of expression (3).

First consider an n-fold ion at a position outside the metal surface, in the absence of any electric field. In this case its standard potential energy can be written in the form:

$$U_n = (H_n - n\varphi^{\mathrm{E}}) - D/x + G/x^t \tag{13}$$

where for notational convenience we use D to denote:

$$D = n^2e^2/16\pi\varepsilon_0. \tag{14}$$

The potential U_n has the shape shown in figure 2(a). Suppose that the potential minimum is at a distance a_n from the metal's electrical surface. It follows that:

$$\partial U_n/\partial x|_{x=a_n} = D/a_n^2 - tG/a_n^{t+1} = 0. \tag{15}$$

Hence:

$$G = Da_n^{t-1}/t. \tag{16}$$

Thus, in the presence of a field (but ignoring the polarisation term) equations (2) and (3) become:

$$U_n(x, F) = (H_n - n\varphi^{\mathrm{E}}) - neFx + D[a_n^{t-1}/tx^t - 1/x]. \tag{17}$$

We next determine the position of the point of inflexion. Let this be at a distance x_{I} from the metal's electrical surface. Clearly, at this distance:

$$\partial^2 U/\partial x^2|_{x_{\mathrm{I}}} = D[(t+1)a_n^{t-1}/x_{\mathrm{I}}^{t+2} - 2/x_{\mathrm{I}}^3] = 0. \tag{18}$$

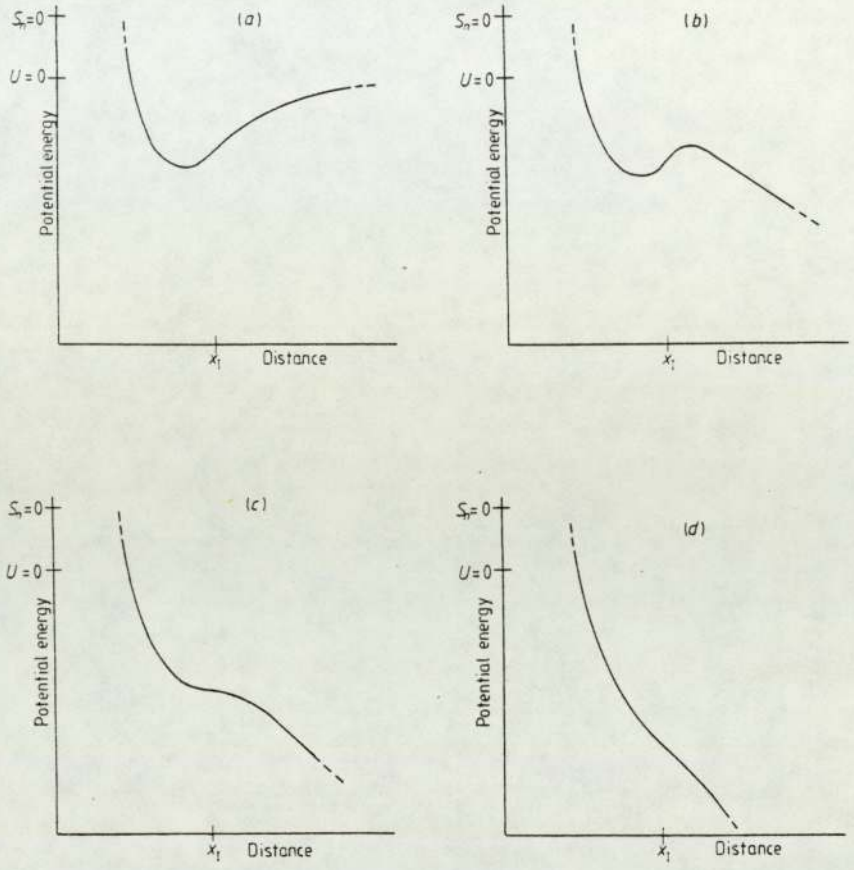


Figure 2. Ionic potential energy curves for several different fields: to illustrate the formation and subsequent disappearance of the Schottky image hump: (*a*) at zero field; (*b*) at low fields, where a hump exists; (*c*) at a higher field where the hump has just disappeared; (*d*) at a much higher field where no hump exists. These curves are schematic, and are not intended as accurate representations of the equations in the text, but note that the level $S_n = 0$ and the position x_I of the point of inflexion are the same at all field strengths.

This leads to the result:

$$x_I = \gamma a_n \tag{19}$$

$$\gamma = [(t+1)/2]^{1/(t-1)}. \tag{20}$$

Note that, in this case where repulsive and attractive forces are measured from the same origin, the value of γ depends only on the exponent assumed for the repulsive power law, and not on the charge number n, nor the field F.

As the field strength is increased from a low value, the height of the Schottky barrier decreases, and the position of the potential maximum moves in towards $x = x_I$. At the critical field F_n^{HD}, at which the barrier 'just vanishes', the potential curve must be horizontal at its point of inflexion, and thus:

$$\partial U_n/\partial x|_{x_I} = D(-a_n^{t-1}/x_I^{t+1} + 1/x_I^2) - neF_n^{HD} = 0. \tag{21}$$

Hence, using equations (19) and (14), we may solve for F_n^{HD}, to give:

$$F_n^{HD} = (1/\gamma^2 - 1/\gamma^{t+1})(ne/16\pi\varepsilon_0 a_n^2) \tag{22}$$

$$\simeq \Gamma(ne/16\pi\varepsilon_0 a_n^2). \tag{23}$$

Γ represents the first bracket in equation (22), and is also a function only of the assumed repulsive exponent, t. Values of γ and of Γ, for values of t lying between 6 and 15, are shown in table 1.

Table 1. Values of the constants γ, Γ, and Θ, defined by equations (20), (23) and (24), respectively, for various values of the repulsive power law exponent t.

t	γ	Γ	Θ
6	1.285	0.433	−1.297
7	1.260	0.472	−1.361
8	1.290	0.506	−1.412
9	1.223	0.535	−1.451
10	1.209	0.560	−1.489
11	1.196	0.582	−1.520
12	1.186	0.602	−1.546
13	1.176	0.619	−1.570
14	1.168	0.636	−1.591
15	1.160	0.650	−1.609

Substituting back into equation (17), we can derive an expression for the absolute level U_n^* of the point of inflexion when the hump just disappears. After some algebraic manipulation we obtain:

$$U_n^* = (H_n - n\varphi^E) + S_n^* \tag{24a}$$

where

$$S_n^* \simeq [(t+1)/t\gamma^t - 2/\gamma](n^2e^2/16\pi\varepsilon_0 a_n) \tag{24b}$$

or

$$S_n^* \simeq \Theta(n^2e^2/16\pi\varepsilon_0 a_n). \tag{24c}$$

The parameter Θ denotes the first bracketed term in equation (24*b*), and again is a function only of the assumed repulsive exponent t; values are given in table 1. The approximate-equality sign is written in equations (24*b*) and (24*c*) because a more complete formulation would include the polarisation term that appears in equation (2); if this is included then S_n^* is given by:

$$S_n^* = \Theta(n^2e^2/16\pi\varepsilon_0 a_n) - \tfrac{1}{2}c_n(F_n^{HD})^2. \tag{25}$$

2.4. Choice of repulsive power law

In order to carry out a numerical estimation of F_n^{HD}, a choice has to be made as to the value of the repulsive power-law exponent t. As there is little to guide us in this, other than well-known simple considerations (see, for example, Dash (1975)), we shall present results for the cases $t = 9$ and $t = 12$. In general, our conclusions are not unduly sensitive to the value assumed for t.

3. Self-consistency test of the simple formalisms

For a given charge state and repulsive power law, the calculated value of the field F_n^{HD}, at which the Schottky hump disappears, depends on the value assumed for a_n. This is clear from equation (23). Figure 3 shows this predicted dependence, for the three charge states $n = 1, 2, 3$. In each figure, plots are drawn for $t = 9$ and $t = 12$.

To test the validity of the simple image-hump formalisms, for a given species, we use chemical thermodynamic data to calculate values of the evaporation field, using equation (12*a*). These values are then plotted on the appropriate figure, at an a_n value corresponding to the atomic radius of the species in question (—as estimated from the nearest-neighbour interatomic distance in the relevant crystal lattice). The species used in this comparison are listed in table 2, in order of descending value of F_1^{ev}; we have included most of the elements commonly employed in the conventional field-ion techniques. The atomic radii have been stated to the nearest $\frac{1}{2}$ pm. Except where otherwise indicated, the thermodynamic data are taken from the tabulation in Tsong (1978a).

Ideally, we would like to compare the basic-formalism predicted evaporation field with the 'true' field at which the Schottky hump disappears, for the species and charge state in question. But, because we do not know exactly where the metal's electrical surface is located, nor the effective radii of the (possibly highly polarised) ions involved, we cannot know the true values of a_n for the various species and charge states. Thus we choose to carry out a comparison between the basic-formalism predicted field F_n^{ev} and the field F_n^{HD} that corresponds to taking a_n equal to the radius (ρ_0) of the neutral atom.

A partial justification for this is as follows. First, we may reasonably assume that the metal's electrical surface is inside the electron charge clouds of the substrate, so that the true value of a_n is somewhat greater than the radius of the external ion. This assumption finds support in the theoretical work of Lang and Kohn (1973) and in the experimental work of Culbertson *et al* (1979). From their papers we tentatively conclude that, for atoms of a size comparable with that of tungsten, the bonding distance a_n should be greater than the radius of the external atom or ion by roughly 50 pm.

Second, we note that with ions in a triply-charged state (or less), the ionic radii, where these are given in the literature (e.g. Moses 1978), are often less than the corresponding neutral-atom radius by approximately 40 to 60 pm. (Though whether these radii are really relevant to the situation of an ion at a highly charged surface is another matter).

Thus, in a loose first approximation, the smaller ionic radii counterbalance the field-penetration effect, and we can take the neutral-atom radius as an estimate of a_n, —certainly in the present context where we are merely trying to judge whether the simple formalisms are generally self-consistent.

Turning to figures 3(*a*) to 3(*c*), we find that in almost all cases the basic-image-hump (BIH) value of F_n^{ev} (from equation (12*a*)) is greater than the field at which the hump disappears (taking $a_n = \rho_0$). Further, a shift of all points by 20 pm to the left (which might be appropriate if we had over-estimated typical ion radii relevant to the present discussion) would still leave nearly all points in the close vicinity of the plotted curves.

Figures 3(*a*)–(*c*) thus give a clear indication that the basic-image-hump formalism based on equation (12*a*) is *unphysical*, certainly for most of the elements commonly employed in the low-temperature field-ion techniques.

Further, because the values of F_n^{ev} predicted from equation (12*b*) are expected to be

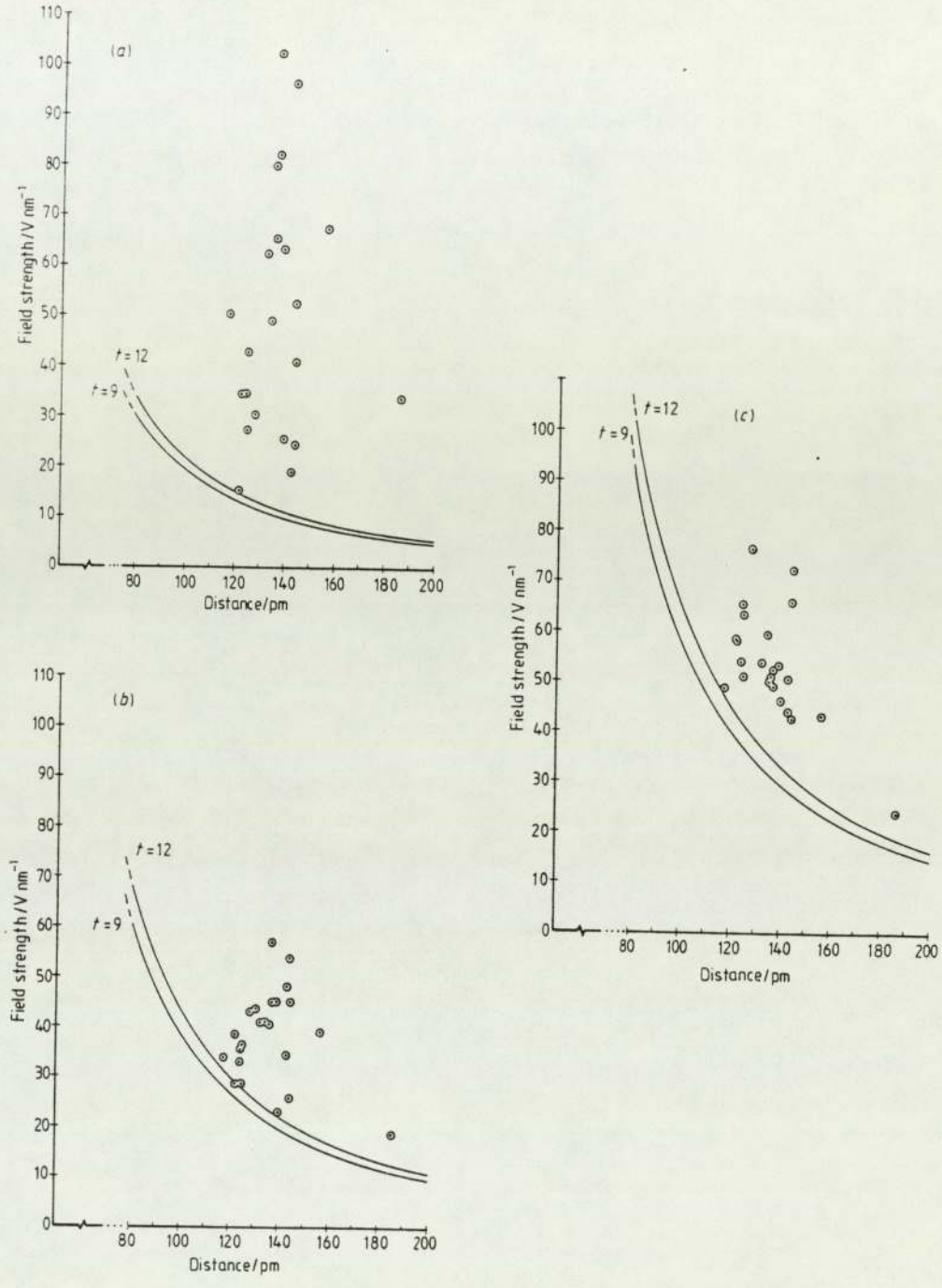


Figure 3. To illustrate the self-consistency or otherwise of the basic image-hump (BIH) formalism. The plotted curves are based on the extended formalism that results in equation (23), and show the field F_n^{HD} at which the Schottky hump disappears, as a function of the presumed zero-field bonding distance a_n, for t values 9 and 12. The plotted points are the BIH-predicted evaporation field values, as listed in table 2, and are plotted at a value of a_n equal to the neutral-atom radius for the species in question. Comparisons are shown for three different charge states: (*a*) $n = 1$; (*b*) $n = 2$; (*c*) $n = 3$.

Table 2. Predicted and observed evaporation fields for selected atomic species. The predicted fields are obtained from equation (12*a*); for Si and Ge the ionisation energies are taken from Moore (1958) and the binding energies and work-functions from Müller and Tsong (1969); otherwise the thermodynamic data in Tsong (1978a) have been used.

Species	Atomic radius (pm)	BIH-predicted evaporation fields			Observed evaporation field	
		F_1^{ev} (V nm^{-1})	F_2^{ev} (V nm^{-1})	F_3^{ev} (V nm^{-1})	Value (V nm^{-1})	Reference
W	137	102	57	52	57	Tsong (1978a)
Ta	143	96	48	44	43	Nakamura (1966)
Re	137	82	45	49	45	
Ir	135.5	80	44	50	46	
Hf	156.5	67	39	43	40	Tsong (1978b)
Mo	136	65	41	51	43	Nakamura (1966)
Pt	139	63	45	53	44	
Ru	132.5	62	41	54	42	Tsong (1978b)
Au	144	53	54	66	35	Tsong (1978a)
Si	117.5	50	34	49	10	Busch and Fischer (1963)
Rh	134.5	49	41	60	41	Nakamura (1966)
Co	125	43	37	63	35	
Fe	124	42	33	54	34	
Ti	144.5	41	26	43	25	
Ni	124.5	35	36	65	35	
Tsong (1978a)						
Ge	122.5	35	29	58	20	Ernst and Block (1975)
La	186.5	34	19	24	—	—
Cu	128	30	43	77	24	Nakamura (1966)
Cr	125	27	29	51	—	—
Sn	140	26	23	46	21	Wada *et al* (1980)
Ag	144.5	24	45	72	20	Nakamura (1966)
Al	143	19	35	50	27	Boyes *et al* (1971)
Ga	122	15	39	59	15	Wada *et al* (1980)

higher than those predicted from equation (12*a*), the addition of a 'polarisation correction' to the basic formalism will only make matters worse.

The general conclusion is that these 'simple' formalisms, that neglect the repulsive term, are in many cases unphysical and are always questionable when applied to the metals typically used in the low-temperature field-ion techniques†. It would thus usually be *wrong* to use them in interpreting field-evaporation data. This conclusion is in line with the reservations often expressed about the simple formalisms, and may be seen as a numerical verification of these reservations.

4. Comparison with observed evaporation fields

Invalidity of the simple image-hump formalisms does not necessarily exclude the pos-

† The only circumstances in which the use of simple image-hump arguments is legitimate is when the Schottky hump is well away from the surface (so the repulsion term is negligibly small). This is not the case under normal conditions of field evaporation at low temperature, though it may have been true of the experiments that Schottky was dealing with.

sibility that field evaporation is occurring by the Müller–Schottky mechanism. So it is instructive to compare observed evaporation fields with values of the field at which the Schottky hump disappears. If it can be shown that an observed evaporation field is greater than F_n^{HD}, then clearly—if our extended image-hump formalism is realistic—the Müller–Schottky mechanism cannot be operating for the species, charge state, and conditions in question.

(Note, however, that the reverse is not true. If the observed evaporation field is less than F_n^{HD} then a Schottky hump exists in the corresponding ionic potential curve, but the nature of the evaporation mechanism then depends on the relative configuration of the neutral and the various ionic curves.)

Results of this comparison are illustrated in figure 4. For simplicity we show the F_n^{HD} versus a_n curves only for the case $t = 12$: as in figures 3(*a*)–(*c*) the $t = 9$ curves lie to the left and lower than the $t = 12$ curves, so comparison with the $t = 12$ curve is the more demanding criterion. The observed evaporation fields are taken from a variety of sources, as indicated in table 2. In cases where different values are indicated in different sources, the lower value has usually been selected; and in cases where the variation of evaporation field with temperature is recorded the value corresponding to a temperature near 80 K has been taken. With tin and with gallium the values taken correspond to the evaporation of an overlayer on a tungsten substrate; with aluminium the value corresponds to evaporation near 55 K; with silicon the value given is the lowest recorded for a doped specimen. As before, there is the problem over the lack of knowledge of the

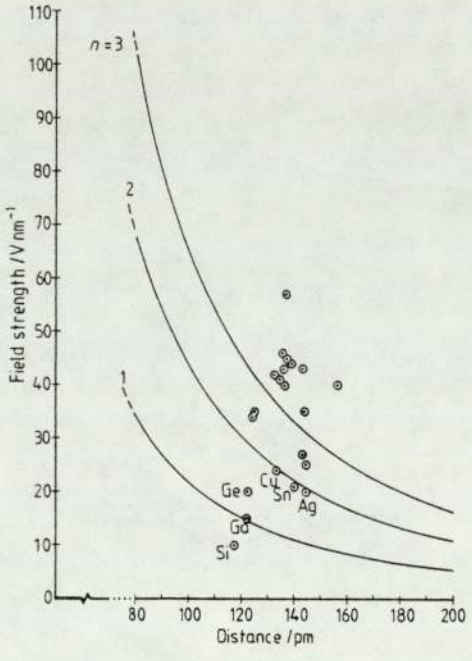


Figure 4. Comparison of experimental evaporation fields with the fields at which the Schottky hump disappears. As in figure 3, the plotted curves show F_n^{HD} as a function of a_n, in the case where the repulsive power-law exponent t equals 12, for the charge states $n = 1$, 2 and 3. The plotted points are the experimental evaporation-field values, as listed in table 2, and are plotted at a value of a_n equal to the neutral-atom radius for the species in question. Species lying below the $n = 2$ curve are individually labelled.

true values of a_n, and consequently the field values are (as before) plotted at the neutral-atom radius of the species in question. This procedure has reasonable validity, we feel, in the context of a discussion of the behaviour of metals in general, though discussion of individual cases might require more careful consideration of the a_n values to be chosen.

It also deserves comment that 'observed evaporation field' is not a uniquely defined quantity, as various criteria are used. (For example, the Nakamura (1966) field values correspond to an evaporation flux of 0.05 layers s^{-1}.) However, because all empirical evaporation-field criteria correspond to evaporation processes with finite activation energy, they must inevitably lead to field values less than that a (hypothetical) empirical zero-Q criterion would produce. There is a sense in which a true comparison of empirical evaporation fields with values of F_n^{HD} should involve an extrapolated (to $Q = 0$) empirical field rather than observed fields. But, since the observed fields are less than the extrapolated field would be, and since the question at issue in figure 4 is whether the empirical field is higher than the relevant F_n^{HD}, it follows that the test conducted in figure 4 is satisfactory. If the observed evaporation fields are higher than the relevant F_n^{HD}, then the extrapolated empirical evaporation field would certainly be. So any positive conclusion reached (on the basis of observed fields), that the Müller–Schottky mechanism is not operating in a particular case, is valid whatever empirical evaporation-field criterion has been used.

Inspection of figure 4 shows that for most materials the plotted points lie above the $n = 2$ curve. But the recent work on post-ionisation, cited earlier, indicates that the low-temperature escape charge state is 1^+ or 2^+, so for the above materials at their evaporation fields no Schottky hump would exist in the ionic potential curve. So—assuming that our extended image-hump formalism is valid, and our choice of ρ_0 as an estimate of a_n broadly reasonable—field evaporation for a large group of metals cannot be taking place by the Müller–Schottky mechanism.

Three materials (Cu, Sn and Ag) lie just below the $n = 2$ curve, but these (and probably Ge) are significantly above the $n = 1$ curve. All the three metals have been reported to field evaporate in a mixture of charge states (Barofsky and Müller 1968, Dixon *et al* 1981, Tsong 1978a) and, with the copper in particular, the proportion of Cu^{++} to Cu^+ ions has been reported to decrease with increasing temperature. This last result is characteristic of an evaporation process in which copper escapes as Cu^+ and then post-ionises to Cu^{++}. Thus figure 4 indicates that the Müller–Schottky mechanism is not operating for copper; this still seems to be the case when the rather low radius of the Cu^{++} ion is taken into account.

The thermodynamic data for silver, that are reflected in the high BIH-predicted value for F_n^{ev} in table 2, strongly indicate that any observed Ag^{++} would be formed by post-ionisation of Ag^+, and so silver too does not escape by the Müller–Schottky mechanism. Nor does Sn^+. However, it is difficult to reach a reliable conclusion about Sn^{++} because direct escape as a Sn^{++} ion is not incompatible with the thermodynamic data. We have no information about the evaporation charge state of germanium.

Thus, of the elements tested, it is only for silicon, gallium and germanium (and possibly Sn^{++}) that we have failed to eliminate (with reasonable sureness) the Müller–Schottky escape mechanism, and none of these is a material used in the mainstream of past discussions concerning evaporation escape mechanism.

In conclusion, we wish to draw special attention to the role that the presumption of the existence of post-ionisation fulfills in our arguments. Casual inspection of figure 4 shows that, if it had to be assumed that ions escaped in the higher charge states sometimes

observed, then the making of arguments against the Müller–Schottky mechanism would be more difficult. Thus it is important that the theoretical calculations concerning post-ionisation do in fact agree, both with detailed measurements of the field dependence of ion abundance, where these exist, and generally with other experimental evidence (see Kingham 1981, 1982). These experimental agreements enable us to have reasonable faith in the broad validity of the theoretical predictions concerning post-ionisation (Kingham 1981, 1982), in the cases where the experimental evidence does not yet exist.

5. Prediction of evaporation fields

Section 3 has shown that the simple image-hump formalisms are physically invalid. It remains to enquire why such formalisms are, nevertheless, able to predict evaporation fields with a fair degree of success.

The answer seems to lie in a result obtained recently by Forbes (1982a). Using a zero-activation-energy criterion for evaporation field, and taking just the electrostatic and image potential terms in equation (7), it may be shown from arguments based on energetics alone that a rough prediction of F_n^{ev} is given by:

$$F_n^{ev} = \sigma_n(16\pi\varepsilon_0/n^3e^3)(\Lambda^0 + H_n - n\varphi^E)^2 \tag{26}$$

where σ_n is a numerical quantity whose value is determined by the value chosen for the neutral-atom binding distance a_0. σ_n is—for the range of values of a_0 that is physically plausible—a relatively weak function of a_0, and has a value in the vicinity of 0.2 for plausible choices of a_0, for most of the metals conventionally employed in the low-temperature techniques, if n is 1 or 2. Equation (26) is capable of predicting observed evaporation fields with moderate success (which we think of as agreement within a factor of 1.5 or better). Predictions, if anything, are a bit on the low side.

Equation (26), however, becomes identical with the basic-image-hump result, equation (12*a*), if we take σ_n equal to 0.25. Since moderate (but slightly on the low side) success is obtained with $\sigma_n = 0.2$ in equation (26), it is numerically inevitable that moderate success is achieved using the basic-image-hump value $\sigma_n = 0.25$. This moderate success is a validation only of the underlying arguments concerning energetics, and of Müller's (1956) assumption that field evaporation is a thermally-activated process. Approximate prediction of evaporation field is not a test as between different thermally-activated evaporation mechanisms, and never implies that the Müller–Schottky mechanism (or any other specific mechanism) is operating.

6. Discussion and conclusions

The basic aim behind this work has been to discredit the Müller–Schottky mechanism as a likely mode of low-temperature field evaporation, so that its existence as a formal possibility does not throw doubt on numerical interpretations of field-evaporation data made assuming a Gomer-type mechanism. Significant progress has been made, but our conclusions should be approached with care.

With reasonable certainty we have shown:

(i) The simple image-hump formalisms (ignoring repulsive forces) are not logically consistent; hence it is dangerous to use them in data interpretation.

(ii) Ability of the simple formalisms to roughly predict observed evaporation fields does not imply that evaporation is occurring by the Müller–Schottky mechanism.

(iii) For most materials employed in the conventional field-ion techniques, observed evaporation fields are higher than the field at which the Schottky hump disappears.

Taken together, these results firmly suggest that discussion of low-temperature field evaporation in terms of the existing simple image-hump formalisms should be discontinued, particularly since existing charge-exchange formalisms seem to provide a moderately satisfactory account of field evaporation data.

It may perhaps be argued that use of the simple Müller–Schottky formula (12*a*) should continue, on grounds of convenience and simplicity. We would strongly oppose this, as likely to perpetuate confusion. It is also unnecessary, as the energetics-based formula (26) is just as simple to use.

The finding that, for most of the conventionally employed metals, the observed evaporation fields seem to be higher than the field F_1^{HD}, also has wider implications. In the context of general discussion as to the escape mechanism for such materials, this result tends to eliminate not only the Müller–Schottky mechanism but also any mechanism that postulates that the evaporating atom is initially bound in the form of a single-charged ion. (It is, of course, assumed that the expression used here for U_1 is a good representation of the binding potential.) This includes: charge-exchange-type mechanisms taking place from an ionic bonding state (as discussed in Forbes (1982b)); mechanisms that postulate ionic bonding followed by a pre-ionisation process as the entity moves away from the surface, followed by escape over a Schottky hump in doubly (or higher) charged state; and the pure Schottky mechanism in which the evaporating entity is initially bound in a singly-charged state and escapes over the Schottky hump without any change in charge state.

Although it seems unlikely that even the refractory metals could bond in a doubly-charged state, we note that for the refractory metals the observed evaporation fields are higher than the field F_2^{HD}. So, by arguments analogous to those above, mechanisms that postulate initial bonding in a doubly-charged state are eliminated.

A caveat must here be entered. Strictly, the arguments above are not sufficient to totally discredit the Müller–Schottky mechanism (or the mechanisms that postulate initial ionic bonding). This is for two reasons. First, our treatment of the ion-surface interaction is based on classical forms of interaction potential, and also assumes that the distance parameter in the repulsive term is to be measured from the metal's electrical surface. Second, we lack good knowledge of the position of the metal's electrical surface, relative to the position of the evaporating-atom nucleus. An exact quantum-mechanical treatment of the interaction between an external ion and a charged metal surface might conceivably lead to revised conclusions (though frankly we doubt it).

We also must point out that, although arguments of the type used in § 4 would apply to high-temperature field evaporation, where the activation energy is markedly greater than zero, if observed high-temperature evaporation fields were used, the data we have used mainly refer to conventional low-temperature (near 80 K) evaporation. Our conclusions here are thus restricted to the low-temperature situation.

The debate over the mechanism of escape, in this low-temperature situation, has now lasted for about twenty five years. The hesitations described above notwithstanding, we believe that this debate has for practical purposes, for the metals conventionally employed in the field-ion techniques, now been settled in favour of a Gomer-type surface charge-exchange process.

Acknowledgments

This work forms part of a research project funded by the UK Science and Engineering Research Council.

References

Aitken K L and Mair G L R 1980 *J. Phys. D: Appl. Phys.* **13** 2165–73
Barofsky D F and Müller E W 1968 *Surface Sci.* **10** 177–96
Boyes E D, Waugh A R, Turner P J, Mills P F and Southon M J 1977 Abstracts of *24th Intern. Field Emission Symposium, Oxford,* (unpublished)
Brandon D G 1963 *Br. J. Appl. Phys.* **14** 474–84
Busch G and Fischer T 1963 *Physik. Kondes. Mater.* **1** 367–79
Clampitt R 1981 *Nucl. Instrum. Meth.* **189** 111–6
Culbertson R J, Sakurai T and Robertson G H 1979 *Phys. Rev.* B **19** 4427–34
Dash J G 1975 *Films on Solid Surfaces* (New York: Academic)
Dixon A, Colliex C, Sudraud P and Van de Walle J 1981 *Surface Sci.* **108** L424–8
Ernst N 1979 *Surface Sci.* **87** 469–82
Ernst N 1979 *Surface Sci.* **87** 469–82
Ernst L and Block J H 1975 *Surface Sci.* **49** 293–309
Ernst N and Bozdech G 1980 *27th International Field Emission Symposium* (Tokyo: University of Tokyo) p 100–5
Forbes R G 1978 *Surface Sci.* **70** 239–54
—— 1982a *Appl. Phys. Lett.* **40** 277–9
—— 1982b *J. Phys. D: Appl. Phys.* **15** 1301–22
Forbes R G, Biswas R K and Chibane K 1982 *Surface Sci.* **114** 498–514
Gomer R and Swanson L W 1963 *J. Chem. Phys.* **38** 1613–29
Haydock R and Kingham D R 1980 *Phys. Rev. Lett.* **44** 1520–3
Kingham D R 1981 *PhD Thesis* University of Cambridge
—— 1982 *Surface Sci.* (to be published)
Lang N D and Kohn W 1973 *Phys. Rev.* B **7** 3541–50
McKinstry D 1972 *Surface Sci.* **29** 37–59
Moore C E 1958 *Circular* 467 (Washington: US National Bureau of Standards)
Moses A J *Practising Scientists Handbooks* (New York: Van Nostrand)
Müller E W 1956 *Phys. Rev.* **102** 618–24
—— 1960 *Adv. Electr. Electron Phys.* **13** 83–179
Muller E W and Krishnaswamy S V 1976 *Phys. Rev. Lett.* **37** 1011–3
Müller E W and Tsong T T 1969 *Field-ion Microscopy: Principles and Applications* (Amsterdam: Elsevier)
Müller E W, Krishnaswamy S V and Tsong T T 1974 *Prog. Surface Sci.* **4** 1–139
Nakamura S 1966 *J. Electron Microsc.* **15** 279–85
Tsong T T 1971 *J. Chem. Phys.* **54** 4205–16
—— 1978a *Surface Sci.* **70** 211–33
—— 1978b *J. Phys. F: Metal Phys.* **8** 1449–52
Wada M, Konishi M and Nishikawa O 1980 *Surface Sci.* **100** 439–52

J. Phys. D: Appl. Phys., 15 (1982) L75–77. Printed in Great Britain

LETTER TO THE EDITOR

An evaporation-field formula including the repulsive ion–surface interaction

Richard G Forbes
Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET, UK

Received 9 March 1982

Abstract. In field-ion emission an evaporation field may be defined by a zero-activation-energy (zero-Q) criterion. Existing formulae used to roughly predict evaporation field do not take into account the repulsive ion–surface interaction. A new formula, including this interaction, has been developed. Slightly higher evaporation fields are predicted. The increase should not exceed 25% in most cases, and may often be substantially less than this.

The prediction of the electric field necessary to initiate field evaporation is of some interest in practical field-ion techniques. Forbes (1982) has presented a treatment based on energetics alone, assuming: (i) the surface atom is bound in a primarily neutral† state; (ii) a zero-activation-energy definition of evaporation field is acceptable; (iii) the field-dependent correction to the neutral-atom binding-energy may be neglected; and (iv) only electrostatic and image-potential terms need be considered in the variable part of the ionic potential energy. This leads to the relationship:

$$neF_n^{ev}x_n^p = K_n - n^2e^2/16\pi\varepsilon_0 x_n^p \tag{1a}$$

$$K_n = \Lambda^0 + H_n - n\varphi^E \tag{1b}$$

where e is the elementary charge, ne the charge on the ion immediately after escape, F_n^{ev} the zero-Q evaporation field for the n-fold ion, x_n^p the distance from the emitter's electrical surface of the 'point of escape' for the n-fold ion, ε_0 the electric constant, Λ^0 the binding-energy of the neutral atom to the surface in the absence of the field, H_n the energy needed to form the n-fold ion from the neutral, in remote field-free space, and φ^E the relevant local work-function of the emitter. K_n is a species and charge-dependent constant.

By assuming that the image-potential term in equation (1a) is, in magnitude, some fraction α_n of the thermodynamic term K_n, i.e.

$$n^2e^2/16\pi\varepsilon_0 x_n^p = \alpha_n K_n \tag{2}$$

Forbes (1982) showed that F_n^{ev} was given by the equation:

$$F_n^{ev} = \sigma_n(16\pi\varepsilon_0/n^3e^3)K_n^2 \tag{3}$$

† I use the phrase 'primarily neutral' in the sense of Biswas and Forbes (1982); surface atoms are actually partially ionic.

0022-3727/82/070075 + 03 $02.00

where σ_n is given by:

$$\sigma_n = \alpha_n(1 - \alpha_n). \tag{4}$$

For the values of n realistically likely, namely 1 or possibly 2, σ_n is in the vicinity of 0.2 (i.e. within about ±25%), for most metals conventionally of interest.

Equation (3) replaces the conventional elementary Müller–Schottky equation (obtained by setting $\sigma_n = \frac{1}{4}$), derived from the basic image-hump treatment (Müller 1956, 1960) of the Müller–Schottky field-evaporation mechanism. The new equation is logically superior, because its derivation does not rely on the basic image-hump (BIH) formalism. A numerical demonstration that the BIH formalism is physically invalid has been given by Biswas and Forbes (1982).

There is, however, a logical anomaly in the present situation. Biswas and Forbes demonstrated the invalidity of BIH arguments by including a repulsive interaction term in the ionic potential energy. But the derivation of equation (3) ignored this repulsive interaction term. The present note aims to rectify the anomaly, and remove any doubts about the theoretical adequacy of the new formula.

It is sufficient to work in the context of a Gomer-type charge-exchange mechanism, in which it is assumed that the neutral and ionic potential-energy curves intersect at the neutral-atom bonding point, a distance a_0 from the emitter's electrical surface. In this case

$$x_n^p = a_0. \tag{5}$$

So, when the repulsive term is included in the ionic potential energy, equation (1*a*) becomes:

$$neF_n^{ev}a_0 = K_n - n^2e^2/16\pi\varepsilon_0 a_0 + G/a_0^t \tag{6}$$

where G is the constant in a tth power repulsive law for the n-fold ion.

Now suppose that at the distance a_0 the magnitude of the repulsive term is some fraction β_n that of the image-potential term, i.e.

$$G/a_0^t = \beta_n(n^2e^2/16\pi\varepsilon_0 a_0). \tag{7}$$

Combining the above relationships gives

$$neF_n^{ev}a_0 = K_n(1 - \alpha_n + \alpha_n\beta_n). \tag{8}$$

Using equations (2) and (5) to replace a_0 then leads to equation (3) again, but with a new value of the coefficient σ, given by

$$\sigma_n' = \sigma_n + \alpha_n^2\beta_n. \tag{9}$$

The form of equation (3) is thus satisfactory when the repulsive-interaction term is included; it remains to estimate the size of the correction term in equation (9).

From equation (16) in Biswas and Forbes (1982), derived by considering the zero-field situation, we have

$$G = (n^2e^2/16\pi\varepsilon_0)a_n^{t-1}/t \tag{10}$$

where a_n is the distance from the emitter's electrical surface of the bonding point for the n-fold ion, in the absence of any applied field. So, from equation (7)

$$\beta_n = (a_n/a_0)^{t-1}/t. \tag{11}$$

The parameter t is conventionally taken as 9 or 12, and for physical reasons the ion

radius a_n must be less than the corresponding neutral radius a_0. So β_n cannot significantly exceed 0.1, and may be substantially less than this. Furthermore, α_n is unlikely to be more than 0.7, if n is 1 or 2, and may be somewhat less than this. So the correction term in equation (9) cannot be much greater than 0.05, (and may be substantially less than this). It follows that, if σ_n is in the vicinity of 0.2, then σ_n' will be in the vicinity of 0.2 to 0.25. That is, the correction to the predicted zero-Q evaporation field F_n^{ev} resulting from inclusion of the repulsive term is of the order of 25% or less.

In physical terms, the argument is as follows. At the ionic bonding point the repulsive energy component must be equal in magnitude to $1/t$ times the image-potential energy; but the relative magnitude of this repulsive component falls off with distance from the surface. The neutral bonding point is outside the ionic bonding point, so this repulsive component must be relatively small there; so it leads to only a small correction.

The accuracy claimed for these simple evaporation-field equations is prediction within a factor of about 1.5. It follows that underestimation of σ in equation (3) by 25% or less would not be expected to cause serious error in the estimation of F_n^{ev}. In reality, the values of σ_n used in Forbes (1982) (namely $\sigma_1 = 0.2$, $\sigma_2 = 0.24$) tended to underpredict observed evaporation fields for the five test materials used, by between 3% and 20%. Neglect of the repulsive interaction energy may have been a part cause of this underestimation. (Neglect of field-dependent terms is another possible cause.)

It is also very understandable, if σ has to be slightly higher than the Forbes (1982) values, that the simple Müller–Schottky equation (in which σ is 0.25, for physically untenable reasons) should give fairly good agreement with observed evaporation fields.

It seems to the author that, at the classical conceptual level, we can now reasonably claim to understand fairly well the physical considerations involved in the approximate prediction of zero-Q evaporation field. It remains to be seen whether quantum-mechanical calculations, of the type advocated (for example) by Kingham (1981), will in effect enable a more accurate prediction of σ.

This work forms part of a research programme funded by the Science and Engineering Research Council.

References

Biswas R K and Forbes R G 1982 *J. Phys. D: Appl. Phys.* **15** 1323–38
Forbes R G 1982 *Appl. Phys. Lett.*, **40** 277–9
Kingham D R 1981 *PhD Thesis* University of Cambridge
Müller E W 1956 *Phys. Rev.* **102** 618–24
—— 1960 *Adv. Electr. Electron Phys.* **13** 83–179

J. Phys. D: Appl. Phys., **15** (1982) L99–104. Printed in Great Britain

LETTER TO THE EDITOR

Towards a criterion for the *a priori* prediction of field-evaporation mechanism

Richard G Forbes
Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET, UK

Received 11 May 1982

Abstract. The nature of the field-evaporation escape mechanism has been debated for over twenty years. Working within the framework of classical representations of the ion–surface interaction, a formula is presented that enables the prediction, from independent atomic and thermodynamic data, of cases where we can be reasonably certain that a Gomer-type surface charge-exchange process is involved. Most metals employed in the conventional low-temperature field-ion techniques fall into this category. However, the formula is indecisive as to the escape mechanism relevant to liquid-metal field-ion sources operating at room temperature and above.

The field evaporation of emitter surface atoms has an established role in the field-ion techniques, and is now discussed as a possible emission process in liquid-metal ion sources (for example, Aitken and Mair 1980, Venkatesan *et al* 1981, Prewett *et al* 1982). These sources have considerable technological potential (Swanson 1980, Clampitt 1981). Greater understanding of the field-evaporation process is important in basic scientific applications, and may in the longer term help to establish precisely how the liquid-metal sources emit.

It is now established (Ernst 1979, Ernst and Bozdech 1980, Haydock and Kingham 1980, Kingham 1981, Kellogg 1982) that evaporation of multiply-charged ions is often a two-stage process: thermally-activated escape into a 1^+ or possibly 2^+ state, followed by one or more post-ionisation events.

For the escape step, two competing mechanisms have been much discussed: the Müller–Schottky mechanism (Müller 1956, 1960), which involves ionisation before escape, and Gomer-type mechanisms (Gomer 1959, Gomer and Swanson 1963), which involve simultaneous ionisation and escape in a surface charge-exchange process. Knowledge of the actual escape mechanism is a prerequisite to detailed discussion or application of field evaporation, because the formulae for the dependence of activation energy and appearance energy on field are different for the different mechanisms (for example, McKinstry 1972).

Suggestions were often made (for example, Brandon 1964, Tsong 1971, McKinstry 1972) that at the high fields used in the conventional field-ion techniques no Schottky hump in the ionic potential would exist (and hence a Gomer-type mechanism must operate). Only recently have numerical estimates been made of the field at which the Schottky hump disappears (Biswas and Forbes 1982). (These estimates are made with a classical modelling of the repulsive ion–surface interaction.)

0022-3727/82/080099 + 06 $02.00

D8—C

If the evaporation field can be measured, and this measured evaporation field is greater than the Schottky-hump disappearance field, then—if emission is in fact a thermally-activated process—we can be reasonably confident that a Gomer-type mechanism is operating. Using literature values of experimental evaporation fields, Biswas and Forbes (1982) have demonstrated this to be the case for many metals conventionally employed in the low-temperature field-ion techniques.

Missing for the last twenty years, however, has been any numerical criterion for the *a priori* prediction (without experiments) of the evaporation mechanism for a specific material or system. Such a criterion would be of some interest in its own right, and could be particularly useful in circumstances where fields on emitters cannot be measured reliably. (For example, when the field calibration is uncertain, or when the shape and/or size of the tip of the emitter is unknown, as is sometimes the case with the liquid-metal ion sources.) This letter points out that, within the framework of classical modelling, a simple formula exists that helps towards *a priori* prediction.

Biswas and Forbes (1982) have shown that, relative to a zero of potential defined by the corresponding neutral entity when stationary in remote field-free space, the level U_n^* of the plateau in the ionic potential when the Schottky hump has 'just disappeared' is given by:

$$U_n^* = H_n - n\varphi^{\mathrm{E}} + \Theta n^2 e^2/16\pi\varepsilon_0 a_n - \tfrac{1}{2}c_n(F_n^{\mathrm{HD}})^2 \tag{1}$$

where: e is the elementary charge; ne the charge on the ion; H_n the work needed to form the ion from the corresponding neutral, in remote field-free space, being given by the sum of the first n free-space ionisation energies; φ^{E} the emitter work-function; a_n the distance of the ionic bonding point from the emitter's electrical surface, in zero applied field; c_n the polarisation-energy coefficient for the n-fold ion; F_n^{HD} the field when the Schottky hump disappears; and Θ a constant depending only on the exponent in the repulsive power law, with value -1.546 for 12th-power repulsion.

Relative to the same zero-level, the corresponding atom when bound has potential energy U_α^{B} given by:

$$U_\alpha^{\mathrm{B}} = -\Lambda^0 - \Delta\Lambda^* \equiv -\Lambda^0 - \tfrac{1}{2}c_\alpha(F_n^{\mathrm{HD}})^2. \tag{2}$$

Λ^0 is the atomic binding energy in zero field, $\Delta\Lambda^*$ the field-induced binding-energy increase for field F_n^{HD}. The suffix α is being used to label quantities appropriate to the atom in its as-bound electronic state, that has to be slightly ionic in order to sustain the external electric field; but, for simplicity, this suffix is omitted from the binding-energy symbols. c_α is the polarisation-energy coefficient for this as-bound state, and is defined by equation (2).

It follows that the energy difference W_n^* between the level of the ionic plateau and the bonding potential energy, at field F_n^{HD}, is given by:

$$W_n^* = (\Lambda^0 + H_n - n\varphi^{\mathrm{E}}) + \Theta n^2 e^2/16\pi\varepsilon_0 a_n + \tfrac{1}{2}(c_\alpha - c_n)(F_n^{\mathrm{HD}})^2. \tag{3}$$

Clearly, W_n^* depends on the basic characteristics of the material or system in question. Figure 1 shows three different cases: (A) $W_n^* \gg 0$; (B) $W_n^* \gtrsim 0$; (C) $W_n^* \ll 0$.

At fields slightly less than F_n^{HD} the ionic potential has a Schottky hump in it, as illustrated in figure 2. The difference in energy between the level of the top of the hump and the bonding potential energy is in this case slightly greater than W_n^*. (It decreases towards W_n^* as the field increases towards F_n^{HD}.) At fields above F_n^{HD} there is no hump in the ionic potential, and the potential-curve configuration (see figure 2) corresponds to charge-exchange.

The Müller-Schottky mechanism presumes that escape occurs by thermal activation, from the initial bonding-state level, over a Schottky hump. It follows that, if the Müller-Schottky mechanism is occurring, the activation energy must (from the considerations above) be greater than the energy difference W_n^*.

But, in practice, with a thermally-activated emission process, the activation energy

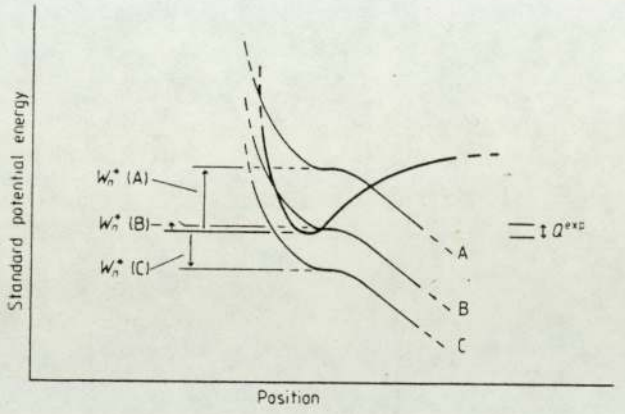


Figure 1. To illustrate three possible configurations of the ionic potential curve, relative to the atomic potential curve, in the situation where the Schottky hump in the latter has just disappeared: curve (A) $W_n^* \geqslant 0$; curve (B) $W_n^* \geqslant 0$; curve (C) $W_n^* \ll 0$. These curves correspond to three different materials, and are intended as schematic rather than accurate potential representations. The energy interval marked Q^{exp} represents an activation energy, as determined from the emission equation, of a size to illustrate the argument in the text.

has to be in accordance with the emission equation:

$$J = n_{hr} A \exp(-Q/kT) \tag{4}$$

where: J is the evaporation flux; n_{hr} the amount of material (count of atoms) at high risk

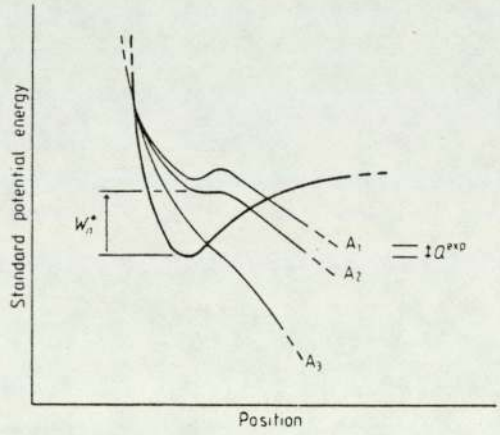


Figure 2. To show three different positions, for a given material, of the ionic potential curve relative to the atomic curve. Curves A_1, A_2, A_3 correspond to increasing values of the external field. Curve A_2 corresponds to the field F_n^{tID} at which the Schottky hump has just disappeared; and curve A_3 corresponds to the situation where field evaporation is taking place with an activation energy Q^{exp}.

of field evaporation; A the field-evaporation pre-exponential; and kT the Boltzmann factor. For conventional low-temperature field evaporation, with flux values of about 0.1 layers/s, this leads to an empirical activation energy (Q^{exp}) requirement of about 0.2 eV.

So, if field evaporation is occurring at a significant rate, but $Q^{exp} \ll W_n^*$, it follows that the Müller-Schottky mechanism cannot be operating. In these circumstances, given values of W_n^* and Q^{exp} such as are illustrated in figure 2, the potential-curve configuration cannot be as in figure 1 curve (A) or figure 2 curve (A_2): rather, the external field must be somewhat higher, so that the configuration of figure 2(A_3) obtains (where the activation energy is suitably small). But this configuration is that corresponding to a charge-exchange process. So we see that the condition $W_n^* \gg Q^{exp}$ implies that a Gomer-type mechanism is operating.

With cases (B) and (C) in figure 1, when $W_n^* \sim Q^{exp}$ or when $W_n^* < 0$, it is difficult to be definitive about mechanism, because this depends on the precise relative configuration of atomic and ionic potential curves. Various mechanisms are possible. But fine details of the shapes of the curves are not sufficiently well known to enable reliable predictions. Thus, for the present, mechanism prediction is limited to identifying those cases for which we can be reasonably certain that a Gomer-type mechanism operates.

The polarisation-energy coefficients c_α and c_n that appear in equation (3) are not well known, though following Brandon's (1964) calculations it has been widely accepted that $c_\alpha \gg c_n$, and hence that the F^2 term in equation (3) is positive in value. With this in mind, we may define an approximate quantity $W_n^*(\mathrm{NP})$ (NP = no polarisation terms) by:

$$W_n^*(\mathrm{NP}) = (\Lambda^0 + H_n - n\varphi^E) + \Theta n^2 e^2/16\pi\varepsilon_0 a_n. \tag{5}$$

If $W_n^*(\mathrm{NP})$ is greater than zero or greater than some value of Q^{exp}, then W_n^* will certainly be greater than the quantity in question.

In evaluating the image-potential term in equation (5), values have to be selected for the ionic zero-field bonding distance a_n. This causes difficulty because we do not clearly know either the location of the electrical surface of the metal, or the value of the relevant ion radius. Biswas and Forbes (1982) argue that there is a counterbalancing effect between two things: on the one hand, the electrical surface is inside the electron charge clouds of the metal substrate atoms (as shown clearly by Lang and Kohn (1973)); on the other, ionic radii are often significantly smaller than the corresponding neutral-atom radius. Thus in this Letter, as elsewhere, a_n is estimated by the neutral-atom radius for the species in question, as listed in table 1.

Values of $W_n^*(\mathrm{NP})$ for various relevant elements, evaluated using the data tabulated in Tsong (1978), are listed in table 1, in order of decreasing $W_1^*(\mathrm{NP})$. Only the $n = 1$ and $n = 2$ cases are listed, because it seems fairly certain that observed higher-charge-state ions are formed by post-ionisation after escape. (In some cases 2^+ ions would be formed from 1^+ in this way.)

Table 1 contains many species with both $W_1^*(\mathrm{NP})$ and $W_2^*(\mathrm{NP})$ greater than 1 eV. At temperatures near 80 K this class of element will certainly escape via a Gomer-type process. The class contains all the metals conventionally used in the low-temperature field-ion techniques. This finding is in line with the results concerning the disappearance of the Schottky hump, cited earlier, and is in line with the conclusion that flux field-sensitivity data for some such metals are most easily explained in terms of a charge-exchange-type mechanism (Forbes *et al* 1982). It thus seems clear beyond reasonable doubt that these conventionally-employed metals escape via a Gomer-type mechanism.

Table 1. Values of the quantities $W_1^*(\mathrm{NP})$ and $W_2^*(\mathrm{NP})$, as defined by equation (5), for selected elements of relevance to field-ion emission. ρ_0 denotes the neutral-atom radius of the species in question, and is rounded to the nearest 0.5 pm. Species for which either of the tabulated quantities is less than 1 eV are marked with an asterisk.

Species	ρ_0 (pm)	$W_1^*(\mathrm{NP})$ (eV)	$W_2^*(\mathrm{NP})$ (eV)
W	137	8.1	9.4
Ta	143	7.9	8.0
Re	137	6.8	6.5
Ir	135.5	6.6	6.0
Hf	156.5	6.3	7.0
Mo	136	5.6	5.3
Pt	139	5.5	6.8
Ru	132.5	5.3	4.9
Au	144	4.8	9.4
Rh	134.5	4.3	5.2
La	186.5	4.0	2.8
Ti	144.5	3.8	1.8
Si*	117.5	3.8	0.8
Co	125	3.4	2.7
Fe	124	3.3	1.6
Ni	124.5	2.6	2.4
Ge*	122.5	2.5	0.0
Cu	128	2.3	4.9
Pb	175	2.2	3.6
Sn*	140	2.1	0.4
Ag	144.5	2.1	7.4
Cr*	125	1.8	0.4
Al	143	1.3	4.4
Cs*	266	0.3	3.7
Ga*	122	0.1	2.8

This has long been hypothesised, and is now well supported numerically. The way ahead now seems clear for closer analysis of Gomer-type mechanisms (see Forbes 1981, Kingham 1981), and for use of formulae based on these mechanisms to derive basic surface atomic information from field-evaporation data.

Elements in table 1 for which either $W_1^*(\mathrm{NP})$ or $W_2^*(\mathrm{NP})$ is less than 1 eV are marked with an asterisk. For these materials the situation is less clear. Broadly, the higher the emission temperature (and hence Q^{exp} for a given emission flux), and the lower the tabulated value of either $W_1^*(\mathrm{NP})$ or $W_2^*(\mathrm{NP})$, the less likely it is that a Gomer-type mechanism operates. But, as indicated earlier, uncertainties in the situation make reliable prediction impossible.

It deserves note that both caesium and gallium fall into this 'uncertain' category. Both these metals are used in liquid-metal ion sources operating near room temperature, and for caesium it has been reported (Aitken and Mair 1980) that the current–voltage characteristics observed with a sharp needle emitter (Aitken and Mair 1980, Clampitt 1981, S Venkatesh and R Clampitt, private communication) have a slope near onset that can be explained in terms of a simple image–hump formalism. (However, other explanations, in terms of liquid electrohydrodynamics and/or space-charge effects, are more plausible for blunt-needle and capillary liquid-metal sources, and may also be possible in the sharp-needle case (G L R Mair, private communication).)

With the high-temperature liquid-metal sources such as tin (operating near 600 K) or copper (operating near 1350 K), where either $W_1^*(\mathrm{NP})$ or $W_2^*(\mathrm{NP})$ is less than or comparable with plausible values of activation energy Q^{exp}, there is again uncertainty over prediction of the field-evaporation escape mechanism.

In general terms, therefore, the simple theoretical criterion presented here is helpful in connection with the conventional low-temperature techniques, but the liquid-metal ion sources require a more sophisticated level of theory. Further investigation, particularly of the high-temperature case, is indicated.

Finally, I would point out that, because the image potential tends to overestimate the true attractive interaction between an external ion and a metal surface (see, for example, Ying 1980), a criterion based on a more realistic interaction potential might favour the Gomer-type mechanisms more than the present formula does.

Acknowledgments

This work forms part of a research programme funded by the United Kingdom Science and Engineering Research Council.

References

Aitken K L and Mair G L R 1980 *J. Phys. D: Appl. Phys.* **13** 2165–73
Biswas R K and Forbes R G 1982 *J. Phys. D: Appl. Phys.* **15** 1323–38
Brandon D G 1964 *Surface Sci.* **3** 1–18
Clampitt R 1981 *Nucl. Instrum. Methods* **189** 111–6
Ernst N 1979 *Surface Sci.* **87** 469–82
Ernst N and Bozdech G 1980 *27th Int. Field Emission Symp.* pp 100–5 (Tokyo: University of Tokyo)
Forbes R G 1981 *Surface Sci.* **102** 255–63
Forbes R G, Biswas R K and Chibane K 1982 *Surface Sci.* **114** 498–514
Gomer R 1959 *J. Chem. Phys.* **31** 341–5
Gomer R and Swanson L W 1963 *J. Chem. Phys.* **38** 1613–29
Haydock R and Kingham D R 1980 *Phys. Rev. Lett.* **44** 1520–3
Kingham D R 1981 *PhD Thesis* University of Cambridge
Lang N D and Kohn W 1973 *Phys. Rev.* B **7** 3541–50
McKinstry D 1972 *Surface Sci.* **29** 37–59
Müller E W 1956 *Phys. Rev.* **102** 618–24
—— 1960 *Adv. Electr. Electron Phys.* **13** 83–179
Prewett P D, Mair G L R and Thompson S P 1982 *J. Phys. D: Appl. Phys.* **15** 1339–48
Swanson L W 1981 in *Proc. Conf. Microcircuit Engineering '80* (Delft: Delft UP)
Tsong T T 1971 *J. Chem. Phys.* **54** 4205–16
—— 1978 *Surface Sci.* **70** 311–33
Venkatesan T, Wagner A and Barr D 1981 *Appl. Phys. Lett.* **38** 943–5
Ying S C 1980, in *Theory of Chemisorption* (Ed. J R Smith) (Berlin: Springer)

Surface Science
North-Holland Publishing Company

THE TEMPERATURE DEPENDENCE OF EVAPORATION FIELD FOR GOMER-TYPE FIELD-EVAPORATION MECHANISMS

K. CHIBANE and Richard G. FORBES

Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET, UK

Received 27 May 1982; accepted for publication 26 August 1982

Theoretical formulae are developed for the temperature dependence of the onset evaporation field F°, assuming a parabolic surface-atom bonding well and a Gomer-type mechanism for the escape process. Recent experimental results for tungsten and molybdenum, when replotted in the form $T^{1/2}$ versus $1/F^{\circ}$, exhibit linear behaviour in the range from about 60 to 150 K. Deviations occur below this (at 50 K for W, 35 K for Mo); the deviation temperatures are compatible with theoretical estimates of the critical temperature at which ion tunnelling effects become important. Values of the zero-Q evaporation field extrapolated from the linear regime (74 V/nm for W, 60 V/nm for Mo) are significantly higher than observed values of F° near 78 K or values of F^{e} derived from the Müller–Schottky formula: the difference can be explained by taking into account repulsive ion–surface and F^{2} potential-energy terms. Our theoretical picture seems generally self-consistent at the classical level, and it is concluded that the low-temperature field-evaporation process for the more field-desorption-resistant metals is, at this level, now basically understood. Field evaporation may now be used to investigate atomic behaviour at charged surfaces, but greater care is needed over the standardization of evaporation conditions and the consistent choice of field calibration.

1. Introduction

Low-temperature field evaporation has long been of significance in the conventional field-ion techniques. High-temperature field evaporation is now of possible relevance to liquid-metal field-ion sources [1]. Thus the temperature dependence of the "evaporation field" may be of some current interest.

It is well known that evaporation field, as defined by some empirical criterion (e.g. a given evaporation-flux level), gets lower as temperature increases. In his original paper [2], Müller reported measurements on the temperature dependence of evaporation field, and attempted to explain these in terms of a basic image-hump formalism. Equivalent experiments, but over a limited range in temperature, were subsequently carried out by various workers [3–7], for a few of the more refractory metals.

On the theoretical side, Brandon [4] introduced a method of inter-relating the partial derivatives of field, temperature and "evaporation rate" with

0039-6028/82/0000–0000/$02.75

respect to each other, and this was applied to the "constant intersection distance" version of Gomer's charge-exchange model [8] by Tsong and Müller [6,9]. In all this early work there is an ambiguity about the distinction between "evaporation flux" and "evaporation rate-constant", and a modified version of Brandon's method, designed to overcome this ambiguity, was put forward by Forbes [10].

With hindsight, other flaws can be seen in the early theoretical discussions. First, it is now clear [11] that the basic image-hump formalism is inadequate because it ignores the repulsive ion–surface interaction term. At observed low-temperature evaporation-field strengths there is no Schottky hump in the ionic potential, as long suspected [12,13]. It is also clear [11,14] that the field evaporation of the more refractory metals almost certainly has as its first stage escape by a Gomer-type surface charge-exchange process. (This may then be followed by one or more post-ionization steps [15,16].)

It also seems advisable [17,18], at least in principle, to distinguish between different types of surface charge-exchange process. The two extreme cases are: (a) *charge-hopping*, when the motion of the departing ion core is relatively fast and electron transfer, if it takes place, occurs in a sharp "hopping" transition; and (b) *charge-draining*, when the motion of the ion core is relatively slow and electron transfer is best described as a slow draining of charge out of the departing entity. The existence of these different possibilities was recognized by Gomer and Swanson [19], but has been largely neglected (and somewhat confused – see ref. [17]) in most subsequent literature. This *class* of process, in which escape and charge transfer are occurring simultaneously, we refer to as *Gomer-type mechanisms*.

It is also now clear [20–23] that the "constant intersection distance" treatment of the charge-hopping mechanism is unsatisfactory, because the variation of the "intersection distance" with field is important in determining the relationship between activation energy and field. This invalidates some of the formulae and numerical conclusions gives in refs. [6,24], particularly those based on derived values of "$x_c + \lambda$" and "α", and leaves us without any satisfactory theoretical discussion of the dependence of evaporation field on temperature.

More recently, on the experimental side, Wada et al. [25] have undertaken systematic studies of the variation of evaporation field with temperature, for tungsten, molybdenum, and gallium/tungsten and tin/tungsten systems. Unfortunately, some of their subsequent analysis [25,26] is based on the "constant intersection distance" approach, and this throws doubt on the validity of some of their numerical conclusions.

The aim of this paper is thus to develop and investigate a new theoretical treatment of the dependence of "evaporation field" on temperature, derived from the new formulae for the dependence of activation energy on field recently put forward [23], for a Gomer-type escape process.

The structure of the paper is as follows. Section 2 sets out some basic theory, and its expected limitations. Section 3 reanalyses the data of Wada et al. [25] in terms of the new theory, and deals briefly with the question of field calibration (which is taken up in more detail in appendix 2). Section 4 discusses some implications of the new analysis. And section 5 then reviews our understanding of the escape process for the more field-desorption-resistant metals, in the light of this paper's achievements. For reference, a summary of the various extant definitions of "evaporation field" is included as appendix 1.

A preliminary account of the main findings of this work was presented elsewhere [27].

2. Theory

2.1. Basic formulae and definitions

The amount of material (or "count of atoms"), J, field evaporated per unit time from the surface of a field-ion emitter is given by the *emission equation:*

$$J = n_{\mathrm{hr}} k_{\mathrm{hr}} = n_{\mathrm{hr}} A \exp(-Q/k_{\mathrm{B}}T), \tag{1}$$

where n_{hr} is the amount of material (or "count of atoms") at high risk of field evaporation, k_{hr} is the *field-evaporation rate-constant* for the high-risk (or "active") sites, A is the corresponding rate-constant pre-exponential, Q is the field-evaporation activation energy at the typical high-risk site, k_{B} is the Boltzmann constant, and T is thermodynamic temperature. As in ref. [10], we regard n_{hr} as having the dimension [amount-of-substance] and J as having the dimensions [amount-of-substance $\times$ time^{-1}], with the "layer" considered as a non-SI unit of amount-of-substance. We call J the *evaporation flux* [22].

For the Gomer-type escape mechanisms, simple approximate expressions for activation energy Q in terms of the *external field* F [28] have recently been derived [23]. These have the form:

$$Q = \Omega[(F^{\mathrm{e}}/F) - 1]^2, \tag{2}$$

where F^{e} is the so-called *zero-Q evaporation field*, i.e. the value of F for which the activation energy Q becomes zero [29]. Ω is a parameter with the dimensions of energy, defined in the context of eq. (2); estimates of Ω can, in principle, be obtained either from experiment or from theory. Ω is in principle a function of field, though (as we shall see later) it can in practice be regarded as a constant.

In the above equations the quantities n_{hr}, A and Q are all, in principle, functions of the temperature T and of the external field F at the typical high-risk site. It follows that $J = J(F, T)$, and that J, F and T are mathematically a set of "thermodynamic" variables [30], in the sense that specifying the

values of any two of them will specify the value of the third. It thus follows that we can invert the relationship and write [31]:

$$F = F(T, J). \tag{3}$$

It follows that, if we define the "onset" of field evaporation by a chosen value of evaporation flux (which we write J° and call the *onset flux*), then for a given temperature T there is a specific value of F corresponding to this choice of T and J°. We call this F-value the *onset evaporation field* and denote it by F°. (Mathematically, the onset labels "o" are not strictly necessary, but they may help to make discussion clearer.)

This onset evaporation field F° should be distinguished from F^{e} and from the so-called "critical evaporation field" F^{c} employed in an earlier discussion [32]. For reference, a summary of the different ways an "evaporation field" can be defined is given in appendix 1.

For the Gomer-type escape mechanisms, a formula for the temperature dependence of onset evaporation field can be obtained by eliminating Q between eqs. (1) and (2). This leads to a relationship that can be written in the reciprocal forms:

$$T^{1/2} = \theta^{1/2}\left[(F^{e}/F^{\circ}) - 1\right], \tag{4a}$$

$$1/F^{\circ} = \left[1 + (T/\theta)^{1/2}\right](1/F^{e}), \tag{4b}$$

where θ is a temperature-like parameter related to Ω by:

$$\theta = \Omega/\{k_{B}\ln(n_{hr}A/J^{\circ})\}; \tag{5}$$

θ can be interpreted as the temperature at which the onset evaporation field becomes equal to half the zero-Q evaporation field – provided form (4) remains adequate at high temperature (which may not be so). Alternatively, eq. (4) can be put in the simplified form:

$$1/F^{\circ} = b + CT^{1/2}, \tag{6}$$

where b and C are parameters related in an obvious way to F^{e} and θ.

2.2. *Specific approximations for* Ω

According to ref. [23], the simplest approximation for Ω is:

$$\Omega = \tfrac{1}{2}\kappa a^{2} \qquad \text{(first approx.)}, \tag{7}$$

where: κ is the force-constant for the evaporating entity in its initial bonding state, where it is hypothesized to be vibrating in a parabolic potential well; and a is the distance of the well bottom from the emitter's electrical surface [33].

In the context of this approximation it is possible to derive from eq. (5) the following relationship between the θ-values corresponding to two different

onset fluxes of J_1° and J_2°:

$$1/\theta_1 = 1/\theta_2 - \left(k_B/\tfrac{1}{2}\kappa a^2\right) \ln\left(J_2^\circ/J_1^\circ\right). \tag{8}$$

Qualitatively, the higher the choice of J°, the higher will be the value of θ, and the higher the value of F° at a given temperature.

A better approximation for Ω is [23]:

$$\Omega = \tfrac{1}{2}\kappa a^2 \left(1 - \pi^e/nea\right)^2 \left(1 - \eta_n'/neF^\circ\right)^{-2}, \tag{9}$$

where e is the elementary (proton) charge, ne the charge on the evaporating ion immediately after escape, η_n is the purely-chemical component [34] of ion–surface interaction energy, η_n' is the partial derivative of η_n with respect to distance, evaluated at the well minimum and for $F = F^e$, and π^e is a parameter with the dimensions of SI dipole moment, related to the field derivatives of η_n and of the total initial-state binding-energy Λ by:

$$\pi^e = (\mathrm{d}\Lambda/\mathrm{d}F)|_{F^e} + (\partial\eta_n/\partial F)|_{a,F^e}. \tag{10}$$

Physically, the first bracketed term in eq. (9) is a correction due to field-induced effects (polarization and partial ionization in the bonding state), and the second comes from correlation and repulsion corrections to the ionic potential energy.

2.3. Limitations

Since Q, F and T are measurable quantities, Ω and θ can in principle be regarded as empirical quantities defined by eqs. (2) and (4) respectively. But, in practice, our theory will be predictively useful only if (for a given value J° of evaporation flux) the quantities Ω and θ (and hence b and C in eq. (6)) are effectively constant. From the theoretical point of view, various conditions have to be satisfied for this to be so.

First, the requirements lying behind the derivation of eq. (2) must be satisfied. These are (see ref. [23]): (a) the field-evaporating atom can be treated as an independent oscillator, vibrating in a parabolic potential well; (b) field evaporation must be a classical thermally-activated process, taking place via a Gomer-type escape mechanism; (c) mathematically, a "curve-intersection" definition of Q must be adequate, and Q must be "not too large". (Even for the parabolic-well approximation, neither eq. (7) nor eq. (9) is strictly correct. And in reality the nucleus is vibrating, not in a parabolic well, but in a well that is better approximated by a Lennard-Jones potential or a Morse potential: the parabolic well is a good approximation to these only at relatively low values of activation energy. For either reason, an increase in evaporation temperature (and hence Q) could lead to apparent field dependence in the parameter Ω in eq. (2).)

Further requirements arise when eq. (5) and the quantities A and n_{hr} are

considered. Classical discussions of thermally-activated field evaporation take A as a constant. But, in reality, ion tunnelling through the activation-energy barrier can occur [19], and at low temperatures this makes the emitted ion flux J greater than would be predicted from the emission equation (1) using the classical value of Q and the high-temperature value of A. Thus, to preserve the form of eqs. (1) and (4) as temperature decreases, one must gradually increase the value of A as it appears in eqs. (1) and (5). This is equivalent to a gradual decrease in the value of θ (increase in the value of C) as the emission temperature decreases. (But at very low temperatures θ as defined by eq. (5) tends to become proportional to T (C proportional to $T^{-1/2}$), and forms (4) and (6) are no longer useful.)

It has conventionally been assumed that these ion-tunnelling effects are not significant at temperatures near 80 K and above. Thus from measurements above this temperature we expect that a "normal" value of θ can be obtained. A "theoretical" plot of F° against T, using this θ-value in eq. (4), would tend to the field F^e as T tends to zero. Ion tunnelling causes the actual plot to fall below this "classical" plot, increasingly so as temperature falls, with the actual plot tending to a limiting value $F^\circ(0, J^\circ)$ as T tends to zero. Behaviour of this type was in fact observed by Wada et al. [25], in the cases of tungsten and molybdenum. It follows from the above that eqs. (4) and (6) are useful only when T is "not too low".

The behaviour of n_{hr} as a function of temperature has never been systematically investigated, but we feel that n_{hr} will decrease somewhat as low temperatures are approached. This could also lead to a requirement that T be "not too low"; (but in practice any effects resulting from change in n_{hr} would probably be masked by uncertainties in the a-priori calculation of ion tunnelling rates).

A final requirement, derived from eq. (9), is that the quantity η_n'/neF° be small and/or slowly varying over the field range of interest. There is some prior expectation [23] that η_n' will be small, because the correlation and the repulsion contributions to this derivative are of opposite sign and thus tend to cancel. If $\eta_n'/neF^\circ \ll 1$, then the binomial theorem enables the second bracket in eq. (9) to be put in the approximate form $(1 + \eta_n'/neF^\circ)^2$; and elimination of Q and Ω between eqs. (1), (2) and (9) leads to the result:

$$T^{1/2} = \frac{(\kappa/2)^{1/2}(a - \pi^e/ne)(1 + \eta_n'/neF^\circ)(F^e/F^\circ - 1)}{k_B \ln(n_{hr}A/J^\circ)}. \tag{11}$$

A plot of $T^{1/2}$ against $1/F^\circ$ will be linear only if (over the field and temperature range in question) the quantities n_{hr}, A and η_n'/neF° are all effectively constant. A plot of this type thus provides an empirical test of the usefulness of the theory derived earlier.

In general terms, given the limitations set out above, the best that we can hope for is linearity of the $T^{1/2}$ versus $1/F^\circ$ plot over a limited range of

temperature that includes 80 K. At temperatures somewhat below this we expect deviation due to increasing relative importance of ion tunnelling; at temperatures somewhat or well above this we expect deviations due to the inadequacy of eq. (2) or breakdown of the assumption that η'/neF° is small and/or effectively constant.

2.4. Status of the present theory

The present theory is essentially a classical theory, assuming a parabolic bonding well and using a curve-intersection formalism. From a wider viewpoint it has to be regarded as an elementary version of an eventual treatment based on quantum-theoretical calculations of the dependence of activation energy on field. But the current treatment has the advantage that it is analytic, and leads to a reltively simple formula, eq. (4) or (6), capable of empirical test.

But, even if such a formula fits the experimental results, the question of how metal evaporation actually occurs still remains. Our current view, also shared by Kingham [18,35], is that the actual process in metal field evaporation near 80 K must be charge draining. But the *accurate* calculation of an explicit Q-formula for the charge-draining mechanism is a very difficult quantum-mechanical problem, possibly somewhat beyond the frontiers of present theoretical methods. We thus take the view that the curve-intersection formalism used here is a first-generation approximation for describing either the charge-hopping mechanism *or* the charge-draining mechanism.

3. The results of Wada, Konishi and Nishikawa

3.1. Tungsten and molybdenum

We now look in more detail at the results in ref. [25]. From an enlarged version of fig. 1, the data points have been measured up, and are shown in fig. 2 replotted in the form $T^{1/2}$ versus $1/F^{\circ}$. For both tungsten and molybdenum the relationship is basically linear.

There is a clear falling away from linearity at low temperatures (below about 60 K), as would be expected if ion tunnelling (or equivalent) effects become important there. But there is no significant evidence of any deviation from linearity at high temperatures, as would be expected if the η'_n term in eq. (11) were large and rapidly varying.

These experimental results are thus compatible with a Gomer-type escape mechanism, as expressed in the elementary formulae (4) and (5), with θ a constant for evaporation temperatures in the range from about 60 K to at least 150 K. The upper temperature limit to the validity of the simple formulae is beyond the range used in the Wada et al. experiments.

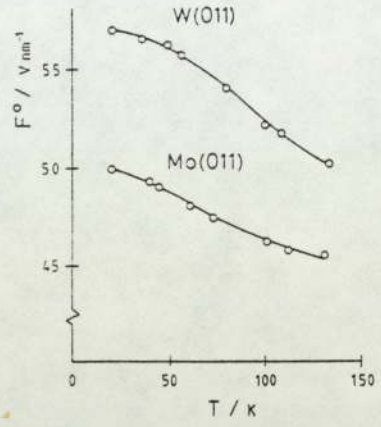


Fig. 1. Plot of evaporation field F° versus temperature T, taken from the work of Wada et al. [25]. The field F° corresponds to the onset flux condition $J^\circ = 0.1$ layers/s. (The field values have not been re-standardized.)

Regression analysis on fig. 2, excluding the lowest-temperature point in the case of molybdenum, and the two lowest points in the case of tungsten, then leads to the values of F^e and θ described in table 1 as the "regressed" values. Using the values for tungsten, substitution of $T = 78$ K into eq. (4b) leads to the value $F^\circ = 53.6$ V/nm. This value agrees with fig. 1, but is somewhat lower than the usually accepted 78 K value of 57 V/nm. The question of standardization thus arises.

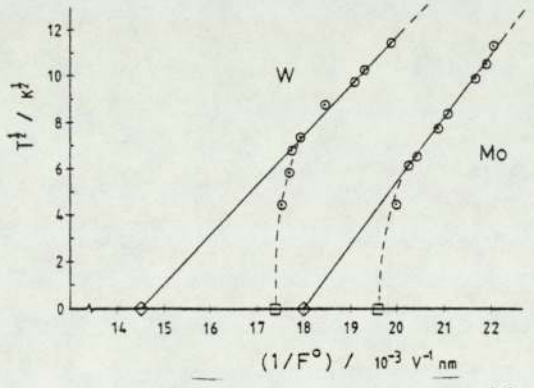


Fig. 2. Data from fig. 1 replotted in the form $T^{1/2}$ versus $1/F^\circ$. The continuous lines are regression fits to the higher-temperature data points; the dashed lines are fitted by eye to the lower-temperature points. The diamonds denote the as-regressed values of zero-Q evaporation field; the squares denote the values of experimental zero-T evaporation field, as estimated from fig. 1. (Field values are not re-standardized.)

Table 1
Values of the parameters θ and F^e, derived from regression on the data in fig. 2 (in the form $1/F^o$ versus $T^{1/2}$), and after re-standardization of the field values (see appendix 2) (calibration errors in the regressed F^e values have not been investigated); for molybdenum the re-standardized value is obtained by assuming that the same correction factor applies as for tungsten; this is not strictly correct, but the error involved should be small

Species	θ (K)	F^e (V/nm)		
		As regressed	Re-standardized	(+ error type)
W	940 ± 120	69.1 ± 1.0	74.2 ± 1.1	(statistical)
			± 10	(calibration)
Mo	2430 ± 190	55.5 ± 0.3	60 ± 12	(est. total)

Virtually all measurements of field strength are based on measurements of voltage. The voltage $V^{\ominus}$ corresponding to a reproducible "standard condition", for which the corresponding "standard field strength" $F^{\ominus}$ is presumed known, is determined; the field strength F^o applicable when a different voltage V^o is applied to the same endform is then given by:

$$F^o = \left(V^o/V^{\ominus}\right) F^{\ominus}. \tag{12}$$

Virtually all experimental field values in field-ion literature thus relate back to one of the two experiments that measure "absolutely" a field strength: the comparison of voltages for field-electron and field-ion emission, by Müller and Young [36], or the method based on the energy deficits associated with the free-space ionization of hydrogen, due to Sakurai and Müller [37].

Neither method is entirely free from uncertainties. So, following Müller and Tsong [24], we shall somewhat arbitrarily take the "standard condition" as field evaporation at a rate of 0.01 layers/s, at a temperature assumed to be 78.0 K precisely, and (for tips of radius greater than about 30 nm) will take the standard field strength as 57.0 V/nm. This choice optimizes consistency with existing theoretical discussions, and is compatible with the experiments. There is, of course, an associated "calibration uncertainty"; this issue is considered in some detail in appendix 2.

Wada et al. used a standard condition different from that above, in that an onset flux of 0.1 layers/s was employed; they also made a different choice of standard field strength, based on the older calibration procedure [36]. We have thus "re-standardized" the regression results, by a procedure described in appendix 2. These re-standardized values of F^e and θ are shown in table 1. The question of the errors on a value of F^o predicted using these data is somewhat complicated, and this issue also is considered in appendix 2.

The re-standardized values of F^e shown in table 1 are somewhat higher than the values of "evaporation field" normally discussed in the literature. We take up this point in section 4.2.

3.2. The Ga / W and Sn / W systems

Wada et al. have also measured the temperature dependence of evaporation field for structures of gallium, and of tin, adsorbed on a tungsten substrate, in several different binding states.

For the Ga/W system, the F° versus T curves have shapes characteristically different from those for tungsten and molybdenum [25]. And when redrawn in the form of a $T^{1/2}$ versus $1/F^{\circ}$ plot, the sets of data points do not conform to the pattern of a basically linear variation with deviation at low temperatures. Some selected results, replotted from fig. 9 in ref. [25], are shown in fig. 3.

This failure of gallium to conform to a formula based on the charge-hopping escape mechanism is not entirely surprising. Gallium is one of the less field-desorption-resistant metals [38], and recent theoretical analyses have shown that it is not a member of the group of metals that can be confidently predicted to escape via a Gomer-type mechanism [11,14]. (The system gallium on tungsten is of course slightly different from the pure-gallium situation envisaged in the analyses; the point is that gallium has already been recognized as a potentially anomalous species.)

Tin is to some extent a borderline material in the analyses [11,14]. The F° versus T and the $T^{1/2}$ versus $1/F^{\circ}$ plots for tin are also broadly intermediate in characteristics between tungsten and molybdenum, on the one hand, and the gallium/tungsten system on the other. But the number of data points on the curves for the Sn/W system is really too few for us to reach any positive conclusions.

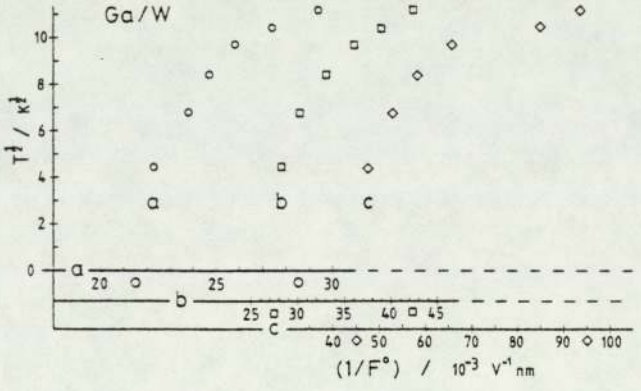


Fig. 3. Plot of $T^{1/2}$ versus $1/F^{\circ}$ for three different bonding situations of gallium on tungsten: (a) Ga on W(111) pseudomorph; (b) Ga on W(011) superstructure; (c) Ga on W(011) overlayer. The data are replotted from ref. [25] without re-standardization of field values.

4. Discussion

For independently established reasons [11,14], tungsten and molybdenum are expected to escape by a Gomer-type mechanism. The demonstrated conformity of the temperature dependence of F° with the simple formula (4) seems a reliable confirmation, both of the general mechanism of escape, and of the hypothesis that the bonding well can be treated as parabolic (certainly for fields corresponding to temperatures in the range up to 150 K).

We thus believe that tungsten and molybdenum field evaporation (in the range 60–150 K) is basically understood. Given this, the experimental results contain some other implications that deserve exploration.

4.1. Low-temperature effects

Fig. 2 provides perhaps the clearest evidence so far that at low temperatures the classical Arrhenius equation fails. This has long been postulated theoretically [19,39]. The evidence is that the experimental plot falls away from the extrapolated straight line that passes through the higher-temperature points. The "deviation temperature" at which departure from linearity occurs is fairly well defined, so (for the first time) we can get good experimental estimates of it; these are listed in table 2.

Table 2
To show experimental estimates of the "deviation temperature" T^{dev} below which the $T^{1/2}$ versus $1/F^{\circ}$ plot departs from linearity, and of the "critical temperature" T^{c} at which the rate-constants for escape by thermal activation over the classical barrier, and escape by tunnelling from the vibrational ground state, are equal

Tungsten		Molybdenum		Details of calculation	
T^{dev} (K)	T^{c} (K)	T^{dev} (K)	T^{c} (K)	Ref.	Barrier-type or method
50 ± 10		35 ± 10		(Fig. 2)	Experimental value
	705			[4]	Field emission analogy
	70		130	[9]	Triangular barrier
	41			[39]	Schottky barrier
	73		91	[40]	Triangular barrier
	53		69	[40]	Rectangular barrier
	66		82	[40]	Parabolic barrier
	40 [a]		45 [b]	[40]	Parabolic approximation to Schottky barrier

[a] The calculation assumes escape as 2^{+} ion, at a field of 56 V/nm.
[b] Obtained by scaling Kingham's tungsten value, using the formula in ref. [40] and taking the molybdenum evaporation field to be 43 V/nm.

The value of 50 K for tungsten is compatible with informal evidence from practical field-ion microscopy that field evaporation near 20 K and below is "different" from evaporation near 80 K. (For example, an endform evaporated near 5 K is often more-sharply defined than the corresponding endform prepared near 80 K.) But it is unclear whether such evidence is specifically associated with a change from an "activation-dominated" to a "tunnelling-dominated" regime.

Various workers [4,9,39–41] have attempted to calculate the "critical temperature" T^c at which the rate-constant for escape by tunnelling from the ground vibrational state is equal to the rate-constant as given by the Arrhenius equation. Estimated values of T^c are listed in table 2, along with information about the potential model employed in the calculations. It is clear that for tungsten most of these calculations produce estimates broadly comparable with the experimental deviation temperature; so in this case the attribution of the deviation to the manifestation, at low temperatures, of ion-tunnelling effects seems a very plausible hypothesis.

On the other hand, the calculations for molybdenum in general agree less well with the experimental deviation temperature. However, as Kingham emphasizes [40,41], the predicted critical temperature is significantly dependent on model and assumptions employed. And there is also the question of whether it is legitimate to exactly identify the critical temperature with the deviation temperature. So further work on this topic seems advisable.

More generally, it is important for overall theoretical consistency that the deviation temperature has been found experimentally to be somewhat less than 80 K, for both tungsten and molybdenum. This is another confirmation that, at temperatures near 80 K, it has been legitimate in many past theoretical discussions to treat field evaporation as a thermally-activated process obeying the Arrhenius equation.

4.2. The value of zero-Q evaporation field

The "re-standardized" values of F^e given in table 1 are substantially higher than the values normally quoted for the 78 K evaporation field of tungsten (57 V/nm) and molybdenum (about 43 V/nm). More to the point, they are substantially higher than the values predicted (for $n = 2$) using either the (now discredited [11]) Müller–Schottky formula which gives $F^e(\mathrm{W}) = 57$ V/nm, or the new energetics-based formula proposed by Forbes [42,43] which gives $F^e(\mathrm{W}) = 55$ V/nm: the discrepancy requires explanation because both these formulae purport to predict the zero-Q evaporation field rather than a 78 K value of F^o.

The Forbes formula, like the Müller–Schottky formula, is based on an approximate representation of the ionic potential in which only the configurational, electrostatic, and image-potential terms are considered. In the corre-

sponding activation-energy expression it is assumed that the image-potential term is some definite fraction of the "thermodynamic" term K_n defined by:

$$K_n = \Lambda^0 + H_n - n\phi^E. \tag{13}$$

As shown in ref. [42], this assumption leads to the formula:

$$F_n^e = \sigma_n \left(16\pi\epsilon_0/n^3e^3\right) K_n^2, \tag{14}$$

where, for clarity, the suffix "n" is added to the symbol for zero-Q evaporation field, to emphasize that the theoretically-predicted field value depends on the postulated ion charge immediately after escape. σ_n is treated as a constant, and ref. [42] recommends the values: $\sigma_1 = 0.2$; $\sigma_2 = 0.24$.

The explanation of the discrepancy, between the restandardized F^e values listed in table 1 and the corresponding values estimated from eq. (14), may lie – at least in part – in the omission of repulsive and field-dependent terms from the elementary formulae. Discussion is easier if we work with the full formula appropriate to the charge-hopping escape mechanism, but use the common elementary representations [34] of the physical interactions involved. In this case the requirement that activation energy be zero leads to the condition:

$$neF_n^e a = K_n - n^2e^2/16\pi\epsilon_0 a + G/a^t + \tfrac{1}{2}\left(c_0 - c_n\right) F_n^{e2}, \tag{15}$$

where G is the constant in a tth power repulsive law, and c_0 and c_n are the polarization-energy coefficients for the "neutral" atom and n-fold ion respectively.

For tungsten $K_2 = 25.6$ eV. The result in ref. [42] would be obtained by disregarding the last two terms in eq. (15) and taking a in eq. (15) to be 140 pm (which gives the image-potential energy as -10.3 eV). Inclusion of the repulsion and polarization terms will clearly increase the estimate of F_n^e. Thus if we estimate the repulsive term as about one tenth the image-potential term in magnitude, say 1 eV, and the polarization term as about 2 eV, we achieve a revised estimate of F_2^e as 65.5 V/nm. This is equivalent to a value for σ_2 of about 0.29, and is much closer to (but still not as high as) the re-standardized value 74 V/nm, which is equivalent to $\sigma_2 = 0.33$.

Obviously, a higher estimate of F_2^e could be obtained if the polarization term were greater than the 2 eV assumed. But another likely possibility is that the image-potential term overestimates the strength of the correlation interaction between the departing ion and the surface; modern surface theories repeatedly show this (ref. [44], for example). Yet another possibility is a general systematic error in the field-strength values used in field-ion literature, owing to undetected error in the "standard field strength" determinations [36,37]. Thus there is no difficulty about giving a plausible theoretical explanation for the high F^e values obtained by extrapolating the Wada et al. experimental results.

The real significance of the above discussion is as follows. The long-known

good agreement, between observed onset fields near 80 K and the predictions of the Müller–Schottky formula, now seems partly spurious. This agreement, more than any other feature of field-evaporation theory, has in the past justified the hypothesis that near-80 K field evaporation is a thermally-determined process obeying the Arrhenius equation. So weakening of this former agreement weakens the validation of the Arrhenius-equation hypothesis. This section has restored the position to some extent.

4.3. Temperatures above 150 K: the Nakamura and Kuroda results

For tungsten and molybdenum, the Wada et al. results justify formula (4), with θ a constant, from 60 K up to about 150 K. But, as already remarked, there must be an upper temperature limit to the validity of this formula. Measurements over a wider range in temperature (from near 20 K to near 300 K) have been carried out by Nakamura and Kuroda [5], for tungsten, molybdenum and tantalum.

Nakamura and Kuroda suggest that, for temperatures in the range from about 180 K to 300 K: "field-evaporation voltage decreased with the nearly square of the emitter temperature". This implies a relationship between onset evaporation field and temperature of the form:

$$F^{\circ} = \gamma_1(1 - \gamma_2 T)^2, \tag{16}$$

where γ_1 and γ_2 are positive constants. For temperatures below about 180 K

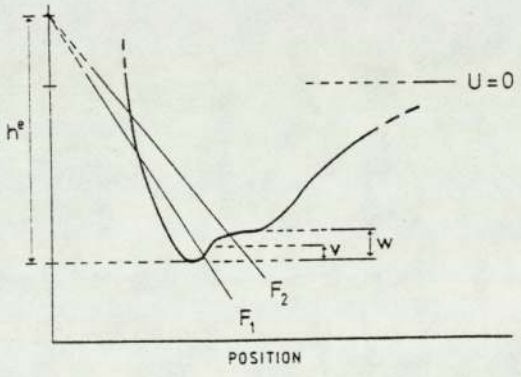


Fig. 4. To illustrate how the form of the dependence of Q upon F could be influenced by local structure in the bonding potential. This diagram is analogous to fig. 1a in ref. [23], and h^e is the "standard pivot height" as defined in ref. [23]. For fields F_1 and F_2 the form of the variation of activation energy with field will be characteristically different. Changes in form might be expected for activation energies equal to v and to w. In the present context the speculation is that v might be about 0.4 eV and w about 1 eV. (This diagram is schematic, and the bold curve is not intended as an accurate representation of bonding-potential structure.)

the relationship diverges from this form [5].

With the help of the binomial theorem, given that $(T/\theta)^{1/2} < 1$, formula (4) can be put in a form analogous to that above, giving:

$$F^{\circ} \sim F^{e}(1 - \theta^{-1/2}T^{1/2}). \tag{17}$$

Clearly the above forms are not equivalent. So the Nakamura and Kuroda results seem to indicate a change in regime at around 180 K.

A change in regime as temperature increases is not implausible in principle. For example, increasing temperature (and hence decreasing F°) might produce a regime change if the curve-intersection point approached an inflection point in the bonding potential curve and moved into a region where the bonding potential is convex upwards. This possibility is illustrated in fig. 4. Since for steady field evaporation the activation energy Q has to be about $25kT$, a regime change near 180 K might imply an inflection point about 0.4 eV above the bottom of the bonding well. This value is not compatible with a simple bonding well of depth 9 eV or so (the zero-field bonding energy for tungsten is 8.66 eV); but it could be compatible with a *localized* bonding well of depth about 1 eV, as shown in fig. 4. And the existence of localized bonding wells of about this depth is entirely compatible with listed diffusion activation energies for tungsten, which are of order 1 eV at a field-free surface [39].

Thus an observed regime change near 180 K might conceivably indicate the existence of short-range structure in the bonding potential; so it is conceivable that measurements of the temperature dependence of onset evaporation field could provide a simple probe of local bonding-well shape, where statistical similarities exist as between different types of kink-site.

An alternative explanation of the regime change in the Nakamura and Kuroda results has, however, been suggested by Kellogg (private communication). He suggests that the field-adsorbed layer of imaging gas, that would have been present at temperatures near 80 K in the Nakamura and Kuroda experiments, would become vacant at higher temperatures: the field-adsorbed helium would affect the tungsten-atom binding energy slightly, and a regime change in the field-evaporation behaviour could be associated with a change from a high-coverage to a low-coverage situation. Such a coverage change could conceivably occur in the vicinity of 180 K.

The interpretation of the Nakamura and Kuroda results is thus uncertain, and further experiments seem to be called for.

4.4. Flaws in the original analyses

Our analysis here of the Wada et al. results differs significantly from those in their original papers [25,26]. Their analyses, which reflect the thinking of the late Professor Müller's research school, are believed to be seriously flawed. In particular:

(1) Eq. (1) in ref. [25] is incomplete in the sense of being dimensionally inconsistent. Their k has the units "layers/s", but their k_0 has the units "s^{-1}". A quantity corresponding to our n_{hr}, tht has the units "layers", is missing. (We also urge, as a matter of notational clarity, that quantities with different dimensions be given different primary symbols, and that the symbol k be allowed to retain its traditional chemical meaning of "rate-constant".)

(2) The simple image-hump formalism, as expressed in eq. (2) of ref. [25], is now known to be mathematically invalid in the circumstances of conventional low-temperature field evaporation [11]. Consequently, it is physically wrong to apply it to the interpretation of field-evaporation data. Either this, or the proven non-applicability of the Müller–Schottky mechanism to the low-temperature field evaporation of tungsten and molybdenum [11,14], may be responsible for the discrepancies reported by Wada et al. when they try to apply a simple image-hump formalism to their data (see p. 451 in ref. [25]).

(3) The whole analysis in sections 2 and 3.1 of ref. [26] is based on the proposition that the "intersection distance" (their x_c) is effectively a constant, independent of field. But this assumption is mathematically erroneous [20–23]: activation energy is a function of field largely because x_c is a function of field. It follows that all the numerical results in ref. [26] are void of physical meaning, notwithstanding the fact that some plausible values of their parameter α can be achieved.

(4) The parameter α, that appears in their analyses and in particular in table 1 of ref. [26], and that corresponds to our $(c_0 - c_n)$, is not a polarizability in the normal textbook sense, because α is not defined in terms of the self-consistent local field acting on the atom in question. Even if the calculations in ref. [26] were legitimate, it would not be legitimate to directly compare their α with the genuine Gaussian polarizabilities listed by Teachout and Pack [45].

There also seems some uncertainty over the width of the zone for the post-ionization of Ga^+ to Ga^{2+}, as determined theoretically. Kingham [18] and Konishi et al. [46] are in disagreement over the interpretation of fig. 5 in ref. [26]; and Kingham's repeat of the post-ionization calculations for Ga^+ in any case seems to indicate a much narrower zone-width than do the Konishi et al. calculations.

Finally, we would stress that experiments of the type performed by Wada et al. *cannot* provide information about the field dependence of surface atom binding energy. (Such information is, however, available from the experiments of Ernst [47,48].)

Notwithstanding these criticisms of their analyses, we wish to emphasize that the Wada et al. experimental results are playing a key role in this discussion.

5. Review of conclusions

In this final section we aim to summarize our present understanding of the escape process for the more field-desorption-resistant metals.

The form of the relationship between evaporation field and temperature depends on the form of the explicit Q-formula, that relates activation energy to field. And the form of this explicit Q-formula depends, in part, on the postulated mechanism of field evaporation, and on the mathematical formalism used to describe it.

This paper arose from a perception that exploring the experimental F° versus T relationship might help consolidate recent progress in field-evaporation theory, achieved through the demonstrations by one of us that the low-temperature field evaporation of the more field-desorption-resistant metals must take place via a Gomer-type escape mechanism, andthat within the framework of such mechanisms simple explicit formulae can be given for the field-dependence of Q.

Such theoretical consolidation seems a useful preliminary to all the applications of field-evaporation theory, be they to the basic physics of charged surfaces, to the emission process for liquid-metal field-ion sources, or to the more conventional field-ion techniques.

The consolidation achieved seems most useful. We have shown that, for temperatures in the range 60 to 150 K, the experimental F° versus T relationship for tungsten and molybdenum conforms well with the theoretical charge-hopping formulae derived here. This tends to confim that field evaporation is thermally activated in this temperature range (and presumably above it), that escape is taking place via a Gomer-type surface charge-exchange process, that the field-evaporating atom can be treated as an independent oscillator, that the bonding well can be treated as parabolic near the well bottom, and that a curve-intersection formalism is adequate to describe the escape process irrespective of whether escape actually occurs via a charge-hopping or a charge-draining mechanism.

Mathematically, this consolidation means that, in the temperature range from about 60 to about 150 K (the exact limits varying from material to material) we expect the escape process for the more field-desorption-resistant metals to be governed by the following set of relatively simple equations:

$$J = n_{\mathrm{hr}} A \exp(-Q/kT), \tag{1}$$

$$Q = \Omega(F^{\mathrm{e}}/F^{\circ} - 1)^{2}, \tag{2}$$

$$1/F^{\circ} = \left[1 + (T/\theta)^{1/2}\right](1/F^{\mathrm{e}}), \tag{4}$$

$$\theta = \Omega/\left\{k_{\mathrm{B}} \ln(n_{\mathrm{hr}} A/J^{\circ})\right\}. \tag{5}$$

The quantities F^{e}, θ and Ω can be regarded as empirical constants in their own

right, associated with a Gomer-type escape process from a parabolic bonding well. In terms of more fundamental atomic and thermodynamic parameters, F^e and Ω can (in the first approximation) be estimated by:

$$F^e = \sigma_n \left(16\pi\epsilon_0/n^3e^3\right) K_n^2, \tag{14}$$

$$\Omega = \tfrac{1}{2}\kappa a^2. \tag{7}$$

And a rather better approximation for Ω can be obtained from eq. (9), with F° set equal to some average or effective value.

It is clearly necessary to think of onset evaporation field as a function of the two variables temperature and onset evaporation flux, and it may be useful to think of this functional relationship as a "thermodynamic" one. Clear specification of field evaporation conditions, in terms of T and J°, is advisable; and questions of standardization of the field-calibration procedure need proper attention.

We have also drawn out, more accurately than in earlier work, experimental estimates of the failure temperature below which the classical Arrhenius equation breaks down. This failure point (near 50 K for tungsten) is roughly where the experimental field-ion microscopists thought it would be, is roughly where theorists hoped it would be (to preserve general consistency of past discussions of evaporation mechanism), and is compatible with the best theoretical derivations of the critical temperature at which ion-tunnelling effects become dominant over activation effects.

A related feature is that empirically-derived values of the zero-Q evaporation field F^e are somewhat higher than the onset evaporation fields observed at low temperatures. This implies that σ_n in eq. (14) must be higher than previously thought, or (less likely) that there is systematic error in field calibration. Various plausible reasons exist for thinking σ_n may have been underestimated in earlier work.

All the above conclusions apply in the cases of tungsten and molybdenum, for which there is prior expectation of field evaporation via a Gomer-type escape mechanism. For the cases of the Sn/W and Ga/W systems, for which there is no such prior expectation, the experimental F° versus T relationships do not provide any clear indication of evaporation mechanism.

Given the above consolidation, we now have confidence that a charge-hopping formalism adequately describes the low-temperature field evaporation of the more field-desorption resistant metals.

Thus we now change viewpoint and think of field evaporation in the form of F° versus T (or related) plots, as a possible technique for exploring the shape of the surface-atom bonding well, for kink-site atoms at a highly-charged surface. Obviously, because of the statistical nature of the field evaporation process, it is only the statistically common (as between different types of bonding site) features of the potential structure that may be explorable in this way.

So we regard the Nakamura and Kuroda results as giving an indication that, at 0.4 eV above the well bottom, the bonding-well shape may be departing from a parabola. Which may conceivably indicate the existence of local structure in the bonding potential.

In general, it seems clear that careful measurement of the relationship between evaporation field and temperature may provide valuable scientific information. Such experiments need not be technologically sophisticated, and we urge that they be carried out fora wide range of temperatures and materials. It would be particularly useful to have more work on tungsten, over as wide a range in temperature as practicable.

Finally, we wish to stress again the change of viewpoint inherent in this paper. When we began, we were discussing experimental data concerning field-evaporation mechanism. As we end, these same experiments are perceived to be using field evaporation to explore the bonding-potential structure.

Acknowledgements

This work forms a part of a research project funded by the United Kingdom Science and Engineering Research Council. In addition, one of us (K.C.) wishes to thank the Ministry of Higher Education and Scientific Research of the Republic of Algeria for personal financial support. We thank Dr. P.D. Prewett and Dr. D.R. Kingham for useful discussions, and thank Dr. D.R. Weaver for assistance and advice about regression procedures.

Note added in proof

Our attention has been drawn to the measurements of Kellogg [54] on the temperature dependence of evaporation voltage, made for tungsten, molybdenum and rhodium using a pulsed laser atom probe. His results show a regime change in the vicinity of 250 K, and indicate that the effects of adsorbed gases need to be taken into account; in the range 50 K to about 200 K his results are compatible with our theory, in that replotting in the form $T^{1/2}$ versus $1/F^{\circ}$ gives reasonably linear plots. His results also bring out the effect on this relationship of varying the chosen flux J°: this is broadly as we expect theoretically.

Appendix 1. The definition of evaporation field

If the external field F at the high-risk sites is sufficiently high then field evaporation occurs from such sites at an appreciable rate. Some value of F for

which the rate is deemed "significant" is called the "evaporation field". There are three basic ways in which this idea of a significant rate can be made quantitative: we may specify a set value of activation energy, a set value of evaporation rate-constant, or a set value of evaporation flux.

In simple theoretical discussions it is convenient to define a "*zero-Q evaporation field*" F^e by the requirement that $Q(F^e) = 0$. This field F^e is a species-related constant, independent of temperature, and its value can be estimated from arguments based on energetics alone [42,43,49]. Empirical evaporation fields (as discussed below) observed in conventional low-temperature field evaporation are necessarily less than F^e. Detachment of surface atoms can in principle occur at fields higher than F^e, but it is then a process limited mainly by the constraints on electron motion, and is not thermally activated.

In single-atom experiments where the evaporation rate-constant is measured directly, such as the pulsing experiments carried out by Tsong [13], it is particularly convenient to define a *critical evaporation field* F^c by the requirement that at this field the rate-constant be equal to 1 s^{-1}. The related theory has been discussed elsewhere [21,32].

However, most experiments involve observation of the evaporation flux (or of the corresponding electric emission current), and it is then convenient to specify an *onset evaporation field* F^o by the requirement that the flux have some set value J^o. In low-temperature field-ion microscope experiments values for J^o of 0.01 layers/s (or slightly greater) are often chosen (e.g. refs. [2,3,6]), these values representing the situation where slow steady change in the field-ion image can just be detected by visual observation. Atom-probe experiments (e.g. ref. [29]) sometimes use a higher J^o value, because of the duty cycle involved in pulse generation, a suitable aim being to detect about one ion per pulse. With the liquid-metal ion sources, where macroscopic electrical measurements are employed, onset is usually defined in terms of a set emission current, for example the 1 nA criterion used by Aitken and Mair [51]. There are difficulties in converting this into a precise flux criterion, because the sources tend to emit in a variety of charge states (and also emit neutrals, clusters and microparticles), but the flux level is much higher than in the low-temperature field-ion microscope context. The latter is of order 10 atoms/s, whereas the 1 nA criterion is equivalent (for singly-charged ions) to a flux of about 6×10^9 atoms/s.

For a given choice of J^o, as temperature is reduced towards zero, the onset evaporation field tends to a fixed value $F^o(0, J^o)$ that is a function of J^o. This *zero-T evaporation field* is always less than the zero-Q evaporation field F^e (see section 2.3).

Appendix 2. Restandardization and conversion procedures

This appendix describes how the currently available experimental data may be used to obtain estimates and error limits for the onset evaporation field $F(T, J)$ corresponding to a given temperature T and flux criterion J. (For notational simplicity we have dropped the onset symbol "°", here.)

From eq. (4b), inverted:

$$F(78, J) = F^e / \left[1 + (78\ \mathrm{K}/\theta_J)^{1/2} \right]. \tag{18}$$

where we have added the suffix "J" to θ, to emphasize that this parameter is a function of the onset flux level chosen. Eliminating F^e between eqs. (4b) and (18) gives:

$$F(T, J) = \alpha(T, J)\, F(78, J) = F(78, J) \left[\frac{\theta_J^{1/2} + (78\ \mathrm{K})^{1/2}}{\theta_J^{1/2} + T^{1/2}} \right], \tag{19}$$

where $\alpha(T, J)$ is a dimensionless parameter defined by this equation and equal to the bracketed term.

Since we need to work with the Wada et al. results, which have a non-standard field calibration (from our point of view), this form is better than eq. (4) or (6), because it avoids employing their field calibration. Because (in the simplest approximation, used here) θ_J is not a function of field, the values of $\theta_{0.1}$ derived from our regressions on the Wada et al. results are not affected by any change in field calibration. These values are listed in table 1, earlier.

A further problem arises because (from our point of view) Wada et al. are using a non-standard choice of flux criterion (0.1 layers/s rather than 0.01 layers/s). Thus, for tungsten, we are presuming that $F(78, 0.01)$ is 57.00 V/nm exactly, but we require to know $F(78, 0.1)$. The conversion can be made [52] using flux field-sensitivity measurements such as those of Tsong [53], and leads to the result: $F(78, 0.1) = 57.59$ V/nm.

We shall continue to use tungsten as an illustration of procedures. From the "regressed" results in table 1 earlier, we have $\theta_{0.1} = 940$ K, so on substituting $T = 0$ K into eq. (19) we obtain $\alpha(0, 0.1) = 1.288$; combining this with the value of $F(78, 0.1)$ just derived gives:

$$F^e = F(0, 0.1) = 74.18\ \mathrm{V/nm}. \tag{20}$$

This is the "re-standardized" value of zero-Q evaporation field F^e, i.e. the value based on the field calibration we have adopted as standard. This value should, in principle, be independent of the flux criterion chosen in the experiments.

If it is wanted to obtain evaporation-field values corresponding to other onset-flux criteria, then it is necessary to first derive the appropriate value of θ_J via eq. (8), which we here write in the slightly more general form (assuming Ω

is a constant):

$$1/\theta_{J2} = 1/\theta_{J1} - (k_B/\Omega)\ln(J_2/J_1). \tag{21}$$

Use of this formula requires an estimate of Ω. Accurate values are not currently available. As an approximation we use the formula, developed from eq. (2) [52]:

$$\Omega = -\tfrac{1}{2}\mu_2 - 2\mu_1 + Q, \tag{22}$$

where μ_2 and μ_1 are the partial energies for rate-constant field sensitivity as used (for example) in ref. [22]. We can approximate these by [52] $\mu_1 \simeq 1.54$ eV and $\mu_2 \simeq -10.37$ eV, and take $Q = 0.18$ eV. This gives $\Omega = 2.28$ eV (probably to an accuracy of about ± 0.2 eV).

As an example, consider the derivation of evaporation-field values appropriate to our standard flux criterion $J^{\ominus} = 0.01$ layers/s. Using the value of Ω just derived, and $J_2/J_1 = 0.1$, we obtain from eq. (21) that $\theta_{0.01} = 869$ K. This value, and the standard value for $F(78, 0.01)$, are then used in eq. (19).

As a test of consistency we can make a second estimate of the zero-Q evaporation field F^e. In eq. (19), $\alpha(78, 0.01) = 1.300$. So we obtain:

$$F^e = F(0, 0.01) = 74.08 \text{ V/nm}. \tag{23}$$

Value (23) ought in principle to be identical with value (20), but is not. This shows that there is a small residual inconsistency somewhere, most probably in the estimation of Ω. But since the difference in values (0.10 V/nm) is much less than the statistical standard deviation associated with either value (about 1 V/nm, see below), this residual inconsistency can be neglected here.

Error limits

Reliable estimation of the error limits on predicted evaporation-field values is difficult, and the following remarks should be understood as a first attempt. It is convenient to write $F(78, J)$ in the form:

$$F(78, J) = (1 + g_J)\, F^{\ominus}, \tag{24}$$

where $(1 + g_J)$ is a correction factor of order unity [52]. From eq. (19) we then obtain:

$$F(T, J) = \alpha(T, J)(1 + g_J)\, F^{\ominus}. \tag{25}$$

The errors or standard deviations associated with these three factors we denote by σ_α, σ_g and σ_{cal} respectively.

Straightforward analysis, based on eq. (19), shows that σ_α is related to the standard deviation $\sigma_\theta(J)$ on $\theta_J^{1/2}$ by:

$$\sigma_\alpha = \left(\theta_J^{1/2} + T^{1/2}\right)^{-2} \left|T^{1/2} - (78 \text{ K})^{1/2}\right| \sigma_\theta(J). \tag{26}$$

σ_α vanishes when $T = 78$ K, as is expected (since we are basing field calibrations on measurements performed at 78 K).

σ_θ itself contains two components: an uncertainty associated with the regression on the experimental results for $J = 0.1$ layers/s; and an uncertainty, related mainly to the uncertainty in Ω, associated with the conversion to a θ-value appropriate to some different value of J. This latter uncertainty will be somewhat reduced when better estimates of Ω become available, and can in any case be eliminated by performing experiments at the chosen value of J; so we shall not consider it in detail here. Continued discussion is restricted to the $J = 0.1$ layers/s case, and uses tungsten as an illustration.

Directly from the regression we have: $\theta_{0.1}^{1/2} = 30.7$ K$^{1/2}$; $\sigma_\theta(0.1) = 2.0$ K$^{1/2}$. To explore the error on F^e, we put these values and $T = 0$ K into eq. (26), obtaining $\sigma_\alpha = 0.019$. The relevant value of α was found earllier to be 1.288, so the percentage error is:

$$\sigma_\alpha/\alpha = 1.5\%. \tag{27}$$

The error in $g_{0.1}$ is difficult to determine precisely, but is unlikely to be more than 30%. So the percentage error in the factor $(1 + g_{0.1})$ is:

$$\sigma_g/(1 + g_{0.1}) = 0.3\%. \tag{28}$$

Combining this in quadrature with value (27), we have that the total statistical error on F^e is 1.5%. So, if there were no calibration error, we should have for tungsten: $F^e = 74.2 \pm 1.1$ V/nm.

There remains the calibration error σ_{cal}. This has two types of component. The first relates to the accuracy with which Sakurai and Müller [37] could derive the "surface field" appropriate to (in this case tungsten) field evaporation. Their estimate is $\pm 3\%$. For other materials, that in practice are often calibrated using a more indirect method involving a comparison between the evaporation voltage and the best image voltage, assuming that the best image field for helium is the same for all materials, the error is expected to be higher – perhaps $\pm 10\%$.

The second type of error arises from questions as to the *relevance* of the Müller and Sakurai surface field to the quantity F used in field-evaporation theory. For example, is the Müller and Sakurai field to be identified with the field at the "typical high-risk site"? And is the "mean field experienced by the atom during escape" necessarily to be identified with their surface field? These questions have rarely been articulated, possibly because of their intractibility in the present state of knowledge. But it is difficult to allocate an uncertainty of less than $\pm 10\%$ to them.

We thus conclude that the uncertainty on the tungsten F^e value is at least $\pm 15\%$, and that on the molybdenum F^e value is at least $\pm 20\%$, with the calibration uncertainty being by far the largest source of possible error.

There are also more general conclusions to be drawn from this discussion.

The precision with which we can estimate evaporation field values, *relative* to the evaporation field for "standard conditions" ($J^{\ominus} = 0.01$ layers/s, $T^{\ominus} = 78$ K), is now relatively good for solid field evaporation. It follows that in experiments care should be taken to: (1) identify the voltage corresponding to standard conditions for the emitter in question: and (2) consistently allocate to these standard conditions the same field value (our choice for tungsten is 57.00 V/nm precisely). It would also be helpful to have direct calibrations of the Müller and Sakurai type for materials other than tungsten and iridium, and for there to be further studies of the relationship between this calibration procedure and the Müller and Young procedure (based on field electron emission), and of the uncertainties involved in the calibration procedures.

References

[1] For example: L.W. Swanson, G.A. Schwind and A.E. Bell, J. Appl. Phys. 51 (1980) 3453; R. Clampitt, Nucl. Instr. Methods 149 (1981) 111.
[2] E.W. Müller, Phys. Rev. 102 (1956) 618.
[3] S. Nakamura, J. Electron Microsc. 15 (1966) 279.
[4] D.G. Brandon, Phil. Mag. 13 (1966) 803.
[5] S. Nakamura and T. Kuroda, Surface Sci. 17 (1969) 346.
[6] T.T. Tsong and E.W. Müller, Phys. Status Solidi (a) 1 (1970) 513.
[7] D.M. Taylor, PhD Thesis, University of Cambridge (1970).
[8] R. Gomer, J. Chem. Phys. 31 (1959) 341.
[9] T.T. Tsong, Surface Sci. 10 (1968) 102.
[10] R.G. Forbes, in: Proc. 7th Intern. Vacuum Congr. and 3rd Intern. Conf. on Solid Surfaces, Vienna, 1977, p. 387.
[11] R.K. Biswas and R.G. Forbes, J. Phys. D (Appl. Phys.) 15 (1982) 1323.
[12] E.W. Müller and T.T. Tsong, Field Ion Microscopy: Principles and Applications (Elsevier, Amsterdam, 1969).
[13] T.T. Tsong, J. Chem. Phys. 54 (1971) 4205.
[14] R.G. Forbes, J. Phys. D (Appl. Phys.) 15 (1982) L99.
[15] N. Ernst and Th. Jentsch, Phys. Rev. B24 (1981) 6234.
[16] D.R. Kingham, Surface Sci. 116 (1982) 273.
[17] R.G. Forbes, Surface Sci. 102 (1981) 255.
[18] D.R. Kingham, Surface Sci. 108 (1981) L460.
[19] R. Gomer and L.W. Swanson, J. Chem. Phys. 38 (1963) 1613.
[20] D. McKinstry, Surface Sci. 29 (1972) 37.
[21] R.G. Forbes, Surface Sci. 70 (1978) 239.
[22] R.G. Forbes, R.K. Biswas and K. Chibane, Surface Sci. 114 (1982) 498.
[23] R.G. Forbes, Surface Sci. 116 (1982) L195.
[24] E.W. Müller and T.T. Tsong, Progr. Surface Sci. 4 (1973) 1.
[25] M. Wada, M. Konishi and O. Nishikawa, Surface Sci. 100 (1980) 439.
[26] M. Konishi, M. Wada and O. Nishikawa, Surface Sci. 107 (1981) 63.
[27] R.G. Forbes, J. Phys. D (Appl. Phys.), November 1982.
[28] That is, the field in space somewhat above the relevant surface atom.
[29] This field F^{ϵ} must be carefully distinguished from empirical zero-temperature evaporation fields.

[30] H. Margenau and G.M. Murphy, The Mathematics of Physics and Chemistry (Van Nostrand, Princeton, NJ, 1956).
[31] This is not to say the inverted relationship is necessarily analytic.
[32] R.G. Forbes, Surface Sci. 46 (1974) 577.
[33] That is, the surface such that (at sufficiently large distances) the electrostatic potential energy of an electron is eFx, where x is the distance from the surface. (This surface can usually be identified with the so-called "image plane".)
[34] R.G. Forbes, J. Phys. D (Appl. Phys.) 15 (1982) 1301.
[35] D.R. Kingham, Vacuum, to be published.
[36] E.W. Müller and R.D. Young, J. Appl. Phys. 32 (1961) 2425.
[37] T. Sakurai and E.W. Müller, Phys. Rev. Letters 30 (1973) 532.
[38] By "field-desorption resistant" we mean that a material has a relatively high (in comparison with other materials) evaporation field, or alternatively a relatively high value of K_n for all values of n. (K_n as in eq. (13).)
[39] G. Ehrlich and C.F. Kirk, J. Chem. Phys. 48 (1968) 1465.
[40] D.R. Kingham, PhD Thesis, Cambridge University (1981).
[41] D.R. Kingham, J. Phys. D (Appl. Phys.), to be published.
[42] R.G. Forbes, Appl. Phys. Letters 40 (1982) 277.
[43] R.G. Forbes, J. Phys. D (Appl. Phys.) 15 (1982) L75.
[44] S.C. Ying, in: Theory of Chemisorption, Ed. J.R. Smith (Springer, Berlin, 1980).
[45] R.R. Teachout and R.T. Pack, Atomic Data 3 (1971) 195.
[46] M. Konishi, M. Wada and O. Nishikawa, Surface Sci. 108 (1981) L463.
[47] N. Ernst, Surface Sci. 87 (1979) 469.
[48] R.G. Forbes and K. Chibane, Proc. 29th Intern. Field Emission Symp. (to be published).
[49] The present paper shows that more attention may need to be paid to F^2 terms.
[50] R.S. Chambers, PhD Thesis, University of Illinois at Urbana–Champaign (1976).
[51] K.L. Aitken and G.L.R. Mair, J. Phys. D (Appl. Phys.) 13 (1980) 2165.
[52] R.G. Forbes, unpublished work. (Details available on request.)
[53] T.T. Tsong, J. Phys. F (Metal Phys.) 8 (1978) 1349.
[54] G.L. Kellogg, J. Appl. Phys. 52 (1981) 5320.

Proc. 29th Intern. Field Emission Symp., 1982 (in press)

THE ESCAPE MECHANISM IN LOW TEMPERATURE FIELD EVAPORATION

Richard G. Forbes

University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET, UK

INTRODUCTION

The last three years have seen substantial gains in the basic theory of field evaporation (FEV). At the classical conceptual level we are beginning to have a consistent, experimentally validated, picture. The intention of this paper is to summarise present understanding of the theory, bringing out the main points on which our improved confidence in it is based.

As space is limited, detailed references to most pre-1979 work are not given. Full references may be found in many of the papers cited here.

FIELD EVAPORATION AS AN ACTIVATED PROCESS

Ever since Müller's original work (1), field evaporation near 80 K and above has been taken to be a thermally-activated process. It obeys an emission equation conveniently written in the form (2) :

$$J = n_{hr} k_{hr} = n_{hr} A \exp(-Q/kT) \qquad (1)$$

where: J is the evaporation flux (as expressed in layers/s); n_{hr} is the amount of material at high risk of FEV; k_{hr} is the FEV rate-constant for the high-risk sites; and the other symbols have their usual meanings, as for example in (2). The traditional proof that FEV is in fact thermally activated has been the ability of the Müller-Schottky (M-S) formula to roughly predict observed evaporationfields near 80 K. However this is now suspect.

POST-IONIZATION AND EVAPORATION CHARGE STATE

Field evaporation is now recognised to be often a two-stage process: (i) thermally-determined escape as a 1^+ or possibly 2^+ ion, sometimes followed by: (ii) one or more "post-ionization steps" in which the ions lose further electron(s) by field ionization.

This realization, however, is recent (since 1979). Historically, Müller (3) speculated about post-ionization, but FEV was long assumed to be a single-stage process. Brandon's criterion concerning thermally-activated escape then predicts a 2^+ observed charge state for many materials. When atom-probe experiments found higher charge states, post-ionization was re-examined by Taylor (in 1970) and by Chambers (in 1975) - but was deduced (wrongly) to be theoretically unlikely.

Then, in 1979, Ernst (4) looked at the dependence on field of the proportion of Rh^{++} to Rh^+ ions produced in FEV, and was able to get a reasonable fit to his results using a 1-D theory of post-(field)-ionization. He also showed from activation-energy measurements that both types of ion derived from the same escape process.

Haydock and Kingham (5,6) and Ernst and Jentsch (7) have confirmed and improved the theory of post-ionization, and Kingham (6) has pointed out deficiencies in the work of Taylor and of Chambers. Experimental work by Kellogg (8) supports post-ionization for various materials.

An important implication of the above work is that the escape charge-state is most unlikely to be greater than 2^+.

BASIC NATURE OF THE INITIAL ESCAPE MECHANISM

The bonding state of a surface atom at a metal field-ion emitter has to be partially ionic in order to sustain the external field. The question arises as to whether a kink-site atom is "primarily ionic" (i.e. simple theory should treat it as a 1^+ ion), or "primarily neutral" (i.e.

simple theory should treat it as a polarised neutral). And if the kinksite atom is primarily neutral, a further question is: does ionization occur before escape (so escape takes place over a Schottky image hump); or do ionization and escape occur together ? We thus reach the three escape possibilities traditionally considered:

(a) Escape mechanisms from ionic bonding;

(b) The Müller-Schottky mechanism;

(c) Gomer-type surface charge-exchange mechanisms (9).

The debate as between (b) and (c) was incompletely resolved for many years, although option (c) was usually favoured.

Biswas and Forbes (10,11) have now, for the metals conventionally employed in the low-temperature field-ion techniques, firmly eliminated possibilities (a) and (b) above. Using classical arguments they showed that at observed evaporation fields near 80 K there is for these metals no Schottky hump in the potential for the escaping ion. Therefore neither ionic bonding nor escape over a Schottky hump can be valid assumptions, and a Gomer-type mechanism must operate for these materials.

This conclusion finds additional numerical support in a criterion developed (12) for the a-priori prediction of escape mechanism from atomic and thermodynamic data. Metals predicted to escape via a Gomer-type mechanism include: W,Ta,Ir,Hf,Mo,Pt,Ru,Au,Rh,La,Ti,Co,Fe,Ni,Cu,Pb,Ag and Al. The criterion is indecisive with respect to some materials of interest to liquid-metal ion-source technology, e.g. Ga,Cs,Sn,Si,Ge.

In general terms it seems that high evaporation-field and escape via a Gomer-type mechanism are both linked with a high value of $(\Lambda^{O} + H_{n} - n\phi^{E})$, where: Λ^{O} is zero-field binding energy, H_{n} the formation energy for an n^{+} ion in remote field-free space, and ϕ^{E} the local work-function.

THE PREDICTION OF NEAR-80K EVAPORATION FIELD

Ref.(11) also shows that the "basic-image-hump" formalism often used to analyse the Müller-Schottky (M-S) mechanism is physically invalid because it predicts evaporation fields higher than the field at which the Schottky hump disappears. This means that the derivation of the familiar M-S formula (eq.(2) below with σ_n put equal to ½) is invalid, notwithstanding its apparent empirical success.

Forbes (13) has thus developed an alternative formula, using the same ionic potential-energy contributions as does the M-S derivation (i.e. the electrostatic and image-potential terms), but basing the argument on energetic considerations alone. This gives an "evaporation field" F^e as:

$$F^e \simeq \sigma_n \, (16\pi\varepsilon_0/n^3e^3) \, (\Lambda^0 + H_n - n\phi^E)^2 \qquad (2)$$

where: e is the elementary charge, ne the charge on the ion immediately after escape; and ε_0 the electric constant. σ_n was (13) a parameter estimated as: $\sigma_1 \simeq 0.2$; $\sigma_2 \simeq 0.24$. It has subsequently been shown (14) that estimates of σ_n and F^e are increased by small amounts, not more than and probably much less than 25%, if the repulsive ion-surface interaction is taken into account.

The empirical success of the M-S formula ($\sigma_n \equiv$ ½) in predicting observed evaporation fields near 80 K thus seemed to have a natural explanation. But eq.(2) is better because its derivation is more satisfactory.

But there is a complication. All the above approaches assume that the observed quantity (near-80K onset field) should be much the same as the calculated quantity (zero-Q evaporation field F^e). However it now seems (see later) that F^e-values derived from experiments are significantly higher than near-80K onset fields. Thus the "success" of the calculations described above is problematical and possibly partly fortuitous.

NEW THEORETICAL FORMULAE FOR GOMER-TYPE MECHANISMS

In the debate about FEV escape mechanism, confusion arose because there was a mathematical false assumption in the standard "constant-intersection distance" treatment of Gomer's intersection model. This error was pointed out by McKinstry in 1972, and others subsequently (e.g. (2)), but only recently has a simple alternative treatment become available.

Forbes (16) has shown that, for Gomer-type mechanisms, the activation-energy formula has the form:

$$Q = \Omega \left[(F^e/F) - 1 \right]^2 \qquad (3)$$

where Ω is a parameter defined in the context of eq.(3). Theoretical estimates of Ω are given by:

$$\Omega = \tfrac{1}{2}\kappa a^2 \qquad \text{(1st apprx)} \qquad (4a)$$

$$\Omega = \tfrac{1}{2}\kappa a^2 \, (1-\Pi^e/nea)^2 \, (1-\eta_n'/neF)^{-2} \qquad \text{(2nd apprx)} \qquad (4b)$$

where: κ is the force-constant for an atom vibrating in a parabolic potential-energy well, a is the distance of the well minimum from the emitter's electrical surface, Π^e is a constant with the dimensions of SI dipole moment, defined in (16), η_n is the purely-chemical (IO) component of ion potential energy, and η_n' is its position derivative evaluated at the well minimum. This last is small, so Ω is expected to be constant or nearly constant.

When eq.(3) is combined with the emission equation (1), an expression can be obtained (15) for the temperature-dependence on the onset evaporation field F^o corresponding to a given choice J^o of onset evaporation flux:

$$1/F^o = (1 + T^{\frac{1}{2}}/\Theta^{\frac{1}{2}}) \cdot 1/F^e \qquad (5)$$

$$\Theta = \Omega/k \, \ln(n_{hr}A/J^o) \qquad (6)$$

Θ is a temperature-like parameter, interpretable as the temperature at which F^o becomes equal to half F^e (if the formula is valid to high temperature). A simplified form of eq.(3) is:

$$1/F^o = b + CT^{\frac{1}{2}} \qquad (7)$$

where b and C are constants related to θ and F^e .

Eq.(3) is valid only for "Q not too large". With eq.(7) one looks for validity over a limited range in temperature that includes 80 K .

A direct confirmation of eq.(3) is to be found in the work of Ernst (4), as discussed in the next paper (17). Confirmation of eq.(7) is found, for tungsten and molybdenum, for the temperature range 60 K to 150 K, when the results of Wada et al. (18) are replotted as shown in Fig. 1.

These agreements tend to confirm that, at temperatures near 80 K and above: (a) kink-site atoms can be treated as independent oscillators vibrating in parabolic wells; (b) escape is a thermally-activated process, occurring via a Gomer-type mechanism; and (c) a curve-intersection formalism is adequate, whatever the actual detailed mechanism of escape (which may be charge-hopping or charge-draining - see later).

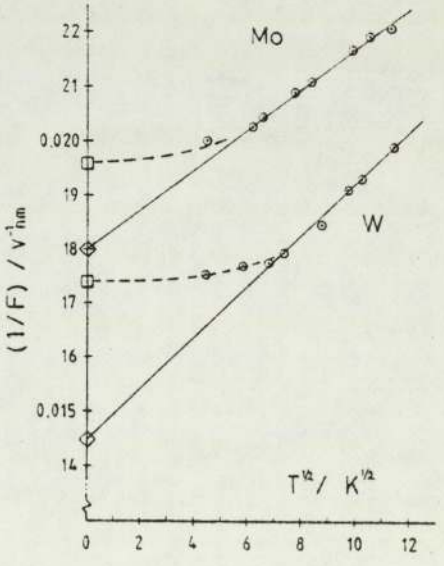


Fig. 1 . Plot of $1/F^o$ against $T^{\frac{1}{2}}$. The squares denote zero-T evaporation field, and the diamonds zero-Q evaporation field.

THE CONSEQUENCES OF HEAVY-ION TUNNELLING

The deviations from linearity observed in Fig. 1 near 50 K for tungsten, near 35 K for molybdenum, probably result from ion tunnelling effects (9,18). If so, these are the first good experimental estimates of the critical temperature below which such effects become dominant. There is broad agreement with recent theoretical estimates of critical temperature (19).

Note also the significant difference between the zero-Q evaporation field (obtained by linear extrapolation of the high temperature results), and the zero-temperature evaporation field (to which the plotted results tend). This difference is a direct result of ion tunnelling.

The zero-Q evaporation field F^e is much higher than observed evaporation fields near 80 K. This high value can in large part be accounted for if F^2-energy terms are taken into consideration: in effect, the value of σ_n in eq.(2) is increased to 0.3 or more.

CHARGE HOPPING AND CHARGE DRAINING

Several variants of surface charge-exchange process can in fact be distinguished, the main ones (20) being: charge-hopping, when ion motion is relatively fast and the electron transfer is pictured as a sharp hop; and charge-draining, when ion motion is relatively slow and charge drains steadily out of the departing entity. The classical theory presented earlier is regarded as a first-generation approximation for both the charge-hopping and charge-draining mechanisms.

Kingham (19) has suggested that the precise shape of the field vs temperature plot in the region of the critical temperature may give an indication of the shape of the effective potential barrier to ion tunnelling, and so of the precise nature of the escape mechanism.

CONCLUSIONS

For many metals field evaporation can now be described by the set of relatively simple equations (1), (3), (5), (6), with the parameters F^e and Ω either determined empirically or estimated from eqns (2) and (4). The nature of the field evaporation process now seems more firmly established than it has been for many years. This should enable FEV theory and experiments to be used with greater confidence in various contexts, in particular in the investigation of atomic behaviour at charged surfaces.

REFERENCES

1. E.W. Müller, Phys. Rev. 102, 618 (1956).
2. R.G. Forbes, R.K. Biswas and K. Chibane, Surface Sci. 114, 498 (1982).
3. E.W. Müller, Adv. Electr. Electron Phys. 13, 83 (1960).
4. N. Ernst, Surface Sci. 87, 469 (1979).
5. R. Haydock and D.R. Kingham, Phys. Rev. Lettrs 44, 1520 (1980).
6. D.R. Kingham, Surface Sci. 116, 273 (1982).
7. N. Ernst and Th. Jentsch, Phys. Rev. B24, 6234 (1981).
8. G.L. Kellogg, Surface Sci. (in press).
9. R. Gomer and L.W. Swanson, J. Chem. Phys. 38, 1613 (1963).
10. R.G. Forbes, J. Phys. D: Appl. Phys. 15, 1301 (1982).
11. R.K. Biswas and R.G. Forbes, J. Phys. D: Appl. Phys. 15, 1323 (1982).
12. R.G. Forbes, J. Phys. D: Appl. Phys. 15 L1 (1982).
13. R.G. Forbes, Appl. Phys. Lettrs. 40, 277 (1982).
14. R.G. Forbes, J. Phys. D: Appl. Phys. 15, L75 (1982).
15. K. Chibane and R.G. Forbes, Surface Sci. (in press).
16. R.G. Forbes, Surface Sci. 116, L195 (1982).
17. R.G. Forbes and K. Chibane, this Proceedings.
18. M. Wada, M. Konishi and O. Nishikawa, Surface Sci. 100, 439 (1980).
19. D.R. Kingham, J. Phys. D: Appl. Phys. (in press).
20. R.G. Forbes, Surface Sci. 102, 255 (1981).

Proc. 29th Intern. Field Emission Symp., 1982 (in press)

THE DERIVATION OF SURFACE ATOMIC INFORMATION FROM FIELD EVAPORATION DATA

Richard G. Forbes and K. Chibane

University of Aston, Department of Physics, Gosta Green, Birmingham B4 7ET, UK

INTRODUCTION

The basic theory of field evaporation (FEV) now seems in a better state than it has been for many years, and hopefully this theory can now be viewed as a scientific tool. With the development of sophisticated experimental techniques involving appearance-energy measurements, e.g. (1), new types of data are available. It seems timely to look again at the nature of the information about atomic behaviour at charged surfaces that can legitimately be derived from field-evaporation experiments.

This paper is intended as a summary of recent investigations that will be reported in detail elsewhere (2,3). We begin by looking again at the electric-field dependence of surface-atom binding energy. We then describe three methods of deriving information from straight-line plots. And we then discuss the resulting parameter values.

THE FIELD DEPENDENCE OF BINDING ENERGY

When an electric field is present, the total binding energy Λ of a surface atom can be expressed as the sum of a zero-field part and a field-induced increase $\Delta\Lambda$. This increase is in part due to polarization of atomic orbitals, in part due to charge transfer ("ct") to the substrate(4,5):

$$\Lambda = \Lambda^{O} + \Delta\Lambda = \Lambda^{O} + \Delta\Lambda(\text{orb}) + \Delta\Lambda(\text{ct}) \qquad (1)$$

The charge-transfer component arises because the surface atom must be partially ionic in order to sustain the external field F (i.e. the field slightly above the surface atom in question).

$\Delta\Lambda$(orb) can be presumed to have F^2 dependence. Ref. (2) argues that $\Delta\Lambda$(ct) must also have F^2 dependence, because in the detachment of the surface atom as a neutral: (a) the work done per unit charge in re-forming the neutral from a partial ion is proportional to F; and (b) from Gauss' theorem the amount of charge moved is proportional to F. Thus both the components, and hence $\Delta\Lambda$, are proportional to F^2. The coefficient involved depends on the crystallographic environment, that is it will be different for a kink site and for an atom diffusing on top of a plane. In the kink-site case we call it c_α, so:

$$\Delta\Lambda = \tfrac{1}{2}c_\alpha F^2 \qquad (2)$$

The experiments of Tsong and co-workers, e.g. (4,6), measure a different coefficient c_δ. And the arguments of Forbes (7) relate to the "orbital-polarization" component of c_α. Most other past methods suggested for obtaining an F^2-term coefficient are thought to have flaws in them.

THE ΛF -PLOT (A METHOD FOR OBTAINING z AND c_α)

It can be shown, e.g. (1,2,8), that total binding energy Λ is related to the onset appearance energy $A^{ons}_{\alpha nr}$, the activation energy $Q_{\alpha nr}$, and the ionic formation energy H_r in remote field-free space by:

$$\Lambda = A^{ons}_{\alpha nr} + Q_{\alpha nr} - H_r + zkT \qquad (3)$$

where kT is the Boltzmann factor, and z is a number of order 10. (The suffix α labels the initial partially-ionic bonding state, n the charge state immediately after escape, and r the charge state observed at the analyser.

In his experiments on Rhodium (1), Ernst showed the FEV mechanism for 2^+ ions to be escape into a 1^+ state, followed by post-ionization to 2^+. It is also clear that

escape is via a Gomer-type mechanism (see (9)). Ernst measured $A^{ons}_{\alpha 12}$ and $Q_{\alpha 12}$ as a function of field F, and has tabulated the temperature of each measurement.

A plot of Λ vs F^2, ignoring the zkT term in the derivation of Λ, is shown in Fig. 1 (circles). Extrapolating back to zero field gives a value of Λ^o less than the known thermodynamic value. Including the zkT term, and adjusting the value of z until regression gives the correct Λ^o-value results in the corrected points (squares). The necessary value of z is 10.5 $\pm$ 3, and the corresponding value of c_α is 1.05 $\pm$ 0.3 meV V^{-2} nm^2 (10), or $4\pi\varepsilon_o \times$ (1.5 $\pm$ 0.5) $Å^3$.

THE QF -PLOT (A METHOD FOR OBTAINING Ω AND F^e)

For Gomer-type escape mechanisms a relationship between activation energy Q and field F has recently been found (9,11). This can be put in the form:

$$Q^{\frac{1}{2}} = (\Omega^{\frac{1}{2}} F^e)/F - \Omega^{\frac{1}{2}} \qquad (4)$$

$$\Omega = \tfrac{1}{2}\kappa a^2 (1-\Pi^e/nea)^2 (1-\eta'_n/neF)^{-2} \qquad (5)$$

$$\Pi^e = (c_\alpha - c_n) F^e \qquad (6)$$

where: F^e is zero-Q evaporation field, Ω an effectively constant parameter defined by eq.(4) and approximately given by eq.(5), κ is the force-constant for an atom vibrating in a parabolic energy well, a the distance of the well minimum from the emitter's electrical surface, e the elementary charge, ne the charge on the ion immediately after escape, Π^e a constant, c_n the coefficient of the F^2 energy term for an n^+ ion, and η'_n a constant presumed small in value (9,11).

A plot of $Q^{\frac{1}{2}}$ vs 1/F, made from Ernst's Rh^{++} data, is shown as Fig. 2. Its linearity helps to confirm the validity of the above formulae (see (9)). By regression we obtain: Ω = 0.26 $\pm$ 0.11 eV; F^e = 61 $\pm$ 20 V/nm.

THE xQ -PLOT (A METHOD FOR OBTAINING κ AND a)

For a parabolic bonding well the relation between activation energy Q and distance x (measured from the emitter's electrical surface) can be put in the forms:

$$Q = \tfrac{1}{2}\kappa(x - a)^2 \qquad (7)$$

$$x = a + (2/\kappa)^{\frac{1}{2}}\, Q^{\frac{1}{2}} \qquad (8)$$

A value of x corresponding to each of Ernst's (1) values of $Q_{\alpha 12}$ can be obtained from the formulae (1,3):

$$x = \left[Y + \{Y^2 - (1-\zeta)e^3F/4\pi\varepsilon_0\}^{\frac{1}{2}}\right] / 2eF \qquad (9)$$

$$Y = A^{ons}_{\alpha 12} - \phi^E - \tfrac{1}{2}c_1F^2 - I_2 + zkT \qquad (10)$$

$$\zeta = (G/x^t) \div (e^2/16\pi\varepsilon_0 x) \qquad (11)$$

where: Y is an energy-like parameter given by eq.(10), ϕ^E is the relevant emitter work-function, I_2 is the Rh second free-space ionization energy, ζ is a parameter given by eq.(11), and G is the constant in a t-th power law representing the repulsive ion-surface interaction. These formulae represent an improved version of a formula given by Ernst and Block (1,12).

A slightly simplified treatment takes ζ as negligible and uses the values: c_1 = 0.15 meV V^{-2} nm^2 (10,13); ϕ^E = 4.8 eV; I_2 = 18.08 eV; z = 10.5 . The resulting plot of x vs $Q^{\frac{1}{2}}$ is shown as Fig. 3. Linear regression leads to the results: a = 0.136 $\pm$ 0.016 nm; κ = 76 $\pm$ 20 eV/nm^2; $\tfrac{1}{2}\kappa a^2$ = 0.71 $\pm$ 0.3 eV.

DISCUSSION OF PARAMETER VALUES DERIVED FOR RHODIUM

Block's parameter z

The parameter z relates to the statistical distribution of initial-state vibrational energy and to the possibility of electron tunnelling into vacant states below the emitter fermi level. Block and co-workers (8,14) suggest that likely values of z are around 10 or less. The value

10.5 found here thus seems generally plausible, although some of the corresponding zkT-values (which go up to 0.54 eV) are higher than Ernst's estimate (0.3 eV) of the likely zkT correction in his experiments.

The F^2 energy-term coefficient c_α

This is the first time a well-defined value has been obtained for this coefficient, for a field evaporating atom, and thus there is little to compare it with. It seems generally compatible with rough theoretical estimates (2), and - in as much as this value is used in discussion of Ω and F^e below without causing problems - fits into the rest of the theory in a generally self-consistent manner.

The zero-Q evaporation field F^e

The value of F^e derived for Rh^+ evaporation, namely 61 $\pm$ 20 V/nm, is markedly higher than observed evaporation fields near 80 K (43 V/nm in ref.(5)) or Ernst's field values, - but is in line with the relatively high values of F^e derived for tungsten and molybdenum from a plot of $1/F^o$ vs $T^{\frac{1}{2}}$ (9). A check on self-consistency can be obtained from eq.(12) below (3,15), with n = 1 :

$$A^o + H_n - n\phi^E = neF^e a + n^2e^2/16\pi\varepsilon_o a - G/a^t - \tfrac{1}{2}(c_\alpha - c_n)F^{e2} \qquad (12)$$

The l.h.s. has value 8.4 eV. The r.h.s. terms have values: + 8.3 eV, + 2.6 eV, - 0.3 eV, - 1.7 eV, respectively, and total 8.9 eV. Agreement is as close as might be expected.

The activation-energy equation parameter Ω

Again, this is the first time a value has been derived directly for Ω , and there is little to compare it with. A self-consistency test can be carried out by putting n=1, n_n'=0, in eq.(5) and using the values of c_α, c_1, F^e, κ, and a, obtained/discussed above. This gives Ω = 0.25 eV (with error so large that it is difficult to assess). The result is extremely close to the regressed value (0.26 eV), but this closeness could be coincidental.

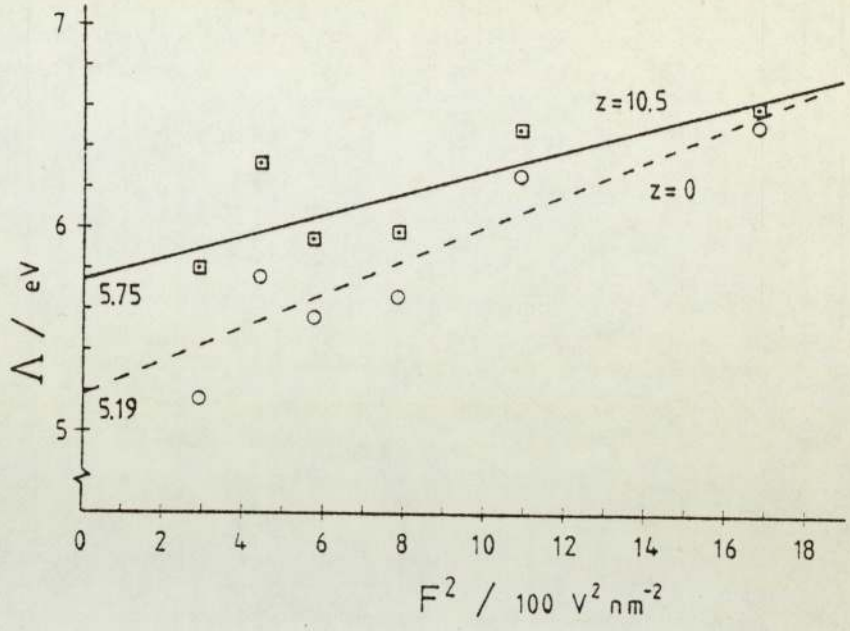


Fig. 1 . Plot of Λ against F^2, showing original (circles) and corrected (squares) points.

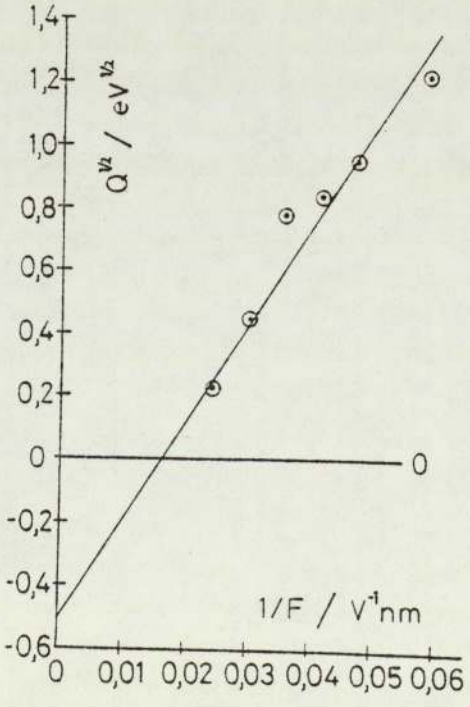


Fig. 2
Plot of $Q^{\frac{1}{2}}$ against 1/F .

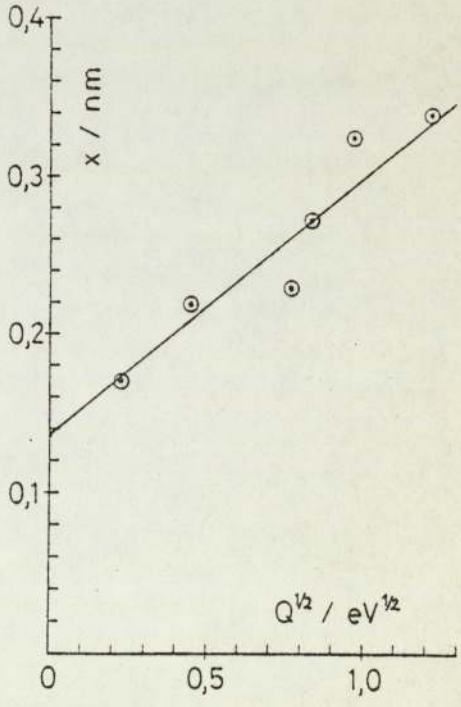


Fig. 3
Plot of x against $Q^{\frac{1}{2}}$

The force-constant κ

This is most easily compared with other data if it is converted to the equivalent vibration frequency ν or temperature θ_C by the equations:

$$\nu = (\kappa/m)^{\frac{1}{2}} /2\pi \qquad (13)$$

$$\theta_C = h\nu/k \qquad (14)$$

where m is the atomic mass, and h and k are the Planck and Boltzmann constants. The resulting values are: $\nu = (1.34 \pm 0.18) \times 10^{12}$ Hz; $\theta_C = 64 \pm 9$ K.

These values are certainly of the order of magnitude expected, though - as Ernst and Block (12) point out - a θ_C-value of this size is significantly less than either the bulk debye temperature for rhodium or the value reported by Chan et al. (16) for vibration of the (111) face. (But a lower value might be expected for a kink-site atom.)

The electrical bonding distance a

For this quantity there is some prior expectation from the work of Lang and Kohn (17) and of Culbertson et al.(18) that a will be greater than the Rh neutral radius (134.5 pm), by a distance of order 50 pm. But the observed value of a (136 pm) is close to the atomic radius. This possibly suggests that, before field evaporation, the atom is sitting in a "maximum coordination" position, and can thus get in closer to the electrical surface than one-dimensional models such as ref.(17) permit. This seems physically plausible, and - if substantiated - would be a firm numerical demonstration of the inadequacy of existing quantum-mechanical models of charged surfaces.

CONCLUSIONS

The Ernst results for Rhodium, when analysed by the techniques outlined earlier, lead to a plausible and generally self-consistent (within their error limits) set of

parameters. This lends support to the underlying FEV theory, and also to the proposition that FEV experiments should provide basic information about surface behaviour.

A major contribution to this result has been the "correlated" nature of Ernst's data: for each experimental field-strength we have values for appearance energy, activation energy, and the temperature corresponding to a given Rh^{++} ion flux. We suggest that detailed experiments of this type should be extended to other materials, and that linked field and temperature values corresponding to a given total flux (all charge states) should also be recorded (so an F^{o} vs T plot can be made, cf.(9)).

REFERENCES

1. N. Ernst, Surface Sci. 87, 469 (1979).
2. R.G. Forbes and K. Chibane, Surface Sci. (in press).
3. R.G. Forbes, K. Chibane and N. Ernst (in preparation).
4. T.T. Tsong and G.L. Kellogg, Phys. Rev. B12, 1343 (1975).
5. T.T. Tsong, Surface Sci. 70, 211 (1978).
6. T.T. Tsong and R.J. Walko, Phys. Stat Sol.(a) 12, 111 (1972).
7. R.G. Forbes, Appl. Phys. Lettrs 36, 739 (1980).
8. E. Hummel, M. Domke, and J.H. Block, Z. Naturforsch. 34a, 46 (1978).
9. R.G. Forbes, this Proceedings.
10. 1 meV V^{-2} nm^2 = 1.602189×10^{-40} J V^{-2} m^2
    = $4\pi\varepsilon_o \times 1.439976$ $Å^3$.
11. R.G. Forbes, Surface Sci. 116, L195 (1982).
12. N. Ernst and J.H. Block, Surface Sci. 91, L27 (1980).
13. D.G. Brandon, Surface Sci. 3, 1 (1964).
14. M. Domke, E. Hummel and J.H. Block, Surface Sci. 78, 307 (1978).
15. R.G. Forbes, J. Phys. D: Appl. Phys. 15, 1301 (1982).
16. C.M. Chan, P.A. Thiel, J.T. Yates and W.H. Weinberg, Surface Sci. 76, 296 (1978).
17. N.D. Lang and W. Kohn, Phys. Rev. B7, 3541 (1973).
18. R.J. Culbertson, T. Sakurai and G.H. Robertson, Phys. Rev. B19, 4427 (1979).

J. Phys. D: Appl. Phys. 15, November 1982

LETTER TO THE EDITOR

Arguments about the emitter shape for a liquid-metal field-ion emission source

Richard G Forbes and G L R Mair

Department of Physics, University of Aston, Gosta Green, Birmingham B4 7ET

Received 27 July 1982

Abstract. Field Evaporation is assumed to be a principal emission mechanism for a liquid-metal ion source, but emitter apex fields calculated on the basis of the conventional shape hypothesis (rounded Taylor cone) do not seem high enough to make field evaporation possible. A numerical estimate is made of the degree of field enhancement provided by the presence of a microprotrusion. Enhancement factors of about 10 or more seem plausible, and these values would be sufficient to render field evaporation a plausible emission mechanism. Alternatively, the Taylor-cone apex may be drawn out into a cusp.

Liquid-phase field-ion emission sources have the potential for significant technological applications, particularly in connection with high-resolution ion lithography and ion implantation techniques (e.g. Seliger et al. 1979, Clampitt 1980). Although it is now clear that space charges and electrohydrodynamic effects may influence current-voltage characteristics (Wagner 1982, Mair 1982), there is still uncertainty about the actual emission mechanism (Venkatesan et al. 1981, Mair and von Engel 1981, Prewett et al. 1982).

Various possibilities have been considered, including: (i) thermally-activated field evaporation; (ii) electron-stimulated desorption of surface atoms as ions; (iii) thermal desorption of neutral atoms, followed by field ionisation. More exotic processes, that lead to the release of neutrals, may also need to be considered.

Field ionisation of neutrals in space above the emitter may account for some part of the emission (e.g. Venkatesan et al. 1981). However, ion energy analysis strongly suggests that the principal emission mechanism is some form of surface ionisation process (e.g. Swanson et al. 1980). Current thinking supports the view that this process is probably some form of field evaporation.

If the emission mechanism is indeed field evaporation, then the field at the emission sites must be comparable with normally expected evaporation fields, even though the emission temperature is much higher than those (near 80K and below) used in the conventional field-ion techniques. This may be demonstrated as follows.

The emission flux J, i.e. the count of atoms evaporated per unit time, is given by the emission equation:

$$J = n_{hr} A \exp(-Q/kT) \qquad (1)$$

where k is the Boltzmann constant, T the thermodynamic temperature, Q the field-evaporation activation energy at the 'high-risk' or 'active' atomic emission sites, A is the rate-constant pre-exponential, and n_{hr} the count of atoms at high-risk sites.

For argument, take the observed emission current to be 10μA. If each emitted atom comes off as a singly-charged ion, then this makes J about 6×10^{13} atoms s^{-1}. Now assume that there are 1000 active sites on an emitter (i.e. $n_{hr} = 1000$ atoms), and that A has its conventional

value of 10^{12} s^{-1} ; substitution into equation (1) gives $Q \sim 3kT$, which at room temperature makes Q about 0.07 eV. The alternative assumption that $n_{hr} = 10^4$ atoms makes Q about 0.13 eV. These values are comparable with (perhaps slightly lower than) activation energies appropriate to conventional field evaporation near 80K. Hence evaporation fields in the ion-source context are expected to be comparable with (perhaps slightly higher than) observed 80K evaporation fields, and may thus be estimated by the Müller-Schottky formula or the alternative formula suggested by Forbes (1982a,b). Thus, for gallium, for example, a field of about 15 V nm^{-1} is expected to be necessary for field evaporation at the above flux value (Tsong 1978).

The question now arises as to whether fields of this size may plausibly exist at the emission sites on a liquid-metal ion source. The original estimates of emitter tip radius were of order 1 μm or greater, these being based on electron micrographs (Aitken 1976, Sudraud et al. 1979, 1980). With tip radii of this size it is difficult to think of any plausible way in which emission site fields of around 15 V nm^{-1} could be produced by the 5-10 kV extraction voltage normally used. However, much sharper tip radii of around 60 nm (600 Å) or less, have been observed in recent experiments (e.g. Swanson et al. 1980, Wagner et al. 1981), and we are led to conclude that the earlier values may well have been invalid.

Even with a tip radius of 60 nm the situation is not entirely clear. General experience with the conventional field-ion techniques suggests that, with a normally-shaped solid emitter, a voltage of 5-10 kV applied to a 60 nm tip could produce fields of 20 V nm^{-1} or more. However, conventional wisdom has it that a liquid-metal emitter has the shape of a rounded Taylor cone, i.e. a half-angle of 49.3° (Taylor 1964). For a conventional solid field-ion emitter, the relationship

$$F = V/kr \qquad (2)$$

is often used, where F is the external field, V the applied voltage, r the tip radius, and k a constant of value about 5 . A cone-shaped emitter of half-angle 49.3° would have a much larger value of k (the exact value depending on the extracting electrode geometry). For example, from a formula in Gomer (1979) we can derive the prediction:

$$k = 0.43\,(R_0/r)^{\frac{1}{2}} \qquad (3)$$

where R_0 "can be thought of as a form factor of the order of the electrode spacing", and for which Gomer suggests the value 1mm . For r = 60nm this formula gives k = 55, and for V = 10kV equation (2) then gives $F = 3\ V\ nm^{-1}$. An alternative, and probably superior, formula (Swatik 1966, Ishitani and Tamura 1981) omits the factor 0.43 from equation (3); this gives k = 130 and $F = 1.3\ V\ nm^{-1}$. Neither of the above field values is high enough to sustain field evaporation.

A related approach is to deduce a field from the stress-balance relationship, as written in the forms:

$$\tfrac{1}{2}\varepsilon_0 F^2 = 2\sigma/r \qquad (4a)$$

$$F = 2(\sigma/\varepsilon_0 r)^{\frac{1}{2}} \qquad (4b)$$

where σ is the surface tension (= $0.71\ N\ m^{-1}$ for gallium), and ε_0 the electric constant. For r = 60 nm this leads to $F \simeq 2\ V\ nm^{-1}$, which is again not high enough to sustain field evaporation.

Our alternatives now are to disbelieve field evaporation, disbelieve the estimation of field, or to assume that the tip endform is not the rounded Taylor cone conventionally assumed.

It is easy to disbelieve the field-value estimations. For example, the assumption of an infinite cone is not reasonable, and the termination of the Taylor cone at a finite base radius would lead to a higher field value. Ishitani and Tamura (1981) find the increase to be at most by a factor of 1.5 ; L W Swanson (private communication 1982) feels that a factor of order 2 could be plausible. But, either way, it seems difficult to get fields of order $15\ V\ nm^{-1}$ from an argument of this kind.

A higher field value would of course be deduced from the above formulae if the apex of the rounded Taylor cone were assumed to be much smaller than has so far been observed. For example, substitution of r = 2 nm into equation (4) leads to the prediction $F = 13\ V\ nm^{-1}$ (which is much the same as the result of Prewett et al. 1982). However, several authors (Ward and Seliger 1981, Kang et al. 1981, Mair 1982, Miskovsky and Cutler 1982) have now shown that the combination of small apex radius and moderate current (even as low as 1 μA) can lead to a severe self-consistency problem for the rounded Taylor-cone geometry:

when the effects of space charges are taken into account it is difficult to obtain a surface field high enough to sustain field evaporation. So, with Ward and Seliger and with Kang et al., we do not believe that ion emission can be explained in terms of a simple rounded-Taylor-cone geometry.

The macrocusp hypothesis. One obvious possibility is that the whole end of the Taylor cone pulls out into a major cusp-like protrusion (or 'macrocusp'). Effects of this type have been observed at moderate to high emission currents (roughly above 60 μA) in the Culham work on caesium (Aitken 1976, Clampitt and Jeffries 1978). More recently, they have been observed for gold by Gaubi et al. (1982): a pronounced cusp is visible down to about 50 μA . (The cusp size depends on current level, and a smaller cusp-like feature is visible down to 20 μA and below.) Similar effects have also been observed with gallium (P Sudraud 1982 private communication).

The change of shape associated with macrocsup formation would certainly increase the field at the emitter apex (for a given extracting voltage). And - if the cusp were sufficiently long, with a sufficiently small terminating half-angle - and if space-charge effects were small or moderate - the field could be high enough to sustain field evaporation.

It is difficult to make any simple estimate of the degree of field enhancement provided by observed cusp shapes. Computer studies based on simplified geometrical representations are in progress (e.g. Ward and Seliger 1981, Kang et al. 1981), but definitive field predictions are not yet available. Nevertheless, we anticipate that the space-charge problem is manageable: for small postulated apex radii the argument would be that, for observed extracting voltages, the predicted Laplace ("no space charge") field would be sufficiently high for the actual surface field to be high enough to sustain field evaporation, even after the field reduction due to space charges; alternatively, one could argue that the necessary fields could be achieved with a somewhat larger apex radius, for which (due to the reduction in postulated current density) space-charge field-reduction effects would be relatively less significant. The analysis of a simplified emitter geometry by Ward and Seliger (1981) supports the view that space-charge limitations need not arise for a macrocusp geometry.

There could, however, be a new form of difficulty if the postulated apex radius is "not small" (say 10 nm or larger). In this case, since we specify that the surface field be high enough to sustain field evaporation, rather than be given by the stress-balance relationship, equation (4) would not hold. For a rounded apex in such circumstances the "outwards" electrostatic forces would be predicted to be markedly greater than the "inwards" surface-tension forces, and the apex region of the liquid metal would be under negative hydrostatic pressure. Such a situation would be severely unstable electrohydrodynamically, and could in practice lead to microdroplet emission and/or to noisy emission current. In circumstances where the observed emission current is almost noise-free the rounded-macrocusp geometry is perhaps not realistic, unless the rounded cusp apex has sufficiently small apex radius.

The microprotrusion hypothesis. An alternative (Weinstein 1975, Prewett et al. 1982) is to suppose that there are a number of relatively small "microprotrusions" on the emitter surface. By "microprotrusion" we mean a feature whose height is significantly less than the radius of the supporting structure (which might be the rounded apex of a Taylor cone, or conceivably the rounded apex of a major cusp-like feature): in this we intend that it should be an adequate first approximation mathematically to consider the microprotrusion as situated on a flat surface. The question of how much features of this type enhance the field has a long history, particularly in connection with vacuum breakdown (e.g. Latham 1981). Thus for a semi-ellipse of revolution, as illustrated in figure 1, the enhancement factor λ is given by (Rohrbach 1971, cited by Latham 1981):

$$\beta = (\lambda^2-1)^{3/2} / \{\lambda \ln [\lambda + (\lambda^2-1)^{1/2}] - (\lambda^2-1)^{1/2}\} \qquad (5)$$

where $\lambda = h/b$. A λ-value of 4 (for example) seems physically plausible; this would give field enhancement by a factor of about 13, which (on the basis of the field estimates above) is more than sufficient to raise the field at the end of the microtip to a level sufficient to induce field evaporation.

Other idealised microprotrusion geometries lead to different estimates of field enhancement (see for example figure 3.3 in Latham 1981), but there seems no real problem in postulating that sufficient

field enhancement can be achieved at the tip of a microprotrusion of some kind.

It may indeed be that such microtips have already been observed in the field-ion microscope. At the 21st International Field Emission Symposium, Polanschutz and Krautz (1974) showed a film of various metals "imaged in their own vapour". The film showed unstable images consisting of numerous short-lived spots, the spots being somewhat larger and far less numerous than in normal field-ion images. The spots seemed not to be images of individual atoms (R G Forbes, E W Müller, private communications). In retrospect it seems possible (though it is certainly not proven) that these were images of microprotrusions, and formed at least in part by field evaporation rather than field ionisation.

We see no problem with postulating that such microtips could be formed at a liquid emitter, though clearly the details of possible formation mechanisms will need to be explored. Overall, we have in mind a process something as follows. A statistical event leads to a small 'bump' on the surface. This then grows outwards under the influence of polarisation forces. As it grows the electric field at its apex gets larger. Finally it reaches a length where field evaporation of the end atoms occurs. One is then in a situation where liquid material is being drawn out continuously through the microtip, and emitted from the end of it. The process is presumably terminated by some statistical event that causes disruption of the microtip (and perhaps the release of neutrals).

For the microprotrusions we feel that the space-charge problem is again manageable, even though the microprotrusion apex radius may be small. For a given extracting voltage, the Laplace field at the microprotrusion apex can be relatively high. And, because of the configuration of the lines of force (as illustrated in Figure 1) the emitted ions will tend to spread sideways, which should alleviate any space-charge problems. (Space-charge problems could also be ameliorated if the microprotrusions were mobile, as envisaged in Weinstein's original proposal.)

It should be clear that this proposal is a speculation. However, it has merit in that: (i) it shows how a normal evaporation field could be present at the emission sites even though the emitter endform

is 'blunt on average'; (ii) it is compatible with the proposition that the average field on the endform is determined by the stress-balance equation; (iii) it is compatible with the observation of relatively narrow energy distributions and small energy deficits - essentially because the microtip is electrically still part of the emitter.

Some further points deserve comment. We assumed earlier that there were 1000 active sites on a 60 nm radius emitter surface. There would probably be several active emission sites on each microtip, but it is not clear how many microtips would be present simultaneously on an endform of basic radius 60 nm; the possibility exists that we might have to consider the number of active sites on the 60 nm emitter to be less than 1000. When the presumed number of active sites reduces to around 100 or less, then - if the presumed emission current is of order 10 μA or more - we cannot necessarily assume that the emission mechanism is thermally-activated field evaporation. The possibility should be considered that the field at the end of the microtips would be so high that the activation-energy barrier disappeared. The mechanism of evaporation in this event would be diferent from that of conventional field evaporation, in that desorbing ions would probably just 'roll off' the end of the microtip, after formation in a process ('super charge-draining') in which electron charge drains out as the field acting increases. The theoretical details of such a process have never been properly explored, but there are no obvious reasons in principle why it should not happen if the field acting increases with time at a sufficiently high rate.

To summarise, we have demonstrated one way in which the field at the emission sites on a liquid-metal ion source could be sufficiently high for field evaporation to occur, and this eliminates the logical difficulty of assuming that the principal emission mechanism is some form of field evaporation from a Taylor cone. We have not, however, in any sense proved that emission takes place from microtips. It may do. Or it may take place from a rounded, but non-Taylor-cone-shaped, emitter tip. Or, if the field calculations leading to equation (2) are seriously in error, it may take place from a rounded Taylor cone. Indeed, the principal mechanism may alter and depend on the prevailing

experimental conditions. The important thing, in our view, is that we have demonstrated that at least one of these possibilities seems to be numerically plausible.

This work forms part of a research programme funded by the UK Science and Engineering Research Council. One of us (GRLM) thanks the SERC for personal financial support.

References

Aitken K L 1976 Proc. Field Emission Day, Noordwijk (Noordwijk: European Space Agency) pp 23-39

Clampitt R 1981 Nucl. Instr. Methods 189 111-6

Clampitt R and Jeffries D K 1978 Nucl. Instr. Methods 149 734-42

Forbes R G 1982a Appl. Phys. Lett. 40 277-9

---- 1982b J. Phys. D: Appl. Phys. 15 L75-7

Gaubi H, Sudraud P, Tence M and Van de Walle J 1982 Proc. 29th Int. Field Emission Symp., Göteborg (to be published)

Gomer R 1979 Appl. Phys. 19 365-75

Ishitani T and Tamura H 1981 Proc. 27th Int. Field Emission Symp. (Tokyo: University of Tokyo Press) pp 194-8

Kang N K, Orloff J, Swanson L W and Tuggle D 1981 J. Vac. Sci. Technol. 19 1077-81

Latham R V 1981 High Voltage Vacuum Insulation: The Physical Basis (London: Academic)

Mair G L R 1982 J. Phys. D: Appl. Phys. 15 (in press)

Mair G L R and von Engel A 1981 J. Phys. D: Appl. Phys. 14 1721-8

Miskovsky N M and Cutler P H 1982 Proc. 29th Int. Field Emission Symp., Göteborg (to be published)

Polanschutz W and Krautz E 1974 in: Abstracts of 21st International Field Emission Symposium (unpublished)

Prewett P D, Mair G L R and Thompson S P 1982 J. Phys. D: Appl. Phys. 15 1339-48

Rohrbach F 1971 CERN Report 71-5/TC-L

Seliger R L, Ward J W, Wang V and Kubena R L 1979 Appl. Phys. Lett. 34 310-2

Sudraud P, Ballongue P, Colliex C, Ohana R and Van de Walle J 1980 Electron Microscopy 80 Vol 1 (Leiden: 7th European Congress on Electron Microscopy Foundation) pp 96-7

Sudraud P, Van de Walle J and Colliex C 1979 J. Physique Lett. 40 L207-9

Swanson L W, Schwind G A and Bell A E 1980 J. Appl. Phys. 51 3453-5

Swatik D S 1969 Charged Particle Res. Lab. Rep. No. CPRL-3-69 (University of Illinois)

Taylor G I 1964 Proc. R. Soc. A 280 383-97

Tsong T T 1978 Surface Sci. 70 211-33

Venkatesan T, Wagner A and Barr D 1981 Appl. Phys. Lett. 38 943-5

Wagner A 1982 Appl. Phys. Lett. 40 440-2

Wagner A, Venkatesan T, Petroff P M and Barr D 1981 J. Vac. Sci. Technol. 19 1186-9

Ward J W and Seliger R L 1981 J. Vac Sci. Technol. 19 1082-6

Weinstein B W 1975 PhD Thesis University of Illinois at Urbaba-Champaign

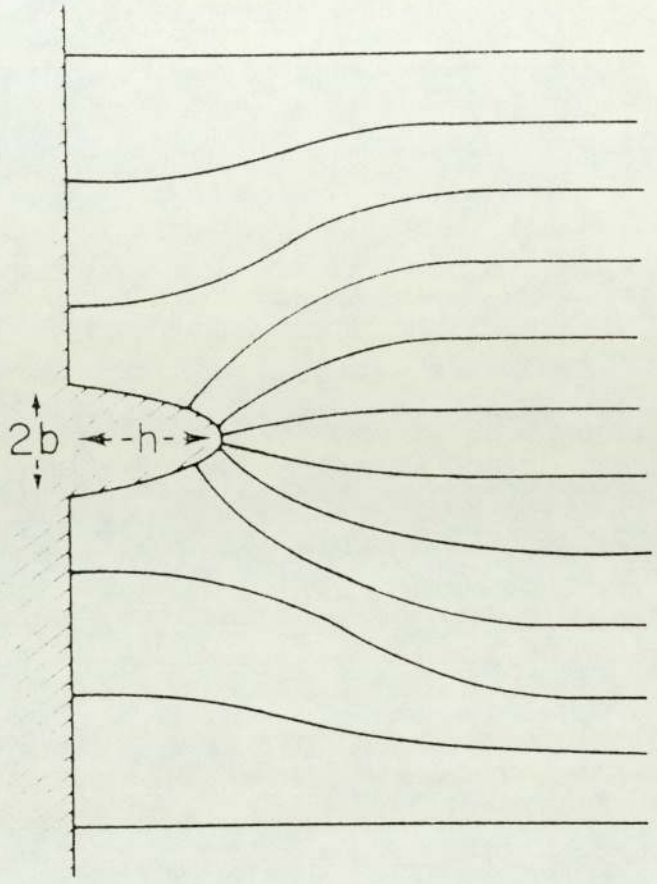


Figure 1. Schematic diagram showing a half-ellipsoid microprotrusion on a locally-flat substrate, and its effect on the field lines. (Reproduced from Latham 1981.)