Some pages of this thesis may have been removed for copyright restrictions.

If you have discovered material in Aston Research Explorer which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our Takedown policy and contact the service immediately (openaccess@aston.ac.uk)
NEURO-ELECTROMAGNETIC IMAGING OF THE HUMAN VISUAL CORTEX

JAMES LEWIS BEDFORD

Doctor of Philosophy

THE UNIVERSITY OF ASTON IN BIRMINGHAM

March 1995

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.
Methods of solving the neuro-electromagnetic inverse problem are examined and developed, with specific reference to the human visual cortex.

The anatomy, physiology and function of the human visual system are first reviewed. Mechanisms by which the visual cortex gives rise to external electric and magnetic fields are then discussed, and the forward problem is described mathematically for the case of an isotropic, piecewise homogeneous volume conductor, and then for an anisotropic, concentric, spherical volume conductor. Methods of solving the inverse problem are reviewed, before a new technique is presented. This technique combines prior anatomical information, gained from stereotaxic studies, with a probabilistic distributed-source algorithm to yield accurate, realistic inverse solutions. The solution accuracy is enhanced by using both visual evoked electric and magnetic responses simultaneously. The numerical algorithm is then modified to perform equivalent current dipole fitting and minimum norm estimation, and these three techniques are implemented on a transputer array for fast computation. Due to the linear nature of the techniques, they can be executed on up to 22 transputers with close to linear speedup.

The latter part of the thesis describes the application of the inverse methods to the analysis of visual evoked electric and magnetic responses. The CIIm peak of the pattern onset evoked magnetic response is deduced to be a product of current flowing away from the surface of areas 17, 18 and 19, while the pattern reversal P100m response originates in the same areas, but from oppositely directed current. Cortical retinotopy is examined using sectorial stimuli; the CI and CIIm peaks of the pattern onset electric and magnetic responses are found to originate from areas V1 and V2 simultaneously, and they therefore do not conform to a simple cruciform model of primary visual cortex.

Keywords: bioelectricity; biomagnetism; evoked response; inverse problem; transputer.
Acknowledgements

The work described in this thesis was carried out by the author, with the following exceptions:

Chapter 7. The author's responsibility was limited to the source analyses. Collection of the data and identification of peak latency was performed by Dr A. Slaven, Mr C. Degg and Dr R.A. Armstrong, to whom the author is grateful. The contour plots were produced by the author, but the data from which these were derived were kindly provided by the above persons.

Chapter 8. Facilities for data recording, signal processing, and averaging were provided by Mr G. Barnes and Dr I. Holliday. Facilities for waveform display, field mapping, global electric and magnetic field power, and magnetic resonance imaging were provided by Dr K. Singh. The author would like to thank these persons for the use of their software and systems.

Signed: James L Bedford

The author would also like to thank his supervisor, Professor G. Harding, for guidance and for making available the research facilities, and his associate supervisor, Dr K. Singh, for support and assistance. The author also gratefully acknowledges the assistance of the Clinical Neurophysiology team at the Department of Vision Sciences, Aston University: Mr P. Furlong, Mrs V. Tipper, Miss A. Edson, Mrs J. Dunne, Dr I. Holliday, Dr S. Anderson, Miss F. Fylan and Mr G. Barnes. Also thanks to Dr R. Hasson of the Open University, Milton Keynes, for his assistance with the transputers in the early stages of the project.
Contents

Figures 8
Tables 15

1. Introduction 16
   Objectives 20

2. The Human Visual System 23
   Introduction 23
   The Retina 24
   Lateral Geniculate Nucleus 29
   Other Subcortical Nuclei 30
   Visual Cortex 31
      Gross Anatomy 31
      Primary Visual Cortex 32
      Histology of Primary Visual Cortex 37
      Physiology of Striate Cortex 41
      Functional Areas of Visual Cortex 45
      Visual Areas in Human 48
   Conclusions 52

3. The Forward Problem 53
   Introduction 53
   The Action Potential 53
   The Postsynaptic Potential 56
   Sources of Evoked Electromagnetic Fields 58
   Fundamental Mathematical Relationships 61
   The Boundary Element Method 66
The Spherical Head Model 70
The Electric Forward Problem 70
The Magnetic Forward Problem 75
Numerical Simulations 78
Conclusions 90

4. The Inverse Problem 91
   Introduction 91
   Equivalent Current Dipole Fitting 92
   Spatio-temporal Dipole Models 99
   Principal Components Analysis 102
   Multiple Signal Classification (MUSIC) 106
   Minimum Norm Estimation 112
   Probabilistic Distributed-source Solutions 116
   The Maximum Entropy Method 120
   The Spatial Filtering Method 124
   Conclusion 127

5. Anatomically Constrained Solutions to the Neuro-electromagnetic 130
   Inverse Problem
   Introduction 130
   Definition of the Source Space 141
   Inversion Technique 145
   Computer Simulations Using a Cruciform Model 154
      Method 154
      Results 157
      Discussion 165
   Computer Simulations Using a Realistic Source Model 168
      No Anatomical Deviation 168
      Variation of the Calcarine Sulcus 169
      Variation of the Whole Brain 169
      Group Averages 173
      Discussion 175
   Conclusions 177

6. Parallel Computing Applied to the Neuro-electromagnetic Inverse 179
   Problem
   Introduction 179
   The Parallel Computer 180
   Transputers 182
   Styles of Parallelism 188

5
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Parallelism</td>
<td>188</td>
</tr>
<tr>
<td>Function Parallelism</td>
<td>188</td>
</tr>
<tr>
<td>Pipelining</td>
<td>188</td>
</tr>
<tr>
<td>The Neuro-electromagnetic Inversion Algorithm</td>
<td>190</td>
</tr>
<tr>
<td>Parallelisation of the Neuro-electromagnetic Inversion Algorithm</td>
<td>192</td>
</tr>
<tr>
<td>Generation of the Gain Matrix</td>
<td>192</td>
</tr>
<tr>
<td>Solution of the Linear System</td>
<td>196</td>
</tr>
<tr>
<td>Process Topology</td>
<td>197</td>
</tr>
<tr>
<td>Performance Characteristics</td>
<td>200</td>
</tr>
<tr>
<td>Scalability</td>
<td>200</td>
</tr>
<tr>
<td>Computation Time</td>
<td>200</td>
</tr>
<tr>
<td>Speedup</td>
<td>200</td>
</tr>
<tr>
<td>Efficiency</td>
<td>201</td>
</tr>
<tr>
<td>Discussion</td>
<td>201</td>
</tr>
<tr>
<td>Minimum Norm Estimation</td>
<td>205</td>
</tr>
<tr>
<td>Numerical Algorithm</td>
<td>205</td>
</tr>
<tr>
<td>Parallel Computation Strategy</td>
<td>207</td>
</tr>
<tr>
<td>Scalability</td>
<td>207</td>
</tr>
<tr>
<td>Computation Time</td>
<td>207</td>
</tr>
<tr>
<td>Equivalent Current Dipole Estimation</td>
<td>208</td>
</tr>
<tr>
<td>Numerical and Computational Algorithm</td>
<td>209</td>
</tr>
<tr>
<td>Scalability</td>
<td>213</td>
</tr>
<tr>
<td>Computation Time</td>
<td>213</td>
</tr>
<tr>
<td>Speedup</td>
<td>214</td>
</tr>
<tr>
<td>Efficiency</td>
<td>214</td>
</tr>
<tr>
<td>Conclusion</td>
<td>214</td>
</tr>
</tbody>
</table>

7. **Source Analyses of Pattern Onset and Pattern Reversal Visual Evoked Magnetic Responses**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>217</td>
</tr>
<tr>
<td>The Pattern Onset Visual Evoked Potential</td>
<td>217</td>
</tr>
<tr>
<td>The Pattern Onset Visual Evoked Magnetic Response</td>
<td>218</td>
</tr>
<tr>
<td>The Pattern Reversal Visual Evoked Potential</td>
<td>222</td>
</tr>
<tr>
<td>The Pattern Reversal Visual Evoked Magnetic Response</td>
<td>223</td>
</tr>
<tr>
<td>Source Analyses</td>
<td>228</td>
</tr>
<tr>
<td>Methods</td>
<td>231</td>
</tr>
<tr>
<td>Results</td>
<td>236</td>
</tr>
<tr>
<td>Group Average Pattern Onset Responses</td>
<td>241</td>
</tr>
<tr>
<td>Group Average Pattern Reversal Responses</td>
<td>243</td>
</tr>
<tr>
<td>Source Analyses of the Group Average Pattern Onset Responses</td>
<td>244</td>
</tr>
</tbody>
</table>
Source Analyses of the Group Average Pattern Reversal Responses
Pattern Onset CIIm Peak in a Single Subject
Pattern Reversal P100m Peak in a Single Subject
Sequential Topography of the Pattern Onset Visual Evoked Magnetic Response
Sequential Source Analyses of the Pattern Onset Visual Evoked Magnetic Response
Conclusions
Cortical Origins of the Pattern Onset Visual Evoked Magnetic Response
Cortical Origins of the Pattern Reversal Visual Evoked Magnetic Response
Effectiveness of the Source Analysis Techniques

8. Cortical Retinotopy Examined Using Visual Evoked Electric and Magnetic Responses
Introduction
Methods
Subjects
Stimulus
Evoked Potential Recording
Evoked Magnetic Response Recording
Peak Identification
Source Analysis
Results
Signals
Global Field Power
Global Source Power
Equivalent Current Dipoles for the Six Subjects
Detailed Analyses for a Single Subject
Minimum Norm Estimates for a Single Subject
Anatomically Constrained Analyses for a Single Subject
Sequential Analyses for a Single Subject
Conclusions
Effectiveness of the Source Analysis Methods

9. Conclusions

References
### Figures

<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Schematic dissection of the geniculostriate visual pathway.</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>Projection of ganglion cell axons through the optic chiasm.</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>The lateral geniculate nucleus.</td>
<td>29</td>
</tr>
<tr>
<td>4</td>
<td>Topography of human visual cortex.</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>The cruciform model of primary visual cortex.</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>Inter-individual variation of the calcarine fissure.</td>
<td>36</td>
</tr>
<tr>
<td>7</td>
<td>Histology of striate cortex.</td>
<td>38</td>
</tr>
<tr>
<td>8</td>
<td>Summary of the principal visual areas and cortical connections in primate.</td>
<td>49</td>
</tr>
<tr>
<td>9</td>
<td>The action potential.</td>
<td>55</td>
</tr>
<tr>
<td>10</td>
<td>The postsynaptic potential.</td>
<td>58</td>
</tr>
<tr>
<td>11</td>
<td>Open-field and closed-field cells.</td>
<td>60</td>
</tr>
<tr>
<td>12</td>
<td>Second order axial gradiometer.</td>
<td>77</td>
</tr>
<tr>
<td>13</td>
<td>Simulated geometry for the calculation of electrical potentials on the surface of a spherical model.</td>
<td>80</td>
</tr>
<tr>
<td>14</td>
<td>Angular distribution of surface potential due to a superficial current dipole.</td>
<td>81</td>
</tr>
<tr>
<td>15</td>
<td>Surface potential distribution due to a superficial current dipole.</td>
<td>82</td>
</tr>
<tr>
<td>16</td>
<td>Angular distribution of surface potential due to a deep current dipole.</td>
<td>83</td>
</tr>
<tr>
<td>17</td>
<td>Surface potential distribution due to a deep current dipole.</td>
<td>84</td>
</tr>
<tr>
<td>18</td>
<td>Simulated geometry for the calculation of magnetic flux density on the surface of a spherical model.</td>
<td>86</td>
</tr>
<tr>
<td>19</td>
<td>Angular distribution of magnetic flux density due to a tangential current dipole.</td>
<td>87</td>
</tr>
<tr>
<td>20</td>
<td>Magnetic flux density distribution due to a tangential current dipole.</td>
<td>88</td>
</tr>
<tr>
<td>21</td>
<td>Geometric construction for calculating the electric potential and magnetic flux density due to two similarly oriented current dipoles of equal moment.</td>
<td>93</td>
</tr>
<tr>
<td>22</td>
<td>Visualisation of principal components analysis.</td>
<td>103</td>
</tr>
<tr>
<td>23</td>
<td>Interpretation of the MUSIC algorithm.</td>
<td>109</td>
</tr>
</tbody>
</table>
24. The minimum norm estimate.  
25. Lead field of a point electrode.  
26. Lead field of a second order magnetic flux gradiometer.  
27. Cruciform model of primary visual cortex.  
28. Right human visual cortex showing the 21 current dipoles which form the right half of the constrained source space.  
29. Six simulated current dipole sources situated within the cruciform model.  
30. Simulated occipital electrode and magnetometer montage.  
31. Occipital distribution of simulated quarter-field visual evoked potentials and visual evoked magnetic fields.  
32. Occipital distribution of simulated quarter-field visual evoked potentials and visual evoked magnetic fields, after addition of 10% noise.  
33. Inverse solutions to the simulated visual evoked potentials.  
34. Inverse solutions to the simulated visual evoked magnetic fields.  
35. Inverse solutions obtained using both electric and magnetic measurements simultaneously.  
36. Inverse solutions to the simulated visual evoked potentials after addition of 10% noise.  
37. Inverse solutions to the simulated visual evoked magnetic fields after addition of 10% noise.  
38. Inverse solutions obtained using both noisy electric and noisy magnetic measurements simultaneously.  
39. Reconstruction of three medial and three calcarine sources, with the sources and source reconstruction space coincident.  
40. Inter-subject variation of the calcarine fissure relative to the occipital lobe: reconstruction of the six current sources at the lowest extreme position.  
41. Inter-subject variation of the calcarine fissure relative to the occipital lobe: reconstruction of the six current sources at the highest extreme position.  
42. Inter-subject variation of the calcarine fissure relative to the inion: reconstruction of the six current sources at the lowest extreme position.  
43. Inter-subject variation of the calcarine fissure relative to the inion: reconstruction of the six current sources at the highest extreme position.  
44. Solution goodness-of-fit as a function of source space rotation for the six sources in the highest and lowest extreme positions.
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>45.</td>
<td>Inter-subject variation of the calcarine fissure relative to the inion: reconstruction of the six current sources at the lowest extreme position after alignment of the source space with the sources by observation of the goodness-of-fit.</td>
</tr>
<tr>
<td>46.</td>
<td>Inter-subject variation of the calcarine fissure relative to the inion: reconstruction of the six current sources at the highest extreme position after alignment of the source space with the sources by observation of the goodness-of-fit.</td>
</tr>
<tr>
<td>47.</td>
<td>Reconstruction of the six current dipoles from the group average of the previous five simulated evoked responses.</td>
</tr>
<tr>
<td>49.</td>
<td>Architecture of a distributed memory parallel machine.</td>
</tr>
<tr>
<td>50.</td>
<td>Uniform communication latency networks.</td>
</tr>
<tr>
<td>52.</td>
<td>Architecture of the IMS T800 transputer.</td>
</tr>
<tr>
<td>53.</td>
<td>Topology of the transputer system used for the work in this thesis.</td>
</tr>
<tr>
<td>54.</td>
<td>Data parallelism.</td>
</tr>
<tr>
<td>55.</td>
<td>Function parallelism.</td>
</tr>
<tr>
<td>56.</td>
<td>Pipeline parallelism.</td>
</tr>
<tr>
<td>57.</td>
<td>Coarse granularity geometric parallelism.</td>
</tr>
<tr>
<td>58.</td>
<td>Fine granularity geometric parallelism.</td>
</tr>
<tr>
<td>59.</td>
<td>Topology of the master-slave process configuration used to implement the fine granularity data parallelism algorithm on a linear pipeline of transputers.</td>
</tr>
<tr>
<td>60.</td>
<td>Topology of a slave process.</td>
</tr>
<tr>
<td>61.</td>
<td>The speedup of the parallel inversion program running on a linear pipeline of transputers.</td>
</tr>
<tr>
<td>62.</td>
<td>The speedup of algorithm steps 1-2 and 4-7.</td>
</tr>
<tr>
<td>63.</td>
<td>The efficiency of the parallel inversion program running on a linear pipeline of transputers.</td>
</tr>
<tr>
<td>64.</td>
<td>The efficiency of algorithm steps 1-2 and 4-7.</td>
</tr>
<tr>
<td>65.</td>
<td>Process topology of the parallel equivalent current dipole program.</td>
</tr>
<tr>
<td>66.</td>
<td>Example of a binocular pattern onset visual evoked potential from a single subject.</td>
</tr>
<tr>
<td>67.</td>
<td>Variation of the pattern onset visual evoked potential with interstimulus interval.</td>
</tr>
<tr>
<td>68.</td>
<td>Effect of stimulus configuration on the pattern onset visual evoked potential.</td>
</tr>
<tr>
<td>69.</td>
<td>Effect of check size on the pattern onset visual evoked potential.</td>
</tr>
<tr>
<td>70.</td>
<td>Example of a binocular pattern onset visual evoked magnetic response from a single subject.</td>
</tr>
<tr>
<td>No.</td>
<td>Description</td>
</tr>
<tr>
<td>-----</td>
<td>---------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>71</td>
<td>Example of a monocular pattern reversal visual evoked potential from a single subject.</td>
</tr>
<tr>
<td>72</td>
<td>Example of a binocular pattern reversal visual evoked magnetic response from a single subject.</td>
</tr>
<tr>
<td>73</td>
<td>Cortical response to left half-field stimulation, as predicted by the cruciform model.</td>
</tr>
<tr>
<td>74</td>
<td>Gross layout of the occipital spherical coordinate system in relation to principal anatomical features and the recording montage.</td>
</tr>
<tr>
<td>75</td>
<td>Left half-field pattern onset visual evoked magnetic response recorded from a single subject.</td>
</tr>
<tr>
<td>76</td>
<td>Group average left half-field, full-field and right half-field magnetic responses to pattern onset and pattern reversal stimulation.</td>
</tr>
<tr>
<td>77</td>
<td>Right half-field pattern reversal visual evoked magnetic response recorded from a single subject.</td>
</tr>
<tr>
<td>78</td>
<td>Equivalent current dipoles for the left half-field, full-field and right half-field group average pattern onset CIIm peak.</td>
</tr>
<tr>
<td>79</td>
<td>Minimum norm estimate for the group average CIIm peak of the response to right half-field pattern onset stimulation.</td>
</tr>
<tr>
<td>80</td>
<td>Minimum norm estimates for the left half-field, full-field and right half-field group average pattern onset CIIm peak.</td>
</tr>
<tr>
<td>81</td>
<td>Anatomically constrained analyses of the left half-field, full-field and right half-field group average pattern onset CIIm peak.</td>
</tr>
<tr>
<td>82</td>
<td>Equivalent current dipoles for the left half-field, full-field and right half-field group average pattern reversal P100m peak.</td>
</tr>
<tr>
<td>83</td>
<td>Minimum norm estimates of the left half-field, full-field and right half-field group average pattern reversal P100m peak.</td>
</tr>
<tr>
<td>84</td>
<td>Anatomically constrained analyses of the left half-field, full-field and right half-field group average pattern reversal P100m peak.</td>
</tr>
<tr>
<td>85</td>
<td>Full-field and right half-field magnetic responses to pattern onset and pattern reversal stimulation, recorded from a single subject using a 42-point montage.</td>
</tr>
<tr>
<td>86</td>
<td>Equivalent current dipoles for the full-field and right half-field pattern onset CIIm peak recorded from a single subject using a 42-point montage.</td>
</tr>
<tr>
<td>87</td>
<td>Minimum norm estimates for the full-field and right half-field pattern onset CIIm peak recorded from a single subject using a 42-point montage.</td>
</tr>
<tr>
<td>88</td>
<td>Anatomically constrained analyses of the full-field and right half-field pattern onset CIIm peak recorded from a single subject using a 42-point montage.</td>
</tr>
</tbody>
</table>
89. Equivalent current dipoles for the full-field and right half-field pattern reversal P100m peak recorded from a single subject using a 42-point montage.

90. Minimum norm estimates of the full-field and right half-field pattern reversal P100m peak recorded from a single subject using a 42-point montage.

91. Anatomically constrained analyses of the full-field and right half-field pattern reversal P100m peak recorded from a single subject using a 42-point montage.

92. Sequential topography of the magnetic response to onset of 9' × 7' checks, recorded from subject 1.

93. Sequential topography of the magnetic response to onset of 9' × 7' checks, recorded from subject 2.

94. Sequential topography of the magnetic response to onset of 77' × 55' checks, recorded from subject 1.

95. Locations of the equivalent current dipoles for the response of subject 1 to 9' × 7' checks.

96. Locations of the equivalent current dipoles for the response of subject 2 to 9' × 7' checks.

97. Locations of the equivalent current dipoles for the response of subject 1 to 77' × 55' checks.

98. Minimum norm estimates for the response of subject 1 to 9' × 6' checks: early latencies.

99. Minimum norm estimates for the complete response of subject 1 to 9' × 6' checks.

100. The goodness-of-fit of an anatomically constrained solution as a function of axis rotation for the response of subject 1 to 9' × 6' checks at a latency of 80ms.

101. Anatomically constrained analyses for the response of subject 1 to 9' × 6' checks: early latencies.

102. Anatomically constrained analyses for the complete response of subject 1 to 9' × 6' checks.

103. Minimum norm estimates for the response of subject 2 to 9' × 6' checks: early latencies.

104. Minimum norm estimates for the complete response of subject 2 to 9' × 6' checks.

105. Anatomically constrained analyses for the response of subject 2 to 9' × 6' checks: early latencies.

106. Anatomically constrained analyses for the complete response of subject 2 to 9' × 6' checks.
107. Minimum norm estimates for the response of subject 1 to 77' × 55' checks: early latencies.
108. Minimum norm estimates for the complete response of subject 1 to 77' × 55' checks.
109. Anatomically constrained analyses for the response of subject 1 to 77' × 55' checks: early latencies.
110. Anatomically constrained analyses for the complete response of subject 1 to 77' × 55' checks.
111. Details of the sectorial stimulus.
112. Occipital electrode montage, showing the relationship of the electrodes to the 10-20 electrode sites.
113. Gradiometer montage, showing its approximate position relative to the 10-20 electrode sites.
116. Visual evoked electric responses to the eight stimuli, recorded from channel 16 on subject AW.
117. Visual evoked magnetic responses to the eight stimuli, recorded from channel 5 on subject AW.
118. Electric global field power for the response of subject AW to the eight stimuli.
119. Magnetic global field power for the response of subject AW to the eight stimuli.
120. Magnetic global field power for the response of subject FF to the eight stimuli.
121. Global source power as a function of latency for the response of subject AW to the eight stimuli.
122. Electric global field power, magnetic global field power, and global source power for the response of subject AW to stimulus 4.
123. Equivalent current dipoles for the response of subject IM to the eight stimuli.
124. Equivalent current dipoles for the response of subject KS to the eight stimuli.
125. Equivalent current dipoles for the response of subject PF to the eight stimuli.
126. Equivalent current dipoles for the response of subject AW to the eight stimuli.
127. Equivalent current dipoles for the response of subject FF to the eight stimuli.
128. Equivalent current dipoles for the response of subject CS to the eight stimuli.

129. Magnetic resonance images of the visual cortex of subject FF, showing the locations of the dipoles for the eight sectorial stimuli.

130. Visual evoked potentials produced in subject IM by the eight stimuli.

131. Visual evoked magnetic responses produced in subject IM by the eight stimuli.

132. Minimum norm estimate of the response of subject IM to stimulus 5, obtained using electric data only.

133. Minimum norm estimate of the response of subject IM to stimulus 5, obtained using magnetic data only.

134. Minimum norm estimate of the response of subject IM to stimulus 5, obtained using both electric and magnetic data.

135. Minimum norm estimates for the response of subject IM to the eight stimuli, obtained using both electric and magnetic data.

136. Goodness-of-fit as a function of source space orientation for the response of subject IM to stimulus 5.

137. Anatomically constrained analyses for the response of subject IM to the eight stimuli.

138. The visual evoked potential elicited in subject IM by stimulus 5.

139. The visual evoked magnetic response elicited in subject IM by stimulus 5.

140. Anatomically constrained analyses for the response of subject IM to stimulus 5.

141. The visual evoked potential elicited in subject IM by stimulus 6.

142. The visual evoked magnetic response elicited in subject IM by stimulus 6.

143. Anatomically constrained analyses for the response of subject IM to stimulus 6.
1. Radii and conductivities of the spherical shells in the 5-sphere electrical conductivity model. 71
2. A comparison of electric potentials and magnetic fields for the case of a spherically symmetric volume conductor. 89
3. Coordinates of the 42 current dipoles which form the source space for visual cortex. 146
4. Coordinates of the six sources in the absence of anatomical variation. 170
5. Variation of the calcarine fissure relative to the occipital lobe: coordinates of the six current sources at the lowest extreme position. 170
6. Variation of the calcarine fissure relative to the occipital lobe: coordinates of the six current sources at the highest extreme position. 170
7. Variation of the calcarine fissure relative to the inion: coordinates of the six current sources at the lowest extreme position. 172
8. Variation of the calcarine fissure relative to the inion: coordinates of the six current sources at the highest extreme position. 172
9. Equivalent current dipole starting locations. 211
10. Details of the gradiometer coils in the 19-channel magnetometer. 304
11. Equivalent current dipole parameters for the six subjects and eight stimuli. 325
1. Introduction

Since 1861, when Pierre Paul Broca first discovered that a specific region of the frontal lobe of the brain was responsible for speech (Broca 1861), scientists have been endeavouring to understand the function of each part of the brain. In 1909, Korbinian Brodmann presented a complete functional map of the cortex, showing the different functional areas as identified by their cytoarchitecture (Brodmann 1909). Although Brodmann’s classification is now known to be incomplete (Clarke and Miklosy 1990), the notion that each area of the brain performs a specific function has endured. The rapidly advancing field of functional localisation aims to determine more accurately the precise role of each part of the brain.

But functional localisation does not merely satisfy a desire to understand the brain: it is the basis of a promising surgical tool. When brain surgeons operate, they need to know exactly which parts of the brain are associated with the vital functions, such as vision, hearing, touch, or movement, so that they do not cause impairment of sight or partial paralysis. To date, this knowledge can only be achieved by exposing a large area of the brain and electrically stimulating it, a procedure which requires the patient to be conscious, or by electrically stimulating one of the limbs and recording the cerebral response using electrocorticography. Both of these techniques are used to localise somatosensory and motor cortex, but neither of them has been used successfully in conjunction with visual cortex, which has a more complex morphology. With functional localisation, however, the functional boundaries of the brain can be established beforehand, thereby enabling the surgeon to plan an access route and to avoid damaging functionally important areas (Harding 1993). The functional localisation procedure itself is completely non-invasive.

The principle behind functional localisation is simple. Firstly, a specific stimulus, such as a visual stimulus characterised by its shape, colour or texture, is presented to a human subject. The resulting cerebral response influences physical quantities outside of the head, and these external phenomena are measured by a scanning technique. A process of deduction is then used to calculate the location of the cerebral response from the observed external events. This enables a link to be established between the stimulus and the part of the brain that it activates. Of course, the stimulus may evoke neural
activity in different parts of the brain at different times after stimulation, so the imaging technique needs to be able to resolve a temporal sequence of events.

The imaging technique, which is used to measure the neural response indirectly and then deduce exactly how it is distributed within the brain, is a crucial stage in the functional localisation process. To date, three main methods have proven to be successful (McCrone 1995). The first, positron emission tomography, makes use of radioactively labelled water - hydrogen combined with oxygen-15 (Fox and Mintun 1989). The labelled water is injected into a vein in the arm. In just over a minute, the radioactive water reaches the brain, the oxygen-15 decaying with a half-life of about two minutes. The decay of the oxygen-15 provides a copious supply of positrons, which scatter from their parent atoms. After travelling a short distance, each positron inevitably collides with an electron, causing an annihilation reaction. This, in turn, produces a pair of gamma-rays of equal energy which travel in diametrically opposite directions. A large ring of gamma-ray detectors surrounding the head records the arrival of the two gamma-rays. The location and time of arrival can be used to deduce the position of the annihilation reaction, which also indicates the position of the labelled water. Thus, by measuring the external gamma-radiation, the distribution of labelled water, and hence blood, can be traced.

Positron emission tomography is used in functional localisation to measure the increase in blood flow associated with a specific stimulus. Before any stimulus is applied, the cerebral blood flow is deduced by the process described above. The subject is then repetitively stimulated and the blood flow again inferred after correcting for the decaying activity of the oxygen-15. The difference between the two blood flows gives a measure of the extra blood flow required by the neurones responsible for processing the stimulus. Statistical measures indicate the significance of the increase in blood flow, relative to the noise level.

Positron emission tomography, as an imaging technique for use in functional localisation of the brain, has several drawbacks. Although the injected sample of oxygen-15 decays completely to a non-radioactive form in about ten minutes and has no harmful effect, the technique cannot be successively repeated in a single subject due to the possibility of radiation overdose. This limits its use in extensive research, where often a number of tests are required. A further drawback of the method is that it has a relatively low spatial resolution of about 5-10mm (Harding 1993). Whilst it can give an indication of the approximate region of the brain associated with a specific function, it cannot pinpoint the active region with high precision. A more serious disadvantage is the low temporal resolution. Positron emission tomography is based on the recording of atomic events which occur at random time intervals. To overcome this inherent randomness, as many events as possible must be recorded, a criterion which is ensured by making the recording time sufficiently long. Since the radiation dose to the patient must be minimised, it is desirable to use as small a radioactive sample as possible, which gives rise to few decay events and thus dictates a long time window. The result of the
long recording time - above ten seconds - is that no information is yielded about the
temporal sequence of cerebral activation: the image represents the average blood flow
over time, there being no distinction between a region which is highly active for a short
time and a region which is less active more frequently.

An alternative technique, which avoids administering a radioactive sample and
has a higher spatial resolution, is functional magnetic resonance imaging. This technique
forms images by measuring nuclear magnetisation (Vennart 1985). When a large static
magnetic field is applied to body tissue, where a large number of protons are found in
such substances as oxygen or protein, the protons precess about the direction of the
applied field, so that their magnetic moments are either directed in the same direction to
it, or in the opposite direction. Under normal equilibrium conditions, there are slightly
more protons oriented in the same direction as the field than protons oriented in the
opposite direction, as the former are in a lower energy state. Thus, there is a small net
magnetisation pointing in the direction of the applied magnetic field. This magnetisation
is removed by applying a pulse of radio-frequency radiation at the Larmor frequency,
after which the magnetisation returns exponentially by a process known as relaxation or
free induction decay (Vennart 1985). The relaxation curve can be measured using a
receiver circuit attached to the radio-frequency source. If a small field gradient is added
to the static magnetic field, the Larmor frequency varies across the scanning area, so
that an electromagnetic pulse at one frequency only activates a narrow slice of the
human subject. The chemical composition of this slice is found by taking the two-
dimensional Fourier transform of the free induction decay signal. A complete three-
dimensional image of the body can be constructed by varying the frequency of the
applied electromagnetic pulse, thereby selecting successive slices of the subject, and then
producing a two-dimensional image of each slice (Lerski 1985).

Functional magnetic resonance imaging measures the increase in oxygenation of
the blood at regions of heightened neural activation (Raichle 1994). It does this by
detecting small differences in the relaxation properties of oxyhaemoglobin and
dehaemoglobin. Inactive neurones consume glucose and oxygen, the latter being
conveyed vascularity by oxyhaemoglobin, which reverts to deoxyhaemoglobin when the
oxygen is expended. Stimulated neurones consume additional glucose but not additional
oxygen. Consequently, when a region of the brain becomes active, the increased blood
flow brings an increased supply of oxyhaemoglobin and glucose, but the oxygen content
of the oxyhaemoglobin is not exhausted. Thus, the small veins draining the active
neurones attain a higher concentration of oxyhaemoglobin and a lower concentration of
dehaemoglobin than normal. This change in concentration is detected by functional
magnetic resonance imaging as a change in the free induction decay process.

It has already been remarked that this technique does not involve a radiation dose
and is therefore preferable to positron emission tomography. The spatial resolution is
also as good as 1-2mm (Harding 1993). The temporal resolution of the latest machines
is in the order of several tens of milliseconds, although the spatial resolution is reduced
at these sampling rates. Importantly, the same machine can be used to produce static anatomical images of the brain and dynamic functional images, making this a very powerful approach to functional localisation.

But both positron emission tomography and functional magnetic resonance imaging both have a fundamental drawback: they both measure vascular quantities as opposed to neural quantities. Neural signals can pass from one part of the brain to another in less than ten milliseconds. However, it takes hundreds of milliseconds for blood flow or oxygenation to change correspondingly (Raichle 1994). Hence, no matter how high the sampling rate of the imaging apparatus, the temporal resolution with respect to neural events (as opposed to vascular events) is always limited to the time taken for the blood flow to reflect the neural activity - hundreds of milliseconds. Thus, the detailed dynamics of neural signal processing in the brain cannot be observed by the methods described so far.

Neuro-electromagnetic imaging overcomes this fundamental limit to temporal resolution. The neurones in the active region of the brain are characterised by intracellular potential differences which give rise to electric potentials on the surface of the scalp. These potentials can be measured using scalp electrodes and high impedance amplifiers (Brown and Smallwood 1981). The potential differences within the active neurones also produce cerebral current flow which generates external magnetic fields. These can be measured using sensitive superconducting quantum interference devices (Clarke 1989). Having measured the external electric potentials and/or magnetic fields, electromagnetic theory is used to "backtrack" to find exactly what neural population is producing the observed signals - a process referred to as the neuro-electromagnetic inverse problem (in contrast to the neuro-electromagnetic forward problem, which is the opposite task of computing the electric and magnetic fields from a known neural population). The only limits to the temporal resolution are the sampling rate and frequency filtration of the recording equipment.

The main drawback of neuro-electromagnetic imaging is that the inverse problem is non-unique. That is to say, many combinations of neural activity all produce the same external electric and magnetic fields, so that, having recorded the external response, it is theoretically impossible to accurately determine the true underlying activity. In practice, however, it is possible to apply various forms of additional information to the problem, so that the active neural population can be identified. The simplest and most common form of additional information is the assumption that one small region of the brain produces the external electric and magnetic fields. However, this may or may not be true. Other assumptions are that several discrete regions of the brain are active, or that certain distributions of activity are more likely than others, and so on.

The spatial resolution of a neuro-electromagnetic image is dependent upon such factors as sensor spacing and geometry, the background noise level, and most importantly, the inverse procedure employed. A distributed-source algorithm, which assumes that any pattern of activity can occur within a two- or three-dimensional region
of the brain, typically provides a spatial resolution of 10-20mm. (It should be noted that it is meaningless to refer to spatial resolution if a single active neural population, which gives rise to a single current dipole, is assumed. This is because the spatial resolution is defined as the closest distance apart that the centres of two active neural populations can come whilst still being correctly identified as two populations. If there is only one active population of neurones, producing one region of neural current, then spatial resolution is obviously undefined. However, the accuracy with which an active neural population can be identified by neuro-electromagnetic imaging, on the assumption of a single current dipole, is about 1-3mm in vitro (Gharib et al. 1995) and 3-4mm in vivo (Sutherling et al. 1988).)

Neuro-electromagnetic imaging is also cost-effective and convenient. A recording system for measuring evoked potentials costs up to £10 000, as opposed to £1 000 000 for positron emission tomography or functional magnetic resonance imaging. Even a neuromagnetometer can be installed for less than £300 000. Both electric and magnetic measurements can be made non-invasively; in the electric case, electrodes must be fixed to the scalp, but in the magnetic case, no patient interference is required at all, although the patient must be co-operative enough to remain completely still. Hence, for reasons of cost-effectiveness, convenience to the subject, and above all, high temporal resolution, neuro-electromagnetic imaging is the most promising of all the functional imaging techniques.

Objectives

This thesis aims to develop a method which can provide a more unique solution to the neuro-electromagnetic inverse problem. The three main ways in which this is achieved are (1) by assuming that the neurones responsible for the evoked electric and magnetic responses are restricted to the cortex only; (2) by analysing both electric and magnetic measurements simultaneously to increase the available information with regard to the neural electric current; and (3) by using parallel computing to speed up the inverse procedure, thereby reducing the computation time to a sufficiently small interval. These principles are described with reference to functional localisation of the human visual system, i.e. using visual evoked electric and magnetic responses. Accordingly, the next chapter is a select review of the anatomy and function of the human visual system. Chapter three then describes the mechanism by which the neurones produce measurable electromagnetic fields. Having established this mechanism, it is possible to formulate the electric and magnetic forward problems mathematically. Some computer simulations are used to illustrate the electric potential distributions and magnetic flux density distributions that would arise from certain neural current sources.

The fourth chapter then discusses existing methods of solving the inverse problem. Particular emphasis is given to the way in which specific prior information is
incorporated into each method to produce a mathematically tractable equation from what is essentially, in its broadest form, an insoluble problem. The prior information ranges from assuming that only one neural current source is present, to assuming that any number are present but that the combination of activity that they form is governed by a probability distribution, so that the most probable current source arrangement can be found. Fourier methods are also briefly considered.

Chapter five introduces a new method of solving the neuro-electromagnetic inverse problem for the case of visual evoked electric and magnetic responses. By assuming that these responses are generated solely within the visual cortex by neurones which lie perpendicular to the cortical laminae, it is possible to reduce the number of unknown parameters in the inverse problem. Secondly, if both electric and magnetic responses are used simultaneously, then the available information with regard to the neural current sources is increased, and when both modalities are used in conjunction with the cortically constrained current region, an overdetermined linear system can be produced, which yields accurate solutions. These principles are illustrated by computer simulations involving a simple geometric model of visual cortex. A realistically shaped model of visual cortex is then developed and used in further computer simulations which show that even if the model brain does not exactly coincide with the brain of each subject, sufficiently accurate inverse solutions can still be obtained.

Chapter six briefly reviews different types of parallel computer and considers general methods by which a conventional computer algorithm can be converted into a parallel program. These concepts are then applied to the neuro-electromagnetic inversion algorithm developed in chapter five, which provides anatomically constrained solutions. The performance of the parallel program running on a transputer network is evaluated. Consideration is then given to ways in which the algorithm can be adapted to perform minimum norm estimation or equivalent current dipole fitting. The parallel implementation of these two methods is described, and their performance is briefly discussed.

The methods developed in chapters five and six are applied to real data in chapters seven and eight. In chapter seven, visual evoked magnetic fields are used to determine the cortical responses to achromatic pattern onset and pattern reversal stimuli. The chapter presents the complete functional localisation procedure: visual stimulation using luminance contrast and form, measurement of the evoked magnetic responses, and deduction of the cortical regions activated by the stimuli. This chapter produces some interesting results to a well-studied but controversial problem, but also provides an opportunity for the analysis methods to be evaluated. Chapter eight further continues these themes, examining cortical retinotopy. The retinotopic organisation of visual cortex is investigated by successively stimulating each sector of the visual field and observing the corresponding region of activity in the brain. In this chapter, both electric and magnetic responses are used to yield an accurate picture of the cortical current. The investigation again provides a forum for the evaluation of the inverse procedures.
This thesis therefore develops new ways of approaching the neuro-electromagnetic inverse problem, with a view to advancing the wider field of functional localisation. The new methods, along with existing methods, are used to determine the cortical responses to achromatic patterned stimuli presented in various regions of the visual field.
2. The Human Visual System

Introduction

The human visual system begins at the retinas, where photoreceptors convert light energy into electric potentials. The optic nerves carry visual information from the retinas to the optic chiasm, and from there a number of neural tracts diverge to various thalamic nuclei and the brainstem. Of the afferent nerves leaving the optic chiasm, the optic tracts are the most significant as they carry some 90% of the visual signals to the lateral geniculate nuclei (Schiller and Malpeli 1977). The lateral geniculate nuclei relay the signals to the primary visual cortex at the occipital region of the neocortex (figure 1). Beyond the primary visual cortex, other visual areas carry out advanced processing of the visual percept.

This chapter provides an overview of the human visual system with a view to showing, in the next chapter, how the visual cortex gives rise to external electromagnetic fields. For completeness, the retina and lateral geniculate nucleus are described, as these carry out preliminary visual processing, but the majority of the chapter is concerned with the anatomy and function of the visual cortex itself.

Much of the present knowledge of the human visual system is extrapolated from studies of primate and cat visual systems. The macaque monkey has a very similar visual performance to that of human (Boynton and Whitten 1970; DeValois et al. 1974) and the visual system of the macaque is therefore a valuable model for studying human vision. However, the primary visual cortex of human is about twice the size of the corresponding cortex in macaque (Van Essen 1985) and each species also has a distinct cortical organisation (Valverde 1985), so care needs to be taken in applying primate results to human. Hence, in this chapter, an attempt is made to concentrate on knowledge which relates explicitly to human experiments, although other species are used where necessary.
Figure 1. Schematic dissection of the geniculostriate visual pathway, viewed axially. After Noback and Demarest (1981).

The Retina

The retina consists of a photoreceptive layer, or neuroretina, and pigment epithelium. It has a total thickness of around 150 - 300μm. In this form, it extends as far as the ora serrata, just posterior to the ciliary body (Polyak 1941). In human, as in all other vertebrates, the retina is inverted: the photoreceptive rods and cones lie closest to the sclera (outermost) while the cells involved with processing and transmitting neural information to the brain are closest to the vitreous humour (innermost). Light must pass through all retinal layers before reaching the rods and cones. At the fovea, the processing layers are spread to the sides, so that light only has to pass through a nuclear layer before reaching the photoreceptive elements. There are five distinct layers, containing six major cell types: three layers of cell bodies (the outer nuclear layer, the inner nuclear layer and the ganglion cell layer), and two layers of synaptic connections (Rodieck 1988). The output from the photoreceptors is projected via synapses to horizontal, bipolar, amacrine, interplexiform and ganglion cells. The axons of the ganglion cells converge to form the optic nerve.

The cells of the pigment epithelium contain a black melanin pigment, fuscin, which is concentrated in granules. The processes of the epithelial cells surround the outer segments of the rods and cones and thus prevent any optical backscatter which would reduce the resolution of the retina. The pigment epithelium also provides metabolic support for the rods and cones (Wald 1968; Boycott and Dowling 1969).

Österberg (1935) examined flat sections of the complete retina, counting the photoreceptors in each region. The fovea was found to be completely free of rods; the first rods appeared around 130μm from the foveal centre. At about 5mm from the
fovea, there was an annulus containing the maximum rod density of 160,000/mm².
There was a reduction in density towards the perimeter of the retina, where the rods
amounted to 23,000 - 50,000/mm². The centre of the fovea contained 147,300 cones
per square millimetre. This reduced to 5,000/mm² at the periphery, with a slight increase
being observed at the ora serrata. There were a total of 110 - 125 million rods and 6.3 -
6.8 million cones. Although more recent studies (Farber et al. 1985; Curcio et al. 1987;
Rodieck 1988) have described significant variations in photoreceptor densities, the
results have been broadly consistent with the estimates given by Østerberg.

Both rods and cones consist of an outer segment, a connecting structure, an inner
segment, a cell body and a synaptic base. Membranous disks in the outer segments of
the rods and cones are responsible for photochemical reactions which transform light
energy into electric potential. Excited photoreceptors release glutamate, an excitatory
neurotransmitter, from their synaptic bodies, resulting in slow postsynaptic
depolarisation of bipolar and horizontal cells. Interestingly, photoreceptors are excited
in the dark and suppressed by light; the presence of light causes them to hyperpolarise
(Svaetichin 1953; Tomita 1963; Werblin and Dowling 1969).

The rods are about 2 μm in diameter at the central retina and about 4 - 5 μm in
diameter peripherally. They have a high sensitivity so that they are effective at low
luminances. Their mean peak spectral sensitivity is 496.3 ± 2.3 nm (Dartnell et al. 1983).
The cones vary in diameter from 1.5 μm at the fovea to 5 - 8 μm at the periphery of the
retina. Cones are specialised for colour vision and resolution of fine detail. They have a
high stimulation threshold and therefore require bright illumination. There are three
types of cones, having spectral sensitivities which are maximal at short, medium and
long wavelengths respectively. Short wavelength cones (S cones) have a mean peak
sensitivity of 419 nm, medium wavelength cones (M cones) have a mean peak sensitivity
of 531 nm, while long wavelength cones (L cones) have a mean peak sensitivity of
around 558 nm (Dartnell et al. 1983). In the retina of the macaque monkey, the relative
densities of rods, S cones, M cones and L cones are typically in the ratio 52:12:47:44
(M. fascicularis), or 22:4:13:13 (M. mulatta) (Hárosi 1987). Most notably, the S cones
are considerably less numerous than the other types of cone (Marc and Sperling 1977;
and are absent from the fovea. Their density increases rapidly with eccentricity and
peaks about 1° from the foveal centre, where they constitute about 20% of the cones.
The S cones appear to be arranged in a regular triangular lattice, while the other cones
fill in the gaps in a random manner (Marc and Sperling 1977).

The simplest rod system consists of a series of rods synapsing onto a rod bipolar
cell which, in turn, synapses with a rod ganglion cell. A cone system contains one or
more cones, a cone bipolar cell and a cone ganglion cell; in central retina, only one cone
is connected to each cone bipolar cell (Schein 1988; Waessle et al. 1989). Beside these
simple systems, however, there are numerous cross-connections. Although rod bipolars
receive only rod input and cone bipolars have exclusively cone input (Boycott and Kolb
1973), some rod bipolars do not send their signals to rod ganglion cells, but to rod amacrine cells instead. These rod amacrine cells connect to cone bipolar cells, so that rod signals merge with cone signals (Kolb and Nelson 1983). Additionally, rods and cones are connected by gap junctions (Kolb 1977). These are anatomical bridges which allow electrical signals to pass from one photoreceptor to another, possibly enabling the photoreceptors to signal light intensity that is beyond the dynamic range of the neurotransmitter mechanism (Attwell et al. 1987). The dynamic range of the neurotransmitter release is low compared to that of the electrical transduction process, and high intensity signals may thus be transmitted to adjacent cells for neurotransmitter release.

Throughout the retina, there are two classes of bipolar cells: ON cells and OFF cells. The former are suppressed by the neurotransmitter released from the photoreceptors, so that they depolarise in the presence of light. The latter are excited by the photoreceptor transmitter and thus depolarise when light is absent, hyperpolarising when the retina is illuminated (Nelson et al. 1981). In other words, ON cells are activated by an illuminated retina and OFF cells are activated in the dark. The result is that there are two parallel pathways, one which signals when light is ON, and the other which signals when light is OFF (Werblin and Dowling 1969). Although the purpose of this replication is not yet understood, it has been suggested by Schiller et al. (1986) that if a single neurone were to signal both an increment and a decrement in luminance, it would have to have an intermediate resting discharge rate, which would be metabolically unacceptable. However, if two neurones are available, one signalling an increment in luminance and the other a decrement in luminance, the resting discharge rate for both of them can be low.

Retinal ganglion cells are also classified into ON and OFF types. However, ganglion cells have centre-surround receptive fields (Hartline 1938; Kuffler 1953). For example, a broadband ON ganglion cell responds maximally to an illuminated circle surrounded by a dim annulus, and a broadband OFF ganglion cell responds to a dim circle surrounded by a bright annulus. Neither ON nor OFF cells respond to even illumination of both centre and surround regions of their receptive fields, since the two regions are antagonistic and their separate responses cancel. Note that the centre-surround receptive field is not a product of the interaction of the separate ON and OFF channels: the centre-surround antagonism occurs entirely within one pathway or the other, while the pathways themselves remain segregated until the cortex (Schiller 1982, 1993).

The ON and OFF categories of ganglion cell are further subdivided into tonic and phasic divisions. Tonic ganglion cells respond in a sustained manner to the persistent presence or absence of light, whereas phasic cells respond transiently to an increment or decrement in luminance (Gouras and Zrenner 1981). For example, if the luminance of a small spot of light is increased abruptly, a phasic ON cell responds transiently to the
increment, whereas a tonic ON cell responds continuously after the increment has taken place (Zrenner 1983).

The centre-surround mechanism of a ganglion cell is also used to code colour (Gouras and Zrenner 1981). Not only are there tonic and phasic ON and OFF cells which possess an broadband spectrum and which therefore respond similarly to light of a variety of wavelengths, but there are also cells which respond maximally when specific colours are presented to, or withheld from, the centre and surround regions of their receptive fields. Of these colour-opponent cells, the most common is a tonic ON cell which responds maximally when a bright red spot is surrounded by an annulus from which green light is excluded (Gouras and Zrenner 1981). The corresponding OFF cell (which is less common than the ON cell) responds maximally when red light is absent from the central part of the receptive field and green light is present in the surrounding annular region.

But perhaps the most important division of retinal ganglion cells is into magnocellular (M, parasol, P_m) cells and parvocellular (P, midget, P_p) cells (Perry et al. 1984). The former account for about 10% of ganglion cells and the latter a further 80% (Schiller and Malpeli 1977). M cells have relatively large receptive fields of about 0.06° (De Monasterio and Gouras 1975) and the three types of cone project in equal proportions to both the centre and surround, so that the spectral response of the cell is broad (Schiller 1993). The M cells are generally phasic, responding transiently to changes in luminance (Schiller and Malpeli 1978). Seventy-five percent of the M cells summate light from every point in their receptive field with perfect linearity (described as an X-type response), whereas the spatial summation of the remaining 25% of the cells is non-linear (Y-type) (Blakemore and Vital-Durand 1981; Shapley et al. 1981; Kaplan and Shapley 1982).

In contrast, P ganglion cells have relatively small receptive fields of about 0.01° diameter (De Monasterio and Gouras 1975). The P pathway contains the colour-opponent ganglion cells. The most common varieties are red-centre ON and OFF cells, green-centre ON and OFF cells, and blue-centre ON cells, while blue-centre OFF cells are rare (Malpeli and Schiller 1978). The P cells are mostly tonic (De Monasterio and Gouras 1975; Schiller and Malpeli 1978), and consequently these cells are less likely to be able to follow flickering stimuli than M cells. Accordingly, Derrington and Lennie (1984) reported that the optimal frequency of stimulation was about 20Hz for M cells, but only 10Hz for P cells. The spatial summation of the P cells is linear (X-type) (Kaplan and Shapley 1982).

Summarising the functions of the magnocellular and parvocellular pathways (Schiller 1993), the M system plays a significant role in fast, low contrast motion perception. The P system, on the other hand, is important for colour vision, stereopsis, and perception of fine form and texture. Both systems detect luminance, coarse form, slow flicker, slow motion, and both are of use for scotopic vision (Schiller et al. 1990).
The axons of the retinal ganglion cells group together at the optic disk, where they exit the eye as the optic nerve. Each optic nerve contains around one million nerve fibres whose thickness ranges from 0.1 to 8.3μm, but whose average diameter is about 1.0μm (Jonas et al. 1990). Both optic nerves converge at the optic chiasm, which is located in the hypothalamic region of the brain, just ventral to the mamillary bodies. There is no exchange or extraction of information here, but the axons of the optic nerves are regrouped. The left optic tract leaves the optic chiasm containing most of the axons from the left half of both retinas, while the right optic tract exits containing most of the axons from the right half of both retinas. In other words, axons from the nasal retina cross over to the opposite hemisphere, while those originating in the temporal retina do not (figure 2). The end result is that the left hemisphere responds to a stimulus presented in the right half of the visual field, and vice versa. The left and right optic tracts pass to the left and right lateral geniculate nuclei, respectively.

Figure 2. Projection of ganglion cell axons through the optic chiasm. TR: temporal retina; NR: nasal retina. The temporal retina of the left eye and the nasal retina of the right eye project to the left lateral geniculate nucleus, while the nasal retina of the left eye and the temporal retina of the right eye project to the right lateral geniculate nucleus. Note that the left half of the retina receives light from the right half of the visual field and vice versa, so that, from the lateral geniculate nucleus onwards, the visual field is represented contralaterally. After Noback and Demarest (1981).
Lateral Geniculate Nucleus

Each lateral geniculate nucleus is a small ovoid projection from the inferior aspect of the posterior thalamus (Horton et al. 1990; figure 3). In higher primates and human, the lateral geniculate nucleus has six layers. The inferior layers (1 and 2) contain large cells and are therefore designated magnocellular, while the remaining four layers consist of small cells and are designated parvocellular (Le Gros Clark 1932; Connolly and Van Essen 1984). The magnocellular layers receive input from the M retinal ganglion cells, while the parvocellular layers receive input from the P ganglion cells (Leventhal et al. 1981; Perry and Cowey 1981). Each layer receives input from one eye only, layers 6, 4 and 1 from the contralateral eye and layers 5, 3 and 2 from the ipsilateral eye (Kaas et al. 1978). The entire lateral geniculate nucleus thus forms two

Figure 3. (a) Coronal section through the midbrain showing the location of the right lateral geniculate nucleus. (b) Section through the central region of the right lateral geniculate nucleus showing the magnocellular layers 1 and 2 and the parvocellular layers 3 - 6. After Williams and Warwick (1980).
retinotopic representations of the contralateral half of the visual field, one obtained from the left eye and the other (in alternate layers) obtained from the right eye. Inputs from both retinas do not merge in any way at the lateral geniculate nucleus (Bishop 1965; Hubel et al. 1977).

The majority of cells in the primate lateral geniculate nucleus are relay cells. Fifteen percent of the magnocellular cells and 95% of the parvocellular cells are interneurones which do not connect directly with the visual cortex (Hamari et al. 1983). Each incoming axon from the retina ends by dividing into no more than five or six terminals which connect with geniculate nerve cells in one lamina. The geniculate cells may connect with terminals from more than one retinal fibre (Minkowski 1913; Le Gros Clark 1941; Glees 1941). Colour-opponent centre-surround, colour-opponent centre-only, and broadband cells have been identified in the parvocellular layers, but colour-opponent centre-surround neurones account for the majority (about 80%) of the parvocellular cells (Wiesel and Hubel 1966; Kruger 1977; Schiller and Malpeli 1978; Creutzfeldt et al. 1979).

Summarising the functional properties of the magnocellular and parvocellular cells of the lateral geniculate nucleus (Schiller and Colby 1983): magnocellular cells lack colour-opponent receptive fields and are generally not colour selective, but respond in a transient fashion to visual stimuli, have large receptive fields, receive input from axons that conduct at high velocity, send high-velocity signals to the cortex, and have greater contrast sensitivity than parvocellular cells. About 75% of the cells have receptive fields whose spatial summation is linear (X-type), and 25% have spatially non-linear receptive fields (Y-type) (Kaplan and Shapley 1982). Most parvocellular cells have colour-opponent receptive fields, respond in a sustained fashion to visual stimuli, have smaller receptive fields, receive input from axons with medium conduction velocities, and conduct at medium velocities to the visual cortex. Around 15 - 20% of parvocellular cells lack obvious colour selectivity (Schiller and Colby 1983). Virtually all (99%) of the parvocellular cells have receptive fields whose spatial summation is linear (Kaplan and Shapley 1982).

It can be seen from the above that relay cells within the lateral geniculate nucleus have very similar properties to retinal ganglion cells. However, minor differences (Hubel and Wiesel 1961; Kaplan et al. 1987) give rise to the suggestion that the lateral geniculate nucleus is not merely a relay station, but a filter of incoming information (Kaplan et al. 1991).

Other Subcortical Nuclei

The M and P retinal ganglion cells account for about 90% of the total population of ganglion cells (Schiller and Malpeli 1977). The remaining 10% of ganglion cells have properties which are different to those of M and P cells (Leventhal et al. 1981), and
their axons pass to the superior colliculus and pretectum after projecting along the optic nerve to the optic chiasm (Schiller and Malpeli 1977). The superior colliculus receives auditory, visual and somatosensory inputs, and contains neurones that discharge before saccadic eye movements (Jacobson et al. 1981). The responses of the cells are complex, being determined by many influences from cortical and subcortical areas (Schiller 1986). It is thought that the superior colliculus performs a complex computational process leading to an action, such as an eye movement (Schiller 1986).

The pulvinar also contributes to subcortical visual processing. It makes efferent connections with occipital, parietal and temporal areas of the cortex (Schiller 1986). It is thought to be the structure where tectal and cortical pathways converge and are relayed back to the cortex. Damage to the pulvinar creates some impairment of directed attention and pattern perception (Schiller 1986).

The claustrum also has a significantly visual function. It projects to visual cortex in a precise, topographically organised fashion (Tigges and Tigges 1985), but it differs from the superior colliculus and pulvinar in that it lacks ascending connections from the retina (Olson and Graybiel 1980). It may therefore be a satellite of the neocortex rather than a relay station.

Visual Cortex

_Gross Anatomy_ Afferent fibres from the lateral geniculate nucleus pass dorsally to the occipital lobe of the cortex, spreading out so as to form the optic radiation. Macular fibres pass directly to the most occipital third of primary visual cortex. The paramacular fibres pass through the optic radiation just above and below the macular fibres to terminate in the middle third of the primary visual cortex, while the peripheral fibres pass to the superior third. The peripheral fibres exit the lateral geniculate nucleus laterally before curving to sweep posteriorly; these form the loop of Meyer.

The earliest definition of the extent of visual cortex was that of Brodmann (1909). Brodmann used a histological method of identifying the visual regions of the brain. By sectioning the brain and staining it with an agent that revealed the cell bodies, he was able to identify the distinct cytoarchitecture of each part of the cortex. It was postulated that differences in cytoarchitecture between adjacent regions of the cortex delineated boundaries between functional areas, a hypothesis which is still held to be true (Zeki 1993). The different functional areas were identified by (1) differences in cell packing densities, (2) differences in cell sizes, and (3) differences in the thickness of the cortex and the laminae of which it was comprised (Brodmann 1914). Using this approach, Brodmann was able to identify three visual areas, which, in his overall classification scheme, were denoted areas 17, 18 and 19. These areas are still widely used.
Area 17, primary visual cortex, lies on the superior and inferior surfaces of the calcarine fissure, extending, by an amount which varies between individuals, onto the external occipital convexity (Brindley 1972). Area 17 is surrounded by area 18, which forms a strip of 5 - 10mm in width, passing posteriorly along the cuneal gyrus, around the occipital convexity, and then back along the lingual gyrus. Area 19, in turn, surrounds area 18, and includes the dorsal, lateral and ventral surfaces. All three visual areas are delimited superiorly by the parieto-occipital sulcus (figures 4a and 4b).

**Primary Visual Cortex** Area 17 is said to be retinotopic: each point in the visual field is represented by a distinct point in the cortex (Holmes 1945; Spalding 1952; Teuber et al. 1960). The left half of the visual field is represented in the right hemisphere, and vice versa. The upper half-field is represented below the calcarine sulcus and the lower half-field is represented above it. The fovea is represented by a point on the occipital pole, with the horizontal meridian passing along the innermost fold of the calcarine sulcus and the vertical meridian lying along the edge of area 17. In accord with this, the macular field is represented inferiorly and the peripheral field is represented superiorly, adjacent to the parieto-occipital sulcus (Drasdo 1980).

A much larger portion of primary visual cortex serves the foveal region of the retina than the periphery, which is necessary to achieve the high visual acuity of the fovea. This feature gives rise to the concept of cortical magnification, $M$, which is defined simply as "the linear extent of cortex concerned with each degree of visual field" (Daniel and Whitteridge 1961). Assuming that cortical magnification is proportional to the square root of retinal ganglion cell density, which has been demonstrated in monkeys (Daniel and Whitteridge 1961), then the reciprocal of the cortical magnification factor can be expressed as:

$$
M^{-1} = M_0^{-1} \left[ 1 + S \left( 1 + 3 \times 10^{-5} \theta^2 \right) \right], \quad (1)
$$

where $M_0$ is the foveal value of $M$, $\theta \ (<70^\circ)$ is the angular retinal eccentricity, and $S$ is a geometric factor which depends upon the meridian containing $\theta$ (Drasdo 1977).

Estimates of foveal cortical magnification in human vary from 7.99mm deg$^{-1}$ (Rovamo and Virsu 1979) to 20 - 25mm deg$^{-1}$ (Tolhurst and Ling 1988), with Drasdo (1991) quoting 14.4mm deg$^{-1}$. Setting $S = 0.59$ provides a general relation for all meridians. Moreover, if $\theta < 30^\circ$, equation 1 can be truncated to the linear term only (Drasdo 1977):

$$
M^{-1} = M_0^{-1} (1 + 0.59\theta). \quad (2)
$$

An integration yields an expression for the fraction, $Q$, of the total cortical area associated with the central $\theta$ ($<30$) degrees of the visual field (Drasdo 1977):

$$
Q = 1 - \exp(-0.0574\theta). \quad (3)
$$
Figure 4. (a) (left-hand column) Lateral (top), medial (centre), and ventral (bottom) aspects of the left occipital lobe, showing the major sulci and gyri. Sulci: stis: superior temporal sulcus; mts: middle temporal sulcus; pos: parieto-occipital sulcus; sIs: superior lateral sulcus; ils: inferior lateral sulcus; cs: calcarine sulcus; cols: collateral sulcus; its: inferior temporal sulcus. Gyri: CG: cuneal gyrus; LG: lingual gyrus, FG: fusiform gyrus. (b) (middle column) Brodmann's visual areas shown on the corresponding views. (c) (right-hand column) Functional visual areas shown on the corresponding views. After Clarke and Miklossy (1990).
From this, it can be seen that the majority of primary visual cortex is devoted to central vision.

Several approximate models of the retinotopic representation of the visual field on primary visual cortex have been proposed (Jeffreys and Axford 1972a; Drasdo 1980); the most commonly used is the cruciform model (Okada 1983b). If the visual field is considered as being comprised of eight octants, the four octants of each half-field are represented, respectively, on the cuneal gyrus, superior surface of the calcarine sulcus, inferior surface of the calcarine sulcus, and the lingual gyrus (figure 5). Note, however, that the cruciform model only describes the retinotopy of area 17, it does not predict the behaviour of area 18 or 19. Thus, if a visual stimulus evokes a cortical response in area 18 or 19, the location of this response may not be that predicted by the cruciform model. Also note that the central 2 - 3° of the visual field are represented on the occipital convexity (Drasdo 1980), and therefore the cruciform model is only reliable for visual stimuli which extend to eccentricities of significantly greater than 2 - 3°.

Since the cruciform model is only intended as an approximation to the layout of primary visual cortex, it is assumed that the calcarine fissures are straight, and that they project radially into the head, which is manifestly not true (Steinmetz et al. 1989). However, if the morphology of area 17 is known, it is possible to construct a more realistic model. This is performed in chapter five of this thesis for the case of a statistically average brain. Accordingly, evidence relating to the inter-individual variation of primary visual cortex is now reviewed, with specific reference to the calcarine sulcus, since this is the major internal landmark around which primary visual cortex is based.

The earliest and most comprehensive statistical survey of occipital lobe morphology was that of Talairach et al. (1967). They sectioned thirty hemispheres and recorded the positions of the major cerebral landmarks relative to the anterior and posterior commissures using stereotactic methods. Inter-individual variations in head size were taken into account by fractionating the brains, i.e. by applying three separate scalefactors to the three dimensions to reduce the brain to the size most commonly encountered (the standard size). Using this method, the calcarine sulcus was shown to begin at the posterior edge of the occipital lobe and extend anteriorly for about 55mm in a direction parallel to the line joining the anterior and posterior commissures. In general, the central portion of the fissure rose slightly. The majority of the inter-individual variation lay not in the orientation of the fissure, but in the superior-inferior dispersion: the most superior and most inferior fissures lay some 20mm apart, with the other fissures lying evenly between them (figure 6a). However, when the most posterior point of the occipital fossa was aligned in each individual brain, thereby aligning the occipital lobe of each brain, the dispersion was found to diminish to about 10mm (figure 6b). This indicated that the dispersion was as much due to the occipital lobe varying relative to the commissures as due to the calcarine sulcus varying relative to the occipital lobe. The anterior and posterior commissures were shown to bear a
Figure 5. The cruciform model of primary visual cortex, used to interpret visual evoked electric and magnetic responses. (a) The visual field considered as eight octants. (b) Medial aspects of the left and right visual cortices. (c) A schematic coronal section through primary visual cortex, showing the cruciform shape adopted by the cuneal and lingual gyri and the calcarine sulci. The four octants of one hemifield are projected onto the contralateral medial cortex, inverted in the vertical direction. The two octants bordering the horizontal meridian fall within the calcarine sulcus. The central visual field is thus represented at the occipital pole and the peripheral field is represented in the most superior cortex, close to the parieto-occipital sulcus. LVF: left visual field; RVF: right visual field; LH: left hemisphere; RH: right hemisphere; CG: cuneal gyrus; LG: lingual gyrus; PO: parieto-occipital sulcus; C: calcarine sulcus. After Okada (1983b).
Figure 6. (a) Inter-individual variation of the calcarine fissure, relative to the line CA-CP joining the anterior and posterior commissures, and the perpendiculurs VCA and VCP which pass through the anterior and posterior commissures respectively. (b) Inter-individual variation of the calcarine fissure when the most posterior point of the occipital fossa is aligned in each subject. Note that the superior-inferior dispersion is reduced. Note also that the majority of the thirty fissures have approximately the same shape and orientation, apart from the superior-inferior dispersion. From Talairach et al. (1967).
predictable relationship to the internal bony landmarks of the skull, so that the brain's variation with respect to the commissures could also be thought of as a variation relative to the internal surface of the skull.

Steinmetz et al. (1989) performed a similar procedure using magnetic resonance imaging. They superimposed 32 sagittal images (32 hemispheres) so that the inions coincided and the nasion-inion lines were collinear. No scaling was applied. The results showed that the anterior end of the calcarine sulcus was more superior to the nasion-inion line than the posterior end, so that the fissure generally made an angle of about 10° with the nasion-inion line. A large degree of superior-inferior dispersion - between 30 and 40mm - was also recorded. When the images were scaled so that the nasion, inion, vertex and pre-auricular points coincided, no change in the morphology of the calcarine sulcus, or reduction in its dispersion, was observed. The large dispersion compared to that reported by Talairach et al. (1967) was attributed to the variability of the inion location relative to the underlying visual cortex: not only was the calcarine fissure varying in position, but the inion, and thus the reference axis, was also varying.

The study of Myslobodsky et al. (1991), also using magnetic resonance imaging, produced similar results. Sagittal images of 28 hemispheres were scaled so that the length of the nasion-inion line measured 170mm. The nasion-inion lines were then superimposed and the perpendicular from the nasion-inion line through the confluence of the retrocalcarine and parieto-occipital sulci was also aligned. The distance of the posterior end of the calcarine sulcus from the inion, measured along the scalp perimeter, varied from 20 to 40mm. The central section of the calcarine sulcus was shown to rise slightly, as found by Talairach et al. (1967). The mean vertical distance of the fissure above the inion was found to be 38 ± 8(SD) mm for the left hemisphere, and 39 ± 12 mm for the right hemisphere. The left fissure rose at a mean angle of 22 ± 9 deg to the nasion-inion line, and the right fissure rose at an angle of 17 ± 4 deg. This study therefore showed that, while there were interhemispheric differences in the course of the calcarine sulcus, these differences were insignificant.

**Histology of Primary Visual Cortex** Area 17, in common with the entire cerebral cortex, is about 2mm thick and divided into six laminae (figure 7). Two stratification schemes were introduced at around the turn of the century: the first, proposed by Ramón y Cajal (1899; DeFelipe and Jones 1988), was based on the Golgi method of staining, while the second, due to Brodmann (1904), used Nissl staining. The latter has been widely adopted and is used in the following description, which relates primarily to the monkey.

Brodmann (1905) divided the primary visual cortex of the monkey into eight alternate light and dark layers, as follows:
I. Lamina zonalis
II. Lamina granularis externa
III. Lamina pyramidalis
   IV. {IVa. Lamina granularis interna superficialis
       IVb. Lamina granularis interna intermedia
       IVc. Lamina granularis interna profunda
   V. Lamina ganglionaris
       VI. {VIa. Lamina triangularis
            VIb. Lamina fusiformis
   Lamina multiformis.

Layer IV was classed as one subdivided layer because the three constituent layers derived from a single, primitive, internal granular layer. Layer IVc is now also subdivided into two sublayers, superficial IVcα and deeper IVcβ (Lund 1973).

Incoming geniculocortical fibres terminate mainly in layer IV. Projections from the magnocellular layers of the lateral geniculate nucleus distribute mainly to layer IVcα, while fibres from the parvocellular laminae terminate in layers IVa and IVcβ. Both systems also project sparsely to layer VI (Hubel and Wiesel 1972; Lund and Boothe 1975; Lund 1981). In Nissl preparations, layers IVa and IVc appear as two dense bands

---

Figure 7. Sections through striate cortex after staining for whole cells (Golgi staining), cell bodies (Nissl staining), and myelin (Weigert staining). Layer 1 is closest to the pial surface and layer 6 is adjacent to the white matter. After Brodmann (1909).
of tightly packed cell bodies, separated by the pale layer IVb. Layer IVb, the stria of
Gennari, from which the name "striate cortex" is derived, can be seen with the naked eye
in an unstained section of occipital cortex as a whitish band, and since this band is
unique to area 17, it provides a useful marker of primary visual cortex. Layer IV
contains almost one half of the total number of neurones in area 17 (O’Kusky and
Colonnier 1982). Layers IVa and IVc contain neurones which do not project outside of
area 17: all of them have intrinsic axons which connect to other cortical levels. These
neurones thus represent the first step in the processing of afferent cortical input
(Valverde 1985).

Layer IVb is characterised by the existence of numerous horizontally-arranged,
myelinated axons of three main types: horizontal collaterals of descending axons of
pyramidal and stellate cells of layer IVa, horizontal branches of axons of neurones
located within layer IVb, and ascending ramifications of spiny stellate cells residing in
layer IVcα (Valverde 1985). Layer IVb contains large pyramidal and multipolar cell
bodies, as well as other neuronal cell bodies of smaller size. The largest components
preferentially occupy the centre of the layer, where they are arranged in a horizontal row
(Valverde 1985). The somata of the large cells in layer IVb are ovoid, triangular, or
pyramidal in shape, and have, in the case of the pyramidal cells, a well-developed apical
dendrite ascending towards layer I. The dendrites of the stellate and pyramidal cells also
run horizontally, contributing to the plexus of fibres which is characteristic of this layer.
Medium-sized pyramidal cells have apical dendrites which reach layer I in most cases,
but the smaller pyramidal cells extend only to layer II or III. The apical dendrites of the
pyramidal cells have collateral branches which are distributed in the lower part of layer
III and also in layer IVa (Valverde 1985). The descending axons of the cells in layer
IVb probably project to layer VI (Lund and Boothe 1975; Lund 1981) and to higher
visual areas (Spatz 1977; Rockland and Pandya 1981). Hence, layer IVb seems to
horizontally distribute the geniculate input relayed through layers IVa and IVc. In
addition, the information collected in layer IVb is relayed to layer III, probably to layer
VI, and also to the association cortex (Valverde 1985).

The upper layers of striate cortex, layers I, II and III, provide important
connections between area 17 and the surrounding association cortex. Descending axons
of pyramidal cells found in these layers enter the white matter and project to areas 18
and 19 (Spatz et al. 1970; Lund et al. 1975; Tigges and Tigges 1982). Area 18 also
sends projections back to area 17, where the fibres terminate mainly in layer I, but also
in all other layers except layer IV (Tigges et al. 1973, 1974, 1977; Wong-Riley 1978).
In the region of area 17 bordering area 18, there are also callosal connections (Zeki
1969b; Spatz et al. 1970; Van Essen and Zeki 1978). In addition to these ipsilateral and
callosal association connections, layer I also receives input from several subcortical
nuclei, such as the pulvinar and the claustrum (Ogren and Hendrickson 1977;
Carey et al. 1979, 1980).
Layer I of striate cortex contains horizontal fibres running for long distances, the terminal axonal arboursisations of some intrinsic neurones located in layer II, and the dendritic trees of pyramidal cells found in layers II, III, IVb and V. Layers II and III contain pyramidal cells of various sizes, together with several types of intrinsic neurones. The pyramidal cells have basal dendrites which extend horizontally, and axons which descend vertically, giving off collaterals into layers II and III. The axons of pyramidal cells in layer III also give off collaterals into layer V (Lund and Boothe 1975). This seems to represent an interlaminar association circuit that carries information from the basal dendrites of layer III to layer V (Valverde 1985).

Below layer IV, the laminar organisation changes considerably. Layers V and VI contain mainly pyramidal cells, with only a small proportion of stellate cells. The medium and large pyramidal cells are output cells: their axons pass out into the white matter, becoming myelinated after the initial section. The pyramidal cells in layer V project to the superior colliculus and the inferior pulvinar nucleus. In addition, pyramidal cells in the deeper part of layer VI project to the magnocellular laminae of the dorsal lateral geniculate nucleus, while those in the upper part project to the parvocellular layers (Lund et al. 1975). The pyramidal cells have apical dendrites which end at differing levels (Lund and Boothe 1975; Lund 1981). Some of these dendrites have collaterals which spread out within the upper part of layer V, others terminate without branching or produce a few terminal branches in layers II and III, and another class terminates with a spray of dendrites in layer IVb or IVc.

The largest pyramidal cells are found in layer V, at the boundary between this and layer VI. These giant pyramidal cells, termed solitary cells of Meynert (Meynert 1867), have stout apical dendrites that ascend to layer I, and long basal dendrites that form a horizontal plexus. The apical dendrite is either unbranched or has a few collaterals in layer V, and at layer I, it ends either without branching or by giving off several slender terminal branches. Neither the apical dendrite nor the basal dendrites contain many dendritic spines. The axons project to temporal association cortex (Spatz 1975). The smaller pyramidal cells have ascending axons which project to layers I, II, III and IVa. Layer V contains a horizontal plexus, involving intrinsic fibres, but also containing the collaterals of layer V pyramidal cells and the collaterals of stellate cell axons from layers IVa and IVc. The plexus also contains the terminal portions of pyramidal cell axons projecting from layers II and III (Valverde 1985).

Incoming fibres entering area 17 also form collateral branches within layers V and VI. These collaterals contain information from the lateral geniculate nucleus (Hubel and Wiesel 1972), from the inferior pulvinar nucleus (Benevento et al. 1975; Ogren and Hendrickson 1977), from the claustrum (Carey et al. 1979), from areas 18 and 19 (Tigges et al. 1974), and from temporal association cortex (Spatz 1977; Rockland and Pandya 1981). Meanwhile, layers V and VI project back to these areas, as described above. To summarise, the pyramidal cells of layers V and VI are influenced by the incoming signals from the subcortical nuclei and the association areas. They also collect
information from the horizontal plexus, which, in turn, receives input from pyramidal cells in layers II and III. Furthermore, the apical dendrites of the pyramidal cells gather information directly from layers I, III and IVb. Thus, the pyramidal cells of layers V and VI "are in the best position to synthesise all information processed in area 17 into a response to a visual stimulus" (Valverde 1985).

*Physiology of Striate Cortex* Cells in layer IVc of area 17, where the geniculo-cortical signals arrive, respond to stimulation of one eye only and have circularly symmetric receptive fields (Hubel and Wiesel 1968). Beyond this layer, however, the visual information is reorganised, so that cells in the cortical layers above and below IVc receive input from both eyes (Arden and Sheorey 1977), have elongated receptive fields (Hubel and Wiesel 1977), and are tuned to orientation, so that they respond maximally to a moving bar of some preferred orientation, but only weakly to a bar moving at right angles to the preferred orientation. Such cells are classified as simple, complex, or hypercomplex (Hubel and Wiesel 1962, 1968).

The receptive field of a simple cell is divided into two adjacent regions, one excitatory and the other inhibitory (Hubel and Wiesel 1962; Pollen and Ronner 1981; Field and Tolhurst 1986). The most effective stimuli are long, narrow bars of light, edges between areas of different luminance, and dark rectangular bars against a light background; these stimuli must be aligned with the orientation of the two regions of the receptive field for maximum response. A moving bar is also an effective stimulus, provided the bar is parallel to the receptive field (Hubel and Wiesel 1968). The response of a simple cell is approximately a linear sum of the contributions from the different parts of the receptive field, so that the response to a given extended pattern can be predicted by mapping the receptive field using a small spot stimulus (Henry 1985). Thus, simple cells exhibit characteristics similar to spatially linear (X-type) lateral geniculate cells (DeValois et al. 1982b).

The receptive field of a complex cell contains many small excitatory and inhibitory regions which are arranged alternately in space (Hubel and Wiesel 1968). Complex cells respond only weakly to a stationary stimulus, but may respond to both increments or decrements in stimulus intensity. Many complex cells are direction-specific as well as orientation-specific: their response to a correctly oriented bar moving in one direction is much stronger than the response to movement in the opposite direction (Hubel and Wiesel 1968).

Simple and complex cells can be conveniently distinguished by applying counterphase-modulated gratings at different positions and orientations (DeValois et al. 1982b). A simple cell responds maximally when a bright bar is centred on the excitatory portion of its receptive field and a dim bar is centred on the inhibitory portion. When the grating is reversed, so that the excitatory portion of the receptive field is dark, and the inhibitory portion is illuminated, the cell does not respond. Thus, a simple cell acts as a half-wave rectifier. A simple cell responds strongly to a grating of a specific
orientation, but does not respond to a grating which is perpendicular to this preferred orientation since the excitatory and inhibitory portions of the receptive field are both illuminated equally. The absence of a response in this latter case demonstrates linearity of spatial summation: both excitatory and inhibitory regions carry equal weight and therefore cancel each other exactly when equally illuminated (DeValois et al. 1982b).

With a complex cell, on the other hand, the predominant response to a counterphase-modulated grating is at twice the modulation frequency, i.e. the response is at the reversal frequency. The complex cell thus acts more like a full-wave rectifier than a half-wave rectifier, as it responds to both ON and OFF stimuli. There is no orientation of a counterphase-modulated grating that produces a complete cancellation of excitation and inhibition so as to give a zero response. This indicates that the spatial summation is non-linear (DeValois et al. 1982b). A complex cell also has a higher resting discharge rate (DeValois et al. 1982b).

Not all cells are inhibited by a grating which is at right angles to the preferred orientation. On the contrary, some neurones are weakly excited by a grating at right angles to the preferred orientation (DeValois et al. 1982a). The degree of orientation selectivity also varies considerably: half-amplitude orientation bandwidths vary from 8° to more than 100° along a continuum, the median bandwidth being about 45° (Poggio et al. 1977; DeValois et al. 1982a). The majority of simple and complex cells in the striate cortex of the monkey are selectively sensitive to spatial frequency, the frequency of maximum sensitivity varying between 0.5 and 8 cycles/deg in a range of cells (Foster et al. 1985). The half-amplitude bandwidth is usually around 1.5 octaves, but can be anywhere between 0.8 and 3.0 octaves (Foster et al. 1985). In general, complex cells are slightly more broadly tuned than simple cells (DeValois et al. 1982b). About two-thirds of neurones in striate cortex have low-pass temporal frequency characteristics (Foster et al. 1985). The remaining third are very broadly tuned with a mean bandwidth of 2.9 octaves.

An interesting feature of orientation tuning is that two simple cells of different preferred orientations can mutually inhibit each other when simultaneously excited (Bishop et al. 1973; Creutzfeldt et al. 1974; Sillito 1975; Sillito et al. 1980). This may explain why cells which respond strongly to a one-dimensional pattern can fail to respond to a pattern involving two orientations, even when their preferred orientation is optically contained within the mixture (Hammond and MacKay 1975, 1977). Complex cells, on the other hand, respond to patterns which involve two different orientations (Hammond and MacKay 1975, 1977). It has been suggested (Burr et al. 1981) that the inhibition of simple cells by complex cells could create simple cells which are specifically sensitive to one-dimensional features such as edges.

Hubel and Wiesel (1965) were the first to notice a third type of neurone in the cat, which was described as hypercomplex. This type of neurone responded strongly to short, optimally-oriented bars or edges, but extending these stimuli beyond some optimal length suppressed the response. This end-stopped inhibition indicated that the receptive
field of the cell contained inhibitory regions beyond the main excitatory region. Schiller et al. (1976) and Rose (1977) reported similar findings in the monkey and cat respectively. Kato et al. (1978) showed that 46% of simple cells and 47% of complex cells exhibited end-stopped inhibition, giving rise to two classes of hypercomplex cell, having, respectively, simple and complex properties. By plotting the response of various cells against the stimulus length, it was found that the response of a conventional simple or complex cell increased steadily with increasing stimulus length, until a plateau was reached. Further increase in stimulus length had no effect on the firing rate of the cell. With a hypercomplex cell of either the simple or complex variety, the response increased dramatically as the bar length was increased, but when the bar length exceeded the optimal length, which was typically 1 - 2° of arc depending on the cell chosen, the response dropped again, levelling out at a plateau. The strength of the response at the plateau was typically between 40 and 100% of the maximum response, having a mean of about 70% for both simple and complex varieties.

Simple, complex and hypercomplex cells are non-uniformly distributed throughout the layers of visual cortex. Hubel and Wiesel (1968) recorded that simple cells were found in layers III, IVa and IVb. In layer IVb, they formed the majority of the neural population, while complex cells were rare. In layer IV, the cells were driven by one eye only. Complex and hypercomplex cells were found in layer II, the upper two-thirds of layer III, and in layers V and VI. Simple cells were virtually absent in these layers, and most cells were binocularly driven. Kato et al. (1978) found that hypercomplex cells resided in all laminae of cat striate cortex, but were mainly present in laminae II, III and V. In laminae II and III, the simple variety of hypercomplex cell was most common, whereas in lamina V, hypercomplex cells were equally divided between the simple and complex families.

The laminar distribution of the different cell types reflects the pattern of synaptic connections (Ferster and Lindström 1983). Incoming geniculate fibres provide input to both simple and complex cells in layer IV (Movshon et al. 1978; Malpeli et al. 1981). These project to layers II and III, which, in turn, project to layer V. Layers II, III and V are the layers where the complex and hypercomplex cells are abundant. Layer VI, which projects back to the lateral geniculate nucleus, contains complex cells and is also the only layer, apart from layer IV, to contain simple cells. These simple cells have greatly elongated receptive fields (Ferster and Lindström 1983; Gilbert and Wiesel 1983).

As well as possessing a laminar organisation, visual cortex has a horizontal organisation. Hubel and Wiesel (1962) noted that cat striate cortex was divided into columns, such that all cells within each column had the same receptive field orientation. The columns extended from the pia to the white matter, and cross-sectional diameters at the surface were around 0.5mm. Although each column contained a variety of simple and complex cells, their receptive fields were located in approximately the same retinal region and overlapped, although the exact position varied slightly. The relative influence of the two eyes was not necessarily the same for all cells in a column. It was
subsequently discovered from multiple parallel microelectrode penetrations (Hubel and Wiesel 1963, 1974) that the columns were parallel to the radial fibre bundles and perpendicular to the cortical layers. The shapes of the columns were found to be highly irregular: some columns were more or less round, while others were long and narrow. In some parts of striate cortex, the columns were arranged in a regular manner, with the receptive field orientation shifting discretely in one direction or the other over long sequences as the microelectrode was moved along the cortex. In other parts, there appeared to be little order in the arrangement of neighbouring columns.

Hubel and Wiesel (1968) also discovered that the cortex was divided into columns according to eye preference. These ocular dominance columns were larger than the orientation columns and the two sets of boundaries were independent. Cells in one column responded preferentially to stimulation of the left eye, while those in an adjacent column responded preferentially to stimulation of the right eye. A pair of left and right ocular dominance columns having the same receptive fields were designated a hypercolumn. (Similarly, a complete set of orientation columns, having the same ocular dominance and receptive field location, but with a complete range of receptive field orientations, was also referred to as a hypercolumn.)

Ocular dominance columns have also been observed using other techniques. By making a small lesion in one layer of the lateral geniculate nucleus, Hubel and Wiesel (1972) were able to induce a pattern of degenerating axonal endings in layer IV of striate cortex; these endings represented input from one eye only. The result was a characteristic pattern of stripes corresponding to alternate ocular dominance columns. Similarly, Hubel and Wiesel (1977) and Shatz et al. (1977) injected one eye with radioactive fucose or proline and stained a tangential section of striate cortex with photoemulsion which responded to radioactive disintegrations, a process known as autoradiography. The terminals of the geniculate neurones supplied by the injected eye were thus identified. By using carbon-14 deoxyglucose autoradiography, the rate of local glucose utilisation was observed by Kennedy et al. (1976) and Hubel et al. (1978).

When one eye was selectively deprived or stimulated with moving vertical stripes, vertical bands of label were found, extending from the pia to the white matter, although the bands of high activity appeared to be most intense in layer IVc, the principal site of termination of geniculocortical fibres. The bands coincided with the ocular dominance and orientation columns.

The columnar organisation of striate cortex is reflected in the pattern of interlaminar connections (Valverde 1985). A tangential Golgi section of the striate area in macaque shows the existence of vertical bundles of axons aligned in parallel rows, separated by zones in which cells appear more abundant. A pattern of parallel stripes can also be observed in a Nissl-stained tangential section. The bands in lower layer III are one-half or less of the width of the ocular dominance bands in sublayer IVc. Thus, a single band in layer IVc appears to give rise to two bands in layer III, the bands themselves being formed by the ascending axons of the stellate cells found in layer IVc.
(Valverde 1985). Hence, the columnar organisation of striate cortex is mediated by a columnar cytoarchitecture.

**Functional Areas of Visual Cortex** Although Brodmann's classification is still widely used, it is now known that there are at least 32 functionally distinct visual areas in the brain of the macaque monkey (Van Essen 1985). These areas occupy all of the occipital lobe, the posterior half of the parietal lobe, the ventral half of the temporal lobe, and even part of the frontal lobe (Felleman and Van Essen 1991). Features used for the delineation of the visual areas have included, in descending order of importance: (1) a specific pattern of connections as revealed by various pathway-tracing techniques; (2) a distinctive topographic map of the contralateral half of the visual field; (3) a distinctive cellular architecture; and (4) a well-defined behavioural deficit after localised lesions. The functions of some of the visual areas are now discussed, firstly for the case of monkey, and then in human.

Area V1, primary visual cortex, possesses a distinct cytoarchitecture and a single, extensive topographic map of the visual field, and therefore coincides exactly with striate cortex, or area 17 as defined by Brodmann (see, for example, Van Essen et al. 1984). The previous description of area 17 therefore applies directly to area V1. The function of V1 is most readily observed using cytochrome oxidase histochemistry, which was first applied to the visual cortex by Wong-Riley (1979). Cytochrome oxidase is a metabolic enzyme and it therefore accumulates at areas of high metabolic activity. A reagent designed to respond to cytochrome oxidase darkly stains layer IVc, lightly stains layer IVb, and gives rise to a row of darkly staining patches within layer IVa. More significantly, however, the method creates a series of darkly stained columns extending from the surface to the white matter, intersecting the tangential pattern at right angles. These columns are most evident in layers II and III, where, in section, they appear as darkly stained "blobs" of diameter around 300μm (Horton 1984; Livingstone and Hubel 1984). The more lightly-staining regions between the blobs are termed "interblobs". The blobs are arranged in parallel rows about 0.5mm apart and are precisely in register with the ocular dominance columns (Livingstone and Hubel 1984).

Unlike the cells found in layers II and III as a whole, the cells found in the blobs lack orientation selectivity (Livingstone and Hubel 1984). About half of the cells are wavelength selective (Kolata 1982); many have double-opponent concentric receptive fields which are different from those of the lateral geniculate cells (Livingstone and Hubel 1984). Such double-opponent cells might respond, for example, to presence of red light and absence of green light in a circular central excitatory zone, and absence of red light and presence of green light in the inhibitory surround zone. This type of receptive field is especially suited to extracting information about colour. Cells which are not wavelength-selective are broadband (Livingstone and Hubel 1984). In contrast, the interblob cells are orientation selective (Livingstone and Hubel 1984) and are capable of responding to contours in wavelength or luminance (Gouras and Krüger
1979; Thorell et al. 1984). Hence, the contrasting functional properties of the cells in the blobs apparently form an extra dimension to the ocular dominance and orientation segregations discovered earlier (Zeki 1993). Each functional unit, containing an ocular dominance column, an orientation hypercolumn and a series of blobs, describes the attributes of one small section of the visual field as presented to one eye. The entire extent of V1 can thus be considered to contain multiple, interleaved maps of the visual field, each map describing a different feature. V1 can therefore be thought of as segregating the different attributes of the stimulus and directing the different types of information to the various specialised areas (Zeki 1993).

Area V2 is a narrow functional area surrounding V1 (see, for example, Van Essen 1985). In the macaque, area V2 accounts for less than half of area 18 (Van Essen 1985). It receives topographically organised input from V1, and its retinotopic representation of the visual field is a mirror image of that in V1 (Zeki 1969a; Allman and Kaas 1971, 1974). That is, the vertical meridian is located along the border with V1, while the contralateral horizontal meridian lies along the external border. Note that the neurones on the superior external border have receptive fields which are adjacent to those of the neurones on the inferior external border: there is a cortical discontinuity along the contralateral horizontal meridian (Allman and Kaas 1974; Gattass et al. 1981). In primates, the second visual area has a somewhat disorderly retinotopic mapping (Van Essen et al. 1979; Gattass et al. 1981; Tootell et al. 1983). Area V2 contains orientation-selective, direction-selective and wavelength-selective cells (Baizer et al. 1977; Zeki 1978c), and it can therefore be inferred that it acts as a non-specialised segregator of stimulus features, like V1 (Zeki 1993). If area V2 is stained to reveal the cytochrome oxidase distribution, a number of dark "stripes" are revealed, running from the cortical surface to the white matter and separated from each other by lightly staining "interstripes". In a tangential section, alternating "thick" and "thin" stipes can be distinguished, so that a cyclic pattern of thick stripe, interstripe, thin stipe, interstripe, is encountered (Livingstone and Hubel 1982; Tootell et al. 1983). In a macaque monkey, the cycle has a period of about 5mm, and the architecture is unique to V2.

Area V3 also lies within area 18 in macaque, and receives a direct input from V1, with the consequence that the visual field is also topographically mapped (Crass 1969; Zeki 1969a, 1978b). The retinotopic representation within V3 is a mirror image of that in V2, so that the horizontal meridian is represented between V2 and V3 and the vertical meridian is represented at the anterior border of V3. The area is generally subdivided into a dorsal region, dV3, and a ventral region, VP, and it has an associated neighbour, V3A. The majority of cells in V3 (and V3A) are responsive to lines of specific orientation (Zeki 1978b), but only 9% of cells in dV3 and 23% of cells in VP are colour-opponent wavelength selective (Felleman et al. 1984; Van Essen 1985). In other words, most cells in V3 (or V3A) respond to a line of a given orientation, regardless of the colour of the line, or the colour of the background against which it is presented. From
these properties, it can be deduced that area V3 is involved in the processing of signals related to the form of a visual stimulus (Zeki and Shipp 1988; Zeki 1993).

On the lateral side of the macaque brain, area V3 is bordered by area V4 (Zeki 1971). V4 receives an input from the foveal area of V1, but its input is mainly from V2 (Zeki 1978a; Zeki and Shipp 1989). The majority of cells in V4 are colour-selective—they respond better to some wavelengths than others (Zeki 1973, 1977). Some of the cells have the property of colour constancy, whereby the colour of a surface is unaffected by the spectral composition of the prevailing illumination. This requires that information is gathered from a large part of the visual field, i.e. it requires that the receptive field is large (Zeki 1983a). Area V4 is evidently involved in the analysis of colour (Zeki 1977), but there are also orientation-selective cells within V4 (Zeki 1983b; Desimone and Schein 1987), possibly forming a subcompartment within it (Zeki 1983b; Zeki and Shipp 1989). Many of the orientation-selective cells are also wavelength-selective to a varying degree (Zeki 1983b; Desimone and Schein 1987; Schein and Desimone 1990). Area V4 is therefore considered to be responsible for analysing colour in association with form (Zeki and Shipp 1988).

In macaque, area V5 lies adjacent to V4, and anterior to it (Zeki 1974). Area V5 is buried in the superior temporal sulcus. It is rich in cytochrome oxidase and heavily myelinated (Van Essen et al. 1981; Jen and Zeki 1983). Area V5 receives input directly from V1, as well as from V2. The cells of V5 are selectively responsive to different kinds of motion (Komatsu and Wurtz 1988; Wurtz et al. 1991), and over 90% are directionally selective (Zeki 1974). The directionally selective cells respond to a visual stimulus moving in one direction, but not in the opposite direction. Most cells respond to spots of light moving in one direction, although some respond optimally to oriented lines. None of the cells is influenced by the colour of the stimulus or of its background (Zeki 1974). Evidently, V5 is involved with motion (Zeki and Shipp 1988).

The pattern of neural connections linking the visual areas is extremely complex (see, for example, Felleman and Van Essen 1991), but several simple trends emerge. As described earlier, the magnocellular stream from the lateral geniculate nucleus terminates in layer IVcα of area V1, and from there, it passes to layer IVb (Lund and Boothe 1975), where the cells are orientation and direction selective (Hubel and Wiesel 1968; Dow 1974). The magnocellular output from V1 leads to the cytochrome oxidase thick stripes of V2 (Livingstone and Hubel 1987), which are characterised by a concentration of directionally selective cells and an absence of wavelength selective cells (DeYoe and Van Essen 1985; Hubel and Livingstone 1985; Shipp and Zeki 1985). The thick stripes, in turn, project to areas V3 and V5 (DeYoe and Van Essen 1985; Shipp and Zeki 1985), most cells in V3 being orientation selective, and almost all cells in V5 being direction selective (Zeki 1974), but neither area being particularly selective for wavelength. The magnocellular output from V1 also leads directly to V3 and V5 without passing through V2 (Lund et al. 1975; Burkhalter et al. 1986). The outputs of V3 and V5 are directed
mainly to posterior parietal cortex in monkey (Maunsell and Van Essen 1983; Van Essen and Maunsell 1983).

In contrast, the parvocellular input to V1 terminates in layers IVa and IVcβ (Livingstone and Hubel 1984). The cytochrome oxidase blobs and interblobs also make an important contribution to the parvocellular stream. Cells in the blobs lack orientation selectivity but are wavelength selective, whereas cells in the interblobs are both orientation and wavelength selective (Livingstone and Hubel 1984; Thorell et al. 1984). The relative contributions of the blobs and interblobs depend upon the sizes or spatial frequencies of the patterns to which they respond: at low spatial frequencies, the blobs are most sensitive, and the subsequent parvocellular pathway contains primarily wavelength information, but at high spatial frequencies, the interblobs are more active, providing luminance information (Tootell et al. 1983). Thus, in the upper layers of V1, the parvocellular system divides into two parallel pathways, one primarily concerned with colour and another primarily concerned with form. These two pathways are maintained beyond V1, the blobs leading to the thin stripes of V2 and the interblobs leading to the interstripes (Livingstone and Hubel 1984; DeYoe and Van Essen 1985; Shipp and Zeki 1985). The thin stripes and interstripes maintain the functional specialisations found in the blobs and interblobs (DeYoe and Van Essen 1985; Shipp and Zeki 1985). Both of the parvocellular streams project heavily to area V4 (DeYoe and Van Essen 1985; Shipp and Zeki 1985), and it is possible that the two streams remain distinct within V4 (Zeki 1983b). Beyond V4, the two streams project to inferotemporal cortex in monkey (Van Essen and Maunsell 1983; figure 8).

There are numerous cross-connections between the magnocellular and parvocellular systems. For example, a number of cross-connections occur in layer IVb and the superficial layers of V1 (Malpeli et al. 1981), and between the thick and thin stripes of V2 (Livingstone and Hubel 1984). Similarly, the thin stripes also have a variable projection to V5 (DeYoe and Van Essen 1985). Area V3 connects to both V4 and V5, and area V4 connects to V5 (Van Essen and Maunsell 1983). In addition, V4 projects to both parietal and inferotemporal cortex (Zeki 1977).

**Visual Areas in Human** Although the preceding discussion applies foremost to the macaque, it has recently been demonstrated that analogous visual areas exist in human. Clarke and Miklossy (1990) sectioned six human brains having unilateral occipital infarctions and stained the sections with silver to reveal the callosal afferents to the intact occipital lobe. They found a band of callosal afferents along the area 17/18 boundary, while areas 17 and 18 were otherwise free of them. More laterally in area 18, and in area 19, patches of callosal afferents alternated with callosal-free regions. From this, it was deduced that V2 was a horseshoe-shaped region around V1. The outer part of area 18 was thought to contain V3, dV3 lying dorsally and VP ventrally. Area V4 was deduced to lie on the fusiform gyrus. The proposed visual areas were found to have distinctive architectonic features. For example, V2 was heavily myelinated and its layer
Figure 8. Summary of the principal visual areas and cortical connections in primate (and, by inference, in human). Numerous reported connections are omitted for clarity. LGN: lateral geniculate nucleus; M: magnocellular; P: parvocellular. After DeYoe and Van Essen (1988).

III contained large pyramidal neurones. Dorsal V3 had similar properties and so was not easily differentiable from V2, but VP lacked the heavy myelination and large layer III pyramids, therefore being easily distinguished from the ventral part of V2. The region lateral to VP was found to be heavily myelinated, having large pyramids in layers III and V, but the myeloarchitecture was not uniform. Most notably, a very heavily myelinated region in the lateral part of area 19, near to the occipito-temporal junction, was considered to be area V5 (figure 4c).

The findings of Clarke and Miklossy (1990) were supported by Zilles and Schleicher (1991). They stained sections of occipital cortex using the Nissl and Gallyas methods and automatically scanned parts of the sections for cell bodies or myelin. They then computed the proportion of each section occupied by such features, using this as a quantitative measure of cyto- and myelo-architecture (Schleicher et al. 1986). They were able to reliably distinguish areas V1, V2, and the heavily myelinated V5. Several distinct visual areas were identified in the region surrounding V2, the identity of which
was not clear, but they were comparable with areas dV3, VP and V4 defined by Clarke and Miklossy (1990).

Although the cyto- and myelo-architecture of visual cortex reveals a number of distinct visual areas, it cannot be used to determine the functional properties and identity of these areas. Numerous positron emission tomography studies have therefore been conducted to demonstrate the functional behaviour of the occipital cortex (see, for example, Fox et al. 1986, 1987; Corbetta et al. 1991; Zeki et al. 1991). The basis of positron emission tomography has already been described in chapter one: a human subject is stimulated by a target which contains a certain feature, such as colour, or motion, the cerebral blood flow is recorded, and the cerebral blood flow for the viewer without the particular stimulus feature is subtracted to give a measure of the extra blood flow elicited by the stimulus (Frackowiak and Friston 1994).

The earliest experiments using positron emission tomography (Fox and Raichle 1984) demonstrated that the frequency at which a visual stimulus was presented influenced the regional cerebral blood flow in striate cortex. A grid of light emitting diodes produced the greatest increase in cerebral blood flow when flashed at a frequency of around 8Hz. The main response lay in area 17. It was subsequently shown (Fox et al. 1986, 1987) that the striate response was retinotopic. Circular, annular and semi-annular red-on-black checkerboards were reversed at a frequency of 10Hz and the increase in blood flow, relative to that recorded from the viewer observing the black background only, were measured. The cortical responses of seven subjects to an upper field stimulus were around 14mm more ventral than the response to a lower field stimulus, demonstrating the inverted topographic representation of the visual field in V1. Because the greatest extent of medial striate cortex was considered to be around 10mm (Talairach et al. 1967; Stensaas et al. 1974), it was deduced that area V2 was also involved in the cortical response. Macular, paramacular and peripheral stimuli produced successively more superior cortical blood flow, demonstrating the topographic organisation of V1 (and V2).

Corbetta et al. (1990, 1991) examined the cerebral blood flow resulting from stimuli involving shape, colour and speed. A stimulus was presented twice and the subject was instructed to decide whether the two stimuli differed in shape, colour, or speed. Four conditions were used: (1) the null stimulus, when the subject viewed the fixation point only; (2) passive viewing, when the subject viewed the stimulus, pressed a button, but made no judgement; (3) selective attention, when the subject judged differences in one stimulus feature only; and (4) divided attention, when the subject had to discriminate differences in any of the three stimulus features. Subtracting blood flows for the different levels of attention gave a measure of the cortical involvement in the different tasks. Passive viewing, relative to viewing the null stimulus, gave a cortical response which was centred in striate cortex, but which also involved the fusiform gyrus. Specific attention to shape, compared to passive viewing, mainly increased the blood flow to the fusiform and parahippocampal gyri. Selective attention to colour
increased the blood flow to the collateral sulcus and dorsolateral occipital cortex, and selective attention to speed mainly activated the lateral occipital gyrus. Divided attention increased the blood flow to the collateral sulcus and lateral occipital gyrus. All of the attentive conditions also increased the blood flow to area 17. The blood flow observed when attending to the specific features, compared to that recorded when attention was divided, showed similar regions of activation. Other regions of activation, outside the visual cortex, were also observed. Note, however, that the experimental paradigm involved a motor response (pressing a button), which could have accentuated the extra-visual activity (Corbetta et al. 1991).

The experiments of Corbetta et al. (1990, 1991) were an attempt to locate functionally specialised regions responsible for the processing of form, colour and motion. But by far the most elegant experiments of this nature were those of Zeki et al. (1991). They carried out cerebral blood flow measurements under three conditions: (1) when the subjects had their eyes closed; (2) when the subjects viewed a monochrome abstract pattern of no recognisable shape; and (3) when the subjects viewed the same pattern in colour, such that the luminances of the coloured pattern elements were the same as the luminances of the previous grey elements. The blood flow occurring when viewing the achromatic scene, compared to that occurring when the eyes were closed, showed heightened blood flow in striate cortex and in the cortex immediately around it, which was presumed to be V2. The cerebral response to the coloured stimulus, compared to that for the achromatic stimulus, showed significant blood flow to the fusiform and lingual gyri, the total blood flow to these parts being more than three times the previous blood flow elicited by the achromatic stimulus. Thus, introducing colour into the stimulus had the effect of activating the fusiform and lingual gyri, with most activity observed on the fusiform gyrus. The latter was therefore deduced to be the site of V4.

A similar experiment (Zeki et al. 1991) compared the cerebral blood flow for stationary and moving stimuli. The three stimulus conditions were (1) eyes closed, (2) the subject viewed a stationary pattern of black-and-white squares, and (3) the subject viewed the same stimulus with it moving, but changing direction every five seconds so that all directions of motion were experienced. The blood flow in response to the stationary pattern, compared to that for the eyes closed, showed that area 17 was active. When the moving pattern was viewed, additional blood flow was observed in an area of lateral cortex at the junction of the occipital and temporal lobes. This lateral region, at the border of Brodmann's areas 19 and 37, was clearly activated by the motion component of the stimulus, and was therefore deduced to be V5. Hence, Zeki was able to demonstrate the presence of specialised functional areas devoted to colour and motion processing in the human visual cortex. These specialised areas corresponded exactly to areas V4 and V5 as suggested by Clarke and Miklossy (1990) and Zilles and Schleicher (1991).
Conclusions

Areas of the cerebral cortex having an exclusively visual function are contained within Brodmann’s areas 17, 18 and 19. Within this region are a number of functional areas, of which V1 to V5 are the most important. The putative locations of V1 to V5 in human are as follows: V1 occupies striate cortex, involving the calcarine sulci and the cuneal and lingual gyri; V2 surrounds V1 in a horseshoe-shaped arrangement, possibly occupying only part of area 18; V2 is surrounded dorsally by dV3 and ventrally by VP; V4 is located ventrally, mainly occupying the fusiform gyrus; and V5 lies laterally at the most anterior extent of area 19, on the lateral occipital gyrus. However, the locations and identities of these visual areas, and their relationship to the analogous primate areas, are not yet well-established. Areas V1 and V2 are functional segregators, separating out the features of the visual stimulus, while V3, V4 and V5 are specialised functional areas. Area V3 is probably concerned with analysing form, V4 is probably involved in processing form in conjunction with colour, and V5 is probably devoted to analysing motion.

Visual cortex has a laminar and columnar organisation. The neuronal axons and dendrites reflect this structure, extending either horizontally along a particular lamina, or vertically across a number of laminae, but rarely diagonally (see figure 7). The columns and layers have functional importance, the columns having distinctive receptive field properties and the layers within each column facilitating the processing of information. The neurones in visual cortex give rise to directional electric currents which, in turn, produce neuro-electromagnetic fields. Mechanisms by which this occurs are discussed in the following chapter.
3. The Forward Problem

Introduction

The most important stage of the neuro-electromagnetic imaging procedure is the solution of the inverse problem. Using the inverse problem, the internal neural electric current can be deduced from measurements of the external electric and magnetic fields. However, solving the inverse problem implicitly requires that the opposite procedure can be performed, that of calculating external electric and magnetic fields from a well-defined neural current. This latter procedure - the forward problem - is much simpler than the inverse problem because it is unique: a given neural current produces one, and only one, electromagnetic field distribution. This chapter therefore establishes the mechanisms by which neurones in the brain produce measurable electric and magnetic fields, and then develops a mathematical description of the forward problem. This formulation becomes a basis for the solution of the inverse problem in subsequent chapters.

The two main field-producing mechanisms are the action potential and the postsynaptic potential; these are considered separately below.

The Action Potential

Nerve axons such as those found in the optic nerve consist of a central conducting fibre and a capacitive myelin sheath. The total diameter of an axon can be anywhere between 0.1 and 8μm, but is typically about 1μm (Polyak 1941; Kupfer et al. 1967; Jonas et al. 1990). The myelin sheath is arranged in sections, each of which has a length of around 1-2mm (Regan 1972). The myelin sections are separated by short, myelin-free notches - Nodes of Ranvier - from which electric current is free to flow. Electrical impulses jump from one node to the next in a process known as saltatory conduction (Brown and Smallwood 1981). The myelin sheath has the effect of increasing the conduction velocity of the unmyelinated axon by an order of magnitude, allowing an electrical impulse to travel along an axon at a speed of around 3ms⁻¹ (Erlanger and Gasser 1938). The precise speed depends upon the diameter and
longitudinal resistance of the axon and the capacitance of the myelin sheath (Rushton 1951; Hodgkin and Huxley 1952a, b).

The mechanism by which a nervous impulse - an action potential - passes along an axon is now described. Surrounding the axon is a semi-permeable membrane whose properties are such that, under normal circumstances, there is a high concentration of potassium ions and a low concentration of sodium ions inside the nerve fibre. These ions are provided by potassium chloride and sodium chloride. The concentration difference produces a voltage of about -90mV within the axon, relative to the surroundings of the axon. The axon is said to be polarised, and remains in this state until some kind of disturbance propagates along the fibre (Brown and Smallwood 1981).

The resting state is an equilibrium involving several forces (Wikswo 1989). In the polarised condition, there are more potassium ions within the cell than without. These ions, therefore, tend to leave the cell by diffusion, but are retained due to the negative potential within the cell. Conversely, there is a higher concentration of sodium ions outside the cell than within, so that they tend to enter the cell by chemical diffusion and by electrical attraction to the negative potential. Restraint of these ions is achieved by the semi-permeable cell membrane, which is around 30 times more resistive to Na+ ions than to K+ ions. In addition, the equilibrium is balanced by a sodium-potassium pump mechanism which actively transports sodium and potassium ions across the cell membrane as required. The pump expends metabolic energy and, in general, removes three Na+ ions out of the cell for every two K+ ions that it moves into the cell. In practice, the Cl- ions found in the sodium chloride and potassium chloride solutions also have some part to play in the transmission of an action potential but, for electrochemical reasons, these ions can be neglected when considering the transmission mechanisms: the Na+ and K+ ions are the only significant species (Carpenter 1990).

If the potential within the axon rises to above the stimulation threshold, which is typically around -50mV, the permeability of the axon membrane to Na+ ions increases by about two orders of magnitude (Wikswo 1989). This results in an influx of Na+ ions and the internal voltage rises to around +20mV. The nerve axon is then said to be depolarised (figure 9). Because of the potential gradient within the axon, intracellular current flows away from the depolarised region in both directions. The effect of positive charge arriving in the region in front of the depolarised section of axon is that the intracellular potential is raised to above the stimulation threshold, the permeability of the membrane increases with respect to Na+ ions, and depolarisation results (Dudel 1985). This is prevented from happening in the reverse direction, for reasons which will shortly be explained. Consequently, the depolarisation spreads in the direction of propagation of the action potential (Hodgkin 1964).

The depolarisation within the nerve cell does not persist indefinitely, since this would render the cell useless for further action potentials. The depolarised section of axon is therefore reset about 0.3ms after depolarisation has taken place. This is achieved in two ways (Wikswo 1989). Firstly, the sudden increase in sodium
permeability is followed by a decrease. This prevents further influx of sodium ions. However, there is already an excess of positive charge within the cell and the decrease in sodium permeability cannot alter this. Hence, the cell membrane undergoes a transient increase in potassium permeability, which allows a sufficient number of K⁺ ions to leave the cell to return the depolarised region of the axon to its original polarised state. This is known as repolarisation.

The variations in transmembrane potential occurring at the end of, or after, repolarisation are referred to as after-potentials. The intracellular potential may become more negative than the resting potential, in which case the after-potential is said to be hyperpolarising. Conversely, if the intracellular potential becomes less negative than the
resting potential, it is referred to as a depolarising after-potential (Dudel 1985). These minor imbalances are corrected by the sodium-potassium pump mechanism. The pump also restores the correct concentrations of sodium and potassium ions within and without the nerve axon. The complete static depolarisation-repolarisation cycle, involving inward Na⁺ current followed by outward K⁺ current and any subsequent adjustment, returns the axon to its resting potential.

For several milliseconds after a region of nerve axon is depolarised by an action potential, it remains inactive, despite the intracellular potential being raised to a much less negative value than the nominal threshold potential. This period is known as the absolute refractory period (Carpenter 1990). This is followed by the relative refractory period, in which the threshold potential is elevated and any action potentials which are successfully generated have a reduced amplitude. The absolute refractory period causes an action potential to propagate in one direction only. Although intracellular current flows in either direction from a depolarised region of axon (figure 9a), causing the potential either side to rise, the current directed in the reverse direction does not induce depolarisation as the section of axon which has been recently traversed is in a refractory state. Consequently, the action potential propagates in the forward direction only.

Viewed dynamically, the transmission of the action potential involves the passage of the depolarisation front along the nerve axon, followed, at a separation of about one millimetre, by the repolarisation phase. Assuming that the action potential propagates along the axon at a constant speed, (an assumption which is not entirely true as the nodes of Ranvier induce saltatory conduction), then the temporal voltage fluctuation at a given point in an axon has a similar form to the spatial voltage variation along the length of the axon (see figure 9). Assuming that the conduction speed of an action potential is 3ms⁻¹, which is a reasonable estimate for an axon of 1μm diameter (Erlanger and Gasser 1938), then in the 0.3ms duration of the depolarisation, the action potential travels around 1mm. Hence, the length of the depolarised region is approximately 1mm. This is important, as the distance separating the two opposed intracellular currents governs the magnitude of the externally observed electromagnetic field. This is discussed later, after we have considered the alternative current generating mechanism - the postsynaptic potential.

The Postsynaptic Potential

A synapse is a junction between two neurones, allowing an electrical excitation to continue along successive neurones. Synaptic mechanisms are evident at several stages in the visual system. In the first stage, at the retina, a photoreceptor activates a bipolar cell by means of a synapse (Murakami et al. 1972, 1975; Wu and Dowling 1978; Slaughter and Miller 1981, 1983). Other examples include relay synapses in the lateral geniculate nuclei and neuronal connections within the visual cortex. The result of
synaptic activity is that postsynaptic potentials are produced in the dendrites of a neurone, these potentials in turn contributing to the generation of an action potential. The postsynaptic mechanism is now considered briefly.

The arrival of a series of action potentials in the enlarged terminus - the presynaptic terminal - of a nerve axon, triggers the release of a number of chemical neurotransmitters. These act upon the dendritic membrane of an adjacent neurone. The permeability of the cell membrane is altered so that the relative concentrations of the sodium and potassium ions are changed, raising or lowering the intracellular potential (Eccles 1964). The neurotransmitter species determines whether the synapse is excitatory, promoting the generation of an action potential, or inhibitory, suppressing it (Regan 1972). In the case of an excitatory postsynaptic potential, the intracellular potential rises. The rising phase lasts for about 2ms, and is followed by a gradual decline, lasting for 10 to 15ms (Curtis and Eccles 1959; figure 10). There may be a sufficient number of excitatory synaptic inputs to raise the intracellular potential above the threshold voltage, so that an action potential is triggered, usually in the region of the cell body.

The resting potential of a dendrite is about -70mV. However, the threshold potential is not constant over the entire neurone, but is considerably less negative in the soma and dendrites than in the axon and axon hillock (the point where the axon leaves the soma). Consequently, if one of the dendrites is depolarised by synaptic activity, the threshold voltage is first exceeded at the axon hillock, and this becomes the starting point of an action potential. Hence, an action potential always originates from the axon hillock, regardless of the original site of dendritic depolarisation (Schmidt 1985). Since the axon hillock merges into the axon, the action potential propagates along the axon to the periphery of the neurone. It may be thought that synapses located near to the cell body contribute greatest to the triggering of the neurone, as these have the greatest influence on the potential at the axon hillock. However, it seems that excitatory postsynaptic potentials originating at the extrema of the dendritic tree are greater in amplitude, compensating for the effect.

If the synaptic behaviour is inhibitory, the dendrite becomes slightly hyperpolarised. This makes it more difficult for excitatory inputs to raise the potential above the trigger threshold, and so action potentials are suppressed (Eccles 1964).

 Whereas the action potential essentially consists of two states, polarised and depolarised, the postsynaptic potential is a continuously varying quantity which is influenced by as many as a thousand synapses on several dendrites. The postsynaptic potential rises and falls more slowly than the action potential, and is spatially more extended (figure 10). It is also characterised by a single direction of current flow, which is important when considering the relative importance of the action potential and postsynaptic potential in generating evoked electric and magnetic responses.
Figure 10. The postsynaptic potential. (a) A schematic population of excitatory synapses producing a depolarisation within a dendrite of a nerve cell. Note the single, extended intracellular current dipole. For simplicity, the multitude of synapses covering the dendrite and soma are omitted and it is assumed that all the synapses are excitatory; inhibitory activity would produce a current dipole in the opposite direction. PT: presynaptic terminal. (b) Intracellular potential measured near to a point of synaptic activity during an excitatory postsynaptic potential (EPSP) and inhibitory postsynaptic potential (IPSP). The magnitudes of the potentials are governed by the number of active synapses. (b) after Schmidt (1985).

Sources of Evoked Electromagnetic Fields

In principle, both the action potential and postsynaptic potential could contribute to the electromagnetic field observed around the head of a human subject upon stimulation of the visual system. However, the postsynaptic potential is far more likely to be the source of the field than the action potential, for reasons which are now discussed.

The action potential is comprised of two opposing currents which travel along an axon at a speed of some 3ms⁻¹, separated by a distance of about a millimetre (Dudel 1985; Carpenter 1990; see figure 9). Each of the two currents can be considered as a current dipole. That is, each is composed of a current source and a current sink which
are separated by a small distance (Plonsey 1981; Tripp 1983). The magnitude of the current dipole is given by the current flow multiplied by the source-sink separation, and it is this quantity which governs the magnitude of any external field (see next section). The reader can see intuitively that the two current dipoles will produce electromagnetic fields which almost cancel. A more detailed analysis shows that in an extensive, homogeneous medium, the electric and magnetic fields fall off with distance away from the source pair as $1/r^3$, as opposed to $1/r^2$ for a single, isolated current dipole (Nunez 1981; Tripp 1981; Roth et al. 1985; Wikswo 1989). It is therefore less likely that action potential fields are observed. Furthermore, action potentials propagating along a bundle of neurones are asynchronous, so that in practice, the individual fields of some of the action potentials cancel, giving a reduced overall field (Romani 1989).

These theoretical considerations are supported by experimental studies which demonstrate the absence of a relationship between axon spikes and surface slow-wave electric potentials (Renshaw et al. 1940). It has also been shown that unitary spike activity can be suppressed while surface slow-wave electric potentials persist (Li et al. 1952; Li and Jasper 1953; Purpura and Grundfest 1956; Amassian et al. 1964).

In contrast, there is every indication that the postsynaptic potential contributes significantly to the observed electric and magnetic fields. The postsynaptic potential is characterised by a single current dipole which extends along a dendrite of a cortical cell; there is no opposing dipole to cancel the external fields. The postsynaptic potential is therefore much more likely to generate a neuro-electromagnetic field than the action potential (Romani 1989).

It transpires that there are two types of cell configuration which undergo postsynaptic excitation. These are referred to as "open-field" and "closed-field" configurations (Lorente de Nó 1947; Llinas and Nicholson 1974; figure 11). An open-field cell, such as a pyramidal cell, has a dendritic tree which is dominated by a single, elongated dendrite. A postsynaptic potential, generated by synapses on the distal dendrites, spreads down the length of the extended apical dendrite (Kuffler and Nicholls 1973). The associated current dipole is long (thus having a large dipole moment), and free from cancelling currents in adjacent dendrites of the same cell. A cell having a closed-field configuration, an example of which is a stellate cell, has a short, spherically symmetric dendritic tree. Synaptic activation of such a cell is likely to result in a series of closely-spaced current dipoles, all acting so as to cancel each other (Lorente de Nó 1947). Needless to say, this type of configuration produces no external electromagnetic field; hereafter, our interest is confined to open-field cells, of which the pyramidal cell is the most significant.

The temporal characteristics of the postsynaptic potential also make it a likely source of neuro-electromagnetic field. A complete postsynaptic activation typically requires 10-20ms, which is exactly in accord with the time taken for a peak of an evoked potential or evoked magnetic field to evolve. In addition, the 10-20ms duration of the postsynaptic potential facilitates the overlap of numerous postsynaptic potentials in
Figure 11. (a) A pyramidal cell, which has open-field properties by virtue of its extended, linear geometry, and (b) a stellate cell, which has closed-field properties due to the symmetrical arrangement of its dendrites. Only open-field cells produce significant external electromagnetic fields. After Lorente de Nó (1947).

adjacent neurones, thus favouring a synchronous response, which is necessary to produce an external field (Okada 1983a; Romani 1989). Observed neuromagnetic fields are of the order of $10^{-13}$T (Okada 1983b; Romani 1989; Harding et al. 1991), while the current dipole moment associated with a single postsynaptic potential is around $10^{-14}$Am (Romani 1989; Hari 1991), therefore approximately $10^9$ neurones are required to produce the observed fields. The mechanism responsible for the observed fields must therefore be synchronous, and the postsynaptic potential fulfils this requirement (Romani 1989).

A number of experimental studies support the postsynaptic potential as the generator of potentials measured on the cortical surface (Li and Jasper 1953; Bremer 1958; Purpura 1959; Lux and Klee 1962; Li 1963; Creutzfeldt et al. 1964; Kostopoulos and Gotman 1984). For example, Eccles (1951) analysed single cell and cortical surface records obtained by direct electrical stimulation of the cortex, by stimulation of afferent fibres leading to the cortex and by antidromic stimulation of efferent fibres passing from the cortex. It was concluded that the surface potentials were generated primarily by postsynaptic potentials. Creutzfeldt et al. (1966) measured evoked responses from the surface of motor cortex in anaesthetised cats after electrical stimulation of the thalamic nuclei and of the cortex itself. The surface responses were compared with the cellular responses from the same cortical area. It was found that the surface evoked potentials were composed of different patterns of excitatory and inhibitory postsynaptic activity. However, the possibility of a contribution from synchronised afferent and efferent fibre activity was not ruled out. Spontaneous cortical potentials were also deduced to be the
result of a sequence of excitatory and inhibitory postsynaptic potentials and cellular excitations. Humphrey (1968a, b) transected the medulla and antidromically stimulated the isolated pyramidal tract. In this experiment, cell discharges (action potentials) and postsynaptic potentials were separated in time, allowing the respective contributions to the cortical surface potential to be accurately assessed. It was discovered that cell discharge contributed negligibly to the cortical surface record, whereas postsynaptic potentials gave relatively large surface signals. It was also found that the disturbances in potential which surrounded sites of postsynaptic activity extended further than those produced by action potentials. It was deduced that pyramidal-shaped neurones were behaving as low-pass filters when potentials were measured at distances of greater than 200μm from the active membrane. Hence, the postsynaptic potential, with its relatively long time constant, dominated cortical surface measurements.

Such experiments enable direct inferences to be made regarding the sources of neuro-electromagnetic signals. The scalp potential is merely a convolution of the cortical surface potential and is therefore closely related to it (Walter 1964; Cooper et al. 1965; Perrin et al. 1987). Hence, if the action potential is immeasurably small at the cortical surface, it will also be absent at the scalp; if the cortical record is dominated by the postsynaptic potential, so also the scalp potential. From these experiments, it can therefore be concluded that evoked scalp potentials are the result of postsynaptic activity within the cortex. This conclusion supports the theoretical and experimental evidence described earlier, which suggests that the action potential does not contribute to the external electromagnetic fields, whereas the postsynaptic potential does. All the evidence thus indicates that evoked electromagnetic fields are the result of postsynaptic activity within the cortex (Okada 1983a).

Pyramidal cells have already been considered as likely generators of neuro-electromagnetic fields because of their open-field properties. Not only are these cells comparatively large, therefore sustaining large intracellular currents, but they are also aligned very regularly throughout the cortex (Valverde 1985; see figure 7). The apical dendrites of the large pyramidal cells in layer 5 extend through many of the more superficial cortical layers and are approximately parallel to each other (Okada 1983a; Valverde 1985). Consequently, synchronous currents within these dendrites tend to summate, producing significant external neuro-electromagnetic fields (Romani 1989). Since the apical dendrites of the pyramidal cells are oriented perpendicular to the pial surface, it is possible to consider the intracellular current in a small region of cortex as a sheet of dipoles with the moments facing towards (or away from) the cortical surface.

**Fundamental Mathematical Relationships**

At any given time, a number of cortical cells have intracellular current flowing along their dendrites. Associated with each intracellular current is a return current,
which spreads out through tissue adjacent to the active cell to complete an electrical circuit (Williamson and Kaufman 1987). Consequently, at any point in the head, there is intracellular current, referred to as primary current, and volume current, which emanates from adjacent cells and approximately obeys Ohm's law (Hämäläinen et al. 1993) (figure 12). The primary current is non-zero only at the sites of cortical activity.

There is some confusion in the literature as to the definition of impressed, primary and volume currents. Impressed current is membrane current which occurs as a result of the ion pump mechanism and is therefore non-ohmic (Tripp 1983; Ilmoniemi 1991). In principle, all other current is ohmic return current. However, as Tripp (1983) points out, the return current can be divided up into intracellular and extracellular flow. The impressed and intracellular currents - the total current relating to the active cell - are grouped together and referred to as primary current. The extracellular return current is termed volume current. It transpires that the impressed current density contributes negligibly to neuromagnetic signals, since (a) a cell membrane has a thickness of less than 100nm, so that the dipole moment associated with the impressed current is negligible, and (b) the circular cross-section of a dendrite ensures that current dipoles situated normally to the cell membrane cancel each other, producing no net dipole moment (Swinney and Wikswo 1980; Williamson and Kaufman 1989). Instead, the intracellular component of the return current is the major contributor to the primary current, as a result of its vorticity (Tripp 1983). To summarise, the longitudinal intracellular current constitutes primary current, and the return flow outside of the active cell constitutes volume current.

Mathematically, the total current density, \( \mathbf{J} \), can be expressed as (Hämäläinen et al. 1993):

\[
\mathbf{J} = \mathbf{J}_p + \mathbf{J}_v = \mathbf{J}_p + \sigma \mathbf{E}
\]

(4)

where \( \mathbf{J}_p \) and \( \mathbf{J}_v \) are the primary and volume currents respectively, \( \sigma \) is the conductivity of the volume conductor and \( \mathbf{E} \) is the electric field in the conductive medium. We make the assumption that the conductivity of the volume conductor is isotropic, so that \( \sigma \) is a scalar. The requirement is to calculate the electric potential, \( V \), and magnetic flux density, \( \mathbf{B} \), resulting from the presence of current density, \( \mathbf{J} \). To do this, Maxwell's equations are invoked (Maxwell 1891):

\[
\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{5a}
\]

\[
\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}, \tag{5b}
\]

\[
\nabla \cdot \mathbf{D} = \rho, \tag{5c}
\]

\[
\nabla \cdot \mathbf{B} = 0. \tag{5d}
\]
Here, \( \mathbf{H} \) is magnetic field intensity, \( \mathbf{D} \) is electric flux density, and \( \rho \) is the density of free electric charge. Also used are the constitutive relations \( \mathbf{D} = \varepsilon \mathbf{E} \) and \( \mathbf{B} = \mu \mathbf{H} \), where \( \varepsilon \) and \( \mu \) are the permeability and permittivity of the conductive medium, respectively. These simple relations are permitted because the volume conductor is linear and isotropic (Smythe 1968). (The medium remains linear and isotropic regardless of whether the conductivity is isotropic or otherwise.) Note that \( \varepsilon = 80\varepsilon_0 \) and \( \mu = \mu_0 \), where \( \varepsilon_0 \) and \( \mu_0 \) are the permeability and permittivity of free space, respectively.

The neuro-electromagnetic fields in which we are interested have frequencies of much less than 1kHz. This enables the quasi-static approximation to Maxwell's equations to be used (Plonsey 1969). If the current density oscillates sinusoidally at some frequency, \( f \), so will all other fields, although they may have some phase difference, \( \phi \) (Tripp 1983). Thus, substituting equation 4 into equation 5b, and assuming that \( \mathbf{J}_p = \mathbf{J}_0 \sin 2\pi ft \), the terms on the right hand side of equation 5b can be written as:

\[
\mathbf{J}_p \sin 2\pi ft + \mathbf{E}_0 \left[ \sigma \sin(2\pi ft + \phi) + 2\pi f\varepsilon \cos(2\pi ft + \phi) \right],
\]

where \( \mathbf{E}_0 \) is the amplitude of the electric field. So long as \( 2\pi f\varepsilon \ll \sigma \), the displacement current term, \( \partial \mathbf{D}/\partial t = 2\pi f\varepsilon \mathbf{E}_0 \cos(2\pi ft + \phi) \), may be neglected in favour of the ohmic term, \( \mathbf{J}_v = \sigma \mathbf{E}_0 \sin(2\pi ft + \phi) \). For frequencies in the range 0-1kHz, \( 2\pi f\varepsilon \) is always less than about 5% of \( \sigma \), so that the displacement current term can be safely ignored (Tripp 1983). Also, since the volume conductor is ohmic, the integral of \( \mathbf{E} \) around a closed loop must vanish, so that \( \nabla \times \mathbf{E} = 0 \) and \( \mathbf{E} = -\nabla \psi \). Hence, simplifying equations 5 and incorporating the constitutive relations, we have:

\[
\begin{align*}
\nabla \times \mathbf{E} &= 0, \\
\nabla \times \mathbf{B} &= \mu_0 \mathbf{J}, \\
\nabla \cdot \mathbf{E} &= \frac{\rho}{\varepsilon_0}, \\
\nabla \cdot \mathbf{B} &= 0.
\end{align*}
\]  

Writing equation 4 as:

\[
\mathbf{J}_p = \mathbf{J} + \sigma \nabla \psi, \tag{7}
\]

and taking the divergence, gives the relationship:

\[
\nabla \cdot \mathbf{J}_p = \nabla \cdot \mathbf{J} + \nabla \cdot (\sigma \nabla \psi), \tag{8}
\]

but, taking the divergence of equation 6b, we find:
\[ \nabla \cdot \mathbf{J} = \nabla \cdot (\nabla \times \mathbf{H}) = 0, \quad (9) \]

so that (Sarvas 1987):

\[ \nabla^2 V = \frac{1}{\sigma} \nabla \cdot \mathbf{J}_p, \quad (10) \]

This is a Poisson equation which must be solved to determine \( V \) within the volume conductor. Two boundary conditions are used (Geselowitz 1967), namely that at an interface between two regions, of conductivity \( \sigma_1 \) and \( \sigma_2 \) respectively, (a) \( V_1 = V_2 \), i.e. potential is continuous across the interface, and (b) \( \mathbf{J}_{n1} = \mathbf{J}_{n2} \), i.e. current density is constant across the interface. Since \( \mathbf{J}_p = 0 \) at boundaries, \( \mathbf{J} = -\sigma \nabla V \), so that this boundary condition can be expressed as \( \sigma_1 \frac{\partial V_1}{\partial n} = \sigma_2 \frac{\partial V_2}{\partial n} \).

This leads to a Fredholm integral equation of the second kind, describing the potential, \( V_k \), at each boundary surface, \( S_k \), within an arbitrary multicompartiment model in which each compartment has a homogeneous and isotropic conductivity (Barnard et al. 1967a; Geselowitz 1967; Meijs et al. 1989):

\[
V_k(\mathbf{r}) = \frac{\sigma_k}{\sigma_k} V^w_k(\mathbf{r}) = \frac{1}{4\pi \sigma_k} \sum_{i=1}^{N} \Delta \sigma_i \int_{S_i} V_i(\mathbf{r''}) \, d\Omega_{ki},
\]

with:

\[
V^w_k(\mathbf{r}) = \frac{1}{4\pi \sigma_k} \int_{\Omega_p} \mathbf{J}_p(\mathbf{r'}) \cdot \frac{\mathbf{r} - \mathbf{r'}}{|\mathbf{r} - \mathbf{r'}|^3} \, d\mathcal{P}
\]

and:

\[
d\Omega_{ki} = \frac{\mathbf{r} - \mathbf{r''}}{|\mathbf{r} - \mathbf{r''}|^3} \cdot dS_i.
\]

The symbols are as follows:

- \( \sigma_p \) is the conductivity of the single compartment in which the region of non-zero \( \mathbf{J}_p \) is located. Note that \( \mathbf{J}_p \) must be non-zero in one compartment of the model only.
- \( \sigma_i \) is the conductivity inside surface \( S_k \).
- \( \sigma^*_i \) is the conductivity outside surface \( S_k \).
- \( \overline{\sigma}_k = 1/2 (\sigma^-_i + \sigma^*_i) \).
- \( \Delta \sigma_k = \sigma^*_i - \sigma^-_i \).

64
\( V_k^c (r) \) is the electric potential which would arise from the source, were it located in an infinite, homogeneous medium of conductivity \( \sigma_p \).

\( d\Omega_{k} \) is the solid angle subtended at a point of observation, \( r \), on surface \( S_k \), by the surface element \( dS \), at \( r' \) on surface \( S_r \).

\( dP \) is a volume element of region \( P \), in which current density \( J_p \) is located at \( r' \).

\( N \) is the total number of compartment boundaries.

If the primary current density, \( J_p \), is a single current dipole of moment \( M \), located at \( r_0 \), equation 11b can be rewritten as (Sarvas 1987):

\[
V_k^c (r) = \frac{1}{4\pi \sigma_p} M \cdot \frac{r - r_0}{|r - r_0|^3}. 
\]  

(11d)

Returning to equations 6, it is possible to derive an expression for the magnetic flux density, \( B \), outside of a multicompartment model. Equation 6d implies that \( B \) can be written as the curl of another vector field, \textit{i.e.}:

\[
B = \nabla \times A, 
\]  

(12)

where \( A \) is referred to as the vector magnetic potential (Cheng 1989). Expanding equation 6b, we find:

\[
\nabla \times B = \nabla \times \nabla \times A = \nabla (\nabla \cdot A) - \nabla^2 A = \mu_0 J. 
\]  

(13)

Choosing \( \nabla \cdot A = 0 \) (the Coulomb condition) gives (Cheng 1989):

\[
\nabla^2 A = -\mu_0 J. 
\]  

(14)

This Poisson equation is the magnetic analogue of equation 10. Solving, with the boundary condition that \( B(r) \to 0 \) as \( r \to \infty \), yields the Biot-Savart law (Sarvas 1987; Cheng 1989):

\[
B(r) = \frac{\mu_0}{4\pi} \int_G J(r') \times \frac{r - r'}{|r - r'|^3} dG, 
\]  

(15)

where \( r \) is a point of observation, \( r' \) is a point at which there is some current density, \( J \), \( G \) is the complete volume conductor and \( dG \) is a volume element thereof. Substituting equation 7 into equation 15, we obtain an expression for \( B(r) \) in terms of primary current only (Sarvas 1987):
\[ \mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \left[ \mathbf{J}_p(\mathbf{r}') - \sigma(\mathbf{r}') \nabla V(\mathbf{r}') \right] \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} dG. \]  \hspace{1cm} (16)

For the specific case of the multicomartment model, this can be written as (Geselowitz 1970; Meijs et al. 1987):

\[ \mathbf{B}(\mathbf{r}) = \mathbf{B}^\ast(\mathbf{r}) - \frac{\mu_0}{4\pi} \sum_{i=1}^{N} \Delta \sigma_i \int_{S_i} \mathbf{V}_i(\mathbf{r}'') \frac{\mathbf{r} - \mathbf{r}''}{|\mathbf{r} - \mathbf{r}''|^3} \times d\mathbf{S}_i, \]  \hspace{1cm} (17a)

where the symbols are as for equations 11 and \(\mathbf{B}^\ast\), the magnetic flux density that would occur if the source region were located in an infinite homogeneous space, is given by:

\[ \mathbf{B}^\ast(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_{P} \mathbf{J}_p(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} dP. \]  \hspace{1cm} (17b)

Equation 17b is identical to equation 15, with \(\mathbf{J}(\mathbf{r}')\) replaced by \(\mathbf{J}_p(\mathbf{r}')\) and the integration now over the primary source region. In other words, in an infinite conductor, only the primary current contributes to the flux density. In the inhomogeneous case, \(\mathbf{B}^\ast\) is the flux density resulting from the source only, so that the second term of equation 17a is the contribution from the volume currents. In the case that the source current density, \(\mathbf{J}_p\), is a current dipole of moment \(\mathbf{M}\), located at \(\mathbf{r}_0\) (Sarvas 1987),

\[ \mathbf{B}^\ast(\mathbf{r}) = \frac{\mu_0}{4\pi} \mathbf{M} \times \frac{\mathbf{r} - \mathbf{r}_0}{|\mathbf{r} - \mathbf{r}_0|^3}. \]  \hspace{1cm} (17c)

Equations 17 are the magnetic counterparts of equations 11. Note that in general, when solving equations 17, the potential distribution over the compartment boundaries must be known. This means that equations 11 must first be solved. This procedure, carried out for an arbitrary, realistically-shaped volume conductor model, is known as the Boundary Element Method.

The Boundary Element Method

A numerical solution to the integral equation 11a can be obtained by decomposing each closed boundary surface, \(S_p\), of the model into \(N_t\) triangular panels. This process is known as triangulation (Lynn and Timlake 1968). In the "centre of mass" approach, the potential on each triangle is considered to be uniform and equal to
the potential at the centre of mass of the triangle (Meijs et al. 1989). Then the integral
\[ \int_{S_i} V_j d\Omega_{ik} \]
is replaced by a discrete approximation:
\[ \int_{S_i} V d\Omega_{ik} = \sum_{j \in S_i} V_j \Delta \Omega_{ij}, \tag{18} \]

where \( V_j \) is the potential at the centre of mass of triangle \( j \) on \( S_i \), and \( \Delta \Omega_{ij} \) is the solid angle subtended by triangle \( j \) at the centre of mass of triangle \( i \) on surface \( S_k \). (When \( i = j \), \( \Delta \Omega_{ii} \) is taken as \(-2\pi\).)

The centre of mass approach leads to numerical inaccuracy, so it is more
common to evaluate the potentials at the vertices of the triangles. Accordingly, \( \Delta \Omega_{ij} \) is
defined, for \( i \neq j \), as one third of the solid angle subtended at vertex \( i \) by the triangles
surrounding vertex \( j \). The value of \( \Delta \Omega_{ii} \) has to be approximated. This method
effectively provides \( 2N^2 \) discretisation points on each surface (Meijs et al. 1989).

The highest numerical accuracy is obtained when the potential distribution across
each triangle is approximately constant, in accord with the assumption made above.
This occurs when the solid angle subtended at the current source by each triangle is
small. To ensure that this is so, local mesh refinement is necessary, so that triangles near
to the current source are smaller than those far from the current source (Meijs et al.
1987). The importance of mesh refinement is greater when the current source is located
near to one of the surfaces of the conductivity model.

In principle, the continuous integral \( \int_{S_i} V_j d\Omega_{ik} \) can now be successfully evaluated
by discrete approximation, but the potentials \( V_j \) are not yet known (these are what we
are trying to find in the first place) and \( V_k \) cannot, therefore, be determined directly.
Instead, it is necessary to form a discrete system describing \( V \) at each point in the
discrete mesh in terms of \( V \) at all other points (Barnard et al. 1967b; Oostendorp and
van Oosterom 1991):
\[ \mathbf{v} = \mathbf{g} + \mathbf{Av}, \tag{19} \]

where \( v_i \) is the potential at the \( i^{\text{th}} \) discretisation point and \( g_i \) is the potential that would
occur at the \( i^{\text{th}} \) discretisation point if the medium were infinite and homogeneous.
Matrix element \( A_{ij} \) is proportional to the solid angle subtended by triangle \( j \) at
discretisation point \( i \).

Matrix \( \mathbf{A} \) is singular, but the singularity can be removed using a technique called
deflation (Lynn and Timlake 1968). The system of equation 19 can then be solved
iteratively using the Gauss-Seidel iteration scheme or Jacobi iteration scheme
(Press et al. 1989). Alternatively, where computer power permits, it is possible to solve
the system directly (Cuppen and van Oosterom 1984): rearranging equation 19 and
defining \( \mathbf{C} = \mathbf{I} - \mathbf{A} \) gives:
\[ Cv = g, \]  
\[ v = C^{-1}g. \]

This latter method has the advantage that \( C^{-1} \) is determined only by the volume conductor geometry, so that once it has been determined, the potential distribution, \( v \), can be evaluated for a variety of current sources by simply postmultiplying \( C^{-1} \) by the source matrix \( g \) (Oostendorp and van Oosterom 1991).

The above method is used to determine the discrete potential distribution over all the surfaces of a triangulated volume conductor model. In general, we are only interested in the potential on the exterior surface, although the potentials on the interior surfaces are a necessary part of the calculation.

Attention is now given to the relatively simple problem of solving equations 17 to find the magnetic flux density which results from a current source within the realistic model. Discretising the integral on the right hand side of equation 17a gives:

\[
\int_S V_i \frac{r - r''}{|r - r''|^3} \times dS_i = \sum_{S_i} V_j \frac{r - r_j}{|r - r_j|^3} \times \Delta S_j ,
\]

where \( V_j \) is the potential at the \( j \)th discretisation point, which is located at \( r_j \) on \( S_j \), and \( \Delta S_j \) is the vector area of the \( j \)th triangle. Provided equation 21 has been solved for the potential distribution, then the values \( V_j \) are known. Consequently, the discrete equivalent of equation 17 can be evaluated directly to yield \( \mathbf{B}(r) \). In particular, if \( \mathbf{B} \) is to be determined at \( M \) positions (Oostendorp and van Oosterom 1991),

\[
\bar{\mathbf{B}} = \bar{\mathbf{B}}^w + Dv.
\]

Note that \( \bar{\mathbf{B}} \) denotes an array of vectors \( \mathbf{B}_i(r_j) \) \((i = 1, M)\) and \( \bar{\mathbf{B}}^w \) an array of vectors \( \mathbf{B}_i^w(r_j) \). \( D \) is proportional to the vector:

\[
\frac{r_i - r_j}{|r_i - r_j|} \times \Delta S_j.
\]

An array of magnetic flux gradients can also be calculated simply (Oostendorp and van Oosterom 1991). Each flux gradient, \( e \), can be considered as a weighted sum of \( N_s \) flux densities, the sum being over all the coils of the gradiometer:
where $\mathbf{h}_k$ is a weighted normal to the gradiometer coil plane. Hence, we have:

$$\mathbf{e} = \mathbf{e}^\infty + \tilde{\mathbf{D}} \mathbf{v},$$

in which $\mathbf{e}$ is now a vector of flux gradients and $\mathbf{e}^\infty$ and $\tilde{\mathbf{D}}$ are weighted forms of $\mathbf{B}^\infty$ and $\mathbf{D}$, respectively, being denoted by:

$$e_i^\infty = \sum_{k=1}^{N_s} \mathbf{B}^\infty(\mathbf{r}_{ik}) \cdot \mathbf{h}_{ik}$$

and:

$$\tilde{D}_{ij} \propto \sum_{k=1}^{N_s} \frac{\mathbf{r}_{ik} - \mathbf{r}_{ij}}{|\mathbf{r}_{ik} - \mathbf{r}_{ij}|^3} \cdot \Delta S_j \cdot \mathbf{h}_{ik}.$$ 

Thus, the boundary element method is an elegant means of calculating a series of flux gradients outside a realistically-shaped volume conductor.

There are several drawbacks associated with the boundary element method. The main disadvantage is that it takes the conductivities of the model to be isotropic. This is certainly not the case for the human head, where the conductivity of the white matter is some ten times greater along the direction of the fibres than transversely (Nicholson 1965). The skull is also thought to possess a tangential conductivity which is about ten times higher than its radial conductivity (Rush and Driscoll 1968). The second major disadvantage is that, in order to be accurate, a large number of triangles must be used, so that the triangles are small and the assumption that potential is constant across each triangle is valid (Meijs et al. 1987). The method therefore takes a considerable amount of computer time and is cumbersome when results are required in a limited time.

Fortunately, the occipital region of the head, in which the visual cortex is located, is reasonably (although not exactly!) spherical. This enables the conductivity of the brain, skull and scalp to be modelled as a series of concentric spheres. The relative simplicity of the geometry permits an analytical solution to be found for the surface potential and external magnetic field, even when the shells of the model are anisotropic. The ability to model the anisotropy of the head is attractive, while the analytical solution is fast compared to the boundary element method. For these reasons, the spherical head model, which is detailed below, is used throughout the work described in this thesis. It
should be noted that neither the boundary element method nor the spherical model are the perfect conductivity model: the boundary element method can model the shape of the head more accurately, but the spherical model can account for the anisotropic conductivity and is faster and simpler.

The Spherical Head Model

This section describes the electric and magnetic forward problems for the case of a spherically symmetric conductivity model. When using a spherical model, the mathematical approaches to the electric and magnetic forward problems are somewhat different, the electric forward problem requiring a more explicit definition of the conductivity profile. For this reason, the two forward problems are considered separately. The electric forward problem is a special case of the general method of de Munck (1988), while the magnetic forward problem is derived from Sarvas (1987). In both cases, it is assumed that the current source is a single dipole; the electric potential and magnetic flux gradient due to multiple dipoles are obtained by superposition.

The Electric Forward Problem

To consider the anisotropy of the head, equation 4 is rewritten as:

\[ \mathbf{J} = J_p + J_v = J_p + \bar{\sigma} \mathbf{E}, \]  

where \( \bar{\sigma} \) is now a conductivity tensor - a 3 \times 3 diagonal matrix whose diagonal elements are the three orthogonal components of the conductivity at some point in the head. Modifying equations 7 and 8 accordingly and noting equation 9, we have the following expression for the monopolar source strength at each point in the medium (Nicholson and Freeman 1975):

\[ s = \nabla \cdot J_p = \nabla \cdot (\bar{\sigma} \nabla V). \]  

It is now assumed that the volume conductor consists of five concentric shells, modelling the scalp, skull, cerebrospinal fluid, grey matter and white matter. The \( j^{th} \) shell has radius \( r_p \) with the outermost surface having radius \( r_1 \) and the surface of the white matter having radius \( r_5 \). Throughout this thesis, the radii of the model spheres are taken to be those of table 1. The radius of the outer surface of the grey matter is taken from a stereotaxic study of 30 hemispheres (Talairach et al. 1967), while the remaining radii are determined by adding or subtracting the thicknesses of the other shells. The grey matter is assumed to have a thickness of 2mm, the cerebrospinal fluid 2mm, the
<table>
<thead>
<tr>
<th>Compartment</th>
<th>Outer radius (mm)</th>
<th>Radial conductivity (Sm⁻¹)</th>
<th>Tangential conductivity (Sm⁻¹)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalp</td>
<td>81</td>
<td>0.435</td>
<td>0.435</td>
<td>Burger and van Milaan 1943</td>
</tr>
<tr>
<td>Skull</td>
<td>75</td>
<td>0.005</td>
<td>0.054</td>
<td>Rush and Driscoll 1968</td>
</tr>
<tr>
<td>CSF</td>
<td>70</td>
<td>0.563</td>
<td>0.563</td>
<td>Crile et al. 1922 Radvan-Ziemnowicz et al. 1964</td>
</tr>
<tr>
<td>Grey matter</td>
<td>68</td>
<td>0.481</td>
<td>0.435</td>
<td>Van Harreveld et al. 1963 Ranck 1963 Freygang and Landau 1955 Crile et al. 1922</td>
</tr>
<tr>
<td>White matter</td>
<td>66</td>
<td>1.250</td>
<td>0.130</td>
<td>Nicholson 1965 Van Harreveld et al. 1963 Crile et al. 1922</td>
</tr>
</tbody>
</table>

Table 1. Radii and conductivities of the spherical shells in the 5-sphere electrical conductivity model.

skull 5mm and scalp 6mm. (An exception to the above occurs in chapter 9, where the scalp radius is determined for each subject individually, and the other radii fixed by assuming the above shell thicknesses.)

The $j^{th}$ layer of the model has radial conductivity $\varepsilon_j$ and tangential conductivity $\eta_j$. Appropriate values are given in table 1; these figures are by no means well-established, and should be treated as reasonable, but approximate. In the case of the white matter, the quoted values are strictly for longitudinal and transverse directions relative to the neural fibres themselves. However, in the occipital region, the optic radiation is predominant, so that most neurones are directed towards the cortex in an approximately radial manner. The longitudinal and transverse components have therefore been used to determine the radial and tangential components, respectively, for the spherical model. Similarly, for conductivity purposes, the grey matter is assumed to cover a spherical region with its surface tangentially oriented. This approach does not account for the presence of numerous sulci and the median fissure, and this needs to be borne in mind when considering the limitations of the model. The cerebellum, which has different conductivities to those of the cerebrum (Crile et al. 1922), effectively forms an eccentric small-radius sphere in the inferior portion of the main spherical model, and may thus be a source of modelling error.

The centres of the spheres coincide with the origin of a spherical coordinate system. A dipolar source of moment $\mathbf{M}$ lies at location $\mathbf{r}_a$ within the innermost shell of
the model. This is somewhat paradoxical as it has been established that the physiological current generators are located within the grey matter. However, considerable portions of the grey matter, namely the medial surfaces and the calcarine fissures, lie within the innermost model sphere, and since the following algorithm applies to sources in one sphere only, all current dipoles are taken as located within the inner sphere. When a source lies in that part of the cortex which coincides with the grey matter sphere of the conductivity model, it is shifted by one millimetre or so to just within the white matter sphere. The algorithm is then applied with all sources lying in the innermost sphere.

In spherical coordinates, equation 27 becomes (de Munck 1988):

$$\frac{1}{\eta} \frac{\partial}{\partial r} \left( r^2 \varepsilon \frac{\partial V}{\partial r} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial V}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 V}{\partial \phi^2} = \frac{r^2 s}{\eta}. \quad (28)$$

This is solved for the case of a monopolar source:

$$s = \frac{M_{mon}}{r^2 \sin \theta} \delta(r - r_0) \delta(\theta - \theta_0) \delta(\phi - \phi_0) \quad (29)$$

by substituting a trial solution of the form:

$$V = \sum_{\eta=1}^{\infty} \sum_{m=0}^{\eta} \sum_{\alpha=0}^{1} \frac{M_{mon} R_n(r_0, r) Y_{n\alpha}(\theta_0, \phi_0) Y_{n\alpha}(\theta, \phi)}{|Y_{n\alpha}(\theta, \phi)|^2}. \quad (30a)$$

The $Y_{n\alpha}(\theta, \phi)$ are spherical harmonics, defined by:

$$Y_{n\alpha}(\theta, \phi) = \begin{cases} P_n^\alpha(\cos \theta) \cos \phi & \text{if } \alpha = 0 \\ P_n^\alpha(\cos \theta) \sin \phi & \text{if } \alpha = 1 \end{cases} \quad (30b)$$

where $P_n^\alpha(\cos \theta)$ are associated Legendre functions (Abramowitz and Stegun 1970).

Equation 28 is satisfied when $R_n$ satisfies the ordinary differential equation for the radial part of the potential distribution, viz.:

$$\frac{d}{dr} \left( r^2 \varepsilon \frac{dR_n}{dr} \right) - n(n+1)R_n \eta = \delta(r - r_0). \quad (31)$$

The particular solution to this equation is found by constructing $R_n$ from two independent solutions to the homogeneous equation. These two solutions, $R_n^{(1)}$ and $R_n^{(2)}$, must satisfy the following equation:
\[
\frac{d}{dr} \left( r^2 \frac{dR_n^{(i)}}{dr} \right) - n(n+1) \eta_i \frac{\eta_j}{\epsilon_j} R_n^{(i)} = 0, \quad r_j \geq r \geq r_{j+1}, \quad i = 1, 2. \tag{32}
\]

The solutions \( R_n^{(i)} \) and \( R_n^{(2)} \) are found to be:

\[
R_n^{(i)} = A_j^{(i)} P_j(r) + B_j^{(i)} Q_j(r), \quad r_j \geq r \geq r_{j+1}, \tag{33a}
\]

where:

\[
\nu_j = \frac{-1 + \sqrt{1 + 4n(n+1) \eta_j/\epsilon_j}}{2}, \tag{33b}
\]

\[
P_j = r^\nu_j, \tag{33c}
\]

\[
Q_j = r^{-\nu_j-1}, \tag{33d}
\]

and \( A_j^{(i)} \) and \( B_j^{(i)} \) are defined by the recursive relations:

\[
\begin{pmatrix}
A_j^{(i)} \\
B_j^{(i)}
\end{pmatrix} = 
\begin{pmatrix}
P_j(r_j) & Q_j(r_j) \\
\epsilon_j P_j(r_j) & \epsilon_j Q_j(r_j)
\end{pmatrix}^{-1}
\begin{pmatrix}
P_j-1(r_j) & Q_j-1(r_j) \\
\epsilon_j P_j-1(r_j) & \epsilon_j Q_j-1(r_j)
\end{pmatrix}
\begin{pmatrix}
A_j^{(i-1)} \\
B_j^{(i-1)}
\end{pmatrix}, \tag{33e}
\]

with the boundary conditions:

\[
A_1^{(i)} = 1, \quad B_1^{(i)} = 0, \quad A_1^{(2)} = -\frac{Q_1'(r_1)}{P_1'(r_1)}, \quad B_1^{(2)} = 1. \tag{33f}
\]

Then:

\[
R_n(r) = \frac{-1}{(2n+1)\epsilon_N B_1^{(2)}} \left\{ \begin{array}{ll}
R_n^{(2)}(r_0) R_n^{(1)}(r) & r \leq r_0, \\
R_n^{(1)}(r_0) R_n^{(2)}(r) & r \geq r_0.
\end{array} \right. \tag{34}
\]

This equation is substituted into equation 30a to provide the potential distribution resulting from a monopolar source. The dipolar potential is then the inner product of the dipole moment and the gradient, with respect to the source point, of the monopolar potential. Hence:

\[
V_{dp} = \sum_{n=1}^{\infty} \frac{R_n^{(2)}(r)}{\epsilon_N B_1^{(2)}} \left\{ M_n R_n^{(1)}(r_0) Y_{n0}(\theta, \phi) + \frac{M_0}{r_0} R_n^{(1)}(r_0) Y_{n0}(\theta, \phi) \right\} \tag{35}
\]
for a current dipole located on the z-axis, with its moment vector lying in the xz-plane, directed towards the direction of decreasing x, such that it has a radial component \( M_r \) and tangential component \( M_\phi \) (de Munck 1988). The spherical harmonics \( Y_{n00} \) and \( Y_{n10} \) can be conveniently calculated using recurrence relations (Abramowitz and Stegun 1970):

\[
\begin{align*}
(n + 1)Y_{n+1,00} &= (2n + 1)\cos\theta Y_{n00} - nY_{n-1,00}, \\
nY_{n+1,10} &= (2n + 1)\cos\theta Y_{n10} - (n + 1)Y_{n-1,10}, \\
Y_{000} &= 1, \quad Y_{100} = \cos\theta, \quad Y_{010} = 0, \quad Y_{110} = -\sin\theta\cos\phi.
\end{align*}
\]

These relationships are known to be numerically stable (Abramowitz and Stegun 1970).

Strictly, equation 35 does not take into account anisotropy of the innermost (white matter) sphere, but assumes that \( \varepsilon_s = \eta_s \). However, Zhou and van Oosterom (1992) have proposed a multiplication factor for the expansion of equation 32 to allow the innermost sphere of a four-sphere model to be anisotropic. The \( n^{th} \) term in equation 35 is multiplied by:

\[
\frac{2n + 1}{\sqrt{1 + 4n(4n + 1)\eta_4 / \varepsilon_4}},
\]

where \( \varepsilon_4 \) and \( \eta_4 \) are the radial and tangential conductivities of the innermost sphere, respectively. A modified factor:

\[
\frac{2n + 1}{\sqrt{1 + 4n(4n + 1)\eta_5 / \varepsilon_5}}
\]

allows full anisotropy to be incorporated into the 5-sphere model under consideration.

In reality, a current source is located at random, rather than on the z-axis. To use equation 35, it is therefore necessary to transform the coordinate system so that the current source lies on the z-axis, directed in the required direction. If the cartesian coordinates of the true current dipole are denoted by \( r_0(x, y, z) \), the cartesian coordinates of the electrode site and the dipole moment vector are premultiplied by the transformation matrix:

\[
\begin{bmatrix}
xz & yz & -r_{xy} \\
x & y & r_0 \\
-x & y & 0 \\
x & z & r_{xy} \\
x & y & z \\
r_0 & r_0 & r_0
\end{bmatrix}
\]
where \( r_{xy} = \sqrt{x^2 + y^2} \). The modified coordinates are those relating to an equivalent system in which the current dipole lies on the z-axis with its moment vector at some angle to the xz-plane. If the new components of the dipole moment are \( M_x, M_y \) and \( M_z \), then the new coordinates of the electrode are premultiplied by a second transformation matrix:

\[
\begin{bmatrix}
M_x & M_y & 0 \\
M_{xy} & M_{xy} & 0 \\
M_y & M_x & 0 \\
M_{xy} & M_{xy} & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

where \( M_{xy} = \sqrt{M_x^2 + M_y^2} \). This gives the coordinates of the electrode in a system in which the current dipole is located on the z-axis with its moment vector lying in the xz-plane, directed towards the direction of decreasing \( x \), as required.

Computer implementation of equation 35 is known to suffer from floating underflow/underflow (Zhou and van Oosterom 1992). This can be overcome by simply scaling the sphere radii to near unity, with appropriate scaling of the current dipole moment. If the 5-sphere model has an outer radius of \( r_i \) and an inner radius of \( r_o \), then \( r_i \) is used to scale the dimensions of the model. If the set of five sphere radii is denoted by \( R \) and the current source is a dipole of moment \( M \) situated on the z-axis at \( z = r_o \), then the scaling mechanism used to evaluate the scalp potential, \( V \), can be described as:

\[
V(R, M, r_o) = V\left(\frac{1}{r_i} R, \frac{1}{r_i^2} M, \frac{r_o}{r_i}\right).
\]  

The expression on the right-hand side is free from floating underflow/overflow. Using this refinement, the author took the first 100 terms of equation 35, these providing in excess of six significant figures accuracy for sources anywhere in the inner compartment of the model.

**The Magnetic Forward Problem**

Before the magnetic forward problem can be formulated for the case of a spherically symmetric volume conductor, it is necessary to observe the simple but important result that volume currents do not contribute to the radial component of the magnetic flux density anywhere outside the spherical conductor. This has been proven for the case of an isotropic volume conductor (Grynszpan and Geselowitz 1973; Sarvas
1987) and, more recently, for an anisotropic volume conductor (de Munck 1989). Consequently, the radial component of the magnetic flux density is given by:

$$B_r(r) = \mathbf{B}^\ast(r) \cdot \frac{r}{|r|} = \frac{\mu_0 M}{4\pi} \frac{r - r_o}{|r - r_o|^3} \cdot \frac{r}{|r|},$$

(38)

if the source is a single current dipole of moment \(M\), located at \(r_o\).

Outside the volume conductor, \(\mathbf{J} = 0\) and therefore, from equation 6b, \(\nabla \times \mathbf{B} = 0\). Thus, it is possible to describe \(\mathbf{B}\) in terms of the scalar magnetic potential, \(U\), such that (Cheng 1989):

$$\mathbf{B}(r) = -\mu_0 \nabla U(r).$$

(39)

Since \(U\) vanishes at infinity, it can be calculated by integration (Sarvas 1987):

$$U(r) = -\int_0^\infty \nabla U(r + te_r) \cdot e_r \, dt$$

$$= \frac{1}{\mu_0} \int_0^\infty B_r(r + te_r) \, dt,$$

(40, 41)

where \(e_r = r/|r|\). Application of equation 39 produces an expression for \(\mathbf{B}(r)\) for the case of a single current dipole (Sarvas 1987):

$$\mathbf{B}(r) = \frac{\mu_0}{4\pi r^2} \left( F \mathbf{M} \times r_o - \mathbf{M} \times r_o \cdot r \nabla F \right),$$

(42a)

where:

$$F = a(r + r^2 - r_o \cdot r),$$

(42b)

$$\nabla F = \left( \frac{a \cdot r}{a} + 2a + 2r + \frac{a^2}{r} \right) r - \left( \frac{a \cdot r}{a} + 2a + 2r \right) r_o,$$

(42c)

$$a = r - r_o,$$

(42d)

$$a = |a|,$$

(42e)

$$r = |r|.$$  

(42f)

Two important points are mentioned here. The first is that this derivation is independent of the number of compartments in the model and the actual value of the conductivity in each compartment. The only prerequisite is that the compartments are spherically symmetric; the conductivity parameters do not influence the external magnetic field. The second point is that expression 42a consists of two terms, each
involving a vector product between $\mathbf{M}$ and $\mathbf{r}_o$. If the dipole is radially oriented, $\mathbf{M} \times \mathbf{r}_o = 0$ and there is no external magnetic field.

In practice, evoked magnetic responses are recorded using a flux gradiometer. The output of an axial gradiometer having $N_g$ coils is given by:

$$ e = \sum_{i=1}^{N_g} T_i \phi_i = \sum_{i=1}^{N_g} T_i \int_{S_i} \mathbf{B}(\mathbf{r}) \cdot d\mathbf{S}_i, $$

(43)

where the $i$th coil of the gradiometer consists of $T_i$ turns and encloses surface $S_i$, which is linked by flux $\phi_i$. ($T_i$ is positive or negative, depending on the direction of the coil winding.) The output of a second order axial gradiometer (figure 12) is equal to:

$$ \Delta^2 \phi = \phi_1 - 2\phi_2 + \phi_3 = \int_{S_1} \mathbf{B}(\mathbf{r}) \cdot \mathbf{n} dS_1 - 2\int_{S_2} \mathbf{B}(\mathbf{r}) \cdot \mathbf{n} dS_2 + \int_{S_3} \mathbf{B}(\mathbf{r}) \cdot \mathbf{n} dS_3, $$

(44)

where $\mathbf{n}$ is a unit vector oriented parallel to the gradiometer axis. The location and orientation of the axial gradiometer are arbitrary, so the surface integrals of equation 43 must be evaluated numerically. Throughout this thesis, this is achieved by means of a

---

Figure 12. Second order axial gradiometer.
24-point Gaussian quadrature method. Local x- and y- axes are set up in the plane of each gradiometer coil, the origin of the axes coinciding with the coil axis. The axial component of the flux density is then integrated over 24 points on each coil surface. The local coordinates of the points of integration on coil \( i \) can be expressed as:

\[
\begin{align*}
  x_{jk} &= p_j R \cos 2\pi q_k, \\
  y_{jk} &= p_j R \sin 2\pi q_k, \\
  (j &= 1..3, k = 1..8), \\
\end{align*}
\]  

(45)

where \( p_j \) is the \( j^{th} \) Gaussian ordinate for a 3-point quadrature, \( q_k \) is the \( k^{th} \) Gaussian ordinate for an 8-point quadrature and \( R \) is the gradiometer coil radius. If the corresponding coordinates in the spherical head system are expressed as \( r_{ijk} \), the gradiometer output is defined as:

\[
e = \pi R^2 \sum_{i=1}^{N_i} T_i \sum_{k=1}^{8} Q_k \sum_{j=1}^{3} P_j \mathbf{B}(r_{ijk}) \cdot \mathbf{n},
\]

(46)

where \( P_j \) is the \( j^{th} \) Gaussian weight for a 3-point quadrature and \( Q_k \) is the \( k^{th} \) Gaussian weight for an 8-point quadrature. This discrete integration can be used to calculate the flux gradient due to a single dipolar current source with an error of less than 0.1% (Roth and Sato 1992).

**Numerical Simulations**

The previous two sections have established formulae describing the electric potential and magnetic flux density around a spherical model of the head. It is instructive to apply these formulae to simple source geometries and observe how the electric potential and magnetic field are distributed spatially. This section therefore describes simulations in which the location and orientation of a single current dipole are varied within a realistically-sized sphere model. The effects of conductivity parameters are also studied briefly.

Mathematical descriptions of the electric forward problem have been proposed for the case of a single spherical shell (Frank 1952), two spherical shells (Geisler and Gerstein 1961), three spherical shells (Rush and Driscoll 1969; Ary et al. 1981) and four spherical shells (Hosek et al. 1978), as well as for an arbitrary number of shells (de Munck 1988). A saline-filled phantom has been used to examine the distribution of electric potential around the head (Rush and Driscoll 1969) and current dipoles have been implanted into monkeys’ heads to observe the external potential (Hosek et al. 1981).
1978). The magnetic field distribution has also been studied mathematically (Cuffin and Cohen 1977) and with the aid of a phantom (Janday and Swithenby 1987).

Some of the most interesting mathematical simulations of the electric forward problem were carried out by Geisler and Gerstein (1961). They considered the angular distribution of electric potential around a single conducting sphere for the case of a single current dipole source of varying eccentricity. They then surrounded the single sphere with a shell whose conductivity was 0.1 times that of the central sphere, to simulate the effect of the skull. The maximum amplitude of the electric potential produced by a radial dipole was found to drop to around 40% of its previous value. Geisler and Gerstein then investigated the effect of a high conductivity shell surrounding the central sphere. The outer shell had a conductivity of 10 times that of the central sphere and was intended to simulate cerebrospinal fluid. Interestingly, the presence of the highly conducting shell caused the maximum potential to drop to only 6% of the previously recorded maximum. The high conductivity shell also caused more angular spread of the potential distribution. These observations indicated the significant influence of the cerebrospinal fluid and skull on externally recorded scalp potentials.

The remainder of this chapter contains a description of comparable simulations carried out by the author. The spherical model described in table 1 was used in the study. Following Geisler and Gerstein (1961), the simulations were performed twice: once with unit isotropic conductivity in all shells and again with the realistic conductivity values of table 1. The unit conductivity model effectively formed a single-sphere homogeneous volume conductor. The value of 1.0Sm⁻¹ was chosen as an order-of-magnitude approximation to the conductivity of the brain. Although not a very exact approximation, it was judged to be adequate given the two orders of magnitude difference between the radial conductivities of the skull and brain. The two conductivity models enabled the effect of the various shells on the electric potential distribution to be illustrated. A single current dipole of 10nAm moment was located in the inner sphere on the z-axis, lying in the xz-plane. Two source eccentricities were used: $r_o = 60\text{mm}$ (superficial source) and $r_o = 30\text{mm}$ (deep source). The source was oriented either radially or tangentially. The surface potential distribution about the y-axis was calculated and the potential distribution over a square grid of side 80mm, centred on the z-axis, was also calculated (figure 13). The results for the superficial dipoles are shown in figures 14 and 15, and those for the deep dipoles are shown in figures 16 and 17.

Referring to figure 14, it can be seen that the maximum potential is lower for both radial and tangential superficial dipoles when realistic conductivities are chosen. In accord with the observations of Geisler and Gerstein (1961), the various anisotropic conductive layers clearly attenuate the surface potentials. The effect is greatest for the radial dipole, where the maximum potential is reduced by a factor of about eight when realistic conductivities are used. In addition to being attenuated, the potential distributions are also spread out by the realistic conductivity shells. This is most noticeable for the tangential dipole, where the extrema of the realistic potential
Figure 13. Simulated geometry for the calculation of electrical potentials on the surface of a spherical model. A current dipole was located on the z-axis, with its moment lying in the xz-plane. The distribution of potential was measured in the xz-plane, with respect to angle a. The potential was also determined across a curvilinear grid centred on the z-axis: a square planar grid was projected downwards onto the surface of the sphere and the potentials were calculated at the intersections of the grid and the sphere. The potentials were plotted against the x- and y-coordinates of the grid points.

distribution have a greater angular separation than those of the unit potential distribution. These effects are borne out by the complete surface potential distributions for the superficial dipoles (figure 15).

The results for the deep current dipoles (figures 16 and 17) differ somewhat from those for the superficial dipoles. Although the potential distribution resulting from a radial source is attenuated slightly by the realistic anisotropic conductivity shells (figure 16a), the distribution from a tangential source (figure 16b) is actually enhanced by the realistic volume conductor by a factor of about four. In neither of these cases does the inhomogeneity produce any significant spreading of the surface potential distribution.

It is concluded from these observations that the choice of conductivity model has a significant effect of the surface potentials emanating from a dipolar current source. It is therefore desirable to use the most accurate volume conductor model available, with the most realistic conductivity parameters. It is difficult to select conductivities, however, owing to the sparse availability of reliable data in the literature.

A comparison is now made between the results for the superficial radial and superficial tangential current dipoles when realistic conductivity values are used. It is
Figure 14. Angular distribution of surface potential due to a superficial current dipole located at \( r_0 = 60 \text{mm} \). Dipole moment 10nAm. (a) Radial dipole, and (b) tangential dipole. Broken lines: model having unit conductivity; solid lines: model having realistic conductivity.
Figure 15. Surface potential distribution due to a superficial current dipole located at \( r_o = 60 \text{mm} \). Dipole moment 10nAm. (a) radial dipole / unit conductivity; (b) radial dipole / realistic conductivity; (c) tangential dipole / unit conductivity; (d) tangential dipole / realistic conductivity. Values in \( \mu \text{V} \).
Figure 16. Angular distribution of surface potential due to a deep current dipole located at $r_0 = 30\text{mm}$. Dipole moment $10\text{nAm}$. (a) Radial dipole, and (b) tangential dipole. Broken lines: model having unit conductivity; solid lines: model having realistic conductivity.
Figure 17. Surface potential distribution due to a deep current dipole located at $r_o = 30\text{mm}$. Dipole moment 10nAm. (a) Radial dipole / unit conductivity; (b) radial dipole / realistic conductivity; (c) tangential dipole / unit conductivity; (d) tangential dipole / realistic conductivity. Values in $\mu\text{V}$. 
noted that for a superficial radial dipole, the maximum surface potential is around 0.6μV, while for a superficial tangential dipole, the potential extrema are approximately ±1.2μV in magnitude. Thus, the tangential dipole elicits potentials which are approximately twice as large as those produced by the radial dipole. Similarly, the maximum surface potential produced by the deep radial source is about 0.7μV, whilst the potential extrema for the deep tangential dipole are around ±2.0μV and are therefore almost three times higher in magnitude than the maximum of the deep radial source. Hence, it is concluded that a tangential dipolar current source produces a greater surface potential than a radial current source, the effect being most marked when the dipole is deep within the volume conductor. This phenomenon can, in fact, be predicted from equation 32, where the first term is the potential contributed by a radial source and the second the potential contributed by a tangential source: the second term is divided by \( r_o \), so that it increases in importance as the dipole approaches the centre of the volume conductor.

Finally, it is instructive to compare the potential distributions produced by the superficial and deep sources. The superficial radial dipole produces a maximum potential of about 0.6μV, while the deep radial dipole produces a maximum potential of 0.7μV, indicating that the potential due to a radial dipole is not affected very much by source depth. The variation in the spread of the distribution with source depth is negligible for the case of a radial source. For the tangential sources, the superficial dipole produces extrema of ±1.2μV, while the deep dipole produces extrema of ±2.0μV. Hence, a tangential source component produces a greater surface potential when the source is deep than when it is superficial. This is a consequence of the division by \( r_o \) in equation 32. There is also a slightly greater angular separation between the extrema of the potential distribution for a deep tangential source, compared to that for a superficial tangential source.

So far, only electric potentials have been considered, but it is also interesting to examine the magnetic flux density arising from a dipolar source in a spherical volume conductor. Hence, the author located superficial \((r_o = 60\text{mm})\) and deep \((r_o = 30\text{mm})\) tangential dipolar current sources on the z-axis within a spherical conductivity model of outer radius 81mm, and the angular distributions of magnetic flux density about the x-axis were calculated (figure 18). The results are shown in figure 19. The complete distribution of magnetic flux density over a square grid was calculated and the results are given in figure 20. Note that as magnetic flux density has both radial and tangential components, the magnitudes of both of these components are shown. Also note that these results are for tangential sources only; a radial source would produce zero magnetic field everywhere outside the volume conductor. Magnetic flux density is also independent of the conductivity profile of the spherically symmetric model, so these results apply to both unit and realistic conductivities.

For both the superficial and deep sources, it can be seen that the radial and tangential components of the magnetic flux density are comparable in magnitude. The
Figure 18. Simulated geometry for the calculation of magnetic flux density on the surface of a spherical model. A current dipole was located on the z-axis, with its moment lying in the xz-plane. The distribution of flux density was measured at the surface of the sphere in the yz-plane, with respect to angle a. The flux density was also determined across a curvilinear grid centred on the z-axis: a square planar grid was projected downwards onto the surface of the sphere and the flux density was calculated at each intersection of the grid and the sphere. The flux densities were plotted against the x- and y- coordinates of the grid points.

The magnitude of the flux density changes considerably with source depth, however. For the superficial dipole at a radius of 60mm, the extreme radial flux density is around 700fT, but for the deep dipole at a radius of 30mm, this drops to about 80fT. Clearly, a superficial source generates a much greater magnetic field than a deep source. This proves to be of great importance in the inverse problem, where a weak, superficial source or a strong, deep source can both explain the same magnetic field distribution.

The superficial source produces a more focal magnetic field distribution than the deep source, the extrema of the radial field distribution occurring at around ±15° as opposed to around ±40° for the deep source. For the superficial source, the zeros of the tangential field distribution occur at around ±20°, but those for the deep source occur around ±50°. Thus, a deeper source produces a field distribution whose extrema are further apart. This can be used as the basis of a crude inversion technique: a current source is deduced to lie midway between the two field extrema of the radial field distribution and the distance between the extrema is used as an indicator of the source depth (Hosaka and Cohen 1976; Ioannides et al. 1990).
Figure 19. Angular distribution of magnetic flux density due to a tangential current dipole of moment 10nAm. (a) Superficial dipole at \( r_0 = 60 \text{mm} \), and (b) deep dipole at \( r_0 = 30 \text{mm} \). Solid lines: radial component of the flux density (positivity indicating flux leaving the sphere); broken lines: modulus of the tangential component of the flux density.
Figure 20. Magnetic flux density distribution due to a tangential current dipole of moment 10nAm. (a) Superficial dipole at $r_o = 60\text{mm}$ / radial component of the flux density; (b) superficial dipole at $r_o = 60\text{mm}$ / modulus of the tangential component of the flux density; (c) deep dipole at $r_o = 30\text{mm}$ / radial component of the flux density; (d) deep dipole at $r_o = 30\text{mm}$ / modulus of the tangential component of the flux density. Values in $\text{fT}$. Positive radial component indicates flux leaving the sphere.
The main features of electric potentials and magnetic fields are summarised in table 2. The reader will notice that the electric and magnetic forward problems are in many ways complementary. For instance, a superficial source produces the greatest magnetic field while a deep source produces the strongest electric field. A radial dipole produces a significant electric potential but no magnetic field, and so on. This section is concluded by noting that the differing properties of the electric and magnetic fields can be used to good effect in the inverse problem, since the two modalities provide complementary information with regard to the underlying sources. This is the basis of subsequent chapters.

<table>
<thead>
<tr>
<th>Electric Potential</th>
<th>Magnetic Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalar.</td>
<td>Vector.</td>
</tr>
<tr>
<td>Depends heavily on conductivity of model.</td>
<td>Independent of volume conductivity.</td>
</tr>
<tr>
<td>Radial dipole produces non-negligible potential.</td>
<td>Radial dipole produces zero magnetic field.</td>
</tr>
<tr>
<td>Tangential dipole produces greater potential than radial dipole.</td>
<td>Tangential dipole produces greater magnetic field than radial dipole.</td>
</tr>
<tr>
<td>Superficial sources generate smaller potentials.</td>
<td>Superficial sources generate larger magnetic fields.</td>
</tr>
<tr>
<td>Deeper sources generate larger potentials.</td>
<td>Deeper sources generate smaller magnetic fields.</td>
</tr>
<tr>
<td>Deeper sources produce potential distributions having greater extremum separation.</td>
<td>Deeper sources produce magnetic field distributions having greater extremum separation.</td>
</tr>
</tbody>
</table>

Table 2. A comparison of electric potentials and magnetic fields for the case of a spherically symmetric volume conductor.
Conclusions

In this chapter, it has been shown that evoked electric potentials and magnetic fields are probably generated by postsynaptic potentials in the cortex of the brain. The pyramidal cells are possible sources of neuro-electromagnetic fields. Primary current sources take the form of current dipoles which, due to the anatomical organisation of the cortex, are directed perpendicularly to the cortical surface.

Mathematical relationships describing the electric potential and magnetic field external to a general inhomogeneous, isotropic volume conductor due to a current dipole source have been discussed. Calculation of the electric potential and magnetic field using the boundary element method has been considered but, for reasons of accuracy, simplicity and computational speed, an anisotropic spherical volume conductor has been chosen as the most appropriate model for the present work. The electric and magnetic forward problems were then discussed in some detail for the specific case of the spherical model. Finally, some simulations based on the spherical model have been used to illustrate the properties of electric surface potentials and external magnetic fields when the source is a single current dipole.

This chapter, therefore, has been concerned with the way in which neural activity in the human brain produces measurable scalp potentials and magnetic fields. The following chapters deal with the opposite procedure: the problem of inferring the locations and directions of the primary current sources from the neuro-electromagnetic fields. Chapter four is a summary of possible methods of solving the inverse problem, while chapter five considers a technique by which the complementary nature of the electric and magnetic fields can be exploited to produce estimates of neural activity with superior accuracy.
4. The Inverse Problem

Introduction

Maxwell's equations indicate that it is possible to determine the current density, $J$, within the brain from measurements of the electric field, $E$, and magnetic field, $B$, around the head. The primary objective of research in neuro-electromagnetism is to use externally measured electric potentials and magnetic fields to deduce the underlying distribution of primary current, $J_p$, which, as was seen in the last chapter, results directly from neural activity. The primary current effectively identifies sites of neural processing, and if its location and strength can be determined, the neurones of the brain can be observed as they process visual information.

Consider the following general (but hypothetical) description of the inverse problem. The electric potential or magnetic flux gradient is measured at about twenty different locations around the head. From these measurements, we would like to calculate the primary current density within each cubic millimetre of the brain. This would provide a detailed image of cortical activity. We therefore write down a set of equations describing the electric potential or magnetic flux gradient measurements in terms of the primary current density in each small cubic volume of the brain.

Immediately there is a problem: there are twenty or so field measurements but there are about a million cubic millimetres in which the current is to be determined. Thus, there are twenty or so equations but about a million unknowns. Clearly, the inverse problem is underdetermined and therefore non-unique. Stating this otherwise, there are an infinite number of primary current distributions which all give rise to the same external electric or magnetic fields, and it is an impossible task to choose between them.

So how is it possible to determine the distribution of primary current density within the brain if the inverse problem cannot be solved? The answer is that by carefully introducing additional physical information in one form or another, it is possible to reduce the general problem described above to one which is mathematically tractable. There are many methods of solving the inverse problem; each of them applies some type of prior assumption, so that the number of unknowns can be reduced below the number of available measurements.
This chapter contains a review of the most important methods of solving the inverse problem. The question of non-uniqueness remains central throughout the chapter. For each method, an indication is given of the way in which prior information is included so that a unique solution can be obtained. To avoid confusion, it is necessary to distinguish between the uniqueness of the general inverse problem and the uniqueness of a particular solution. For example, in the above argument, there are an infinite number of possible current distributions which are consistent with the data; the general solution is non-unique. However, if it were hypothesised that the required solution was the smoothest, or most focal distribution, then a particular - unique - solution could be obtained. Uniqueness generally only occurs when a particular solution is sought; it does not imply that the complete problem is unique.

All of the following methods apply equally to neuro-electric or neuro-magnetic measurements, but it is worth noting that it is not possible to determine the position or magnitude of a radial dipolar current source when the volume conductor is spherical and only magnetic data are being used. This is because a radial current dipole produces no external magnetic field (Roth and Wikswa 1986; de Munck and van Dijk 1991), so that the dipole location and magnitude and the field are mutually independent. Similarly, a loop of primary current is "silent" if only electric data are used in conjunction with a spherical conductivity model (Roth and Wikswa 1986). However, such a loop is physiologically implausible and therefore rarely, if ever, encountered. In solving the inverse problem, it is normal to assume that the magnitude of a radial current dipole is zero. This is effected explicitly in some of the methods, but in others it is a by-product of the numerical algorithm. While it is almost certainly possible to generalise the methods to handle both electric and magnetic data simultaneously, it is assumed that only one modality is being used; a technique for combining electric and magnetic measurements is reserved for the next chapter. To avoid continual reference to electric or magnetic fields, these will sometimes be simply referred to as external measurements. Initially, we are concerned with measurements taken at one particular time instant; for example, at one particular latency during an evoked response.

The simplest and most common approach to solving the inverse problem is to fit an equivalent current dipole to the external measurements.

Equivalent Current Dipole Fitting

Suppose that there are two populations of neurones which form two current dipoles of moment $\mathbf{M}$, separated by a distance $2a$ (figure 21). To make things easier, these are assumed to lie within an infinite, homogeneous medium of isotropic conductivity, $\sigma$. The potential, $V$, at a point, $P$, lying in the same plane as the two current dipoles and displaced from the first current dipole by a vector, $\mathbf{r}_1$, and from the second by a vector, $\mathbf{r}_2$ (figure 21), is then given by (Sarvas 1987):
Figure 21. Geometric construction for calculating the electric potential and magnetic flux density at P, due to two similarly oriented current dipoles of equal moment. Underscores denote vectors, which are denoted by bold type in the text.

\[ V = \frac{1}{4\pi\sigma} \frac{M \cdot r_1}{r_1^3} + \frac{1}{4\pi\sigma} \frac{M \cdot r_2}{r_2^3}, \]  

(47)

\[ = \frac{M}{4\pi\sigma} \left[ \frac{\cos A}{r_1^2} + \frac{\cos B}{r_2^2} \right], \]  

(48)

where \( r_1 = |r_1|, r_2 = |r_2| \) and \( M = |M| \). By making use of the sine and cosine rules, it can be shown that:

\[ V = \frac{Mr \cos C}{4\pi\sigma} \left[ \left( a^2 + r^2 + 2ar \sin C \right)^{\frac{3}{2}} + \left( a^2 + r^2 - 2ar \sin C \right)^{\frac{3}{2}} \right]. \]  

(49)

where \( r = |r| \). Expanding this as a power series in \( 1/r \) produces the following result:

\[ V = \frac{M \cos C}{2\pi\sigma} \left[ \frac{1}{r^2} + \frac{3}{2} \frac{\left( 5\sin^2 C - 1 \right) a^2}{r^4} + \frac{15}{8} \frac{\left( 21\sin^4 C - 14\sin^2 C + 1 \right) a^4}{r^6} + \ldots \right]. \]  

(50)

This is the multipole expansion of the potential at P (Panofsky and Phillips 1955). A similar expression can be derived for the magnetic flux density at P. Referring to figure 21, the flux density at P is given by (Sarvas 1987):

\[ \mathbf{B} = \frac{\mu_0}{4\pi} \frac{M \times r_1}{r_1^3} + \frac{\mu_0}{4\pi} \frac{M \times r_2}{r_2^3}. \]  

(51)

Since P lies in the plane of the current dipole moments, both of the two components of this expression are co-oriented. Then:
\[ |\mathbf{B}| = \frac{\mu_0}{4\pi} M \left[ \frac{\sin A}{r_1^2} + \frac{\sin B}{r_2^2} \right]. \tag{52} \]

from which trigonometry yields the following:

\[
|\mathbf{B}| = \frac{\mu_0}{4\pi} M \left[ \left( \left( a^2 + r^2 + 2ar \sin C \right)^{-2} - r^2 \cos^2 C \left( a^2 + r^2 + 2ar \sin C \right)^{-3} \right)^{1/2} + \left( \left( a^2 + r^2 - 2ar \sin C \right)^{-2} - r^2 \cos^2 C \left( a^2 + r^2 - 2ar \sin C \right)^{-3} \right)^{1/2} \right]. \tag{53} \]

The multipole expansion of this expression is:

\[
|\mathbf{B}| = \frac{\mu_0}{2\pi} M \left[ \frac{\sin C}{r^2} - \frac{3}{2} \left( 5\cos^4 C - 7\cos^2 C + 2 \right) \frac{a^2}{\sin C} \frac{1}{r^4} + \ldots \right]. \tag{54} \]

In both the electric potential and magnetic flux density expansions, the $1/r^2$ term is referred to as the dipolar term, the $1/r^4$ term as the octupolar term, and so on (Nunez 1981). Other terms, which in this case are zero because of the specific source geometry employed, are the monopolar term in $1/r$ and the quadrupolar term in $1/r^3$ (Nunez 1981).

Now consider what happens when the distance $r$ is large compared to the separation of the dipoles, $2a$. The quotients $a^2/r^4$, $a^4/r^6$ etc. in equations 50 and 54 become small compared to $1/r^2$, so that terms in $1/r^4$ and higher can be neglected. Then the potential at P is given by:

\[
\phi = \frac{1}{4\pi \sigma} \frac{2M \cos C}{r^2}, \tag{55} \]

\[
= \frac{1}{4\pi \sigma} \frac{2M \cdot r}{r^3}. \tag{56} \]

The magnitude of the flux density is given by:

\[
|\mathbf{B}| = \frac{\mu_0}{4\pi} \frac{2M \sin C}{r^2}, \tag{57} \]

and since $\mathbf{B}$ is perpendicular to the plane in which the dipole moments lie,

\[
\mathbf{B} = \frac{\mu_0}{4\pi} \frac{2M \times r}{r^3}. \tag{58} \]
Comparing equation 56 with equation 47, and equation 58 with equation 51, it is seen that the electric potential and magnetic flux density at P are approximately equal to that which would be produced by a single current dipole of moment $2\mathbf{M}$, located midway between the two original current dipoles. Hence, this simple analysis shows that two distinct current dipoles can produce distant electric and magnetic fields which appear as though they were generated by a single equivalent current dipole (Nunez 1986). This is, in fact, a more general phenomenon: any assembly of current dipoles possesses such an equivalent current dipole (Nunez 1981). (This is the reason for treating a small population of co-oriented neurones as a single current dipole.) However, the accuracy with which the equivalent current dipole matches the external fields depends upon the extent to which the quadrupolar and higher terms can be neglected. If the location and orientation of the equivalent current dipole can be found, it can be used to account for the electric and magnetic fields. This is the basis of the equivalent current dipole method.

It is very important for the reader to understand that the equivalent current dipole is not necessarily physiologically meaningful. For example, in the above two-dipole example, the equivalent current dipole lies at neither of the two original dipole sites, but midway between them. If the two original dipoles represent sites of cortical activity, the equivalent current dipole has missed them both and, worse still, indicates that there is activity somewhere where there is not! The only time when an equivalent current dipole has a physiological interpretation is when the true generator of the external fields is also a single current dipole (or a small population of neurones which can be treated as such). The quadrupolar and higher terms in the multipole expansion then vanish, and the equivalent current dipole coincides with the true source.

In the above description of the multipole expansion, it was assumed that the true current sources were known, so that the electric potential and magnetic flux density could be calculated and expanded and the equivalent current dipole then determined. In practice, the true sources are unknown, so the equivalent current dipole must be derived directly from the external measurements. This is achieved by finding a single current dipole whose external field (calculated using the forward problem) most closely matches the measured field (Schneider 1972; Henderson et al. 1975; Kavanagh et al. 1978; Ary et al. 1981). Since the external measurements are not linear functions of the dipole position parameters, $r$, $\theta$, $\phi$, a non-linear estimation technique has to be used to find the complete set of dipole parameters, $r$, $\theta$, $\phi$, $M_r$, $M_\theta$, $M_\phi$ (Stok 1987; Oostendorp and van Oosterom 1991). There are several ways of doing this. The most direct is to employ a six-dimensional minimisation procedure to determine all of the six parameters simultaneously, using the chi-square as the function to be minimised (Stok 1987):
\[ \chi^2 = \sum_{i=1}^{N} \left( \frac{f_i - \tilde{f}_i(r, 0, \phi, M_r, M_\theta, M_\phi)}{\sigma_i} \right)^2, \]  

where \( f_i \) is the \( i \)th external measurement, \( \tilde{f}_i(r, 0, \phi, M_r, M_\theta, M_\phi) \) is the \( i \)th computed measurement and \( \sigma_i \) is the standard deviation of the \( i \)th measurement. Several methods are available to do this; suggestions are Powell's method (Powell 1964; Press et al. 1989), which does not require the derivatives, with respect to the six dipole parameters, of the minimisation function, and the Marquardt algorithm (Marquardt 1963), which does require derivatives, but converges faster. It should be noted at this point that if magnetic data are being used, \( \tilde{f}_i = \tilde{f}_i(r, 0, \phi, M_\theta, M_\phi) \), so that \( M_r \) is neglected and the remaining five parameters are searched for (Stok 1987). In the final solution, it is assumed that \( M_r = 0 \).

A more elegant method of finding the equivalent current dipole is to use a combination of non-linear and linear estimation. This method is used by the author in subsequent chapters. A current dipole is assumed to have some arbitrary starting position. The vector of external measurements, \( b \), can then be described in terms of the dipole moment, \( m = (M_r, M_\theta, M_\phi)^T \), as follows:

\[ b = Gm, \]  

where \( b_i = f_i / \sigma_i \) (\( i = 1...N \)) and \( G_{i1} = \tilde{f}_i(r, 0, \phi, 1, 0, 0) / \sigma_i \), \( G_{i2} = \tilde{f}_i(r, 0, \phi, 0, 1, 0) / \sigma_i \) and \( G_{i3} = \tilde{f}_i(r, 0, \phi, 0, 0, 1) / \sigma_i \). This linear system is solved to calculate the dipole moment:

\[ m = G^*b, \]

where \( G^* \) is the Moore-Penrose pseudoinverse of \( G \). For the given dipole position, this provides the dipole moment which best fits the set of external measurements in the sense that it minimises the chi-square function. The dipole location is then adjusted, using a non-linear parameter search, in such a way that it reduces the chi-square for the dipole. At the new location, the linear inversion is repeated to find the new dipole moment. Thus, the method proceeds with alternate non-linear and linear steps. Eventually, the chi-square cannot be minimised further and the dipole is taken to be the equivalent current dipole. A more rigorous analysis shows that the current dipole which minimises the chi-square statistic is in fact the maximum-likelihood estimate of the current source, given the noise in the data (Hämäläinen 1987).

A novel method for dipole parameter estimation has recently been reported by Del Gratta et al. (1993). This overcomes the nonlinear dependence of the external fields on the dipole position by using a linear iterative estimation procedure. In common with the methods already described, the iteration is started with an initial set of dipole parameters. These are then successively adjusted so that the forward field matches the
measured field. The novel feature of the algorithm is that at each stage in the iteration, the forward field is compared with the actual field to produce a residual field distribution. This residual field is expanded by a Taylor series, retaining only the first order term:

\[
\Delta f_i = \frac{\partial \tilde{f}_i}{\partial r} \Delta r + \frac{\partial \tilde{f}_i}{\partial \theta} \Delta \theta + \frac{\partial \tilde{f}_i}{\partial \phi} \Delta \phi + \frac{\partial \tilde{f}_i}{\partial M_r} \Delta M_r + \frac{\partial \tilde{f}_i}{\partial M_\theta} \Delta M_\theta + \frac{\partial \tilde{f}_i}{\partial M_\phi} \Delta M_\phi, \quad i = 1...N,
\]

where \( \Delta f_i \) is the (scalar) residual at the \( i \)th measurement location:

\[
\Delta f_i = f_i - \tilde{f}_i(r, \theta, \phi, M_r, M_\theta, M_\phi).
\]

The partial derivatives are those of the forward field with respect to the dipole parameters, and \( \Delta r, \Delta \theta, \Delta \phi, \Delta M_r, \Delta M_\theta, \Delta M_\phi \) are the first-order corrections to the dipole parameters. Equations 62 constitute a linear system of \( N \) equations in six unknowns. This can be solved by taking the pseudoinverse of the matrix of partial derivatives; the result is the least squares estimate of the corrections which need to be applied to the dipole to remove the residual. The corrected dipole parameters are used as the start of the next iteration. Hence, this algorithm eliminates the need for a conventional parameter search, replacing it with an iterative linear inversion procedure. The method does, of course, require that the partial derivatives of the forward field can be evaluated with respect to the dipole parameters. Del Gratta et al. (1993) claim that the method converges after 3 to 7 iterations, providing an accuracy comparable to that of a conventional dipole fit, but with considerably increased speed.

It is necessary to quantify how well an equivalent current dipole accounts for the measured data, as this indicates whether the single dipole source model is appropriate or not. The definition of the goodness-of-fit of an equivalent current dipole is somewhat arbitrary, but there are two methods in common use. The first measure of goodness-of-fit is defined as:

\[
g = 1 - \frac{\sum_{i=1}^{N} (f_i - \tilde{f}_i)^2}{\sum_{i=1}^{N} f_i^2},
\]

where \( f_i \) is the \( i \)th external measurement and \( \tilde{f}_i \) is the corresponding reconstructed field as produced by the equivalent current dipole. This measure is analogous to that used in linear regression analysis (Kaukoranta et al. 1986). If \( g = 1 \), the model exactly reproduces the data, and if \( g = 0 \), the model does not describe the data any better than if the source were completely absent and the forward field were zero everywhere.

The above criterion is robust and easy to compute, but describes how well the model fits the data including the noise, which is not entirely appropriate. The value of \( g \) is reduced by the presence of noise in the measurements, so that even if the true source
is a current dipole and the equivalent current dipole exactly coincides with it, $g$ is less than one. A more meaningful measure of goodness-of-fit is obtained from the minimised chi-square for the model (equation 59). Noise in the data can cause chance fluctuations in the value of the minimised chi-square, and provided the measurement errors are normally distributed, the probability distribution for these values is given by the chi-square distribution for $N - X$ degrees of freedom, where $N$ is the number of measurements and $X$ is the number of model parameters. The probability that the minimised chi-square exceeds the observed value by chance is then a useful measure of goodness-of-fit (Press et al. 1989). Unit goodness-of-fit indicates that the model is appropriate and a goodness-of-fit approaching zero indicates the converse. The only difficulty with this method is that the goodness-of-fit changes rapidly from zero to one when the standard deviations, $\sigma_i$, of the measurements increase. The standard deviations therefore need to be assessed carefully to avoid reaching false conclusions with regard to the validity of the fitted model (Supek and Aine 1993b). However, this method has the advantage of measuring the goodness-of-fit of the model without being unnecessarily influenced by the noise in the data: a residual which is large because of noise, rather than because of an inappropriate model, is recognised as such. This is because the terms of the chi-square are normalised by the standard deviations, $\sigma_i$, so that a large residual value, $f_i - \tilde{f}_i$, is considered relative to a large $\sigma_i$.

The goodness-of-fit of an equivalent current dipole needs to be interpreted carefully. If the goodness-of-fit is low, the equivalent current dipole is worthless, as it does not adequately describe the data. A poor fit shows that the multipole expansion of the field contains significant quadrupolar and higher terms, which, in turn, indicates the presence of multiple sources. If the fit is high, the field distribution is dominated by the dipolar term of the multipole expansion, in which case the current source is generally a single, localised current dipole or several closely spaced current dipoles, to which the equivalent current dipole accurately corresponds. Hence, the goodness-of-fit is an important indicator of the physiological significance of a current dipole. However, it should be noted that the goodness-of-fit is not infallible. It is possible for several spatially separate sources to produce an external field which is dominated by the dipolar term of the multipole expansion. This type of field appears to have been generated by a single current dipole, and an equivalent current dipole can be fitted with a high goodness-of-fit. Two side-by-side dipoles in a spherical conductor are an example of this phenomenon. The two dipoles produce an external magnetic field which can be closely modelled by a deeper, stronger equivalent current dipole. The goodness-of-fit of this equivalent current dipole is high, indicating only a very small modelling error (Hari et al. 1988). A similar effect occurs for a distribution of side-by-side dipoles along a line. Thus, it is possible to obtain a high goodness-of-fit even when the equivalent current dipole does not coincide with the true sources: the equivalent current dipole is actually a good model to describe the data, but it is not the original source configuration!
The advantage of equivalent current dipole fitting is that there are only six parameters (five in the case of magnetic data) to be determined from at least twenty measurements. The equivalent current dipole is generally unique since the number of measurements considerably exceeds the number of unknowns. Moreover, because there are so few parameters to be determined, the solution is relatively robust to noise. The method is popular because of these advantages, but a multiple-dipole model must be used if it is suspected or known that there are several sources.

Spatio-temporal Dipole Models

In principle, if there are 20 measurements, it is possible to fit up to three current dipoles to the data, giving 18 unknowns. However, in practice, this is unsatisfactory as the chi-square function contains a multiplicity of local minima, many of which are almost as prominent as the overall minimum. Hence, a fitting procedure may find a solution, but this may not be the best solution. A typical problem encountered with this type of method is when the source model contains two dipoles. The minimisation procedure may locate the dipoles close together, such that they are oriented in opposite directions and have large, but almost cancelling, moments. The solution will be significantly affected by the noise in the data, and will therefore be unreliable (Press et al. 1989).

By making use of the external field distribution at multiple time instances, it is possible to increase the number of data available for inversion. This facilitates the use of a more comprehensive dipole model. The more dipoles there are in an inversion procedure, the more likely it is that all possible generators of the external fields are modelled closely and accurately. The spatio-temporal approach uses as many as five dipoles to model the current sources; the increased number of measurements which are necessary to determine all the parameters come from a temporal sequence of spatial field distributions (Scherg and Von Cramon 1986; Scherg 1990; Mosher et al. 1990; Supek and Aine 1993a). The method is based on the premise that an evoked response is generated by several distinct populations of neurones, each of which can be modelled by an equivalent current dipole. The neurone populations can be active separately or simultaneously: they are mutually independent. They may represent, for example, several visual areas, or several spatially distinct regions of the primary visual area. Because the regions have fixed locations, the corresponding current dipoles are also assumed to be stationary. The dipoles are allowed to rotate to accommodate changes in primary current direction brought about by activity shifting between adjacent, differently oriented, populations of neurones (Scherg et al. 1989; Scherg 1990).

Having established the physical model, it is necessary to determine the locations and moments of the constituent dipoles. First, a range of latencies are selected from the evoked electric or magnetic field (Aine et al. 1993). The latencies may span all of the major peaks of the recorded waveforms, or some subset of these peaks. The larger the
latency range, the more accurately the dipole parameters can be determined, as the
available information increases. However, care must be taken to avoid analysing periods
of noise which, besides not being of interest, may cause the dipole locations to be in
error. Next, the dipole locations must be found. The procedure for doing this is similar
to that of the single equivalent current dipole method. The \( D \) dipoles are assumed to
have arbitrary starting positions. These are typically found by performing single-dipole
analyses at various times within the selected latency range to give an idea of the possible
locations involved in the response (Aine et al. 1993; Supek and Aine 1993a). The \( D \)
dipoles are located, as far as possible, at all sites of activity.

An iterative process is used to refine the locations of the dipoles (Mosher et al.
1990; Supek and Aine 1993a). For a particular set of dipole locations, a linear system
describing the spatio-temporal data in terms of the current dipole moments is created:

\[
\mathbf{F} = \mathbf{GM},
\]  

(64)

where \( \mathbf{F} \) is an \( N \times T \) matrix representing the \( N \) field measurements at each of \( T \) latencies:
\( F_{it} = f_{it}/\sigma_i \) (\( i = 1...N, t = 1...T \)) where \( f_{it} \) is the measured field at location \( i \) and time \( t \),
and \( \sigma_i \) is the standard deviation of the measurements at the \( i \)th location. Matrix \( \mathbf{G} \),
which has dimensions \( N \times 3D \) (\( N \times 2D \) in the case of magnetic data), describes the field
at each of the \( N \) measurement sites due to each orthogonal unit component of the \( D \)
dipoles; the \( i \)th row is divided by \( \sigma_i \). Matrix \( \mathbf{M} \) is a \( 3D \times T \) matrix giving the \( 3D \) dipole
components at each of the \( T \) time instants. This system is the spatio-temporal equivalent
of equation 60. It is solved as in equation 61, viz.:

\[
\mathbf{M} = \mathbf{G}^* \mathbf{F},
\]  

(65)

where \( \mathbf{G}^* \) is the pseudoinverse of \( \mathbf{G} \). The solution is a maximum-likelihood estimate of
the set of dipole moments at each of the \( T \) time instants, for the given dipole locations.
The estimate minimises the chi-square:

\[
\chi^2 = \sum_{i=1}^{N} \sum_{t=1}^{T} \left( \frac{f_{it} - \tilde{f}_{it}(\mathbf{M})}{\sigma_i} \right)^2 ,
\]  

(66)

where \( f_{it} \) is the external measurement at location \( i \) and time \( t \), and \( \tilde{f}_{it}(\mathbf{M}) \) is the
corresponding reconstructed measurement as produced by the \( D \) dipole moments.

The locations of the \( D \) dipoles are successively adjusted in a \( 3D \)-dimensional
parameter search, so as to minimise the chi-square (Supek and Aine 1993a). When the
chi-square cannot be reduced further, the dipoles are deduced to be located at the sites
responsible for the measurements. The dipole moments have already been calculated
during the last step of the iteration, as the matrix \( \mathbf{M} \) provides the moment for each
dipole at each of the time instants. Therefore, the spatio-temporal model, which in terms of solution method is an extension of the single dipole approach, conveniently provides the fixed locations of a set of dipoles, and the varying orientations and magnitudes of these dipoles over a relevant period of time.

The obvious advantage of this method is that the number of available measurements increases from \( N \) to \( T \times N \) with the use of temporal information. The number of current sources can therefore be correspondingly greater, ensuring that the model accurately represents the neural activity. The major difficulty, however, lies in the execution of the parameter search. Because the external field is not linear with respect to the dipole locations, the chi-square is not quadratic throughout the 3D-dimensional space of dipole position coordinates (Lewis et al. 1992). Instead, it is a complex surface which contains many local minima. The only fail-safe means of searching for the global minimum is to exhaustively search all of the parameter space, thereby considering all possible combinations of dipole location. The time required for this is prohibitive. The practical alternative is to start a series of parameter searches from a variety of starting points in parameter space and declare the global minimum to be the one which is found by the majority of searches, a procedure which is generally inconsistent (Lewis et al. 1992). Hence, this method, as with all the inversion methods, has its weaknesses as well as strengths.

The model described above is the most common spatio-temporal model. An alternative model has been described by Scherg (Scherg 1984; Scherg and Von Cramon 1985a,b). In this model, both the dipole location and dipole orientation are fixed throughout the time interval. In addition, the timecourse of the magnitude of each dipole is constrained to be either monophasic, biphasic or triphasic. A biphasic discharge pattern can be defined by the amplitudes and latencies of the two peaks, and the latencies of the discharge onset and offset: a total of six parameters. Similarly, a monophasic waveform can be defined by four parameters and a triphasic waveform by eight. These temporal waveforms are synthesised using spline functions. Thus, the entire spatio-temporal activity of a current dipole can be typically defined by eleven parameters: three location parameters, two direction parameters and six timecourse parameters (assuming a biphasic waveform). A source model containing \( D \) dipoles therefore consists of \( 11D \) unknowns, to be determined from the \( N \times T \) measurements. A reduction in the number of unknown parameters can be achieved by assuming, for instance, that the sources are located symmetrically about the medial plane of the head. In this case, only 19 parameters are required to define two dipoles. Further reductions in the number of parameters can be made by assuming that the timecourses of a pair of symmetric dipoles are mutually dependent.

Each combination of the \( 11D \) parameters produces a particular spatio-temporal forward field. The parameters are determined by using a least-squares fitting procedure (Deuffhard and Apostolescu 1980) to minimise the difference between the model and recorded waveforms at all measurement sites and all latencies simultaneously (Scherg
and Von Cramon 1985a). If $\mathbf{F}$ is the $N \times T$ spatio-temporal data matrix, and $\hat{\mathbf{F}}$ is the corresponding reconstructed matrix produced by the 11D parameters, then the solution minimises $\|\mathbf{F} - \hat{\mathbf{F}}\|_F$, where $\|\cdot\|_F$ denotes the Frobenius norm. Starting values for the dipole locations, orientations and temporal activations are found by hypothesis testing.

This is an interesting variant of the spatio-temporal model, but it is difficult to justify a particular choice of temporal dipole activation. It is also difficult to justify any specific assumptions made regarding the bilateral nature of the sources, or the mutual dependence of their timecourses. The spatio-temporal model described previously makes no such assumptions and is therefore a more general and reliable model, but the larger number of unknown parameters makes parameter searching more difficult.

**Principal Components Analysis**

The difficulties associated with searching for the parameters of a conventional spatio-temporal source model can be overcome by using a principal components analysis. This utilises a source model consisting of a series of current dipoles whose locations and orientations are fixed throughout a time interval, but whose strengths are variable. The method determines the number of uncorrelated current sources required to describe the spatio-temporal data matrix and finds their locations, orientations and timecourses (Maier et al. 1987). The ability of an algorithm to select the minimum number of current sources that are required to account for the external data is important. If there are too few current dipoles, the model deviates from reality and does not adequately describe the data. If there are too many dipoles of fixed location in a spatio-temporal model, the problem is not so severe, but a conventional parameter search becomes more difficult and the results are likely to be less reliable (Supek and Aine 1993b). Hence, a method such as principal components analysis, which can determine the number of current sources, is of value.

The concept of principal components analysis is as follows (John et al. 1964). Any recorded waveform sampled at $T$ time instants can be viewed as a vector in $T$-dimensional space. $N$ of such waveforms span a subspace of dimension $N$ or less, depending on the degree to which the waveforms are correlated. (If $N > T$, the maximum dimension of the subspace is $T$.) Principal components analysis determines the dimensionality of the subspace, which indicates the number of uncorrelated current sources. It then computes an orthonormal basis for the waveform subspace, the base vectors being the principal components. Hence, any of the $N$ recorded waveforms, $f_i(t)$, can be expressed as a linear combination of the principal components:

$$f_i(t) = \sum_{j=1}^{r} a_{ij} P_j(t), \quad i = 1 \ldots N, \ t = 1 \ldots T,$$  \hspace{1cm} (67)
where \( r \) is the dimensionality of the signal subspace, \( P_j(t) \) is the \( j \)th principal component, and \( a_y \) is the magnitude of the projection of \( f_i(t) \) onto \( P_j(t) \) (figure 22). Conceptually, the principal components are fitted to the \( N \) waveforms by least squares fitting (John et al. 1964). The first principal component, \( P_1(t) \), is the best fit of a single waveform to the entire set of measured waveforms in the sense that it maximises the sum of the squared projections of the waveforms onto \( P_1(t) \), i.e. it maximises

\[
\sum_{i=1}^{N} a_i^2.
\]

The second principal component is the best fit to the residual from the first fit; this factor maximises

\[
\sum_{i=1}^{N} a_{i2}^2,
\]

subject to the constraint that \( P_2(t) \) is orthogonal to \( P_1(t) \). The third principal component is the best fit to the residual from the second fit, and so on (see figure 22).

Figure 22. Visualisation of principal components analysis. Two waveforms, \( f_1(t) \) and \( f_2(t) \), each sampled at \( T \) time instants, can be represented as vectors in \( T \)-dimensional space. The principal components \( P_1(t) \) and \( P_2(t) \) form an orthonormal basis for the space spanned by \( f_1(t) \) and \( f_2(t) \). The waveforms \( f_1(t) \) and \( f_2(t) \) can therefore be constructed by a linear combination of \( P_1(t) \) and \( P_2(t) \). The first principal component, \( P_1(t) \), maximises \( a_{11}^2 + a_{21}^2 \), while the second principal component, \( P_2(t) \), maximises \( a_{12}^2 + a_{22}^2 \) and is orthogonal to \( P_1(t) \). After John et al. (1964).
Once the principal components of the spatio-temporal data matrix have been obtained, it does not follow that each linearly independent component is associated with a single current source. In other words, the principal components are not necessarily physiologically meaningful (Maier et al. 1987). Hence, it is necessary to apply some rotation to the principal components to make them so; this rotation can be visualised as a rotation of the orthonormal basis in T-dimensional space. The rotation is chosen according to any specific prior information which may be known about the sources. The usual assumption is that the sources are dipolar (Maier et al. 1987), although other rotations are possible (Wood and McCarthy 1984). The result is a new set of components, $P_j^t(t)$, and corresponding weights, $a^t_j$, such that:

$$f_i(t) = \sum_{j=1}^{r} a^t_j P_j^t(t), \quad i = 1...N, \quad t = 1...T,$$

(68)

where the $P_j^t(t)$ are now not necessarily orthonormal. Then the vector $(a^t_{ij}, a^t_{ij}, a^t_{ij}, \cdots, a^t_{ij})^T$ represents the time-invariant spatial distribution of the $j^{th}$ ($j = 1...r$) physiologically meaningful data component and $P_j^t(t)$ represents the timecourse of the corresponding source. If the assumption underlying the rotation is that the sources are dipolar, the rotation is chosen so that a minimal residual is obtained when a single current dipole is fitted to each of the $r$ spatial distributions (Maier et al. 1987).

In reality, the above procedure is slightly more complicated as the spatio-temporal data include noise. The analysis can be accomplished conveniently using matrix methods, as follows (Maier et al. 1987):

1. The $N \times T$ spatio-temporal data matrix, $F$, is first decomposed by singular value decomposition:

$$F = U \Sigma V^T,$$

(69)

where $U$ and $V$ are orthogonal matrices and $\Sigma$ is a diagonal matrix containing the singular values of $F$. The dimension of $\Sigma$, and hence the maximum dimensionality of the waveform space, is $S = N$ if $N \leq T$ and $S = T$ if $N > T$. The columns of $U$ are the spatial distributions of the principal components, and the columns of $V$ are the corresponding time functions. The value $\Sigma^2_j$ can be interpreted as the power of the $j^{th}$ ($j = 1...S$) principal component (Maier et al. 1987). (Note that if some of the waveforms are linearly dependent, some of the components may contribute zero power.)

2. Principal components which do not contribute significantly to the total power of the data are defined as noise. These components form a noise subspace which is
orthogonal to the signal subspace. The partition between the signal and noise subspaces is determined from the matrix of singular values, such that:

\[
\sum_{j=1}^{r} \Sigma_{jj}^2 < (1 - \gamma) \sum_{j=1}^{S} \Sigma_{jj}^2,
\]  

(70)

where \( r \) is the dimensionality of the signal subspace, which represents the number of independent sources (Maier et al. 1987). The value \( \gamma \) is the fraction of a recorded waveform which is attributable to noise; this figure is typically around 10%. The first \( r \) values of \( \Sigma \) are retained and the remaining \( S - r \) values are set to zero (Mosher et al. 1992). The resulting truncated matrix is referred to as \( \Sigma' \), and the columns of the matrix \( W = \Sigma \Sigma' \) are then said to be the "factor loadings." The first \( r \) columns of \( U \) represent the spatial distribution of the useful principal components, while the first \( r \) columns of \( V \) are the corresponding time functions.

3. A single dipole is fitted to each of the \( r \) rotated principal components of the signal subspace. The \( j \)th dipole location and orientation are chosen as a least squares fit to a linear combination of the factor loadings. That is, the dipole location and orientation minimise the measure (Mosher et al. 1992):

\[
|G_j(L_j)M_j - WC_j|^2,
\]

where \( G_j(L_j) \) is the forward gain matrix for a dipole at location \( L_j \), and \( M_j \) is the dipole moment. \( C_j \) is the \( j \)th column of the \( S \times S \) rotation matrix, \( C \). The term \( WC_j \) is a linear combination of the columns of \( W \), the coefficients for the combination being contained in \( C_j \). Hence, \( WC \) represents the rotated spatial distribution of principal components. The value of \( C \) can be precomputed according to a prior source assumption or according to a desired rotation. For example, Achim et al. (1988) chose \( C_{jj} = 1 \) and \( C_{jk} = 0 \) for \( j \neq k \). In this case, a single dipole was fitted to each column of \( W \). Others have examined the possibility of Varimax rotation (Wood and McCarthy 1984; Möcks and Verleger 1986). More successfully, the columns \( C_j \) can be determined during the minimisation procedure, such that they reduce the residual between the dipole forward field and the linear combination of factor loadings (Maier et al. 1987). The rotated principal components then have the property of being those which are the most dipolar in nature.

4. The rotated version of \( V \) is calculated as:

\[
V' = (C^{-1}V)^T.
\]  

(71)
The columns of $V'$ are then the time functions of the sources.

This method has the advantage of being able to determine the number of sources as well as their locations, orientations and time functions. The complete spatio-temporal problem is broken down into the simpler problem of fitting a single current dipole to each of $r$ spatial component distributions, each of which consists of $N$ values. The disadvantage is that the rotation matrix has to be determined either before or during the fitting procedure, in addition to the dipole parameters. Although Maier et al. (1987) appeared to achieve this successfully, it has been shown theoretically (Mosher et al. 1992) and by simulations (Achim et al. 1988) that incorrect choice of rotation matrix can result in the mislocation of the current sources. Nevertheless, principal components analysis is a useful alternative to conventional spatio-temporal parameter estimation techniques.

**Multiple Signal Classification (MUSIC)**

Another spatio-temporal source estimation method, which is closely related to principal components analysis, is the MUSIC algorithm (Mosher et al. 1992). This algorithm was first used to determine the number and direction of a series of wavefronts arriving at an array of antennas (Schmidt 1986). It is therefore natural to use it to determine the number of dipoles responsible for an external field, and the locations of these sources.

A spatio-temporal source model can be expressed as a linear system (Lewis et al. 1992):

$$f_i(t) = \sum_{j=1}^{D} G_i(\beta_j) M_j(t), \quad i = 1 \ldots N, \quad t = 1 \ldots T,$$

where $f_i(t)$ is an external measurement at location $i$ and time $t$, $M_j(t)$ is the magnitude of the $j^{th}$ dipole moment at time $t$, and $G_i(\beta_j)$ is a gain factor, which is a function of the $i^{th}$ measurement location and the parameters, $\beta_j$, of the $j^{th}$ dipole. This can be encoded in matrix form:

$$
\begin{bmatrix}
  f_1(t) \\
  f_2(t) \\
  \vdots \\
  f_N(t)
\end{bmatrix} =
\begin{bmatrix}
  G(\beta_1) & G(\beta_2) & \cdots & G(\beta_D)
\end{bmatrix}
\begin{bmatrix}
  M_1(t) \\
  M_2(t) \\
  \vdots \\
  M_D(t)
\end{bmatrix},
$$

or:

$$
\begin{bmatrix}
  f_1(t) \\
  f_2(t) \\
  \vdots \\
  f_N(t)
\end{bmatrix} =
\begin{bmatrix}
  G(\beta_1) & G(\beta_2) & \cdots & G(\beta_D)
\end{bmatrix}
\begin{bmatrix}
  M_1(t) \\
  M_2(t) \\
  \vdots \\
  M_D(t)
\end{bmatrix}.
$$
\[ F(t) = G(\beta)M(t), \quad (73b) \]

where \( F \) is the spatio-temporal data matrix, \( M \) is a matrix of dipole moments and \( G \) is a gain matrix (Lewis et al. 1992). \( G(\beta) \) is a column vector representing the external measurements that would result from a unit dipole having parameters \( \beta_j \). It is a function of the measurement locations and the parameters, \( \beta_j \), of dipole \( j \).

Any type of spatio-temporal model can be encoded into this form (Mosher et al. 1992). Examples are (a) multiple dipoles which are free to move and rotate, (b) dipoles which are fixed in location throughout a time interval but unconstrained in orientation, and (c) dipoles whose location and orientation is fixed. Strictly speaking, mathematical expression of model (a) involves that matrix \( G(\beta) \) is a function of time, as the parameters \( \beta \) vary with time. However, the model can be incorporated into the above mathematical framework by dividing it into its constituent sequence of time instants and analysing each time instant separately (Mosher et al. 1992). The MUSIC algorithm is capable of determining the number of dipoles, \( D \), responsible for a spatio-temporal data set, and the locations and orientations of the dipoles throughout the time interval in question. The complete method is complex, but the following geometrical interpretation (Schmidt 1986) provides an insight into the operation of the algorithm. For simplicity, it is assumed that external measurements are obtained at one time instant only. The spatio-temporal source model then reduces to a simple instantaneous multiple-dipole model.

Suppose that \( D \) dipoles produce a set of \( N \) field measurements, \( f_n \), so that:

\[
\begin{bmatrix}
    f_1 \\
    f_2 \\
    \vdots \\
    f_N
\end{bmatrix} =
\begin{bmatrix}
    G(\beta_1) & G(\beta_2) & \cdots & G(\beta_D)
\end{bmatrix}
\begin{bmatrix}
    M_1 \\
    M_2 \\
    \vdots \\
    M_D
\end{bmatrix},
\quad (74a)
\]

or:

\[ F = GM, \quad (74b) \]

where \( G(\beta) \) is a column vector giving the gain of the \( j^{th} \) \( (j = 1...D) \) dipole and \( M_j \) is the strength of the \( j^{th} \) dipole (Schmidt 1986). The measured vector, \( F \), can be visualised as a vector in \( N \)-dimensional space. The gain vectors \( G(\beta) \) can also be viewed in this way. Equation 71a states that \( F \) is a particular linear combination of the gain vectors, with the elements of \( M \) being the coefficients of the combination. The measurement vector, \( F \), must lie within the range space of \( G \), which is commonly referred to as the signal subspace. For example, if there are three dipoles, so that \( G \) has three columns, the
range of $G$ is a three-dimensional subspace within the $N$-dimensional space, and $F$ must lie within the subspace. An array manifold is defined as the continuum of all field vectors that can be generated by a single source located anywhere within the brain, or scan region. As a consequence of the non-linear dependence of the field on the parameters $\beta$, the array manifold is a contorted, non-linear region. In terms of this geometrical picture, the problem of numbering and locating multiple sources consists of locating the intersections of the array manifold with the signal subspace. The MUSIC algorithm finds the signal subspace from the measured data, determines its dimensionality, which defines how many sources are present, and then obtains its intersections with the array manifold, to determine the source locations (Schmidt 1986).

In simple terms, the MUSIC algorithm proceeds by scanning a single current dipole through the whole of the scan region (Lewis et al. 1992). For each point in the scan region, the forward field is calculated; this field corresponds to a point on the array manifold in $N$-dimensional space. The field is projected onto the signal subspace and the magnitude of the projection is normalised and squared to provide a measure of source presence (Lewis et al. 1992). This measure ranges from zero at a location that produces a forward field vector which is orthogonal to the signal subspace, to one at a true source location, where the forward field due to the single current dipole lies within the signal subspace. The complete scan defines the array manifold and its intersections with the signal subspace.

The principles of the method are illustrated elegantly by Lewis et al. (1992). Two tangential dipoles are located within a spherical conductor at $\theta = 12^\circ$, $\phi = -180^\circ$ and $\theta = 8^\circ$, $\phi = 0^\circ$, respectively, relative to an arbitrary spherical coordinate system. The first source has a strength of two units and the second, one unit. Both sources are located at a radius of 0.85 units. The radial component of the magnetic field is computed at 13 locations on the surface of the sphere, which has unit radius. The measurement sites are located between $\theta = 30^\circ$, $\phi = -180^\circ$ and $\theta = 30^\circ$, $\phi = 0^\circ$, at 5$^\circ$ intervals (figure 23a). A single dipole is scanned between $\theta = 45^\circ$, $\phi = -180^\circ$ and $\theta = 45^\circ$, $\phi = 0^\circ$ at the source radius. A least-squares measure is used to quantify the goodness-of-fit of the computed field to the simulated measurements. A value of one indicates an exact fit; zero indicates a poor fit. It is found that a single dipole scan using a least squares measure successfully locates the stronger of the sources, but not the weaker (figure 23b). However, when a MUSIC scan is applied, the MUSIC measure peaks at the angles of both of the sources: it is capable of determining the multiple sources correctly (figure 23c). A second simulation consists of a single unit current dipole at $r = 0.60$ units, $\theta = 5^\circ$, $\phi = 0^\circ$. Two measurements are simulated from this source, one at $\theta = 10^\circ$, $\phi = -180^\circ$ and the other at $\theta = 10^\circ$, $\phi = 0^\circ$ (figure 23d). The two-dimensional signal space can thus be conveniently viewed (figure 23e). This figure shows the projection of the computed field of the scan dipole onto the signal subspace. The signal subspace is a straight line through the origin in this case, as varying the strength of the single simulated source merely increases or decreases the two simulated
Figure 23. Interpretation of the MUSIC algorithm. (a) Thirteen measurements are simulated from two dipolar current sources, $S_1$ and $S_2$, which have moments of two units and one unit, respectively. A single current dipole is then scanned through an angle of 90° to locate the sources using the measurements. (b) Scan using a conventional least-squares measure (LSM), which locates the strongest source, $S_1$, but fails to locate $S_2$. (c) Scan using the MUSIC measure (MM), which peaks at the locations of both $S_1$ and $S_2$. (d) Second simulation using one source, $S$, and two simulated measurements. A single dipole is again scanned through 90° at the source depth. (e) Normalised measurement space, showing the MUSIC measure as the projection of the array manifold vector onto the signal subspace. This projection has unit magnitude at the location of $S$, when the array manifold intersects the signal subspace. (f) The corresponding MUSIC scan. After Lewis et al. (1992).
measurements proportionately. The figure clearly shows that the projection of the array manifold onto the signal subspace is maximal when the array manifold and signal subspace are coincident.

It is worth noting that when an array manifold vector projects entirely onto the signal subspace, none of it projects onto an orthogonal subspace. Such an orthogonal subspace is referred to as a noise subspace; the projection of the array manifold onto this subspace vanishes at the site of a source. The projection onto the noise subspace can therefore be used as an alternative to the signal subspace projection (Lewis et al. 1992).

The preceding illustrations have indicated the mode of operation of the MUSIC algorithm for the simple case of a single time instant. As was mentioned earlier, the algorithm applies more fully to a complete spatio-temporal model. The various steps of the method are therefore outlined for the full spatio-temporal case (Mosher et al. 1993):

1) A singular value decomposition is performed on the spatio-temporal data matrix, such that:

\[ \mathbf{F} = \mathbf{U}\Sigma\mathbf{V}^T. \]  \hspace{1cm} (75)

The rank, \( r \), of the signal subspace is determined from the sudden drop in the singular value spectrum, \( \Sigma \). Matrix \( \mathbf{U} \) is partitioned into a basis, \( \mathbf{U}_s \), for the signal subspace and a basis, \( \mathbf{U}_N \), for the noise subspace, such that:

\[ \mathbf{U} = [\mathbf{U}_s, \mathbf{U}_N]. \]  \hspace{1cm} (76)

2) The forward problem is calculated for three orthogonal unit current dipoles at each point in a fine grid covering the scan region of the head. This produces a gain matrix, \( \mathbf{G} \). Its singular value decomposition is found:

\[ \mathbf{G} = \mathbf{U}_G\Sigma_G\mathbf{V}_G^T. \]  \hspace{1cm} (77)

For each location in the scan grid, the MUSIC measure is calculated:

\[ K_f = \lambda_{\text{max}} \left\{ \mathbf{U}_G^T \mathbf{U}_s \mathbf{U}_s^T \mathbf{U}_G \right\}, \]  \hspace{1cm} (78)

which is the maximum eigenvalue of the bracketed expression. \( K_f \) is a measure of where fixed-orientation sources are located within the scan grid.

3) At each maximum of \( K_f \), a further measure is used to determine whether several closely spaced orthogonal dipoles actually summate to produce a rotating dipole. The measure is:
\[
K_r = \frac{\|U_5 U_G \Sigma_G\|_F^2}{\|\Sigma_G\|_F^2},
\]

where \(\cdot\|_F\) is the Frobenius norm. For locations where \(K_r = K_f = 1\), it is concluded that several fixed-orientation dipoles produce a rotating dipole. Using the fixed dipole measure, \(K_f\), and the rotating dipole measure, \(K_r\), a total of \(r\) sources are found, where a fixed dipole counts as one source, and a rotating dipole as two.

In summary, the MUSIC method is a means of locating a number of current dipoles whose locations are fixed, but whose orientations can be either fixed or rotating. A source model consisting of dipoles which are unconstrained in both location and orientation can be accommodated by applying the MUSIC method to each time instant consecutively (Mosher et al. 1992).

The advantages of the method are obvious: it will determine (a) how many sources are needed to describe the data, and (b) where they are located. It does this within a well-defined mathematical framework. However, despite the ability of MUSIC to select the number of sources, the number of dipole parameters cannot exceed the number of available data, \(T \times N\), nor can the number of dipoles exceed the length of the singular value spectrum of the spatio-temporal data matrix, which in turn is equal to the number of external spatial measurements, \(N\). In particular, as the signal-to-noise ratio decreases, the rank of the signal subspace decreases, so that the number of dipoles used to describe the data also decreases.

There are also several ambiguities which can arise during the MUSIC procedure (Lewis et al. 1992). The first is when the normalised projection onto the signal subspace reaches unity somewhere other than at a true source location. This occurs when a single source at the scan location generates the same external field distribution as that produced by a linear combination of two or more true sources. Essentially, the single source at the scan location could describe the measured data as well as a series of true sources. The ambiguity is detected when there are more peaks in the projection measure than dimensions in the signal subspace. The difficulty cannot be overcome, apart from recording extra measurements.

A second type of ambiguity occurs when the timecourses of the \(D\) dipoles are not linearly independent, so that a combination of \(D - 1\) timecourses equals the \(D^{th}\) timecourse (Lewis et al. 1992). An example of this is when two sources are synchronous. Under these circumstances, the rank of the singular value spectrum is less than the true number of sources and the complete signal subspace cannot be determined. The problem cannot be resolved by altering the number of measurements or the geometry of the measurement locations. Instead, it must be overcome by making the source model explicitly synchronous. Despite these difficulties, however, the MUSIC
algorithm probably represents the most advanced of the spatio-temporal inversion methods.

**Minimum Norm Estimation**

It was suggested in the previous section that it is important for a current source model employing discrete dipoles to contain at least as many dipoles as there are active neurone populations. This is necessary if the current reconstruction is to accurately reflect the true primary current. One means of ensuring that this requirement is always fulfilled is to make no prior assumption about the number and spatial extent of the current sources, but to model the primary current as a spatially continuous distribution. The magnitude of the current at each point in the brain is then completely arbitrary, and because there is no constraint on the spatial arrangement of sources, any number of sources can be accommodated.

The difficulty with this type of distributed-source model is that an infinite number of parameters are required to specify a continuous current distribution. Therefore inversion of a given data set requires that an infinite number of parameters are determined from a finite number of measurements, as remarked in the introduction to this chapter. The current distribution clearly cannot be determined uniquely, so some additional constraint must be applied to the problem.

The minimum norm estimate (Hämäläinen and Ilmoniemi 1984) is the original and most common method of estimating a primary current distribution from a set of electric or magnetic measurements. Of the infinity of current distributions which are compatible with the data set, this method selects the one having the minimum Euclidean norm.

Minimum norm estimation is based on the concept of the lead field of an electrode or magnetic flux gradiometer. If there is some distribution of primary current, \( \mathbf{J}_p(r) \), throughout a current source region, \( G \), then the lead field, \( \mathbf{L}_i(r) \), of detector \( i \) is defined by:

\[
 f_i = \int_{G} \mathbf{L}_i(r) \cdot \mathbf{J}_p(r) dG,
\]

where \( f_i \) is the output of detector \( i \) (Hämäläinen and Ilmoniemi 1984). The lead field, \( \mathbf{L}_i(r) \), describes the sensitivity of detector \( i \) to primary current \( \mathbf{J}_p(r) \). Equation 80 states that the contribution of \( \mathbf{J}_p(r) \) to the output of detector \( i \) is equal to the projection of \( \mathbf{J}_p(r) \) onto the lead field at \( r \). In other words, if there is unit current density at \( r \), flowing in the direction of \( \mathbf{L}_i(r) \), then the magnitude of \( \mathbf{L}_i(r) \) is the resulting output of detector \( i \). The lead field is calculated using the forward problem and is thus a function of the conductivity model employed and the sensor characteristics.
A given primary current distribution can be represented as a vector in an infinite-dimensional space (Hämäläinen and Ilmoniemi 1984). Each dimension in the space represents one of the three orthogonal components of the current density at one of the infinity of points in the current region. The lead fields of the $N$ detectors span, in general, an $N$-dimensional subspace within the infinite space. The detectors measure the projection of the primary current onto this subspace; they yield no information about the component of the current which is orthogonal to the subspace. A consequence of this is that only currents which lie entirely within the lead field subspace can be fully detected. The principle of minimum norm estimation is that since no information is available about the component of the current which is orthogonal to the lead field subspace, only the projection of the current onto the subspace is reconstructed. A minimum norm estimate therefore lies in the subspace spanned by the detector lead fields and is a linear combination of the lead field vectors (Hämäläinen et al. 1993; figure 24). Denoting the minimum norm estimate by $\mathbf{J}^*(r)$, we have:

$$
\mathbf{J}^*(r) = \sum_{j=1}^{N} w_j \mathbf{L}_j(r),
$$

(81)

Figure 24. A distribution of primary current density in real space can be represented as an infinite-dimensional vector, $\mathbf{J}_p$, in current space (here represented by a three dimensional space indexed by $J_1$, $J_2$, and $J_3$). The lead fields, $\mathbf{L}_1$ and $\mathbf{L}_2$, of a pair of detectors span a two-dimensional subspace within the infinite-dimensional space. The minimum norm estimate, $\mathbf{J}^*$, is the projection of $\mathbf{J}_p$ onto this subspace. $\mathbf{J}^*$ can be constructed by a linear combination of $\mathbf{L}_1$ and $\mathbf{L}_2$. Underscores denote vectors, which are denoted by bold type in the text. After Hämäläinen et al. (1993).
where the $w_j$ are scalars to be determined from the measurements (Hämäläinen et al. 1993). Since $J^*(r)$ is required to account for the measured signals, the following applies:

$$f_i = \int_G L_i(r) \cdot J^*(r) dG, \quad i = 1...N. \quad (82)$$

Substituting equation 81 into equation 82 gives the set of linear equations:

$$f_i = \sum_{j=1}^{N} w_j \int_G L_i(r) \cdot L_j(r) dG, \quad i = 1...N, \quad (83a)$$

or:

$$F = Pw, \quad (83b)$$

where $F = (f_1, ..., f_N)^T$, $w = (w_1, ..., w_N)^T$ and $P$ is an $N \times N$ symmetric matrix whose elements are given by:

$$P_{ij} = \int_G L_i(r) \cdot L_j(r) dG. \quad (83c)$$

The elements $P_{ij}$ can be calculated by dividing the source region, $G$, into a fine grid and evaluating:

$$P_{ij} = \sum_{G} L_i(r) \cdot L_j(r) \Delta G, \quad (83d)$$

where $\Delta G$ is the volume associated with a single grid point. In general, the lead fields are linearly independent, so that $P$ is non-singular (Hämäläinen et al. 1993). Then $w$ can be determined by premultiplying equation 83b by $P^{-1}$:

$$w = P^{-1}F, \quad (84)$$

and the minimum norm estimate found from equation 81.

Unfortunately, some of the lead fields are almost linearly dependent, so that large errors can be introduced into the computation of $w$, and hence $J^*$ (Hämäläinen et al. 1993). More significantly, small changes in the measurement set, $F$, can cause the minimum norm estimate to vary considerably. This means that the current estimate is affected dramatically by the presence of a small amount of noise in the measured data.
The problem has to be overcome by a process known as regularisation (Sarvas 1987). Matrix \( P \) is decomposed by singular value decomposition:

\[
P = V \Sigma V^T,
\]

in which \( V \) is an orthogonal matrix and \( \Sigma \) is a diagonal matrix containing the (non-zero) eigenvalues, \( \lambda_1, \ldots, \lambda_N \), of \( P \) in descending order of magnitude. Then the regularised inverse of \( P \) is calculated:

\[
P^+ = V \Sigma^+ V^T,
\]

where \( \Sigma^+ = \text{diag}(\lambda_1^{-1}, \ldots, \lambda_k^{-1}, 0, \ldots, 0) \) and the cut-off value, \( k (k < N) \), is chosen so that the resulting minimum norm estimate is not excessively affected by noise. The minimum norm estimate no longer describes the data exactly, but the misfit between the measured and computed measurements is no larger than the uncertainty due to noise in the measured data (Sarvas 1987). Note that singularity in \( P \) implies that there is more than one minimum norm solution. However, \( P^+ \) can always be found, so the regularised minimum norm estimate is unique.

Several general points are worthy of mention here. Suppose that the measured data are free of noise and that \( P^{-1} \) can be computed accurately to produce a unique minimum norm solution. Regularisation is inappropriate, as it would produce an unjustifiable misfit between the measured and computed fields. Unlike the methods already described, unregularised minimum norm estimation makes no attempt to reduce the number of free parameters below the number of measurements. Instead, it can be considered to "solve" the non-unique inverse problem for the infinity of solutions and then select one of these - that of minimum Euclidean norm. The major drawback of the method is quite simply that, of all the possible primary current distributions which could explain the measured data, the true source distribution may not be the one having the minimum norm. Hence, there is no guarantee that the minimum norm solution will successfully reconstruct the original current distribution. In fact, it is highly unlikely that it will, because the minimum norm solution is only one current distribution out of an infinity of possible distributions that could have produced the measured data. Thus, under optimal, noise-free conditions, the method is unreliable.

Fortunately, with real, noisy data and regularisation, this problem diminishes somewhat, as a regularised solution tends to be a simple distribution of current (Gorodnitsky et al. 1992) so that, provided the original source has a relatively simple morphology, the minimum norm estimate is accurate. The problem with regularisation though, is that if a source distribution is complex, the reconstructed solution is still simple: the method finds one or more equivalent current dipoles.

To summarise: under noise-free conditions, the method is scarcely capable of correctly determining the true primary current distribution, because of the non-
uniqueness of the overall inverse problem. Under noisy conditions, however, it can successfully estimate the primary current distribution responsible for the measured data, so long as the true source distribution is reasonably simple.

One interesting feature of this method is that it naturally zeros the radial component of the primary current if magnetic data are being used. This is because the lead field of a magnetometer is purely tangential in a spherical conductivity model. The minimum norm estimate is a linear combination of lead fields (equation 81), and since all the magnetometer lead fields are tangential, the resulting current reconstruction is also tangential.

It was shown in the previous chapter that a deep tangential current dipole produces a much greater surface potential than a superficial current dipole. Conversely, a deep tangential current dipole produces a much smaller external magnetic field than a superficial current dipole. Minimum norm estimation selects the "smallest" current distribution which can describe a given set of external data. Consequently, if electric measurements are used, the method will reconstruct a deep source distribution, since this can explain the data more efficiently than a superficial distribution of larger magnitude. In the case of magnetic data, the minimum norm estimate is generally a superficial current distribution, as superficial sources can account for the data more efficiently than deep sources (Gorodnitsky et al. 1992). In terms of lead fields, what is actually happening is that the electric lead fields increase in magnitude with depth, whereas the magnetic lead fields decline with depth. Since the minimum norm estimate is merely a weighted sum of lead fields, the reconstructed current density for electric data must also increase in magnitude with depth, and that for magnetic data decrease with depth. The easiest and most common way to overcome the problem is to define a two-dimensional source region which lies at one radius only, within a spherical conductivity model (e.g. Ahlfors et al. 1992). However, this is generally physiologically implausible as the true primary currents are known to be distributed throughout a three-dimensional volume, so that a three-dimensional reconstruction is really required.

Another possible method of overcoming the depth bias of a minimum norm estimate is to introduce a corrective weight which promotes superficial sources when electric data are used and deep sources when magnetic data are used. This is discussed in the next section. A third alternative is to combine electric and magnetic data so that the solution is a weighted sum of both electric and magnetic lead fields. In this case, the depth bias should diminish because some of the lead fields increase in magnitude with depth and others decrease with depth. This is the hypothesis of the next chapter.

Probabilistic Distributed-source Solutions

The problem of selecting a single distributed-source solution to the inverse problem from an infinity of possible solutions requires the introduction of a probabilistic
algorithm which can select the "most probable" of the possible solutions. A mathematical framework for this has been proposed by Clarke (Clarke and Janday 1989; Clarke 1989). A prior probability is defined for each possible continuous source distribution, so that the complete set of possible source distributions are described by a probability density function. A posterior probability density function is then derived by incorporating the measured data according to Bayes' theorem:

\[ P(J_p \cap F) = P(J_p)P(F | J_p), \]  

(87)

where \( P(J_p) \) is the prior probability of obtaining the source distribution \( J_p \), \( P(F | J_p) \) is the probability of obtaining a measurement set \( F \) from \( J_p \), and \( P(J_p \cap F) \) is the posterior probability of obtaining both \( J_p \) and \( F \) (Hämäläinen 1987). The most probable source distribution is found by taking the expectation of the posterior probability density function.

Using this theoretical foundation, Ioannides et al. developed a probabilistic inversion algorithm (Clarke et al. 1989; Ioannides et al. 1989, 1990). Defining the lead field, \( L_i(r) \), of a detector as in equation 80, the expected primary current density is given by:

\[ \langle J_p(r) \rangle = \sum_{j=1}^{N} w_j L_j(r) p(r), \]  

(88)

where the \( w_j \)'s are the lead field coefficients to be determined from the \( N \) measurements, and \( p(r) \) is a spatial prior probability density for \( J_p(r) \). For instance, if the source is known to be confined to a small region of cortex, \( p(r) \) can be set to zero outside of this region. Similarly, if it is more likely that a source distribution is deep rather than superficial, \( p(r) \) can be graded from a low value at the surface of the conductivity model to a high value at the centre. Substituting equation 88 into equation 80 gives the linear set of equations:

\[ f_i = \sum_{j=1}^{N} P_{ij} w_j, \quad i = 1...N, \]  

(89a)

where:

\[ P_{ij} = \int_G L_i(r) \cdot L_j(r) p(r) dG, \]  

(89b)

which, apart from the inclusion of \( p(r) \), is identical to equation 83a. So long as the lead fields are linearly independent, equations 89 can be solved for the \( w_j \) values, which can
then be used to construct \( \langle J_p(r) \rangle \) according to equation 88. The result, for the case of 
\( p(r) = 1 \) throughout the source space, corresponds exactly to the minimum norm estimate, \( J'(r) \). However, the lead fields are often almost linearly dependent, so that the solution is unstable; under these circumstances, regularisation has to be applied (Clarke et al. 1989).

The regularisation used by Ioannides et al. (1990) imposes a finite standard deviation, \( \beta \), on the prior probability density function for the primary current density at a single point in the source region, on the basis that large primary currents are physiologically unexpected. As with the regularisation described in equation 86, this produces a misfit between the measured and computed fields, but this is within the standard deviation, \( \sigma \), of the measured data. Instead of using all \( N \) lead fields in the computation of \( \langle J_p(r) \rangle \), only \( t < N \) are used, so that:

\[
\langle J_p(r) \rangle = \sum_{j=1}^{t} w_j L_j(r) p(r).
\]

(90)

Then equation 89a is replaced by:

\[
\tilde{f}_i = \sum_{j=1}^{t} \tilde{P}_{ij} w_j, \quad i = 1 \ldots t,
\]

(91a)

where:

\[
\tilde{f}_i = \sum_{j=1}^{N} P_{ij} f_j,
\]

(91b)

\[
\tilde{P}_{ij} = \sum_{k=1}^{N} P_{ik} P_{kj} + \zeta P_{ij},
\]

(91c)

\[
\zeta = \frac{\sigma^2}{\beta^2}.
\]

(91d)

Increasing the value of \( \zeta \) increases the degree of smoothing present in the solution. The advantage of this type of regularisation is that the matrix \( \tilde{P} \) is positive definite. It can therefore be inverted by a fast method such as Cholesky's method to find the \( w_j \) values, which are then used to evaluate the \( t \) terms of equation 90. The result is the most probable solution compatible with the data, within the limits of uncertainty due to noise, and given the assumption that the prior probability density function of primary current at each point in the source space is Gaussian, with standard deviation \( \beta \) (Ioannides et al. 1990).

A typical use of this probabilistic algorithm is to overcome the tendency for an electrical solution to be biased towards the deepest region of the source space, and a
magnetic solution to be superficially biased. Using the above algorithm, it is possible to choose \( p(\mathbf{r}) \) such that \( \mathbf{L}(\mathbf{r})p(\mathbf{r}) \) has a uniform magnitude at all depths of the source space. For example, Ribary et al. (1991) used a weighting function which increased monotonically with source depth to differentiate between deep and superficial sources using magnetic measurements alone. Similarly, Gorodnitsky et al. (1992) defined \( p(\mathbf{r}) \) for a given point in the source region to be a function of the average distance of that point from each of the magnetometer sites. However, Jeffs et al. (1987) noted that it is difficult to theoretically predict a weighting function. Moreover, the weighting function is highly dependent on the detailed geometry of the detectors.

The choice of whether to weight a distributed-source solution or not is a difficult one. Applying a prior probability density function which is weighted towards the deep part of the source space, for instance, implies that there is prior knowledge that the sources are located deep within the source space. This prior knowledge, in general, does not exist. Assuming that the morphology of the cortex is not known, so that sources are not constrained to lie on specific surfaces, but are allowed to lie anywhere within a three-dimensional source volume, then, prior to obtaining measurements, there is a uniform likelihood that a source lies at any point in that source volume. In other words, \( p(\mathbf{r}) = 1 \) throughout. Hence, using a graded \( p(\mathbf{r}) \) in order to overcome the depth bias of a source reconstruction is philosophically unjustifiable. Moreover, there are practical difficulties associated with weighting the prior probability density. Suppose that a set of simulated magnetic measurements are produced from a deep current dipole source of known location. An unweighted minimum norm or probabilistic algorithm is applied to the simulated measurements and produces an erroneous superficial source reconstruction. A weighted algorithm is applied, with \( p(\mathbf{r}) \) increasing with depth, and the original source is correctly reconstructed. It is tempting to say at this point that the weighted algorithm is the one to use. However, the two algorithms have highlighted a potential difficulty: there are two source reconstructions - one superficial and the other deep - which are both compatible with the measurements. Suppose that the location of the true source were not known. It would then be impossible to know which of the two reconstructed sources was the correct one. Unfortunately, this is the situation which occurs when real data are analysed. Essentially, the problem is a manifestation of non-uniqueness and is caused by the distributed-source model being grossly underdetermined. It is ultimately impossible to choose between unweighted and weighted algorithms but, on the basis of the foregoing philosophical argument that \( p(\mathbf{r}) \) should equal one when there is no prior information with regard to the locations of the sources, the author has chosen not to use a biased prior probability. A more natural and elegant means of determining the correct depth profile of a source distribution is to use both electric and magnetic data simultaneously. The magnitude of an electric lead field increases with depth, while that of a magnetic lead field decreases with depth, so that a combined solution, which is a linear combination of both electric and magnetic lead
fields, has a more uniform distribution. This approach is preferable to that of prior probability weighting, and is considered more fully in the next chapter.

It is difficult for a distributed-source method to reconstruct a localised source such as a discrete current dipole (Ioannides et al. 1990). However, several researchers have used the probability weight, \( p(\mathbf{r}) \), as the basis of an iterative algorithm which can correctly locate point sources (Ioannides et al. 1990; Gorodnitsky et al. 1992). The principle of the method is to repeatedly apply a distributed-source algorithm to the data, successively updating the prior probability weight. For a given iteration, \( p(\mathbf{r}) \) is chosen to be proportional to the primary current density resulting from the previous iteration. In qualitative terms, the first iteration produces a diffuse estimate of the current source. Setting \( p(\mathbf{r}) \) to be proportional to this estimate effectively reduces the source space to a small region surrounding the point source. The second iteration can therefore produce a more accurate estimate of the point source within this small region, and so on. When a test source is used, a focal solution is typically obtained after four or five iterations (Ioannides et al. 1990; Gorodnitsky et al. 1992). With real data, however, it is questionable whether the number of iterations can be fixed, as it is never known how localised a current source is. In any case, if it is believed that the primary current distribution consists of a localised current dipole, it would probably be more reliable to apply a single dipole or spatio-temporal model.

The Maximum Entropy Method

Radio-astronomy gives rise to an inverse problem which is in many ways analogous to that of neuro-electromagnetism. In radio-astronomy, it is desirable to reconstruct a map of the incident radiation intensity from an incomplete set of noisy data. The measured data constitute the Fourier transform of the intensity map, so one means of solving the inverse problem is to take the inverse Fourier transform of the measurement set. However, since there are never sufficient measurements to completely define the Fourier transform, and since the measurements are noisy, some other form of analysis is useful. Such an alternative method has been described by Gull and Daniell (1978). Their method selects a single, general-purpose map of the sky from the series of maps which are compatible with the data. The analogy with the neuro-electromagnetic inverse problem should now be apparent: in astronomy, the measured data are the Fourier transform of the intensity to be determined, inverse solution yields a series of maps which are all consistent with the measured data, and the most probable map is the required result; in neuro-electromagnetism, the measurements are related to the underlying current source distribution by the lead field functions, inverse solution also provides a series of source distributions which are consistent with the measurements, and the most probable distribution is also required. The author therefore proposes the following probabilistic algorithm, which is a direct adaptation of the maximum entropy
method of Gull and Daniell (1978). The difference between this method and that described in the previous section is that this method provides the most probable solution in the mode sense, rather than the expected solution or mean. A different, but related method was examined by Jeffs et al. (1987).

The primary current density at each point in a three-dimensional source space can be defined by the three Cartesian components of the current density vector. If each vector of the Cartesian basis is divided into two sub-components of opposite direction, the current density can be described in terms of six positive components. This construction allows us to treat primary current density as a distribution of positive quantities, which is necessary for the following arguments. The task is to find the most likely distribution of primary current density which is compatible with the measured electric or magnetic field. This is achieved as follows:

The source space is divided into $S/6$ small elemental regions and the $S$ positive parameters of the current density throughout these regions are denoted $x_j (j = 1 \ldots S)$. (There are six positive parameters associated with each element or voxel.) The parameters $x_j$ are quantised. For example, the $x_j$s may be chosen as the primary current density in units of $10^{-10}$Am$^{-2}$, so that they are large positive integers. Each of the $N$ external measurements is denoted by $f_i (i = 1 \ldots N)$, and is assumed to have a Gaussian error with standard deviation $\sigma_i$. Also, $\tilde{f}_i$ represents the forward field at the $i^{th}$ measurement site due to a given primary current distribution, which is defined by a set of $x_j$s. In these terms, we require the most likely combination of $x_j$s to satisfy the $\chi^2$ test:

$$\sum_{i=1}^{N} \left( \frac{\tilde{f}_i - f_i}{\sigma_i} \right)^2 = N.$$  

(92)

Setting the chi-square to equal the number of measurements, $N$, is a means of ensuring that the solution is not fitted to the noise as well as the useful component of the data (Gull and Daniell 1978). A large value of $\chi^2$ indicates that the solution is a poor fit to the data, but a very small value signifies that the solution fits the noise as well. Selecting $\chi^2 = N$ is a means of regularisation which extracts the meaningful data from the external measurements while overcoming the noise.

The probability of a given primary current density distribution is obtained by considering the $S$ unknown parameters as a collection of boxes, into which are thrown $X$ balls, $X$ being much larger than $S$. The balls are, of course, quanta of current density. The probability of the first $x_1$ balls landing in box 1, the second $x_2$ in box 2, etc. is given by:

$$\left( \frac{1}{S} \right)^{x_1} \times \left( \frac{1}{S} \right)^{x_2} \times \left( \frac{1}{S} \right)^{x_3} \times \cdots \times \left( \frac{1}{S} \right)^{x_S}.$$
But there are \( \binom{X}{x_1, x_2, x_3, \ldots, x_S} \) ways of picking the groups \( x_1, x_2, x_3, \ldots, x_S \) out of \( X \), so the probability of a given distribution of balls is:

\[
P = \frac{X!}{x_1!x_2!x_3!\cdots x_S!} \left( \frac{1}{S} \right)^S \quad (93a)
\]

\[
= \frac{X!}{x_1!x_2!x_3!\cdots x_S!} \left( \frac{1}{S} \right)^X . \quad (93b)
\]

In real terms, \( P \) is the probability of obtaining the distribution \( x_1, x_2, x_3, \ldots, x_S \) of quantised current density.

In principle, to find the most probable solution which is compatible with the measured data, \( P \) is maximised subject to the \( \chi^2 \) constraint. However, a distribution which maximises \( P \) can be found by maximising \( \log P \), and since it is easier to maximise \( \log P \), this is the way to proceed. From equation 93b,

\[
\log P = \log X! - \sum_{j=1}^{S} \log x_j! - X \log S . \quad (94)
\]

Since \( X \) and \( x_j \) are large, \( \log X! \) and \( \log x_j! \) can be evaluated by Stirling's theorem, giving:

\[
\log P = X \log X - X - \sum_{j=1}^{S} \left( x_j \log x_j - x_j \right) - X \log S , \quad (95a)
\]

\[
= X \log \left( \frac{X}{S} \right) - \sum_{j=1}^{S} x_j \log x_j . \quad (95b)
\]

This must be maximised subject to the chi-square constraint:

\[
\sum_{i=1}^{N} \left( \frac{\bar{f}_i - f_i}{\sigma_i} \right)^2 - N = 0 . \quad (96)
\]

Using Lagrange's method, we form the function:

\[
F = X \log \left( \frac{X}{S} \right) - \sum_{j=1}^{S} x_j \log x_j + \lambda \sum_{i=1}^{N} \left( \frac{\bar{f}_i - f_i}{\sigma_i} \right)^2 - \lambda N , \quad (97)
\]

122
where \( \lambda \) is the Lagrange undetermined multiplier. The \( x_j \)'s are then determined by satisfying the conditions:

\[
\frac{\partial F}{\partial \lambda} = 0, \\
(98a)
\]

and:

\[
\frac{\partial F}{\partial x_j} = 0, \quad j = 1 \ldots S. \\
(98b)
\]

The first of these conditions is satisfied when:

\[
\sum_{i=1}^{N} \left( \frac{\tilde{f}_i - f_i}{\sigma_i} \right)^2 = N, \\
(99a)
\]

and the second when:

\[
x_j = \frac{X}{S} \exp \lambda \sum_{i=1}^{N} \frac{2(\tilde{f}_i - f_i)}{\sigma_i^2} \frac{\partial \tilde{f}_i}{\partial x_j}, \quad j = 1 \ldots S. \\
(99b)
\]

Equations 99 constitute \( S + 1 \) simultaneous equations in \( S + 1 \) unknowns, the unknowns being \( \lambda \) and the \( x_j \)'s. It can be shown that there is a unique solution to equations 99b for all positive \( \lambda \), and that this solution maximises the probability, \( P \), of the quantised current density distribution (Gull and Daniell 1978). Equations 99b can therefore be solved by applying the \( x = g(x) \) iteration in \( S \) dimensions, treating \( \lambda \) as a constant. During the iteration, \( \lambda \) is increased so that the \( \chi^2 \) condition (equation 99a) is satisfied (Gull and Daniell 1978). The derivatives \( \partial \tilde{f}_i / \partial x_j \) can be calculated easily, since \( \tilde{f}_i \) is a weighted sum of the quantised current densities, \( x_j \):

\[
\tilde{f}_i = \sum_{j=1}^{S} G_{ij} x_j, \quad i = 1 \ldots N, \\
(100)
\]

where \( G_{ij} \) is a gain term giving the contribution of \( x_j \) to \( \tilde{f}_i \). Then:

\[
\frac{\partial \tilde{f}_i}{\partial x_j} = G_{ij}, \quad (101)
\]
An insight into the physical interpretation of this algorithm can be gained by observing what happens when there are no measurements, in which case the method gives the most probable prior distribution of primary current density. Setting $N = 0$ in equation 99b gives $x_j = X/S \ (j = 1 \ldots S)$. In other words, in the absence of measurements, the quantised current distribution is completely uniform, with the $X$ quanta divided evenly throughout the $S$ parameters. Thus, the algorithm can be considered to provide the smoothest distribution of primary current density which is consistent with the measurements (Gull and Daniell 1978). The advantage of this is that the method does not construct a feature in the source distribution unless there is evidence for it in the data. The resolution of the source reconstruction is as high as possible, given the noise in the data; if the data are sufficiently noise-free, then the image will have a high resolution (Gull and Daniell 1978).

The division of each Cartesian base vector into two opposite vectors, whilst arising out of the need to specify the current density by positive scalars, also ensures that the algorithm is stable: although the distribution of $X$ quanta throughout the $S$ parameters of the source space produces a positive level of quantisation throughout the source space, only the difference in quantisation between pairs of parameters constitutes a net current density. This removes the dependence of the solution on the value of $X$.

In the case of a uniform distribution of quanta, each pair of opposing current density vectors cancels exactly, so that no current density results.

The method cannot, of course, overcome the overall non-uniqueness of the inverse problem. In this respect, the algorithm still only provides a single solution from an infinite number of possible solutions and the result cannot therefore be considered as the solution. Neither can the method reconstruct a complex distribution of sources if the data are too few or too noisy. Because of this, the author has not pursued the practical implementation of the algorithm any further, but has concentrated, in the next chapter, on reducing the number of parameters required to describe a source distribution, so that a unique solution becomes more readily available. However, the above maximum entropy method provides the most probable solution and, by virtue of its smooth images and resistance to noise, it may prove to be a useful approach.

The Spatial Filtering Method

A distribution of current can be determined uniquely if it is confined to a single plane in space and measurements of the magnetic flux density are recorded in a spatially continuous fashion from all points in another plane which is parallel to the current space (Roth et al. 1989). This is the motivation for attempting to reconstruct a current image using Fourier or spatial filtering techniques. The general principle of the method is to take the Fourier transform of the measurements and invert this using a transformed version of Maxwell's equations, this being equivalent to a spatial filtering operation. The
result is the Fourier transform of the current density, which is then inverse transformed to provide a current image (Dallas 1985; Roth et al. 1989; Tan et al. 1990). Note that in the present context, we are referring to total current, rather than primary current. The total current (primary current plus volume current) must be confined to a single plane.

Suppose that the reconstruction space is located in the \( z = 0 \) plane and has a small thickness, \( d \). For mathematical simplicity, it is assumed that only the \( x \)-component, \( B_x(x,y,z) \), of the magnetic flux density is measured; this is produced by the \( y \)-component, \( J_y(x',y') \), of the current density (Roth et al. 1989):

\[
B_x(x,y,z) = \frac{\mu_0 dz}{4\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{J_y(x',y')}{[(x-x')^2 + (y-y')^2 + z^2]^{3/2}} dx'dy'. \tag{102}
\]

Here, \( J_y(x',y') \) is convolved with the Green's function:

\[
G(x-x',y-y',z) = \frac{\mu_0 dz}{4\pi} \frac{1}{[(x-x')^2 + (y-y')^2 + z^2]^{3/2}}, \tag{103}
\]

so that application of the convolution theorem produces the following expression for the two-dimensional Fourier transform of the magnetic flux density:

\[
b_x(k_x,k_y,z) = g(k_x,k_y,z) \cdot j_y(k_x,k_y), \tag{104}
\]

where \( g(k_x,k_y,z) \) and \( j_y(k_x,k_y) \) are two-dimensional Fourier transforms of the Green's function and current density respectively, and \( k_x \) and \( k_y \) are the components of the spatial frequency (Roth et al. 1989). The two-dimensional Fourier transform is defined as:

\[
j_y(k_x,k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_y(x,y) \exp(i(k_x x + k_y y)) dx dy. \tag{105}
\]

The Fourier transform of the Green's function is required; this is given by the analytical expression:

\[
g(k_x,k_y,z) = \frac{\mu_0 d}{2} \exp \left(-\left(k_x^2 + k_y^2\right)^{1/2}\right). \tag{106}
\]

This expression indicates that the \( x \)-component of the magnetic flux density is just a low-pass filtered version of the \( y \)-component of the current density (Roth et al. 1989).

The inverse problem can be solved by simply evaluating \( j_y(k_x,k_y) \):
\[ j_y(k_x, k_y) = \frac{b_x(k_x, k_y, z)}{g(k_x, k_y, z)}, \]  

(107)

where \(1/g\) is the inverse filter function. The current density, \(J_y(x', y')\), is then found by the inverse Fourier transform of \(j_y(k_x, k_y)\):

\[ J_y(x, y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} j_y(k_x, k_y) \exp(-i(k_x x + k_y y)) dk_x dk_y. \]  

(108)

In order to determine both \(J_x(x', y')\) and \(J_y(x', y')\) from \(B_s(x, y, z)\) only, use is made of the continuity equation (equation 9):

\[ \nabla \cdot \mathbf{J} = 0, \]  

(109)

the Fourier transform of which is:

\[ -ik_x j_x(k_x, k_y) - ik_y j_y(k_x, k_y) = 0. \]  

(110)

However, under some circumstances, the filter function, \(g\), tends to zero, so that the inverse filter cannot be applied; this has to be overcome by explicitly measuring two components of the magnetic flux density (Roth et al. 1989).

Current distributions produced by this method are limited in spatial resolution by the size of the gradiometer coils, the distance of the gradiometer from the current plane, and the level of noise in the data (Roth et al. 1989). The separation of the measurement sites also limits the high spatial frequency components of the source reconstruction, thereby limiting the degree of detail which can observed in the current image. The relatively low spatial frequency of gradiometer arrays currently in use would severely limit the achievable image resolution.

Although this method has been used to examine axonal current flow (Tan et al. 1990) and current flow in a small flat coil (Roth et al. 1989), it is not clear from the literature whether it can be extended to the neuro-electromagnetic inverse problem in which we are interested. Firstly, a two-dimensional volume conductor is unrealistic, but a full three-dimensional inversion using the same methods is non-unique and therefore unworkable (Roth et al. 1989). Secondly, the method provides an estimate of the total current density. Whether it can resolve this into primary and volume currents is uncertain: having written down the Green's function in terms of primary current only, so that the problem becomes one of determining primary current rather than total current, it is questionable whether the inverse filter function could still be evaluated. Only if this were possible, would it be feasible to use a two-dimensional primary current
reconstruction region within a three-dimensional volume conductor - the very least that is required from an inversion method. The method does have the advantage, however, of only requiring division by the filter function, \( g \), to achieve inverse solution. The Fourier transform and its inverse can also be computed with considerable speed.

Conclusion

This chapter has examined some of the ways of determining the locations and strengths of current sources within the brain from external measurements of the electric or magnetic field. The principal message of the chapter is that the inverse problem is non-unique. This occurs primarily because of the limited number of noisy measurements, but also because some current sources do not produce an external field. Therefore, whatever technique is chosen for the solution of the inverse problem, prior information of some kind must be incorporated. The prior information may take one of many forms: it may be the assumption that the source has some particular spatial form, such as a single current dipole, or it may be a prior probability density function of primary current distributions, but it must inevitably be included. The order in which the different techniques have been discussed has been intended to emphasise the range of prior assumptions. On the one hand is the equivalent current dipole; the assumption that the source is confined to a single point in space reduces the number of unknown parameters to just five or six. On the other hand are the probabilistic algorithms, which model the primary current density as a continuous vector field having an infinite number of unknown parameters. These methods utilise a prior probability density describing the likelihood of obtaining a particular distribution of primary current, to select a single solution as the most probable. The principal features of the foregoing techniques are summarised below:

**Equivalent Current Dipole Fitting** Assumes a single, point source to reduce the unknown parameters to five or six, which is fewer than the number of measurements, facilitating a unique solution. May have little physiological meaning as the true source may be less localised. Generally solved by non-linear parameter estimation.

**Spatio-temporal Dipole Models** Use the temporal dimension of the data to increase the number of available measurements, thus facilitating a more involved multiple-dipole source model. A series of dipoles are assumed to be fixed in location throughout the time interval. Their moments may be fixed or rotating. Solved by non-linear parameter estimation, which can be unreliable due to the presence of local minima in the chi-square function.
**Principal Components Analysis** is based on a spatio-temporal source model consisting of dipoles which are fixed in location and orientation. The analysis determines the number of sources, their locations and orientations, and the temporal variation of their strengths. This is achieved by fitting single current dipoles to the independent components of the data matrix.

**Multiple Signal Classification** As for a conventional spatio-temporal dipole model, but uses an exhaustive search throughout the source space to determine the dipole locations, overcoming the difficulty associated with non-linear parameter estimation. Will determine (a) how many sources are present, (b) how many of these are fixed in orientation and how many are rotating, and (c) the dipole locations and moments. The number of dipoles is limited by the signal-to-noise ratio of the data.

**Minimum Norm Estimation** Allows any continuous distribution of primary current throughout a predefined source space. The number of unknown parameters exceeds the number of measurements, so that there are an infinite number of possible solutions. The distribution having minimum Euclidean norm is selected as the best solution; this distribution is all that can be deduced from the measurements. The estimate can be regularised to overcome noise in the data, but this limits the resolution. Suffers from depth bias problems due to the lead fields being uneven in magnitude with depth.

**Probabilistic Distributed-source Solutions** These assume that the probability density of primary current strength at each point in the source space is Gaussian, i.e. large primary currents are unlikely. The method finds the expectation of the posterior probability density, the result being identical to the minimum norm estimate. Probability weighting can be used to overcome depth bias problems and to incorporate additional prior spatial information.

**The Maximum Entropy Method** Uses randomness to define a prior probability distribution. The method finds the mode of the posterior probability density function by maximising the configurational entropy of the primary current, subject to the measured data. The result is the smoothest distribution which is consistent with the measurements. Only reconstructs a source if there is justification for it in the data.

**The Spatial Filtering Method** The data are Fourier transformed and filtered in the spatial frequency domain, then inverse transformed to give the estimated current density distribution. The method extracts all of the available information from the data, but is non-unique in three dimensions.

All of the distributed-source methods provide only one solution from an infinity of possible solutions, and that single solution may not be the correct one. If these
methods are to be really useful, some additional prior information must be added to reduce the range of possible solutions. One way of doing this is to discretise the primary current into a series of current dipoles which lie on a surface representing the cortex. The locations and directions of the current dipoles are fixed, so that only the magnitudes of the dipole moments have to be found from the data. The result is a constrained distributed-source model which has features in common with both spatio-temporal and probabilistic models. Most notably, the model is either overdetermined or only just underdetermined, so that the range of possible solutions is reduced and a unique solution becomes feasible. This approach is presented in detail in the next chapter.
5. Anatomically Constrained Solutions to the Neuro-electromagnetic Inverse Problem

Introduction

Distributed-source estimates are becoming increasingly important in neuroelectric and neuromagnetic source analysis, for several reasons. The first reason is that a source model consisting of a distribution of current sources most naturally reflects the true neural current sources within the brain, which are spread around the folded cortical lamina (George et al. 1993). The second reason is that a distributed-source model can be used at all times, as opposed to the equivalent current dipole method, which can only be successfully applied when the current distribution is dipolar. Throughout the course of an evoked electric or magnetic response, there are typically periods when the external data can be represented accurately by an equivalent current dipole, and other periods when they cannot. This does not prevent the application of a distributed-source model, as it can successfully image both single or distributed sources, having greater success with the latter (Ioannides et al. 1990). Probably the most significant reason for the popularity of distributed-source models, however, is that they produce spatially continuous images similar to those of positron emission tomography or functional magnetic resonance imaging. Continuous images are necessary if neuro-electromagnetism is to rival these other methods. Steps have already been taken to promote neuro-electromagnetism as an alternative imaging technique, with Ribary et al. (1991) using a distributed-source method which they referred to as "magnetic field tomography." Hence, distributed-source analyses are important to the future of neuro-electromagnetism.

However, the previous chapter showed that there are several severe limitations to the success of distributed-source estimation. Regularised estimates are diffuse and therefore have low resolving power (Jeffs et al. 1987; Gorodnitsky et al. 1992). Under some circumstances, regularisation has the effect of simplifying a complex source distribution, so that an equivalent current source is reconstructed instead. These problems are the result of using a limited number of noisy measurements to determine a
large number of unknown parameters. An even more severe shortcoming is manifested if the data are free of noise: regularisation is no longer justified and it becomes necessary to select one solution from an infinity of possible solutions. The single solution - typically that of minimum Euclidean norm - may not be the current distribution which produced the data. Unfortunately, as magnetometry and signal processing capabilities improve, this situation will inevitably occur more often, and so the future of distributed-source estimation is limited.

Fortunately, there is a way forward. It was established in chapter three that evoked electric potentials and magnetic fields are produced by postsynaptic potentials in large elongated cells within the cortex. Thus, when carrying out a distributed-source analysis, it is not necessary to scan the whole of the brain for sources, but only to scan the cortex in those areas of the brain which could possibly be involved in generating the measured signals. For instance, when analysing a visual evoked electric or magnetic response, it is necessary only to look for sources in the visual areas of the cortex. In distributed-source terms, this means that the source space can be restricted to these regions, so that there are fewer parameters to be determined and the solution can be made more accurate.

A further restriction can be made on current density orientation. It was shown in chapter three that large elongated cells, such as pyramidal cells, are predominantly oriented perpendicular to the pial surface. Postsynaptic potentials in these cells therefore produce primary currents which are also perpendicular to the pial surface. Hence, the primary current in the restricted source space of an anatomically constrained model can be constrained to flow towards the cortical surface, or away from it. This restriction further divides the number of unknown current density parameters by three. It also has the additional advantage of enabling current components which would otherwise be indeterminate to be calculated. For example, in an unconstrained analysis, the radial component of a current dipole cannot be determined if only magnetic measurements are used, as a radial dipole is "silent." However, if the primary current is constrained in orientation, then fixing the tangential component of the primary current also fixes the radial component (so long as the current is not constrained to be purely radial).

A convenient means of defining the location of the anatomically constrained source space, and the orientation of the primary current within it, is to consider the current flow as a distribution of dipoles. The dipoles are located at regular intervals throughout the regions of cortex which could possibly have produced the measured data. The orientations of the dipoles are chosen so as to be perpendicular to the pial surface. Hence, the distributed-source problem can be broken down into one which involves a discrete distribution of current dipoles of known location and orientation. The dipoles can be chosen to be as close or as far apart as desired. For example, if a continuous distribution of sources is appropriate, the dipoles can be positioned close together to approximate continuous current. Conversely, the dipoles can be spaced
out, so that each one is effectively the equivalent current source for a small region of cortex. The increased spacing then means that fewer current dipoles are required to define the source space, and the magnitudes can therefore be determined more accurately from the measurements. This is the approach of the present chapter.

Several authors have examined the above anatomically constrained model. Jeffs et al. (1987) suggested that it was a means of reducing the dimensionality of the unconstrained distributed-source problem. They proposed that the shape of the cortex could be derived from a computerised brain atlas, stating that "a model using the major features and landmarks of the cortex shape is sufficient for an initial study." They also described a simple model in which the brain was treated as a series of concentric ellipsoids. The model was divided into discrete voxels and a current dipole of fixed orientation was assigned to each, the dipole orientation being chosen to represent the local orientation of the brain. Kaufman et al. (1991) employed a cruciform model of the visual cortex and located it within a spherical volume conductor. A current dipole was positioned within each square millimetre of the model, giving a total of 1386 dipoles. The dipoles were located normally to the respective surfaces of the model. Simulated evoked magnetic fields were generated by considering some of the current dipoles to be mutually parallel and the remainder to be randomly oriented (i.e. either parallel or antiparallel). The magnitudes of the dipole moments were all randomly chosen. A similar cruciform model was also used by Garnero et al. (1993) to compare the effectiveness of electric and magnetic source analyses.

The most extensive studies of constrained distributed-source estimation are those of Wang (Wang et al. 1992, 1993a,b; Wang 1993). Simulated data were inverted to examine the feasibility of source estimation using a constrained source region. The early studies (Wang et al. 1992) utilised an L-shaped source region whose two surfaces were perpendicular to the surface of a half-space volume conductor. Specific source geometries were used to generate simulated magnetic measurements over a square planar grid. The measurements were then used to reconstruct the original sources. Each of the two surfaces of the source space was 50mm x 50mm in size and typically contained a grid of 14 x 14 current dipoles, giving a total of 196 dipoles for each surface. The measurement surface was of side 120mm, and the detector spacing was 10mm, giving a total of 169 measurements. Because there were fewer measurements than dipole magnitudes to be determined, it was necessary to apply an unweighted regularised minimum norm algorithm to reconstruct the sources. Hence, the images suffered from the problem of diffuse reconstruction that occurs with conventional minimum norm estimation. However, the source space restriction was sufficient to prevent the solutions from being biased towards the superficial part of the source space. This may have been aided by the fact that the simulated sensors were point magnetometers, measuring the magnetic field, as opposed to its gradient. The sensitivity of a magnetometer falls off with distance less rapidly than that of a gradiometer, so that the lead field is more evenly distributed with depth. The sensors
were also spread over a large area relative to the width of the source space. It was found that the original sources could be correctly reconstructed from both noise-free and noisy measurements. The spatial resolution of an image was found to be influenced by the sensor spacing, the dipole grid spacing, and the source depth.

These simulations were also extended to an L-shaped source space within a spherical conductivity model (Wang 1993). A total of 331 sensors, distributed over a spherical region, were used to reconstruct sources within a source space consisting of two approximately square regions of side 50mm, containing a total of 356 discrete current dipoles. Besides confirming the earlier half-space simulations, these analyses showed that the number of sensors could be reduced to 37, with corresponding loss of detail in the source image. A superficial bias occurred when the number of sensors was reduced. It was also found that if the reconstruction space differed from the location of the true source by up to 5mm, a reasonable image was still obtained.

The most recent study (Wang et al. 1993b) attempted to discriminate between current on opposite walls of a sulcus. For a sulcus whose opposing walls were 2mm apart, it was found that a source on one wall could be located correctly, except that the activity appeared to extend to both walls. This was the result of attempting to determine 750 dipole strengths from 217 measurements.

Anatomically constrained distributed-source estimation has also been used by George et al. (1991, 1993) for analyses of visual evoked magnetic responses, and by Fuchs et al. (1993) for analyses of auditory evoked magnetic fields. Both of these studies made use of magnetic resonance images to define the source space. The cortex was triangularised into a series of discrete planes and current dipoles were located normally to some of the individual segments. Fuchs et al. (1993) found that positioning a current dipole in each of the 200 000 segments of the triangularised cortex gave too many dipoles for efficient source reconstruction. They therefore limited the source space to a region of diameter 140mm and depth 100mm, situated below the gradiometer. This reduced the number of discrete dipoles to around 50 000. The dipoles were further "thinned out" to a spacing of around 7mm, which gave a total of 1000 to 2000 grid points. Minimum norm estimation was then used to reconstruct current sources within the three-dimensional cortical source space. Both George et al. (1991) and Fuchs et al. (1993) concluded that anatomical constraints improved the fidelity of the reconstructions, but because there were so many discrete dipoles in the source space, the reconstructions still suffered from the problems associated with unconstrained algorithms: the source distributions tended to be diffuse and superficially biased.

Even when using a restricted source space, it is difficult to obtain sufficient measurements to produce an overdetermined situation, and hence a unique solution (Jeffs et al. 1987). The number of measurements can be increased by using both electric potentials and magnetic fields simultaneously to determine the underlying source distribution. If an evoked response is measured by twenty scalp electrodes and
twenty magnetic flux gradiometers, the number of measurements can be doubled by using the electric and magnetic records simultaneously. But besides this, the two types of data can provide complementary information about the sources. Equations 111 and 112 give the electric potential, \( V \), and magnetic flux density, \( \mathbf{B} \), resulting from a primary current distribution, \( \mathbf{J}_p \), in an infinite, homogeneous volume conductor of isotropic conductivity, \( \sigma \) (see equations 11b and 17b):

\[
V(r) = \frac{1}{4\pi\sigma} \int_G \mathbf{J}_p(r') \cdot \frac{r - r'}{|r - r'|^3} \, dG, \tag{111}
\]

\[
\mathbf{B}(r) = \frac{\mu_0}{4\pi} \int_G \mathbf{J}_p(r') \times \frac{r - r'}{|r - r'|^3} \, dG, \tag{112}
\]

where \( r \) is an observation point and \( r' \) is a point in the source region, \( G \). It can be seen that the electric potential at \( r \) is a weighted integral of an inner product between \( \mathbf{J}_p(r') \) and a vector, \( r - r' \), from the source point to the observation point. In contrast, the magnetic flux density at \( r \) is a weighted integral of a vector product between \( \mathbf{J}_p(r') \) and \( r - r' \). Hence, the electric potential measures the component of \( \mathbf{J}_p(r') \) in the direction of \( r - r' \), while the magnetic flux density measures the component of \( \mathbf{J}_p(r') \) orthogonal to \( r - r' \). A specific example of the above effect is provided by a current dipole. In this case, the electric potential and magnetic flux density are given by (see equations 11d and 17c):

\[
V(r) = \frac{1}{4\pi\sigma} \mathbf{M} \cdot \frac{r - r_0}{|r - r_0|^3}, \tag{113}
\]

\[
\mathbf{B}(r) = \frac{\mu_0}{4\pi} \mathbf{M} \times \frac{r - r_0}{|r - r_0|^3}, \tag{114}
\]

where \( \mathbf{M} \) is the dipole moment and \( r_0 \) is the dipole location. In this situation, the electric potential measures the component of \( \mathbf{M} \) in the direction of \( r - r_0 \), and the magnetic flux density measures the orthogonal component of \( \mathbf{M} \). The above arguments generalise easily to an inhomogeneous volume conductor (Ilmoniemi 1991). Thus, electric and magnetic measurements reflect independent components of a source and, used simultaneously in an inverse procedure, should therefore yield complementary source information.

The independent nature of electric and magnetic measurements can also be observed by examining the lead field of an electrode and magnetometer. Figure 25 shows the lead field of a point electrode situated on the x-axis, in contact with a spherical model as described in chapter three. The lead field is calculated over a
volume bounded by the spherical coordinates \( r = 25\text{mm}, r = 65\text{mm}, \theta = 45^\circ, \phi = 135^\circ, \phi = -45^\circ, \phi = 45^\circ \). This volume constitutes a pseudo-cubic region within the sphere. (For simplicity, the electrode is assumed to measure absolute potential. If the electrode were to measure the potential relative to some reference point, then the resulting lead field would be equal to the absolute lead field of the measurement electrode minus the absolute lead field of the reference electrode.) Figure 26 is the lead field of a second order gradiometer modelled on the BTi 601 single channel system. The gradiometer has a baseline of 50mm, a separation between the middle and rear coils of 50mm, a coil diameter of 20mm and a dewar base of thickness 12mm. In the simulation, the dewar tail is against the scalp on the x-axis, so that the front coil of the gradiometer is 12mm from the scalp. It can be seen that the electric lead field increases in magnitude with depth, and is directed towards the electrode in a radial manner. The magnetic lead field decreases in magnitude with depth, and is directed around the magnetometer site in a circular fashion. Hence, the electric and magnetic lead fields have maximum magnitude at differing depths, and at a given depth they are perpendicular and hence complementary in orientation.

The value of simultaneous electric and magnetic recording becomes obvious when there is a radially-oriented dipolar current source. The magnetic measurements alone cannot detect the source as it produces no external magnetic field, but the electric measurements can successfully detect it. Conversely, a loop of primary current circling a single electrode such that it is perpendicular to the lead field at each point in its path cannot be detected by the electrode, whereas this type of current would be optimally detected by a magnetometer at the same site as the electrode. Hence, simultaneous use of both electric and magnetic records can double the number of available measurements and provide complementary information about the source distribution.

Several authors have discussed methods by which electric and magnetic data can be analysed simultaneously. Stok (1987) determined equivalent current dipoles using both electric and magnetic simulated measurements. The dipole parameters were determined by simply minimising the combined chi-square for the electric

Figure 25 (overleaf). Lead field of a point electrode situated at \( \theta = 90^\circ, \phi = 0^\circ \). For simplicity, the electrode is assumed to measure absolute potential. The length of an arrow is proportional to the tangential component of the lead field, and the diameter of a circle is proportional to the outwardly-directed component of the lead field, the constant of proportionality being the same for both arrows and circles.

Figure 26 (2 pages ahead). Lead field of a second order magnetic flux gradiometer situated at \( \theta = 90^\circ, \phi = 0^\circ \). The gradiometer is located radially to the head, with its front coil 12mm from the scalp.
measurements, $e_i$ and $M$ magnetic measurements, $m_i$:

$$\chi^2 = \sum_{i=1}^{E} \left( \frac{e_i - \bar{e}_i}{\sigma_{e,i}} \right)^2 + \sum_{i=1}^{M} \left( \frac{m_i - \bar{m}_i}{\sigma_{m,i}} \right)^2,$$

(115)

where $\bar{e}_i$ is the $i$th forward-computed electric measurement and $\bar{m}_i$ is the $i$th forward-computed magnetic measurement. $\sigma_{e,i}$ and $\sigma_{m,i}$ are the standard deviations of the $i$th electric and $i$th magnetic measurement, respectively. It was found that the estimates based on magnetic data alone were generally better than those based on electric data alone, but a further improvement in accuracy was obtained by using electric and magnetic measurements together. This was also concluded by Ishiyama and Kanai (1992), using a similar analysis.

The above chi-square minimisation is, in fact, a special case of a more general method, described by Greenblatt (1993). The total of $N$ electric and magnetic measurements are modelled as:

$$f(t) = Gq(t) + n(t),$$

(116)

where $f(t)$ is the measurement vector and $q(t)$ is a vector of $S$ dipole magnitudes, representing, for example, three orthogonal components of a single current dipole, or the magnitudes of $S$ dipoles in a constrained source space. $G$ is a gain matrix in which $G_{ij}$ ($i = 1...N, j = 1...S$) is the sensitivity of detector $i$ to dipole component $j$, and $n(t)$ is zero-mean Gaussian noise. If there are $T$ time samples, the covariance matrix for the data can be estimated as:

$$C = \frac{1}{T} \sum_{t=1}^{T} (f(t) - f_0)(f(t) - f_0)^T,$$

(117)

where $f_0$ is the mean measurement vector. If $q(t)$ represents the moment of a single equivalent current dipole, then the solution to equation 116 may be found by minimising the least squares expression:

$$X = \sum_{i=1}^{N} \left( f_i - \sum_{j=1}^{S} G_{ij} q_j \right)^2,$$

(118)

$$= (f - Gq)^T (f - Gq),$$

(119)

but since the electric components of $f$ are approaching ten orders of magnitude larger than the magnetic components, the minimisation procedure will be dominated by the
electric measurements. Hence, it is more appropriate to weight the minimisation function by the covariance matrix (Sarvas 1987):

\[
\tilde{X} = (f - Gq)^T C^{-1} (f - Gq).
\]  

(120)

If \( C \) is diagonal, this is equivalent to minimising the chi-square:

\[
\chi^2 = \sum_{i=1}^{N} \left( \frac{f_i - \sum_{j=1}^{3} G_{ij}q_j}{\sigma_i} \right)^2
\]

(121)

where \( \sigma_i \) is the standard deviation of \( n_i(t) \). The covariance matrix can be estimated from a set of measurements which does not contain the signal to be localised. It may be close to singular, but \( C^{-1} \) can be calculated using an appropriate regularisation method, such as singular value decomposition (see equation 86) (Sarvas 1987).

If \( q(t) \) contains the magnitudes of \( S \) dipoles of fixed orientation, where \( S > N \), then the minimum norm solution is defined as:

\[
q^* = \left( G^T G \right)^{-1} G^T f.
\]

(122)

Again, because of the differing signal and noise scales between the electric and magnetic measurements, it is more appropriate to compute:

\[
q^* = \left( G^T C^{-1} G \right)^{-1} G^T C^{-1} f,
\]

(123)

\( C^{-1} \) being calculated as for the single dipole case.

This approach is closely related to the method of Hasson (1991). The latter is based upon the probabilistic algorithm described by equations 90 and 91 of the previous chapter. Recall that the lead field matrix, \( P \), is defined as:

\[
P_{ij} = \int_G L_i(r) \cdot L_j(r) p(r) dG,
\]

(124)

where \( L_i(r) \) and \( L_j(r) \) are the lead fields of detectors \( i \) and \( j \), respectively, \( p(r) \) is a probability weight and \( G \) is the source space. In the case that the measurement vector, \( f \), contains \( M \) magnetic and \( E \) electric measurements, \( P \) can be divided into smaller matrices:

\[
P = \begin{pmatrix}
A & C \\
C^T & B
\end{pmatrix},
\]

(125)
where $A$ is an $M \times M$ matrix whose elements involve scalar products between magnetic lead fields, $B$ is an $E \times E$ matrix whose elements involve scalar products between electric lead fields, and $C$ is an $M \times E$ matrix containing scalar products between magnetic and electric lead fields. The magnetic measurements and lead fields are then multiplied by a factor, $\gamma$, to make them compatible with the scale of the electric measurements:

$$\gamma = x_1 \times \sqrt{\frac{M \text{Tr} B}{E \text{Tr} A}}, \quad (126)$$

where $x_1$ is a parameter which can be adjusted to alter the relative contributions of the electric and magnetic sensors. Hasson (1991) demonstrated by separate electric and magnetic inversions that an electric solution was considerably more localised than a magnetic solution, although this was undoubtedly due to the choice of the electrical conductivity model. A simultaneous analysis of both electric and magnetic measurements yielded a solution which was very similar to that for electric data alone, but which was slightly more focal.

Hasson (1991) also considered a second method of combining electric and magnetic measurements. In this method, a primary current image is produced using electric data only. This image is then used as the probability weight for an inversion using magnetic data only, such that:

$$p(r) = |\mathbf{J}_e(r)|^{x_2}, \quad (127)$$

in which $\mathbf{J}_e(r)$ is the primary current reconstructed from the electric data, and $x_2$ is an adjustable parameter which alters the influence of the electric solution on the second, magnetic, inversion. A solution using this algorithm was found to be less precise than a solution from electric data only, but more precise than a solution from magnetic data only.

In this chapter, a description is first given of two anatomically constrained source models for visual cortex. The first is a simple cruciform model for simulation purposes, and the second is a realistic model suitable for analysing real data. The inversion algorithm by which the dipole magnitudes are determined is then described. Two simulations follow, the first uses the simple cruciform model to demonstrate the value of using both electric and magnetic measurements simultaneously in conjunction with a restricted source space, while the second simulation uses the realistic source model to address the question of what happens when the source space does not coincide with the true morphology of the cortex. It is shown that a single average source model can be used for source analyses in a variety of individuals with differing anatomies. It should be noted that while the simulations described in this chapter relate specifically to visual cortex, the general method can also be applied to other areas of
neuro-electromagnetic research. If this is required, only the source model and volume conductor need to be changed. A preliminary report of the work in this chapter is given in Bedford (1992).

Definition of the Source Space

The physical arrangement of primary visual cortex is often approximated by the cruciform model (Jeffreys and Axford 1972a; Okada 1983b). Because of its geometrical simplicity, the cruciform model is ideal for a preliminary definition of the layout of primary visual cortex. Hence, this model has been used as a source space with which to perform simple computer simulations. The model is shown diagrammatically in figure 27. Because of the need to show all of the constrained current dipoles, it has been necessary to show the model in three separate sections, but in principle, it consists of eight surfaces, each of which corresponds to one of the eight distinctly-oriented surfaces of primary visual cortex (see figure 5). The 6mm-wide vertical gap represents the interhemispheric fissure, while the two 4mm-wide horizontal gaps represent the calcarine fissures. Three current dipoles are located on each surface, having x-coordinates of 25mm, 40mm and 55mm, respectively. The dipoles are oriented perpendicular to the surfaces of the model, to represent perpendicular neural primary current.

Clearly, this type of simple model is not suitable for analysing real data, as it is a gross simplification. Some other method of defining the source space must be found. The most sophisticated method is to use magnetic resonance images to define the locations and orientations of the various cortical surfaces. Current dipoles can be arrayed throughout a region of cortex using a segmentation method (George et al. 1991, 1993; Fuchs et al. 1993). However, this approach has several disadvantages. The main drawback is that it is costly to perform a magnetic resonance scan on each individual subject for whom a neuro-electromagnetic inverse solution is required, and this would preclude the algorithm from being clinically useful. Another difficulty is that of registering the magnetic resonance images with the locations of the neuro-electromagnetic sensors. If the relationship between the sensors and the underlying cortex is not modelled correctly, then the source space will be displaced from the true sources and systematic reconstruction errors could be introduced.

An alternative strategy, which is adopted in this thesis, is to define a standard or generic brain model, which is the same for all individuals. This overcomes the need for magnetic resonance images and is therefore inexpensive and clinically viable. However, the possibility of there being a discrepancy between the locations of the true sources and the location of the standard source space needs to be addressed properly, as each individual subject will have an anatomy which deviates from that on which the standard
source space is based. This question is considered further in the second set of simulations, later in this chapter. Derivation of a standard source space is now described.

It was seen in chapter two that visual areas V1 to V5 are most prominent in the processing of visual signals. However, the precise locations of areas V1 to V5 have only recently been identified in human (Clarke and Miklossy 1990), so it is common in discussing the origins of evoked responses to refer to the Brodmann classification, and this is the approach that has been adopted in the following work. The author also wished to use a stereotaxic atlas (Talairach and Tournoux 1988) which was based on Brodmann's classification, and this was an additional reason for using the Brodmann system.

Previous source analyses of visual evoked electric and magnetic responses indicated that the sources were limited to areas 17, 18 and 19 (Jeffreys and Axford 1972a,b; Richer et al. 1983; Butler et al. 1987; Maier et al. 1987). These results were
consistent with those of positron emission tomography experiments, where a visual stimulus elicited a response which was mainly confined to areas 17 and 18 (Fox and Raichle 1984, 1985; Fox et al. 1986, 1987), but extended, in some cases, to area 19 (Corbetta et al. 1991). Moreover, it has recently been shown that areas V1 to V5, which are the most important visual areas, are situated within areas 17, 18 and 19 (Clarke and Miklossy 1990). Hence, the visual source model was confined to areas 17, 18 and 19. In the light of the above experimental evidence, modelling a larger area of cortex would have been unnecessary, and would have reduced the effectiveness of the restricted source space, as it would have increased the number of source parameters to be determined.

A number of model and preserved brains were first studied to determine the number and approximate position of the dipoles required to represent primary current within areas 17, 18 and 19. Reference was also made to several statistical studies describing the inter-individual variability of the calcarine sulcus (Steinmetz et al. 1989; Myslobodsky et al. 1991). The dipoles were arranged so that each one represented around 100-400mm$^2$ of cortex. This spacing was chosen so that the major fissures such as the calcarine fissure were incorporated, but minor sulci and gyri were neglected. A further reason for this choice of spacing was that it provided optimal solution accuracy and resolution. If a restricted source space contains a large number of closely-spaced current dipoles, the problem is very underdetermined and an accurate solution is difficult to achieve because of the non-uniqueness: the method suffers from the same difficulties as unconstrained estimation (George et al. 1991, 1993). The close dipole spacing does not improve the resolution, because the resolution is limited by the diffuse, distributed nature of the regularised source reconstruction. For example, Wang (1993) reduced the spacing of the current dipoles in the source plane and the resulting image did not show any greater detail. The observable detail is limited by the sensor geometry, noise level and regularisation process.

The converse, however, is that if the dipoles in the source space are decreased in number, the detail in the solution remains similar, but the underdetermined nature of the problem is reduced. This is clearly desirable. Pursuing this approach to the extreme yields a source model containing so few dipoles that two sources cannot be resolved as there are insufficient dipoles to properly represent them. In other words, the dipole spacing exceeds the resolution of the inversion process and becomes a limit to the detail that can be observed in the solution. Hence, it is necessary to choose the dipole spacing for the source model to be comparable with the inherent resolution of the inversion process, so that it enables an overdetermined problem to be obtained while not itself becoming a limit to the source resolution. This optimal dipole spacing was estimated at 10-20mm on the basis of a variety of published distributed-source analyses (e.g. Ioannides et al. 1990; Ahlfors et al. 1992; Wang 1993).

The precise location and orientation of each dipole was determined using the stereotaxic atlas of Talairach et al. (1967), which provided average sections based on
thirty hemispheres. This prevented the source model from representing a single atypical brain. However, where the atlas was insufficient, the atlas of Talairach and Tournoux (1988) had to be used; this was based on a single subject. Equidistant dipole sites were marked on a template placed over a single sagittal section of the brain. The template served to define the overall layout of the dipoles. However, it was recognised that not all of the dipoles should lie in the same sagittal plane, so the template was treated as an approximate guide only. A given dipole was located on the appropriate coronal section. The lateral location of the dipole was determined on this section. Invariably, the lateral coordinate was different to that of the single sagittal section used to find the overall dipole layout. Hence, the new sagittal section was found, the original template was laid over this and dipole location transferred onto it. After several iterations of this process, a location was found for the dipole, such that it had approximately the correct location in the sagittal plane, as given by the overall scheme on the template, and also the correct lateral position, as given by the corresponding coronal section. The position coordinates of the dipole, in the anterior-posterior commissure system, were then recorded. The dipole site was taken to be on the border between the grey and white matter. The dipoles on the opposite surfaces of the calcarine fissure were located 4 mm apart, as the fissure was shown only as a single line in the atlas of Talairach et al. (1967). A perpendicular to the cortex, passing through the dipole site, was then drawn on both the sagittal and coronal sections, enabling the dipole orientation to be specified. This complete process was repeated until all of areas 17, 18 and 19 had been covered with equidistant dipoles. An exception to the procedure occurred for two lateral dipoles, which could not be represented properly on a sagittal section. A separate axial template was therefore made for these two dipoles, and they were located by iteration between the axial and coronal planes.

Finally, the dipole coordinates were converted into the coordinate system of the spherical conductivity model. This was achieved by drawing a circle of 68 mm radius on the outer surface of the statistical average brain of Talairach et al. (1967). The centre of this circle formed the origin of the occipital coordinate system, and the x-axis was chosen to coincide with the inion. The z-axis was constructed in the medial plane, and the y-axis was constructed perpendicular to it, directed towards the right side of the head. After coordinate transformation, it was found that several dipoles had a radius greater than that of the white matter conductivity sphere, which was 66 mm. Since calculation of the electric scalp potentials required that the source current dipoles lay within the innermost volume conductor sphere, it was necessary to relocate the outlying dipoles to slightly within the white matter. Hence, the entire location procedure was repeated for these dipoles only, this time taking the dipole site to be just within the white matter. This yielded corrected coordinates which were compatible with the forward problem calculation method. The final coordinates of the 42 dipoles required to represent areas 17, 18 and 19 are given in table 3, and the source model is shown diagrammatically in figure 28.
Inversion Technique

Having defined the source space, it is possible to describe the instantaneous electric potential difference, $e_i$, between the $i^\text{th}$ point of an $E$-electrode montage and an arbitrary reference site, as:

$$e_i = \sum_{j=1}^{S} I_q^i q_j, \quad i = 1...E, \quad (\text{Volts})$$  \hspace{1cm} (128a)

while the instantaneous magnetic flux gradient, $m_i$, obtained at each of $M$ magnetometer sites, is:

$$m_i = \sum_{j=1}^{S} I_q^m q_j, \quad i = 1...M, \quad (\text{Webers})$$  \hspace{1cm} (128b)

The values of $q_j$ are the signed magnitudes (hereafter magnitudes) of the $S$ current dipoles in the source space. (A negative magnitude denotes a current source oriented in the opposite direction to the source space current dipole.) The detector sensitivity factors $I_q^i$ and $I_q^m$ are determined using a spherical conductivity model of the occipital region of the head, as described in chapter three.

The reference electrode must be sited so that it lies on the surface of the spherical conductivity model. Since the spherical model was chosen to represent only the occipital region of the head, the reference must also lie in the occipital region. Throughout this chapter and the following chapters, the inion is used as the reference site. The factors $I_q^i$ are therefore calculated by finding the absolute potential at the $i^\text{th}$ point in the electrode montage, due to a unit dipole at the $j^\text{th}$ source space site, and subtracting from it the corresponding absolute potential at the inion. Using the inion as a reference site does not mean that evoked potentials have to be recorded relative to the inion; recording can be carried out relative to a frontal reference and the

Table 3 (overleaf). Coordinates of the 42 current dipoles which form the source space for visual cortex. $dx$, $dy$ and $dz$ are the Cartesian components of the unit vector describing dipole orientation (arbitrary units). LH: left hemisphere; OP: occipital pole; SC: superior surface of calcarine fissure; IC: inferior surface of calcarine fissure; CG: cuneal gyrus; LG: lingual gyrus; D: dorsal surface; V: ventral surface; L: lateral surface; RH: right hemisphere.
<table>
<thead>
<tr>
<th></th>
<th>x (mm)</th>
<th>y (mm)</th>
<th>z (mm)</th>
<th>dx</th>
<th>dy</th>
<th>dz</th>
</tr>
</thead>
<tbody>
<tr>
<td>LH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OP</td>
<td>63</td>
<td>-8</td>
<td>15</td>
<td>-1.0</td>
<td>0.0</td>
<td>0.04</td>
</tr>
<tr>
<td>SC</td>
<td>27</td>
<td>-16</td>
<td>20</td>
<td>-0.57</td>
<td>0.0</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>-16</td>
<td>24</td>
<td>0.14</td>
<td>0.0</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>-16</td>
<td>22</td>
<td>0.14</td>
<td>0.0</td>
<td>0.99</td>
</tr>
<tr>
<td>IC</td>
<td>29</td>
<td>-16</td>
<td>16</td>
<td>0.57</td>
<td>0.0</td>
<td>-0.82</td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>-16</td>
<td>20</td>
<td>-0.14</td>
<td>0.0</td>
<td>-0.99</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>-16</td>
<td>18</td>
<td>-0.14</td>
<td>0.0</td>
<td>-0.99</td>
</tr>
<tr>
<td>CG</td>
<td>36</td>
<td>-3</td>
<td>36</td>
<td>0.0</td>
<td>-1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>-3</td>
<td>31</td>
<td>0.0</td>
<td>-1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>58</td>
<td>-3</td>
<td>24</td>
<td>0.0</td>
<td>-1.0</td>
<td>0.0</td>
</tr>
<tr>
<td>LG</td>
<td>38</td>
<td>-3</td>
<td>18</td>
<td>0.0</td>
<td>-1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>-3</td>
<td>16</td>
<td>0.0</td>
<td>-1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>58</td>
<td>-3</td>
<td>13</td>
<td>0.0</td>
<td>-1.0</td>
<td>0.0</td>
</tr>
<tr>
<td>D</td>
<td>41</td>
<td>-17</td>
<td>48</td>
<td>-0.57</td>
<td>0.0</td>
<td>-0.82</td>
</tr>
<tr>
<td></td>
<td>52</td>
<td>-13</td>
<td>38</td>
<td>-0.69</td>
<td>0.0</td>
<td>-0.72</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>-8</td>
<td>26</td>
<td>-0.81</td>
<td>0.0</td>
<td>-0.59</td>
</tr>
<tr>
<td>V</td>
<td>26</td>
<td>-29</td>
<td>-2</td>
<td>-0.44</td>
<td>-0.43</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>41</td>
<td>-25</td>
<td>3</td>
<td>-0.11</td>
<td>-0.33</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>-21</td>
<td>3</td>
<td>-0.11</td>
<td>-0.14</td>
<td>0.99</td>
</tr>
<tr>
<td>L</td>
<td>40</td>
<td>-43</td>
<td>29</td>
<td>-0.60</td>
<td>0.70</td>
<td>-0.40</td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>-29</td>
<td>19</td>
<td>-0.86</td>
<td>0.41</td>
<td>-0.31</td>
</tr>
<tr>
<td>RH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OP</td>
<td>63</td>
<td>8</td>
<td>15</td>
<td>-1.0</td>
<td>0.0</td>
<td>0.04</td>
</tr>
<tr>
<td>SC</td>
<td>27</td>
<td>16</td>
<td>20</td>
<td>-0.57</td>
<td>0.0</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>16</td>
<td>24</td>
<td>0.14</td>
<td>0.0</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>16</td>
<td>22</td>
<td>0.14</td>
<td>0.0</td>
<td>0.99</td>
</tr>
<tr>
<td>IC</td>
<td>29</td>
<td>16</td>
<td>16</td>
<td>0.57</td>
<td>0.0</td>
<td>-0.82</td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>16</td>
<td>20</td>
<td>-0.14</td>
<td>0.0</td>
<td>-0.99</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>16</td>
<td>18</td>
<td>-0.14</td>
<td>0.0</td>
<td>-0.99</td>
</tr>
<tr>
<td>CG</td>
<td>36</td>
<td>3</td>
<td>36</td>
<td>0.0</td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>3</td>
<td>31</td>
<td>0.0</td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>58</td>
<td>3</td>
<td>24</td>
<td>0.0</td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td>LG</td>
<td>38</td>
<td>3</td>
<td>18</td>
<td>0.0</td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>3</td>
<td>16</td>
<td>0.0</td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>58</td>
<td>3</td>
<td>13</td>
<td>0.0</td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td>D</td>
<td>41</td>
<td>17</td>
<td>48</td>
<td>-0.57</td>
<td>0.0</td>
<td>-0.82</td>
</tr>
<tr>
<td></td>
<td>52</td>
<td>13</td>
<td>38</td>
<td>-0.69</td>
<td>0.0</td>
<td>-0.72</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>8</td>
<td>26</td>
<td>-0.81</td>
<td>0.0</td>
<td>-0.59</td>
</tr>
<tr>
<td>V</td>
<td>26</td>
<td>29</td>
<td>-2</td>
<td>-0.44</td>
<td>0.43</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>41</td>
<td>25</td>
<td>3</td>
<td>-0.11</td>
<td>0.33</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>21</td>
<td>3</td>
<td>-0.11</td>
<td>0.14</td>
<td>0.99</td>
</tr>
<tr>
<td>L</td>
<td>40</td>
<td>43</td>
<td>29</td>
<td>-0.60</td>
<td>-0.70</td>
<td>-0.40</td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>29</td>
<td>19</td>
<td>-0.86</td>
<td>-0.41</td>
<td>-0.31</td>
</tr>
</tbody>
</table>
Figure 28. (a) Medial and (b) axial aspects of right human visual cortex showing the 21 current dipoles which form the right half of the constrained source space. Arrows represent current dipoles in the plane of the page; circles represent current dipoles directed into the page. Each current dipole describes the primary current in 100-400mm² of cortex. PO: parieto-occipital sulcus; C: calcarine sulcus; IH: interhemispheric sulcus.

potential at the inion subtracted from the potential at all the other electrodes at each time instant to convert the measurements to an inion-referenced form. This type of conversion alters the appearance of the recorded waveforms and topographic maps, but does not alter the information content of the data, as it only amounts to a spatially uniform shift in potential for each time instant.

Equations 128a and 128b together form one linear system. This linear system is the discrete equivalent of the lead field representation of the forward problem:

\[
e_i = \int_G L_i^e(r) \cdot J_p(r) d^3r, \quad i = 1...E,
\]

\[
m_i = \int_G L_i^m(r) \cdot J_p(r) d^3r, \quad i = 1...M,
\]

where \( G \) is now a continuous source space, \( L_i^e(r) \) and \( L_i^m(r) \) are the electric and magnetic lead fields respectively, and \( J_p(r) \) is the continuous current density at \( r \). It can be seen from equations 128 and 129 that \( L_i^e \) is the component of the \( i \)th electrode lead field in the direction of the \( j \)th discrete current dipole, while \( L_i^m \) is the component of the \( j \)th magnetometer lead field in the direction of the \( j \)th discrete current dipole; i.e. if the \( j \)th dipole is located at \( r_j \) and is oriented in the direction of the unit vector \( n_j \), then \( L_i^e = L_i^e(r_j) \cdot n_j \) and \( L_i^m = L_i^m(r_j) \cdot n_j \).

Solving the inverse problem now consists of finding the \( S \) current dipole magnitudes which best fit the given electric and magnetic field distributions. The linear system of equations 128 is ill-conditioned as a result of the ten orders of magnitude
difference between the electric potentials, \( e_i \), and magnetic flux gradients, \( m_i \). Each of the \( E + M \) equations is therefore divided by the standard deviation of the associated external measurement to reduce the ill-conditioning:

\[
\frac{e_i}{\sigma_e} = \sum_{j=1}^{S} \frac{I_{ij}^e}{\sigma_e} q_j, \quad i = 1 \ldots E,
\]  
(130a)

\[
\frac{m_i}{\sigma_m} = \sum_{j=1}^{S} \frac{I_{ij}^m}{\sigma_m} q_j, \quad i = 1 \ldots M.
\]  
(130b)

The electric standard deviation, \( \sigma_e \), is assumed to be constant over all the electric measurements and the magnetic standard deviation, \( \sigma_m \), is assumed to be constant over all the magnetic measurements. In matrix form, equations 130 are as follows:

\[
\begin{bmatrix}
\begin{bmatrix}
\frac{e_1}{\sigma_e} & \frac{e_2}{\sigma_e} & \ldots & \frac{e_E}{\sigma_e} \\
\frac{m_1}{\sigma_m} & \frac{m_2}{\sigma_m} & \ldots & \frac{m_M}{\sigma_m}
\end{bmatrix}
\end{bmatrix}
= 
\begin{bmatrix}
\begin{bmatrix}
I_{11}^e & I_{12}^e & \ldots & I_{1E}^e \\
I_{M1}^m & I_{M2}^m & \ldots & I_{Mm}^m
\end{bmatrix}
\end{bmatrix}
\cdot
\begin{bmatrix}
q_1 \\
q_2 \\
\vdots \\
q_{S-1} \\
q_S
\end{bmatrix}
\]  
(131a)

or simply:

\[
f = Lq.
\]  
(131b)

This linear system can be overdetermined, exactly determined, or underdetermined, depending on whether the total number of measurements, \( E + M \), is greater than, equal to, or less than the number of dipoles, \( S \), in the source model. Non-uniqueness can exist under any of these conditions.

Consider the case when the linear system is overdetermined. It is unlikely, due to the presence of noise in the measured data, that all the equations in the system will be consistent: if there are \( S \) sources in the source space, then a solution obtained using \( S \) rows of the system will be different to that obtained using a different \( S \) rows. It is therefore necessary to find a solution, \( \tilde{q} \), which satisfies, as closely as possible, all the
equations of the system; i.e. it is necessary to find \( \tilde{q} \) such that \( |f - L\tilde{q}|^2 \) is minimal. This is equivalent to the chi-square minimisation:

\[
\chi^2 = \sum_{i=1}^{E} \left( \frac{e_i - \sum_{j=1}^{S} L_{ij} \tilde{q}_j}{\sigma_e} \right)^2 + \sum_{i=1}^{M} \left( \frac{m_i - \sum_{j=1}^{S} L_{ij} \tilde{m}_j}{\sigma_m} \right)^2 = \min .
\]  

(132)

However, even in this overdetermined case, the solution can become non-unique, as there can be several source distributions which all minimise the chi-square to the same minimum value. Typically, this occurs when the source space allows the existence of "silent" sources or source distributions. In this case, it is possible to find a solution which minimises the chi-square, but any multiple of the silent source distribution can be added without altering the forward field and thus without increasing the chi-square, implying that the chi-square solution to the linear system is non-unique.

If, on the other hand, the total number of independent measurements is less than the number of dipoles in the source space, the linear system is underdetermined. Again, there is no unique solution. Insufficient measurements are the most common reason for the existence of non-uniqueness.

Whenever the solution is non-unique, and for whatever reason, it is necessary to choose a single solution from the variety of current distributions which all minimise the chi-square equally well, and it is reasonable to seek the most probable of these possible distributions. Hence, following the approach of Clarke and Janday (1989) and Ioannides et al. (1990), a Gaussian prior probability density function is assigned to the magnitude of each dipole in the source space:

\[
P(q_j) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left( -\frac{q_j^2}{2\sigma^2} \right), \quad j = 1...S,
\]  

(133)

where \( \sigma \) is the standard deviation of the function and is taken to be equal at all points in the source space. The probability density function is chosen to be Gaussian primarily to simplify the resulting mathematics, but this choice can be justified on the ground that the probability of an extremely large dipole moment somewhere in the source space is low, whilst the probability of zero dipole moment (reflecting neural inactivity) is clearly much greater. A Gaussian probability density function has also been used to model the amplitude histogram of a postsynaptic potential (Turner and West 1993), and is therefore a reasonable function to use for primary current, which results directly from postsynaptic activity. It follows that the probability density function for the occurrence of some arbitrary combination of current dipole magnitudes is given by:
\[ P(q) = P(q_1, q_2, \ldots, q_s) = \prod_{j=1}^{s} \frac{1}{\sigma \sqrt{2\pi}} \exp \left( -\frac{q_j^2}{2\sigma^2} \right), \]

\[ \propto \exp \left[ -\frac{1}{2\sigma^2} \left( q_1^2 + q_2^2 + \cdots + q_s^2 \right) \right]. \] (134)

Maximising this expression, it is found that the most probable current distribution is that which minimises the Euclidean norm of the dipole magnitude vector, \( q \). Thus, of all the solutions which minimise the chi-square, the most probable (in both the mean and mode senses) is the minimum norm estimate. Because this estimate has minimum Euclidean norm of all the solutions which minimise the chi-square, it is sometimes referred to as the minimum-norm least-squares estimate (Wang et al. 1992).

In practice, both conventional least squares solutions and minimum norm solutions are unstable in the presence of noise. That is, small variations in the electric and magnetic data, due to noise, cause large changes in the solution. A solution produced from a set of noisy data is therefore likely to be grossly different to the true solution which would have been obtained had the data been free of noise. Hence, some form of regularisation must be applied to the solution.

Conveniently, singular value decomposition can be used to provide a stable solution in both the unique and non-unique cases. A suitable computer algorithm is given by Press et al. (1989). Although singular value decomposition is defined for any matrix, this algorithm requires that matrix \( L \) has at least as many rows as columns. Hence, if \( E + M < S \), matrix \( L \) is augmented with zeros, so that it becomes a square matrix of dimension \( S \times S \), and \( f \) is similarly augmented so that it has \( S \) elements. The number of rows in the augmented matrix \( L \) is then denoted by \( R \): \( R = S \) when \( E + M < S \) and \( R = E + M \) when \( E + M > S \). The augmented matrix \( L \) is decomposed, such that:

\[ L = U \Lambda V^T, \] (135a)

where \( U \) is an \( R \times S \) orthogonal matrix and \( V \) is an \( S \times S \) orthogonal matrix. \( \Lambda \) is an \( S \times S \) diagonal matrix containing the singular values of \( L \) ranked in descending order:

\[ \Lambda = \text{diag} [\lambda_1, \lambda_2, \cdots, \lambda_j, \lambda_{j+1}, \cdots, 0, \cdots, 0], \quad \lambda_j \geq \lambda_{j+1}. \] (135b)

The pseudoinverse of \( \Lambda \) is then defined as:

\[ \Lambda^* = \text{diag} [\lambda_1^*, \lambda_2^*, \cdots, \lambda_j^*, \cdots, 0, \cdots, 0], \] (136a)

such that:

\[ \lambda_j^* = \frac{1}{\lambda_j}, \quad \lambda_j > \lambda_1. \]
\[ \lambda_j^* = 0 \quad \text{otherwise,} \quad j = 1 \ldots S, \]  

\[ \hat{q} = V \Lambda^* U^T f. \]  

(137)

where \( \gamma \) is the maximum signal-to-noise ratio of the data set, obtained from the largest element in the data matrix, \( f \). The regularised solution, \( \hat{q} \), is calculated according to the formula:

The threshold for zeroing elements of \( \Lambda^* \) is set according to the signal-to-noise ratio of the data in \( f \) in order to produce a stable solution. The result is that no attempt is made to fit the solution to the data more exactly than the noise level. Despite the misfit, the solution is representative of the true source distribution and is not excessively affected by noise present in the measurements.

An interpretation of the above algorithm is now provided. Any combination of \( S \) current dipole magnitudes can be viewed as a vector in \( S \)-dimensional current space. Similarly, the \( R \) data values in the augmented matrix \( f \) can be considered as a vector in \( R \)-dimensional data space. The matrix \( L \) describes a linear mapping between a vector, \( q \), in current space to another vector, \( Lq \), in data space. \( L \) has \( S \) columns, so the range space of \( L \), which is the space of possible vectors \( Lq \), is an \( S \)-dimensional subspace within the data space; this subspace may or may not fill the entire data space, depending on whether the linear system is over- or under-determined. Equation 137 states that the solution, \( \hat{q} \), is a linear combination of the columns of \( V \). In fact, matrix \( V \) provides an orthonormal basis with which to construct the solution. The columns of \( U \) provide an orthonormal basis for the range space of \( L \), so that \( L\hat{q} \) is a combination of the columns of \( U \).

First consider the case when the linear system given by equations 131 is overdetermined and a unique solution can be found. In this case, \( R > S \) and the range of \( L \) is an \( S \)-dimensional subspace which does not fill all of the data space. A particular measurement set, \( f \), may lie within the range of \( L \), but it is more likely that it lies outside of it, in which case no current vector, \( q \), can satisfy the linear system exactly. Singular value decomposition constructs an exact solution, \( \hat{q} \), if one is available, but otherwise constructs a solution such that the distance (in data space) between \( \hat{Lq} \) and \( f \) is minimal, given that \( \hat{Lq} \) lies within the range of \( L \) and \( f \) lies outside. In other words, the solution minimises \( |f - Lq|^2 \); i.e. it minimises the chi-square.

The next case to consider is when the system is exactly determined, i.e. when \( R = S \). The range of \( L \) now completely fills the \( S \)-dimensional data space. The measurement vector, \( f \), must lie within the range of \( L \), so an exact solution, \( \hat{q} \), can be found, such that \( L\hat{q} = f \).

The underdetermined case, when \( E + M < S \), can be treated as the exactly determined case, as the matrices are augmented with zeros so that \( R = S \). However, the square augmented matrix \( L \) will be singular, so that a unique solution cannot be
obtained. Since non-uniqueness can occur when the linear system is overdetermined, exactly determined, or underdetermined, it is now considered in detail.

Non-uniqueness is manifested when there are current vectors, \( \mathbf{q}_n \), in the current space, which are mapped to the zero vector in data space; \( \text{i.e.} \ L\mathbf{q}_n = \mathbf{0} \). Such vectors \( \mathbf{q}_n \) form a subspace within the current space; this subspace is referred to as the nullspace. Suppose that \( R > S \) and a least squares solution, \( \mathbf{q} \), has been found, and it is mapped to \( L\mathbf{q} \). A vector \( \mathbf{q} + \alpha \mathbf{q}_n \), where \( \alpha \) is a scalar, is mapped to \( L(\mathbf{q} + \alpha \mathbf{q}_n) = L\mathbf{q} + \alpha L\mathbf{q}_n = L\mathbf{q} \) since \( L\mathbf{q}_n = \mathbf{0} \). Hence, if \( \mathbf{q} \) is a least squares solution to the linear system, \( \mathbf{q} + \alpha \mathbf{q}_n \) is also a solution. Similarly, if \( R = S \) and \( \mathbf{q} \) is an exact solution to the linear system, \( \mathbf{q} + \alpha \mathbf{q}_n \) is also an exact solution. Since there are a whole series of vectors \( \mathbf{q}_n \) and \( \alpha \) can be any scalar, there are a whole series of solutions which satisfy the linear system: it is non-unique.

Singular value decomposition overcomes this situation readily. The number of zero diagonal elements in \( \Lambda \) indicates the dimensionality of the nullspace. This dimensionality is referred to as the nullity. The last columns in \( V \), which correspond to the zero elements of \( \Lambda \), form an orthonormal basis for the nullspace. The non-uniqueness is countered simply by not including any null vectors in the solution: the solution vector, \( \mathbf{q} \), is constructed from the first columns of \( V \), while setting \( \lambda_j^* = 0 \) when \( \lambda_j = 0 \) ensures that the solution does not span the nullspace. The result is that the solution has the minimum norm of all the possible solutions. Hence, whenever there is non-uniqueness, singular value decomposition returns the minimum norm solution.

In practice, \( \lambda_j^* \) is set to zero when the corresponding values of \( \lambda_j \) are small, but non-zero. A small value of \( \lambda_j \) implies that only a small multiple of the \( j^{th} \) orthonormal column of \( U \) is required to construct the data vector, \( f \). Hence, neglecting that orthonormal component of \( U \), thereby constraining vectors \( L\mathbf{q} \) to lie in a subspace of dimensionality less than that of the true range space of \( L \), only introduces a small increase into the chi-square, \( f - L\mathbf{q} \). A solution vector, \( \mathbf{q} \), which is mapped to the reduced data subspace, is provided by application of equation 137, neglecting the \( j^{th} \) column of \( V \) when \( \lambda_j \) is small. If \( \lambda_j^* \) were not set to zero for small values of \( \lambda_j \), \( \lambda_j^* \) would be large, and the solution would contain a large multiple of the \( j^{th} \) column of \( V \); such a large multiple would change dramatically with changes in \( f \) due to noise: the solution would be unstable. Hence, the criterion \( \lambda_j^* = 0 \) when \( \lambda_j < \lambda_j^*/\gamma \) ensures that the solution is stable, despite noise in the data, at the expense of a small increase in the chi-square or decrease in the goodness-of-fit.

Generally, a solution is required for each of \( T \) time instants, as opposed to just one time instant. In this case, the equations to be solved are:

\[
\mathbf{f}(t) = L\mathbf{q}(t), \quad t = 1...T.
\]  

t (138)
Matrix $\mathbf{L}$ is decomposed just once, as described above, and the spatio-temporal solution is obtained from the relation:

$$\tilde{q}(t) = \mathbf{V}\Lambda^*(t)\mathbf{U}^T\mathbf{f}(t), \quad t = 1...T,$$

(139)

where $\Lambda^*(t)$ is the pseudoinverse of $\Lambda$. Although $\Lambda$ is only produced once, $\Lambda^*(t)$ has to be computed for each time instant, as the signal-to-noise ratio of the data is effectively different at each time instant:

$$\Lambda^*(t) = \text{diag}[\lambda_1^*, \lambda_2^*, \lambda_3^*, \ldots, \lambda_s^*, \ldots, 0, \ldots, 0],$$

(140a)

where the elements are now defined by:

$$\lambda_j^* = 1/\lambda_j, \quad \lambda_j > \frac{\lambda_1}{\gamma(t)}$$

$$\lambda_j^* = 0 \quad \text{otherwise}, \quad j = 1...S, \quad t = 1...T.$$  

(140b)

Thus, the generation and decomposition of matrix $\mathbf{L}$ have to be performed only once, and the entire sequence of solutions can then be found simply by matrix multiplication.

In summary, the above algorithm provides the following features:

1. If the system is overdetermined, it returns the closest stable approximation to the least squares solution, provided that this is unique.

2. If the system is overdetermined but there are several equally good least squares fits, the algorithm returns the closest stable approximation to the fit having the minimum norm.

3. If the system is exactly determined, the algorithm returns the closest stable approximation to the exact solution.

4. If the system is underdetermined, the algorithm returns the closest stable approximation to the minimum norm solution.

The goodness-of-fit of the solution is defined as the probability that the chi-square statistic could equal or exceed the observed minimal chi-square by chance, due to noise in the measurements. This probability is given by the chi-square distribution for $v$ degrees of freedom; $v$ being the difference between the number of independent measurements and the number of free parameters in the inversion. In this case, the
number of free parameters is provided by the number of non-zero elements in $\Lambda^*$. The
goodness-of-fit, $g$, is computed numerically using the incomplete gamma function, $P$:

$$g = 1 - P\left(\frac{\sqrt{\chi^2}}{2}, \frac{\nu}{2}\right).$$

(141)

As with all goodness-of-fit methods, it is necessary to define what is a "good fit" and
what is a "bad fit." Press et al. (1989) suggested that if the above goodness-of-fit
is larger than about 0.1, then it is acceptable, and if it is larger than $10^3$, then it may be
acceptable if the errors are non-normal or have been conservatively estimated. If the
goodness-of-fit is less than $10^3$, then the model can be called into question. Supek and
Aine (1993b) applied the method to multiple-dipole simulations and used a rejection
level of $10^{-2}$, but also noted that highly improbable values of goodness-of-fit could be
caused by underestimation of the noise level. When used in conjunction with the linear
inversion procedure given above, the goodness-of-fit is reduced by the regularisation
process. Hence, the convention adopted in this thesis is that a goodness-of-fit of
greater than 0.1 (10%) is very good, between 0.1 and $10^{-4}$ (10% - 10^{-3}%), is acceptable,
and less than $10^{-4}$ (10^{-9}%) is unlikely. This scheme is reliable even when noise levels
have to be estimated subjectively as opposed to being specified exactly in simulation
studies.

**Computer Simulations Using a Cruciform Model**

A number of computer simulations were carried out to assess the value of
restricting the distributed source space and to examine the effect of using electric and
magnetic data simultaneously. The cruciform-shaped 24-dipole source model described
in the section headed "Definition of the Source Space" was used as the basis of these
studies, as it enabled the physical principles to be established without the geometric
complications of a realistic source model.

**Method** A hypothetical source distribution was set up, consisting of six active current
dipoles in the left upper quadrant of the source model (figure 29). The six active
dipoles, each of 10nAm dipole moment, were intended to represent the pattern of
activity that might occur on stimulation of the lower right quadrant of the visual field.
Using the spherical conductivity model of chapter three, the electric potential and
magnetic flux gradient were calculated over a simulated 20-point occipital montage.
The montage consisted of four rows of five points; the points were spaced at 30mm
intervals and the middle point of the bottom row coincided with an imaginary inion
(figure 30). The electrodes were assumed to be point contacts, and the potentials were
Figure 29. Six simulated current dipole sources situated within the cruciform model.

Figure 30. Simulated occipital electrode and magnetometer montage. The broken lines indicate the position of the underlying source space.
calculated relative to the ion. The simulated gradiometer was a single channel device having a baseline of 50mm, a distance of 50mm between the middle and rear coils, and a coil diameter of 20mm. It was located radially to the head, within a dewar whose tail was 12mm thick, so that the front coil lay 12mm from the scalp. The resulting distributions of electric potential and magnetic flux gradient are given in figure 31. These distributions were used to reconstruct the original sources. In the case of the electric distribution, only the 19 potentials surrounding the ion were input to the inverse procedure, as the ion itself was being used as reference. (Supplying the zero ion potential to the inversion algorithm would have resulted in a null equation, $0 = 0$.) No regularisation was applied in the first instance, as the data were noise-free. However, the standard deviations of the electric and magnetic data were set to $10^{-12}$V and $10^{-22}$Wb, respectively; these figures were approximately $10^6$ times the data values. The purpose of this was to achieve a small degree of regularisation, just sufficient to overcome any computational instability due to round-off error.

In the first three simulations, only the electric data were used. In the first simulation, the entire 24-dipole source model was used as the source space. In the second simulation, the source space was restricted to the 12 dipoles of the left half of the cruciform model. The third simulation used a source space which consisted of only the 6 dipoles of the left upper quadrant. These inversions were chosen so that the linear system ranged from underdetermined to strongly overdetermined. This procedure was then repeated using the magnetic distribution alone. To examine the effect of using the electric and magnetic data simultaneously, the three simulations were repeated using both the electric and magnetic data together.

Figure 31. (a) Occipital distribution of simulated quarter-field visual evoked potentials. Contours are marked in units of $\mu$V. (b) The corresponding distribution of simulated quarter-field visual evoked magnetic fields. Contours are marked in units of fWb.
To evaluate the ability of the inversion algorithm to successfully analyse noisy electric and magnetic data, the above simulations were repeated after addition of Gaussian noise to both the electric and magnetic distributions. The noise values for addition to the electric (respectively magnetic) data were normally-distributed random numbers whose standard deviation was equal to 10% of the electric (respectively magnetic) signal having the largest modulus. This level of noise was considered to provide a realistic test of the algorithm's performance. The resulting electric and magnetic distributions are given in figure 32. In the inverse procedure, regularisation was applied as detailed in the previous section.

Results The outcome of the inversion based on noise-free electric measurements only is given in figure 33. In this figure and in the figures following, the three layers of the source space (see figure 27) are shown superimposed, and eccentricity is used to represent source radius. The eight dipole sites at $x = 25\text{mm}$ are shown in the diagrams as least eccentric and the eight dipole sites at $x = 55\text{mm}$ are most eccentric. It can be seen that the inversion using a 24-dipole source space yielded an accurate solution, but low-level spurious current was reconstructed in wrong quadrants. The dipolar distribution was determined correctly when both the 12- and 6-dipole source spaces were used. Similar solutions were obtained when only magnetic measurements were used for inversion (figure 34). When both electric and magnetic distributions were inverted simultaneously (figure 35), all three source spaces were found to give an accurate reproduction of the original dipolar distribution.

![Diagram](a)

Figure 32. (a) Occipital distribution of simulated quarter-field visual evoked potentials after addition of 10% noise. Contours are marked in units of $\mu\text{V}$. (b) The corresponding distribution of simulated quarter-field visual evoked magnetic fields, also after addition of 10% noise. Contours are marked in units of fWb.
Figure 33. Inverse solutions to the simulated visual evoked potentials, obtained (a) using all 24 dipoles as the source space, (b) using the 12 left dipoles as the source space, and (c) using the 6 left upper dipoles as the source space. The three layers of the source space are shown superimposed, with the deepest dipole sites shown as least eccentric and vice versa. The dipoles in (c) are 10nAm in magnitude; arrow width is proportional to dipole magnitude.
Figure 34. Inverse solutions to the simulated visual evoked magnetic fields, obtained (a) using all 24 dipoles as the source space, (b) using the 12 left dipoles as the source space, and (c) using the 6 left upper dipoles as the source space.
Figure 35. Inverse solutions obtained using both electric and magnetic measurements simultaneously. The results were produced (a) using all 24 dipoles as the source space, (b) using the 12 left dipoles as the source space, and (c) using the 6 left upper dipoles as the source space.
The solutions obtained by inverting noisy electric data are given in figure 36. The solution produced using the 24-dipole source space is approximately correct, except that the primary current extends over the medial surface of the right upper quadrant and the superior surface of the left lower quadrant, instead of being limited exclusively to the left upper quadrant. The source reconstruction is also biased towards the deeper region of the source space. Some spurious activity is observed in the lower and right regions of the source space. The result obtained using the 12-dipole source space shows similar effects, although the sources on the right upper medial surface and the spurious sources on the right calcarine fissure and right lower medial surface are now absent as the source space does not include the right hemisphere. The bias towards the deeper region of the source space appears to be more pronounced than when the 24-dipole source model was used. On the left upper medial surface, this is largely a visual effect caused by the adding together of sources which were previously spread across both medial surfaces. There is, however, a genuine increase in the depth bias on the left calcarine fissure, where the most superficial layer of the source space now contains only low-level spurious activity. When the 6-dipole source model is used, the source distribution is constrained to lie in the correct quadrant, so the inversion yields a reasonable representation of the true sources. The deep sources are over-emphasised and the superficial sources are under-emphasised, although this bias is comparable to that in the previous case, given that the deep source on the calcarine fissure was spread over two surfaces when the 12-dipole source space was used, but is now reconstructed as a single source of twice the magnitude. It is clear that the depth bias is most prominent on the calcarine fissure, where superficial activity is now entirely absent.

Solutions obtained using noisy magnetic data (figure 37) show similar features to the solutions from noisy electric data, notably that primary current spreads to adjacent surfaces when permitted by the source space, and low-level spurious current appears in inactive regions of the source space. There is also a depth bias, but this is now such that deep sources are underestimated. The bias is less severe than that seen in the electric solutions. When the 24-dipole source space is used, the superficial sources are reconstructed with approximately the correct magnitude if spread across adjacent surfaces is taken into account, while the intermediate and deep sources are progressively underestimated. However, in the 12- and 6-dipole solutions, both the superficial and intermediate sources are determined correctly; only the deep sources are underestimated. This is clearly an improvement on the electric inversions, where the superficial sources are underestimated and the deep sources are overestimated.

The simultaneous solutions to the noisy electric and magnetic measurements are given in figure 38. Current sources within these solutions still spread onto adjacent, parallel, surfaces whenever permitted by the source space, but the level of spurious activity, for a given size of source space, is slightly lower than in the separate electric and magnetic solutions. Most significantly, though, the depth bias seen in the previous
Figure 36. Inverse solutions to the simulated visual evoked potentials after addition of 10% noise, obtained (a) using all 24 dipoles as the source space, (b) using the 12 left dipoles as the source space, and (c) using the 6 left upper dipoles as the source space.
Figure 37. Inverse solutions to the simulated visual evoked magnetic fields after addition of 10% noise, obtained (a) using all 24 dipoles as the source space, (b) using the 12 left dipoles as the source space, and (c) using the 6 left upper dipoles as the source space.
Figure 38. Inverse solutions obtained using both electric and magnetic measurements simultaneously. The measurements included 10% noise and the results were produced (a) using all 24 dipoles as the source space, (b) using the 12 left dipoles as the source space, and (c) using the 6 left upper dipoles as the source space.
solutions is almost removed. When the 24-dipole source space is used, the superficial sources are slightly underestimated; when the 12-dipole source space is used, the underestimation is limited to the superficial source on the left upper medial surface, with the intermediate source on the left calcarine fissure appearing slightly diminished; but when the 6-dipole source space is used, all six sources are uniform in magnitude, as desired.

Discussion The performance of the anatomically constrained inversion procedure can be summarised as follows:

1. Solution accuracy is highest when the number of dipoles used in the source space is lowest.

2. Spurious current levels are lower when fewer dipoles are used in the source space.

3. The presence of noise in the electric and magnetic data causes the solution to be more distributed than the true source geometry. This smoothing effect is observed mainly on opposing surfaces of fissures and can only be removed by excluding one of the surfaces from the source model.

4. A solution using magnetic data only is more accurate than a solution using electric data only, as the magnetic solution suffers less from depth bias effects.

5. Solution fidelity and spurious current levels are optimised when electric and magnetic data are used simultaneously.

6. Over-emphasis of either deep or superficial sources can be reduced significantly by analysing electric and magnetic data simultaneously.

The results of the noise-free simulations simply illustrate that when the number of independent measurements is less than the number of unknowns, there is some error in the solution, but when the number of measurements exceeds the number of unknowns, the solution is correct. Hence, if nineteen electric or twenty magnetic measurements are used, a 24-dipole analysis yields a solution which has some erroneous sources, whereas a 12-dipole or 6-dipole analysis produces an entirely accurate solution. If nineteen electric and twenty magnetic measurements are used simultaneously, then 24-, 12- and 6-dipole analyses are all accurate, because there are 39 independent measurements. Perhaps the most surprising result of these simulations is that the spurious current level remains remarkably low when the linear system is underdetermined, and the true source distribution is reconstructed virtually perfectly.
Earlier inversions of noise-free data (Bedford 1992) showed that an unregularised solution is very poor when twenty measurements are used to determine thirty unknown dipole moments. The accuracy here must therefore be attributed to the fact that there are only four or five more unknowns than measurements.

The noise-free results also show that anatomically constrained source estimation is a worthwhile objective. As environmental noise levels are reduced, due to improved recording and signal processing techniques, noise-free measurements will become more readily available. Under these conditions, conventional minimum norm estimates cannot be regularised, and the resulting solution, while being a valid interpretation of the data, may not be accurate. In fact, unregularised minimum norm estimates of as few as thirty parameters are completely unreliable if there are only twenty measurements (Bedford 1992). Hence, the ability of the anatomically constrained procedure to produce an unambiguous solution by means of an overdetermined system, is valuable.

The inversions of noisy measurements show several interesting effects. The first is the spread of primary current across adjacent, parallel surfaces such as the opposing surfaces of a fissure. In the simulations, the dipoles on opposite surfaces of the interhemispheric fissure are 6mm apart, and those on opposite surfaces of the calcarine fissure are 4mm apart. The "smearing out" of a current source onto both opposing surfaces is due to the regularisation procedure, which is why the effect is absent in the noise-free simulations. This effect was also noted by Wang et al. (1993b) in similar simulations. The phenomenon occurs because both surfaces of a fissure, when active, provide similar external field distributions. The difference between the distributions is often less than the level of noise which is present in the measured fields. The algorithm is effectively unable to distinguish between a subtle difference in source location and variation of the data due to noise. In other words, the resolution of the inverse procedure is inadequate to distinguish two sources on adjacent surfaces. This reflects the comment made in the introduction, that decreasing the dipole spacing does not necessarily yield higher resolution, because resolution is determined primarily by the regularisation process. Unfortunately, the problem of "smearing" across fissures cannot be overcome, apart from removing one of the surfaces from the source space (or reducing the noise level). If this cannot be done, then no conclusions should be drawn about which side of a fissure a source is located on.

The simulations in this study use far fewer dipoles than in previous studies, for the above reason. Using a large number of dipoles does not improve the resolution, but produces problems of non-uniqueness and diffuseness that are normally associated with conventional unconstrained algorithms. For example, Wang (1993) used a grid of 356 dipoles to invert 37 simulated magnetic measurements from a single current dipole, and found that the source was reconstructed at the superficial edge of the source space. In the simulations presented above, the solution produced by analysing twenty magnetic measurements using a 24-dipole source space does have a superficial bias, but it is not
that pronounced, and certainly does not deposit all of the source current in the superficial layer of the source space. Hence, the restricted source space used here helps to overcome problems of depth bias.

The solutions of Wang (Wang et al. 1992; Wang 1993) were diffuse because of the regularisation procedure, and a fine grid spacing did not produce any more detail in the source image than a coarse spacing (Wang 1993). Similarly, the solutions produced by Fuchs et al. (1993) using magnetic resonance images suffered from large "whirls" surrounding the constrained current distribution: although the source space was "constrained," it was so extensive and finely-spaced that the solution amounted to a conventional planar minimum norm estimate. Such problems of diffuseness do not exist in the author's simulations, simply because the dipoles are sufficiently far apart that they do not permit a diffuse solution, except across opposing surfaces of fissures. It is therefore concluded that anatomical constraint is most effective when the source space consists of a few strategically-positioned current dipoles, rather than a fine grid of closely-spaced dipoles.

Spurious current sources around the perimeter of a source space have been observed by others (Ioannides et al. 1990; Wang et al. 1992; Wang 1993). This appears to be a characteristic of a distributed-source estimate. However, in the above study, the spurious current levels are low due to the overdetermined nature of the problem. Restricting the source space sufficiently generally reduces the problem altogether.

On the subject of depth bias, it can be seen from the simulation results that solutions produced using electric measurements suffer most, the superficial sources being under-emphasised and the deep sources over-emphasised. In contrast, magnetic measurements can produce reasonable solutions, with only a small level of superficial bias. In this case, the superficial sources are reconstructed correctly, with only the deep sources being slightly underestimated. The difference in the severity of the bias between electric and magnetic measurements occurs purely as a consequence of the lead field distributions (figures 25 and 26), which have differing rates of fall-off with depth. It is suggested that it is feasible to carry out a constrained distributed-source analysis using magnetic data alone, but not electric data alone. It is preferable to use both electric and magnetic data simultaneously, as this virtually removes any bias problems, as well as doubling the number of measurements, thus facilitating the generation of an overdetermined system, with consequent improvement in spurious source levels and solution fidelity. However, the 24-dipole analysis using both modalities does show a slight underestimation of superficial sources, and this may be due to the fact that the depth bias of the electric component of the solution is greater than that of the magnetic component, resulting in a solution which more closely resembles an electric solution than a magnetic solution. The bias is marginal, though, and certainly much less significant than in the separate electric and magnetic solutions.

So far, it has been assumed that the phantom sources lie within the source
space. However, in reality, the possibility exists that the surfaces of the source space do not coincide with the true cortical surfaces. This could introduce errors into the inversion technique, so the question is addressed in the following simulations.

**Computer Simulations Using a Realistic Source Model**

A realistic source model based on a statistically average brain can be used to carry out source analyses in a variety of individuals without the need for prior magnetic resonance scanning. An average source model is therefore clinically convenient and cost-effective. However, each individual has a brain whose morphology deviates from that of the standard source model. Hence, the standard model does not correctly represent the relationship between the cortical sources and the external electric potentials and magnetic fields, but introduces a systematic error into the inverse procedure. The following simulations therefore evaluate the accuracy of anatomically constrained distributed-source estimates when the standard source model deviates from the true brain shape.

Six dipolar current sources of dipole moment 10nAm, situated in right visual cortex, were simulated. Between the different experiments, the positions of these sources were altered to simulate anatomical variations. The electric potentials and magnetic flux gradients resulting from the six dipoles were calculated over the occipital montage described in the last section (see figure 30). The spacing of the rows and columns of the montage was again 30mm. The electric potentials were referred to the inion, while the simulated gradiometer, which was the same as that described in the last section, was located radially. Gaussian noise was added to the simulated potential distribution, such that the standard deviation of the noise was 10% of the maximum potential. Ten percent noise was added to the flux gradient distribution in the same way. Finally, the inversion technique was applied to the electric and magnetic distributions simultaneously, in conjunction with the realistic source space described in the section headed "Definition of the Source Space."

**No Anatomical Deviation** To test the performance of the inversion algorithm in the absence of anatomical variation, the coordinates of the six sources were set equal to those of six of the dipoles in the standard source space. These six dipoles were the three dipoles of the right cuneal gyrus and the three dipoles on the superior surface of the right calcarine fissure. The sources then exactly coincided with the source reconstruction space (table 4).

The results of the inversion are shown in figure 39. In this figure, and those to follow, each arrow represents the average dipole moment of a group of three sources, the groups being those of table 3, with the exception of the lateral sources, which are grouped together with the source on the occipital pole. The algorithm has
reconstructed the sources correctly, except that the regularisation process has caused the current to appear smoothed out across adjacent surfaces of the source space, as observed in the earlier simulations. This effect requires that results relating to opposing surfaces of fissures are interpreted carefully. Spurious sources of low amplitude also occur, as noted in the previous simulations.

*Variation of the Calcarine Sulcus* The extreme variation zone for the calcarine fissure relative to the occipital lobe was determined using a stereotaxic statistical study (Talairach *et al.* 1967). The extreme lowest position of the fissure was used to define the coordinates for the six source dipoles. The three medial dipoles were positioned symmetrically along the cuneal gyrus, which now had a larger area due to the low position of the calcarine fissure. The y-components of both the position and direction vectors of the three dipoles representing the calcarine fissure were adjusted randomly to simulate folding of the fissure (table 5). This was repeated for the extreme highest position of the calcarine fissure, the coordinates for which are given in table 6.

The result of the inversion for the lowest extreme position of the calcarine fissure is given in figure 40. It can be seen that the sources on the right cuneal gyrus have been reconstructed correctly, except that they "spread" across to the left cuneal gyrus. The sources on the right calcarine fissure also have a similar spread. The right calcarine sources have been underestimated: the sum of the average dipole moments over the superior and inferior surfaces of the right calcarine fissure is around 6nAm, as opposed to 10nAm for the true sources. The loss of dipole moment is accounted for in the spurious sources of average moment 3nAm on the right ventral surface of the brain. Thus, the discrepancy between the real and model brains has introduced low-magnitude erroneous sources. Spurious activity levels are otherwise slightly higher than for the previous simulation with no anatomical deviation.

The result of the inversion for the highest extreme calcarine fissure is given in figure 41. In this case, both the cuneal and calcarine sources have been estimated approximately correctly. Spurious current levels are only slightly higher than in the absence of anatomical deviation.

*Variation of the Whole Brain* The occipital lobe does not bear the same relationship to the interior of the skull in all individuals (Talairach *et al.* 1967), as was assumed above. Variation of the occipital lobe position therefore needs to be taken into account when considering the extreme variation of the calcarine fissure. Moreover, geometric studies of the brain relative to the nasion and inion (Steinmetz *et al.* 1989; Myslobodsky *et al.* 1991) show the brain as having greater inter-subject variability than when it is studied relative to internal brain and skull landmarks (Talairach *et al.* 1967). This difference in variability is primarily due to the variation of the inion location. Since the location of
<table>
<thead>
<tr>
<th>x (mm)</th>
<th>y (mm)</th>
<th>z (mm)</th>
<th>$q_x$ (nAm)</th>
<th>$q_y$ (nAm)</th>
<th>$q_z$ (nAm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>16</td>
<td>20</td>
<td>-5.7</td>
<td>0.0</td>
<td>8.2</td>
</tr>
<tr>
<td>44</td>
<td>16</td>
<td>24</td>
<td>1.4</td>
<td>0.0</td>
<td>9.9</td>
</tr>
<tr>
<td>60</td>
<td>16</td>
<td>22</td>
<td>1.4</td>
<td>0.0</td>
<td>9.9</td>
</tr>
<tr>
<td>36</td>
<td>3</td>
<td>36</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
<tr>
<td>46</td>
<td>3</td>
<td>31</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
<tr>
<td>58</td>
<td>3</td>
<td>24</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 4. Coordinates of the six sources in the absence of anatomical variation. $q_x$, $q_y$, and $q_z$ are the Cartesian components of the current dipole moments.

<table>
<thead>
<tr>
<th>x (mm)</th>
<th>y (mm)</th>
<th>z (mm)</th>
<th>$q_x$ (nAm)</th>
<th>$q_y$ (nAm)</th>
<th>$q_z$ (nAm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>16</td>
<td>12</td>
<td>-3.8</td>
<td>1.5</td>
<td>9.1</td>
</tr>
<tr>
<td>45</td>
<td>14</td>
<td>15</td>
<td>0.3</td>
<td>-1.4</td>
<td>9.9</td>
</tr>
<tr>
<td>60</td>
<td>10</td>
<td>16</td>
<td>-1.6</td>
<td>3.2</td>
<td>9.3</td>
</tr>
<tr>
<td>31</td>
<td>3</td>
<td>31</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
<tr>
<td>43</td>
<td>3</td>
<td>29</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
<tr>
<td>56</td>
<td>3</td>
<td>25</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 5. Variation of the calcarine fissure relative to the occipital lobe: coordinates of the six current sources at the lowest extreme position.

<table>
<thead>
<tr>
<th>x (mm)</th>
<th>y (mm)</th>
<th>z (mm)</th>
<th>$q_x$ (nAm)</th>
<th>$q_y$ (nAm)</th>
<th>$q_z$ (nAm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>16</td>
<td>20</td>
<td>-6.8</td>
<td>1.5</td>
<td>7.2</td>
</tr>
<tr>
<td>39</td>
<td>14</td>
<td>28</td>
<td>-1.2</td>
<td>-1.3</td>
<td>9.9</td>
</tr>
<tr>
<td>57</td>
<td>10</td>
<td>23</td>
<td>2.6</td>
<td>2.7</td>
<td>9.3</td>
</tr>
<tr>
<td>35</td>
<td>3</td>
<td>40</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
<tr>
<td>47</td>
<td>3</td>
<td>35</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
<tr>
<td>59</td>
<td>3</td>
<td>29</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 6. Variation of the calcarine fissure relative to the occipital lobe: coordinates of the six current sources at the highest extreme position.
Figure 39. Reconstruction of three medial and three calcarine sources, with the sources and source reconstruction space coincident. The diagram on the left shows the relationship between the six sources and the surfaces of the standard anatomical source space. The diagram on the right is the corresponding source reconstruction, viewed from the back of the head. Each arrow gives the average magnitude of one of the groups in table 3. (The dipole at the occipital pole and the two lateral dipoles are treated as one group.) Goodness-of-fit: 12.7%.

Figure 40. Inter-subject variation of the calcarine fissure relative to the occipital lobe: reconstruction of the six current sources at the lowest extreme position. Goodness-of-fit: 92.3%.

Figure 41. Inter-subject variation of the calcarine fissure relative to the occipital lobe: reconstruction of the six current sources at the highest extreme position. Goodness-of-fit: 11.4%.
the inion governs the position of the chosen x-axis, the brain may be effectively rotated about the y-axis as a consequence of inion variation. To simulate these effects, the coordinates of the six dipolar sources were deduced from the extreme variation zone for the calcarine fissure relative to internal skull landmarks, then rotated about the y-axis by an amount corresponding to 10mm inion shift. The y-coordinates of the three dipoles representing the calcarine fissure were adjusted randomly, as in the previous section, to simulate folding of the fissure, and all six dipoles were shifted towards the left by 3mm to simulate a lateral shift of the whole brain. Tables 7 and 8 contain the coordinates for the lowest and highest extreme sources respectively, after rotation.

This simulation contained three anatomical deviations: (1) the inherent variation of the calcarine fissure within the occipital lobe of the brain, as described in the previous subsection, (2) the variation of the occipital lobe relative to an internal skull landmark or average inion location, and (3) rotation equivalent to the inion (and associated external field measurements) rotating relative to the brain. The simulated brain was therefore grossly atypical and there was a large discrepancy between the location of the simulated source distribution and the location of the standard source space. It should be emphasised that this discrepancy was chosen to be an extreme

<table>
<thead>
<tr>
<th>x (mm)</th>
<th>y (mm)</th>
<th>z (mm)</th>
<th>q_x (nAm)</th>
<th>q_y (nAm)</th>
<th>q_z (nAm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>13</td>
<td>6</td>
<td>-2.5</td>
<td>1.6</td>
<td>9.5</td>
</tr>
<tr>
<td>45</td>
<td>11</td>
<td>6</td>
<td>1.8</td>
<td>-1.7</td>
<td>9.7</td>
</tr>
<tr>
<td>61</td>
<td>7</td>
<td>5</td>
<td>-0.3</td>
<td>3.5</td>
<td>9.4</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>23</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
<tr>
<td>52</td>
<td>0</td>
<td>17</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
<tr>
<td>63</td>
<td>0</td>
<td>12</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 7. Variation of the calcarine fissure relative to the inion: coordinates of the six current sources at the lowest extreme position.

<table>
<thead>
<tr>
<th>x (mm)</th>
<th>y (mm)</th>
<th>z (mm)</th>
<th>q_x (nAm)</th>
<th>q_y (nAm)</th>
<th>q_z (nAm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>13</td>
<td>19</td>
<td>-8.6</td>
<td>1.8</td>
<td>4.8</td>
</tr>
<tr>
<td>28</td>
<td>11</td>
<td>33</td>
<td>-4.4</td>
<td>-1.3</td>
<td>8.9</td>
</tr>
<tr>
<td>48</td>
<td>7</td>
<td>36</td>
<td>-1.9</td>
<td>2.7</td>
<td>9.5</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>45</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
<tr>
<td>36</td>
<td>0</td>
<td>43</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
<tr>
<td>47</td>
<td>0</td>
<td>42</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 8. Variation of the calcarine fissure relative to the inion: coordinates of the six current sources at the highest extreme position.
condition so as to rigorously test the inversion procedure; the author considers it very unlikely that a given subject would have a calcarine fissure that happened to be situated ventrally within the occipital lobe, as well as an occipital lobe that happened to be shifted downward and laterally within the skull, as well as an inion that happened to be rather high!

The inverse solution for the sources at the lowest extreme is shown in figure 42. Because of the low position of the medial sources, they have been reconstructed on the lingual gyrus, as opposed to the cuneal gyrus, with an almost correct average amplitude of about 9nAm. Meanwhile, the sources on the calcarine fissure are considerably diminished, the sum of the average dipole moments over both the superior and inferior surfaces being reduced to around 5nAm. Spurious sources of average moment 4nAm appear on the right ventral surface of the brain. The goodness-of-fit of this solution has fallen by two orders of magnitude, compared to the first case where anatomical deviation was absent, indicating a discrepancy between the true sources and the source model.

The corresponding solution for the sources at the highest extreme position is given in figure 43. This solution is approximately correct, except for spurious sources of average dipole moment 3nAm, located on the dorsal surface of the left hemisphere. Spurious activity is otherwise low. The goodness-of-fit has dropped to 10^-4%, however, indicating that the source model is doubtful.

A novel method was applied to overcome the rotational discrepancy between the sources and the source model. It has been shown previously that when a two-dimensional source space coincides with an actual source distribution, an increase in the solution goodness-of-fit occurs (Ioannides et al. 1990). This principle was applied to the present case and it was found that if the standard source space was rotated about the y-axis to simulate the brain varying relative to the inion, a clear gradient was obtained in the solution goodness-of-fit (figure 44). The degree of source space rotation was limited to the equivalent of ±20mm inion displacement, i.e. the maximum angle of rotation was 20/81 radians, where the scalp radius was 81mm. The solutions for which the goodness-of-fit was greatest were found to be much closer to the original source distributions (figures 45 and 46). In these solutions, spurious current is as low as when the source model coincided exactly with the true anatomy (c.f. figure 39). The only inaccuracy is that the recovered solution for the sources in the highest extreme position (figure 46) slightly overestimates the cuneal sources at around 13nAm. The goodness-of-fit of the recovered solution for the sources in the lowest extreme position is 15.8%, and that for the sources in the highest extreme position is 30.8%, indicating that the source model is registered correctly with the true sources.

*Group Averages* Producing an inter-subject group average response is one common means of improving the signal-to-noise ratio of an evoked electric or magnetic response. To investigate the effect of group averaging on the inversion performance,
Figure 42. Inter-subject variation of the calcarine fissure relative to the inion: reconstruction of the six current sources at the lowest extreme position. Goodness-of-fit: 0.2%.

Figure 43. Inter-subject variation of the calcarine fissure relative to the inion: reconstruction of the six current sources at the highest extreme position. Goodness-of-fit: 10^{-4}%. 

Figure 44. Solution goodness-of-fit as a function of source space rotation for the six sources in the highest extreme position (line A) and the lowest extreme position (line B). The rotation is measured in millimetres of inion shift, the inion being 81mm from the origin.
the previous five simulated evoked responses were averaged at each point on the
occipital montage and the inversion algorithm was applied to the resulting electric and
magnetic distributions. The standard deviation, \( \sigma_g \), of the noise in each group average
distribution was taken to be:

\[
\sigma_g = \frac{\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2 + \sigma_5^2}}{5},
\]

(142)

where \( \sigma_1 \ldots \sigma_5 \) were the standard deviations used in the previous five simulations. The
reconstructed sources (figure 47) are seen to be a good reproduction of the original
sources, and spurious current levels are only slightly higher than for the original case of
no anatomical deviation (c.f. figure 39). In particular, there is no evidence of erroneous
source distributions produced by the individual atypical anatomies. The goodness-of-fit
of 63.7% shows the solution to be well matched to the "group average sources"
underlying the group average field distributions.

Discussion The simulations in which the calcarine fissure is varied relative to the
occipital lobe show that the standard source model produces reasonably accurate
results even when the true brain deviates from it by up to about 5mm. This level of
source model discrepancy introduces erroneous current sources whose magnitudes
remain considerably less than those of the true sources. The effect is worse when an
individual calcarine fissure is inferior to that of the source model. This is because the
calcarine fissure is only about twenty millimetres away from the ventral surface of the
brain, and co-oriented with it, so that the inversion algorithm has difficulty in
distinguishing whether the given measurements were generated by the calcarine fissure
or ventral surface.

Gross anatomical variations, such as simultaneous variation of the calcarine
fissure, the occipital lobe and the inion, cause the algorithm to produce erroneous
results under some circumstances. The inverse procedure locates current on the
surfaces of the model which are nearest to the true locations of the sources, but if there
is a spatial discrepancy between the real brain and the brain model, these surfaces may
not be the ones representing the active region of the brain. Again, this effect is most
noticeable when the anatomical variation is such as to lower the calcarine fissure
relative to the inion. The external measurements then appear to have been produced by
the ventral surface of the brain, and the current sources are thus reconstructed
ventrally. In the simulation involving a severely lowered calcarine fissure, the cuneal
sources coincide with the lingual gyrus of the standard source model, and the external
measurements therefore appear to have been generated by sources on the lingual gyrus.
Consequently, in the solution, current appears on the lingual gyrus rather than the
cuneal gyrus. The solution for the calcarine fissure in the highest extreme position is
surprisingly accurate, however, with only a small increase in spurious current. This is
Figure 45. Inter-subject variation of the calcarine fissure relative to the inion: reconstruction of the six current sources at the lowest extreme position after alignment of the source space with the sources by observation of the goodness-of-fit. Goodness-of-fit: 15.8%.

Figure 46. Inter-subject variation of the calcarine fissure relative to the inion: reconstruction of the six current sources at the highest extreme position after alignment of the source space with the sources by observation of the goodness-of-fit. Goodness-of-fit: 30.8%.

Figure 47. Reconstruction of the six current dipoles from the group average of the previous five simulated evoked responses. Goodness-of-fit: 63.7%.
attributed to the fact that there are no suitable adjacent surfaces in the source model on which the current sources could be erroneously reconstructed. The goodness-of-fit successfully indicates the inadequacy of the source model when there are gross anatomical deviations. The goodness-of-fit also enables a more appropriate model orientation to be found, from which an accurate solution can be calculated.

An inversion of a group average response has been shown to produce an accurate solution. Most importantly, the solution reflects the source distribution which is common to each individual brain, rather than being merely a superimposition of erroneous current sources. Although the group average electric and magnetic responses may be produced by a series of grossly differing anatomies, the outcome of the inversion is a source distribution which accurately represents the activity of any one of these individual brains. The accuracy of the group average inversion is presumably because the average response represents an average source distribution which deviates negligibly from the standard anatomical source space.

Simulations of the type described in this section have not yet been reported very widely in the literature. Wang (1993) reported the effects of position error when using a constrained source space consisting of a pair of orthogonal planes in which the current dipoles were fixed in orientation. It was found that if the simulated current source, a single dipole, was removed from one of the source planes by up to 5mm, the resulting image had a greater spread, but retained a correct position. In this simulation, though, there was no alternative surface on which the displaced source could be erroneously located. Spurious current sources, whose magnitudes were in some cases greater than those of the correctly reconstructed sources, also appeared at the boundaries of the source space. The simulations described in this section are more realistic than those of Wang (1993), and are also much more challenging to the inversion procedure, as the anatomical variations are much larger.

Conclusions

This chapter has discussed the use of an anatomically constrained distributed-source model in the interpretation of neuro-electromagnetic measurements. The motivation for using an anatomically constrained model is that it reduces the number of unknown source parameters, enabling a more accurate solution to be obtained. Moreover, if both electric and magnetic measurements are analysed simultaneously in conjunction with an anatomically constrained model, the number of available data can be increased, and problems such as depth bias, which are associated with single-modality analyses, can be overcome.

Two source models were presented for visual cortex; the first was a simple cruciform model and the second was a realistic model based on an average brain. A description was then given of an inversion algorithm by which the magnitudes of the
constituent current dipoles could be determined. Simulations using the cruciform model then illustrated, for both noise-free and noisy data, that it is desirable to restrict the source space as much as possible and use both electric and magnetic measurements together. Further simulations showed that a source model based on a statistically average brain can be used to analyse evoked responses in a variety of individuals whose brains deviate from the average brain. It was shown, for the case of visual evoked electric and magnetic responses, that despite the inter-subject variation of sources located in the calcarine and interhemispheric fissures, a standard anatomical source space is still appropriate. Variation of the calcarine fissure within the occipital lobe did not prevent the algorithm from correctly locating sources in these regions using a standard visual cortical source space. Even when the spatial variation of the occipital lobe and the inion were taken into account, the method was able to locate current sources correctly, although a degree of innovation was required to overcome a relative shift between the brain and inion. A group average response was shown to be a sound basis for source analysis using this type of approach.

It is worth noting here that although an average source space was used for the realistic simulations, as opposed to an individual source space based on magnetic resonance images, the results are still applicable to the latter. If magnetic resonance images are used, there is always some degree of error in co-registering the coordinate systems of the magnetic resonance scan and the neuro-electromagnetic scan. Hence, the source space as defined by the magnetic resonance images may be as much as 5mm in error, relative to the neuro-electromagnetic measurement sites. The simulations of this chapter therefore provide an assessment of the errors that could occur in this situation.

No mention was made of the computer implementation of the inversion algorithm. Due to the linear nature of the anatomically constrained distributed-source estimation procedure, it transpires that the inversion algorithm can be conveniently implemented on a parallel computer, which performs many calculations simultaneously, with a consequent increase in speed. This subject is addressed in the next chapter.
6. Parallel Computing Applied to the Neuro-electromagnetic Inverse Problem

Introduction

The anatomically constrained distributed-source estimation technique presented in the previous chapter is based upon the generation and solution of a linear system of equations. This linear system describes the measured electric and magnetic data in terms of a discrete set of dipoles of fixed location and orientation. Central to the linear system is a gain matrix, which gives the contribution of each discrete dipole to the various electrode and magnetometer signals. The elements of the gain matrix are calculated using a spherical conductivity model. These calculations are all independent: each one involves a separate solution of the forward problem. Hence, the elements of the gain matrix may be calculated concurrently on a number of discrete processing units. The parallel generation of the gain matrix is efficient because the independent calculations require few inter-processor communications, so that little time is lost in passing data.

The solution to the linear system is found by postmultiplying the regularised pseudoinverse of the gain matrix by the data vector. When a number of solutions are required from the same measurement montage, such as when a millisecond by millisecond source derivation is required, this backsubstitution process must be repeated many times. The multiple backsubstitutions can be divided up and executed in parallel. Hence, the linear system can be both generated and solved using parallel processing. Anatomically constrained distributed-source analyses can thus be fast and convenient.

This chapter begins by outlining the essential features of parallel computer architecture, with specific reference to transputers. The concepts of parallel program design are also briefly reviewed. Once these foundations have been laid, consideration is given to methods by which the anatomically constrained distributed-source method can be implemented in parallel, and the performance characteristics of the optimal parallel program are described. It is then shown how the linear system approach and its parallel implementation can be easily adapted to perform minimum norm estimation and equivalent current dipole fitting. This illustrates the versatility of the linear technique. A
preliminary report of the work in this chapter is given in Bedford (1993). To the author's knowledge, the only other report of the use of parallel computing in neuro-electromagnetism is that of Ioannides et al. (1990), in which an array of transputers is used to obtain probabilistic distributed-source solutions to the neuromagnetic inverse problem (see chapter 4).

The Parallel Computer

Before the neuro-electromagnetic inversion algorithm can be executed in parallel, it is necessary to define the type of parallel computer on which it is to run. There are two types of parallel computer: the shared memory parallel machine and the distributed memory parallel machine. The first of these contains a series of discrete processors, all of which have access to a common memory (figure 48). This machine suffers from several drawbacks. Firstly, as the number of processors is increased, the routing device handling the transfer of data between the processors and memory saturates because the demand of the processors becomes excessive. This typically starts to occur when between six and twenty processors are used. Consequently, this type of machine cannot be expanded very easily; a system upgrade requires that the routing device (and ultimately the complete machine) is replaced. Moreover, a shared memory parallel machine can suffer from bus contention, where two or more processors attempt to use the communication mechanism simultaneously. This type of machine therefore requires bus arbitration logic to control access to the memory.

![Diagram of a shared memory parallel machine](image)

Figure 48. Architecture of a shared memory parallel machine. $P_1, P_2, ..., P_N$: parallel processors.
A more satisfactory parallel computer is the distributed memory parallel machine. This consists of a series of discrete processors, each having its own memory for its exclusive use (figure 49). A routing system connects the processors together for start-up purposes and data distribution. Since the majority of processing occurs within the individual processors and memory units, the load on the routing bus is much reduced. Consequently, it does not saturate, and the system can be expanded ad infinitum without difficulty. As the number of processors is increased, the total quantity of memory also increases, so that the memory is naturally upgraded with the processing power. Memory contention is also eliminated in a distributed memory parallel machine as only one processor has access to each local memory.

A distributed memory parallel machine is connected together by a communication network, of which there are two types. These are the uniform and non-uniform communication latency networks. All processors - nodes - within a uniform communication latency network communicate with the same latency or communication time. There are numerous examples of this type of machine (figure 50), but the simplest is where all the processors (each with its own local memory) are connected to a single or multiple bus. The advantage of a uniform communication latency machine is that there are only two levels of communication time: local access time required by a given node to access its own local memory, and remote access time required to access the memory of any other node using the communication bus. The disadvantage of this type

![Figure 49. Architecture of a distributed memory parallel machine. Processors P₁, P₂,...,Pₙ all have independent memories M₁, M₂,...,Mₙ. Each processor and its associated memory is referred to as a node.](image-url)
Figure 50. Uniform communication latency networks. (a) Single bus network; (b) multiple bus network. P: processing nodes. All communications have the same latency.

of machine is that it is not very scalable. If many nodes are used, the communication latency becomes excessive and/or the communication bus becomes complex, resulting in high cost.

All nodes within a non-uniform communication latency network are able to communicate with each other, but latency may vary considerably. Examples of this type of network are linear, ring, mesh and hypercube processor configurations (figure 51). The main disadvantage of this architecture is that there are many levels of access time: the time required to access the memory of a neighbouring node is short, but the time required to access the memory of a distant node is high. However, the advantage of a non-uniform communication latency network is that it is very scalable. If more nodes are added, the communication latency between highly separated nodes increases, but the communication latency between neighbouring nodes remains low. Because of the modular structure of the network, the overall architecture can be modified simply and nodes can be added at a reasonable cost.

Transputers

A transputer network is a specific example of a non-uniform communication latency network. A transputer is a 16- or 32-bit microprocessor containing a central processing unit, a small quantity of local memory and four fast serial communication ports for connection to adjacent transputers (Inmos 1987). A complete distributed memory parallel machine can be constructed by connecting together transputers, the connections being simple enough to create using a twisted wire pair. Transputers are designed for point-to-point communications and their performance is therefore not limited by accesses to shared memory via a common bus; the total communication bandwidth of the system increases linearly as more transputers are added to it. System cost is low as there is no contention for the communication mechanism and thus no requirement for bus arbitration logic.
Figure 51. Non-uniform communication latency networks. (a) Linear, (b) ring, (c) mesh, and (d) hypercube network topologies. Some communications have to traverse more nodes than others, so communication latency varies. (Note that all the illustrated topologies can be constructed using transputers, as no processor requires more than four connections.)

Transputers are manufactured solely by Inmos Limited, a member of the SGS-Thomson Microelectronics Group. There are currently several groups of transputer, referred to as the T2, T4 and T8 classes (Inmos 1990). A very recent addition to the range, in a class of its own, is the T9000 transputer. The transputers in all of the classes have a similar architecture, permitting software portability between any member of the family. The T2 and T4 families do not explicitly support floating-point arithmetic, but this can be performed using software floating point libraries. Meanwhile, the members of the T8 family have a 64-bit floating point unit which operates in parallel with the central processing unit. The T9000 has a speed of five to ten times that of the T8 transputers and faster communication times. Its principal advantage is that it enables data to be sent between two non-neighbouring processors without the intermediate processors having to explicitly receive the data and pass it on. In short, it permits through-routing of data without the intermediate processors being diverted from their main task of computation.
The transputer is the unit by which a parallel network is constructed, but is itself a parallel machine (figure 52). The central processing unit, the floating point unit and the link interfaces all operate concurrently, ensuring high speed. Peak processing power of a T8 transputer is typically 25 million instructions per second and 3.5 million floating point operations per second (Inmos 1990). T8 devices have a 32-bit architecture, except for the floating point unit, which is 64-bits wide. Since many instruction codes are only a single byte long, up to four instructions can be fetched by the central processing unit simultaneously. The 4kbytes of static random access memory allows single cycle access to instructions and data and acts as a scratchpad for the central processing unit. However, this memory is not sufficient to store all of the program code and data, so an external memory interface is provided, which, in the T8 transputers, can address up to 4Gbytes of external memory. This external memory is located outside of

Figure 52. Architecture of the IMS T800 transputer. All bus structures are 32-bits wide.
the transputer itself, typically on a small circuit board called a transputer module. The external memory is local to a given transputer, as opposed to being shared by other transputers.

Parallel processes executing concurrently on a network of transputers need to exchange data. Each transputer therefore has four serial links which act as direct memory access channels for communication between transputers. The links are simple pairs of wires which support communications with a combined bandwidth of 80-100 Mbits/s. The two wires in a pair always carry data in opposite directions. Each message, regardless of its length, is transmitted as a sequence of single byte communications and the event logic provides an asynchronous handshake between the two processors involved in the communication.

The parallel system used for the work in this thesis is a network of 22 IMS T805-20 transputers. These are incorporated in transputer modules, each of which contains either 1Mbyte or 4Mbyte of dynamic random access memory. The transputer modules are mounted on three Transtech TMB08 transputer motherboards which are housed in the application slots of an AST 386 personal computer. Each motherboard has space for ten transputers and contains the necessary connections with which to construct the parallel network. The design of the motherboard is such that the transputers are hard-wired together into a pipeline configuration. Thus, each transputer has two of its links permanently connected into the pipeline, while the other two are available for connection to any other arbitrary processor. These optional connections are achieved by means of an Inmos IMS C004 link switch, controlled by an IMS T212 transputer. All of the spare transputer links are connected to the link switch, and it enables any transputer link to be connected to any other. The IMS C004 switch also has eight connections with the edge of the motherboard; these connections are for routing to the IMS C004 switches on adjacent motherboards, which, in turn, connect to another set of transputers. Thus, motherboards can be cascaded and a transputer on one board connected to any other transputer on any other board, subject to the limitations of only having eight edge connectors. (Note that certain processor topologies, such as mesh topologies, are difficult to achieve, as they require all ten transputers to be connected to the ten transputers on the next motherboard.) The link switches are controlled by Transtech Network Configuration Software, which merely enables the user to state in a high-level form which links are to be connected where. For the work in this thesis, the three motherboards are cascaded, and the transputers are used in their hard-wired pipeline form (figure 53). In this configuration, the first transputer in the pipeline, which is referred to as the root transputer, communicates with the host computer via an Inmos IMS C012 link adapter. This converts between the 8-bit bi-directional format of the personal computer port and the format of a transputer serial link. The graphics processor is located on the last motherboard and forms the last processor in the pipeline, although the pipeline connection to it is not used and therefore simply ignored. Instead, the processor is connected via the motherboard edge connectors to the root transputer.
Figure 53. Topology of the transputer system used for the work in this thesis. H: host computer; R: root transputer (IMS T805 with 4Mbytes of external memory); G: Transtech TTG1 graphics transputer (IMS T800 with 1Mbyte of external video memory); unlabelled transputers are IMS T805 devices with 1Mbyte of external memory. The transputer modules are mounted on motherboards M0, M1 and M2.
The graphics processor itself is a Transtech TTG1 transputer module having an IMS T800-25 transputer and 1Mbyte of video random access memory. Half of the memory is used to store program code and data and the other half supports a maximum of two images. Each image requires 256kbytes of memory and consists of a series of single bytes representing the colours of an array of $512 \times 512$ pixels. One of the images is continuously output as a video picture to a multi-sync monitor.

The transputer system is programmed in *Occam* 2. *Occam* is a high-level language which is designed to reflect the architecture of the transputer (Inmos 1988). In fact, the transputer can be considered as an *Occam* machine. The *Occam* language is based on a programming model known as Communicating Sequential Processes, wherein a number of processes all perform sequential tasks at the same time, communicating as necessary. The communications take place over channels, which are uni-directional conceptual links between processes. Communication is synchronous, although the receiving process may not be ready to receive a message until after the transmitting process is ready to send it, or *vice versa*. However, it is not possible for a message to be lost: the first process to become ready waits for the other process to communicate with it (Pountain and May 1988).

Once the parallel program has been designed in software, the logical *processes* must be mapped onto physical *processors*. This is achieved by a pseudo-*Occam* network configuration language. Any number of processes can be mapped onto a single transputer and executed concurrently, although, in real terms, they only function in quasi-parallel as the central processing unit of a single transputer can only execute one instruction at a time. Processes mapped onto differing transputers genuinely operate in parallel. The transputer links are the only limit to the mapping of logical processes onto physical processors. An *Occam* channel must have a physical transputer link by which to pass data. Each transputer link consists of two one-way wires, transmitting in opposite directions, and therefore a maximum of two oppositely-directed channels can be assigned to one link.

Program building (compilation, linking, debugging *etc.*) is provided by the Inmos IMS D7205A toolset, working under the host operating system. A feature of this toolset is that the tools actually execute the building process on the root transputer. Primitive drawing operations for the graphics transputer are provided by the Transtech TTGS graphics server. This is an *Occam* process which is mapped onto the graphics transputer. Drawing operations are carried out by sending commands to the server process over normal *Occam* channels.
Styles of Parallelism

The neuro-electromagnetic inversion algorithm must be decomposed into a series of parallel processes. This section therefore takes an objective look at the three main ways in which a computation is parallelised.

Data Parallelism Suppose that a computation has to be performed repeatedly on a series of different data. The data are decomposed into pieces and each parallel task then performs the same computation on its own data (figure 54). The level of parallelism is proportional to the amount of data: the more data there are, the greater the scope for dividing them up into pieces. This type of parallelism is therefore suitable for scientific computations on large and regular data structures, such as vectors and matrices. Usually, the temporal communication pattern is regular.

Function Parallelism If an extensive computation is to be performed only once on an entire set of data, then data parallelism is no longer appropriate as the computation is not repetitive. Instead, the computation is divided into pieces or functions, which, as far as possible, proceed in parallel (figure 55). The data are taken to the tasks as and when they are needed, and the temporal communication pattern therefore tends to be irregular. This method is suitable for irregular data which cannot be divided into comparable blocks requiring equal computation. However, the level of parallelism is limited by the number of functions into which the computation can be divided. If the computation consists of, say, ten functions, then the maximum number of processes is also ten. Hence, programs using function parallelism are not very scalable.

Pipelining This is valuable when a large and complex computation must be performed repetitively on a series of different data. Firstly, the computation is divided into functions which form the parallel processes. The first item of data is then operated on by the first function. The item of data is then passed to the second function for further processing, and so on, until the entire computation has been carried out on that item. So far, no time advantage has been gained over a conventional serial computation. However, as each function releases the first item of data, it begins processing the second item, and when it finishes the second item, it begins processing the third. Consequently, once the first item of data reaches the last function, all of the functions are operating on a separate item of data, and a performance increase over a sequential computation is realised. The overall effect is that all of the individual data items are passed along the pipeline of functions, gradually being processed as they go. This is the computational equivalent of a production line (figure 56). Although this is an elegant means of parallelisation, the level of concurrence is limited by the number of functions into which the overall computation can be divided. Since there are a series of data items to be
Figure 54. Data parallelism. Processes $P_1$, $P_2$ and $P_3$ each perform the same computation but on different data.

Figure 55. Function parallelism. Processes $P_1$, $P_2$ and $P_3$ each perform a different part of the overall computation, fetching data as required.

Figure 56. Pipeline parallelism. The data are divided up and passed sequentially through processes $P_1$, $P_2$ and $P_3$, and the computations are performed as they go.
processed identically, it is sometimes possible to use data parallelism, without the limitation on scalability introduced by decomposing the computation into functions.

Whichever style of parallelism is used, there are several principles which must be kept in mind:

1. Computational load must be balanced between the processors. It is desirable to have all processors working all of the time; if one or more processors are idle during the computation time, the parallel computer is not being used optimally.

2. The ratio of communication to computation must be minimised. As much time as possible should be spent computing rather than exchanging data: communication time is "wasted" time.

3. Sequential "bottlenecks" should be minimal. A bottleneck occurs when part of a computation cannot be parallelised. The parallel processes must terminate for their output to be processed by one sequential process. Parallel operation then resumes. A bottleneck is inefficient as some processors become idle during the sequential computation. Furthermore, it interrupts the independent operation of the parallel tasks.

4. A parallel program should be scalable. That is, it should be able to run efficiently on any number of nodes with only trivial modifications.

To some extent, these criteria must be balanced against each other. For example, a scalable program often requires more communication than one which is not. The "best" parallel program is simply that which completes the given computation in the shortest time.

**The Neuro-electromagnetic Inversion Algorithm**

Full details of the anatomically constrained distributed-source method were given in the previous chapter. However, since the steps of the mathematical procedure are of importance to this chapter, they are summarised below:

1. Generate the gain matrix using the electric and magnetic forward problems:
\[
L = \begin{bmatrix}
I_{11}^e & I_{12}^e & I_{13}^e & \cdots & I_{1S}^e \\
\sigma_e & \sigma_e & \sigma_e & \cdots & \sigma_e \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
I_{E1}^e & I_{E2}^e & I_{E3}^e & \cdots & I_{ES}^e \\
\sigma_e & \sigma_e & \sigma_e & \cdots & \sigma_e \\
I_{11}^m & I_{12}^m & I_{13}^m & \cdots & I_{1S}^m \\
\sigma_m & \sigma_m & \sigma_m & \cdots & \sigma_m \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
I_{M1}^m & I_{M2}^m & I_{M3}^m & \cdots & I_{MS}^m \\
\sigma_m & \sigma_m & \sigma_m & \cdots & \sigma_m 
\end{bmatrix}
\]

where \( I_{ij}^e \) \((i = 1...E, j = 1...S)\) is the electric potential difference between the \(j\)th electrode and the reference electrode, due to a unit dipole at the \(j\)th dipole site, and \( I_{ij}^m \) \((i = 1...M, j = 1...S)\) is the corresponding magnetic flux gradient at the \(j\)th magnetometer site. \( \sigma_e \) and \( \sigma_m \) are the electric and magnetic standard deviations, respectively. If necessary, augment matrix \( L \) with zeros, so that it has at least as many rows as columns.

2. Form the combined data vectors:

\[
f(t) = \begin{bmatrix}
e_1(t) \\
\sigma_e \\
\vdots \\
e_E(t) \\
\sigma_e \\
m_1(t) \\
\sigma_m \\
\vdots \\
m_M(t) \\
\sigma_m
\end{bmatrix}, \quad t = 1...T,
\]

where \( e_1(t) \ldots e_E(t) \) are the electric measurements at time \( t \) and \( m_1(t) \ldots m_M(t) \) are the magnetic measurements at time \( t \). If necessary, augment matrix \( f \) with zeros so that it has as many rows as the augmented version of \( L \).

3. Decompose matrix \( L \) by singular value decomposition:

\[
L = U \Lambda V^T,
\]

(143, 144, 145)
where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal and $\Lambda$ is a diagonal matrix of singular values, $\lambda_j$.

Then, for each value of $t$ in turn:

4. Identify the element of $\mathbf{f}(t)$ which has the largest modulus. This defines the maximum signal-to-noise ratio, $\gamma(t)$, for the given time instant.

5. Evaluate the pseudoinverse of $\Lambda$:

$$\Lambda^*(t) = \text{diag}[^{\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \frac{1}{\lambda_3}, \cdots, \frac{1}{\lambda_j}, \cdots, 0, \cdots, 0}],$$

such that $\lambda_j = 1/\lambda_j$ when $\lambda_j > \lambda_i / \gamma(t)$, and $\lambda_j = 0$ otherwise.

6. Backsubstitute to obtain the solution vector:

$$\hat{\mathbf{q}}(t) = \mathbf{V}\Lambda^*(t)\mathbf{U}^\dagger\mathbf{f}(t).$$

7. Calculate the chi-square for $\hat{\mathbf{q}}(t)$, and hence the goodness-of-fit:

$$g = 1 - P\left(\frac{v \chi^2}{2}, \frac{2}{2}\right),$$

where $P$ is the incomplete gamma function and $v$ is the total number of measurements, $E + M$, minus the number of non-zero elements in $\Lambda^*$.

Parallelisation of the Neuro-electromagnetic Inversion Algorithm

The inversion algorithm is naturally divided into two parts, the first being the generation of the linear system (steps 1 and 2) and the second being the solution of the linear equations (steps 3 to 7). The main computational burden lies in the forward problem calculations required to produce the $I_q$ and $L_q$ values of the gain matrix, $\mathbf{L}$.

Once this matrix has been generated, the subsequent steps are relatively fast.

Generation of the Gain Matrix A conventional sequential algorithm for the generation of matrix $\mathbf{L}$, expressed in Occam pseudo-code, is:

```
SEQ j = 1 FOR s
SEQ
reference.potential := scalp.potential(reference electrode, unit.dipole[j])
SEQ i = 1 FOR e
  L[i][j] := (scalp.potential(electrode[i], unit.dipole[j]) - reference.potential) / O_e
SEQ i = 1 FOR m
L[e + 1][j] := magnetic.flux.gradient(magnetometer[i], unit.dipole[j]) / O_m
```

192
The total time required to execute this sequentially is the time to perform the scalp.potential function \((E + 1) \times S\) times plus the time to perform the magnetic.flux.gradient function \(M \times S\) times. The elements of \(L\) are calculated by just two functions, the first operating repetitively on \((E + 1) \times S\) data sets and the second similarly operating on \(M \times S\) data sets. This situation clearly elicits the use of data parallelism, whereby each processor in the parallel system executes the same operation but on a different set of data. Since each element in the matrix is associated with a different combination of points in space (i.e. a different combination of sensor location and dipole location), this type of data parallelism is specifically referred to as geometric parallelism: the division of data is according to spatial geometry. Geometric parallelism was used by Ioannides et al. (1989) for the neuromagnetic inverse problem.

Because there are two types of calculation - electric and magnetic - there is some degree of flexibility in the way in which geometric parallelism is implemented. One could, for instance, allocate electric forward problems to some of the processors and magnetic forward problems to the rest. However, this is not ideal, as the electric and magnetic forward problems inevitably require different lengths of time, so that load balancing is difficult. It is better to include some electric and some magnetic forward problems on each processor, so that each one has to perform the same tasks and the computational load is balanced.

The question now arises as to how to divide up the data from which the electric potentials and magnetic flux gradients are to be calculated. In other words, which elements of the gain matrix are to be generated by each parallel processor? First note that each \(L_{ij}\) value is equal to the absolute potential at the \(i^{th}\) electrode site minus the absolute potential at the reference site. The reference potential need only be calculated once for each unit dipole or column of the gain matrix, as expressed in the above sequential pseudo-code. Hence, the electric elements of \(L\) are generated by calculating absolute potentials at the measuring electrodes only, and the absolute reference potentials are considered as an extra virtual row on the matrix. Then, once the matrix, including the virtual row, has been generated, the virtual row is simply subtracted from all of the electric rows to provide the true gain matrix. For this reason, in the computational description which follows, the gain matrix is assumed to have \(E + 1\) electric rows, each element of which is produced by a single electric forward problem.

The most obvious strategy for calculating matrix \(L\) is to calculate one electric row and one magnetic row on each parallel processor. However, unless the number of electric rows, \(E + 1\), and the number of magnetic rows, \(M\), are both exact multiples of the number of processors, \(N\), some of the processors will be idle during the overall computation time. The maximum idle time is the time required to calculate a complete electric row of the gain matrix, because the electric forward problems take about five times as long as the magnetic forward problems (figure 57). Clearly, this strategy uses the parallel system inefficiently: the granularity of the data is too coarse.

An alternative strategy for dividing the data is as follows. The \((E + 1) \times S\) data
Figure 57. Computation time, \( t \), as a function of processor number, \( p \), for an \( N \)-processor parallel system generating the gain matrix using coarse granularity geometric parallelism. Each processor nominally deals with one electric and one magnetic row of the gain matrix. The number of electric rows in the gain matrix, \( E + 1 \), may exceed the number of processors, \( N \), so that the first few processors have to generate two electric rows. The same occurs for the \( M \) magnetic rows, except that, in order to balance the computational load across the network of processors, a different set of processors is chosen to execute any excess rows. This method is inefficient, because some of the processors are idle for a period \( I \), which is equal to the time taken for one processor to calculate one electric row of the gain matrix. The time required for the parallel program to terminate is \( T \).

with which to calculate the \( L_{ij}^e \) values are divided into \( N \) sets of as near equal size as possible and these sets are distributed over the \( N \) processors. This can be visualised by dealing a pack of cards into \( N \) sets, where the number of cards is not necessarily a multiple of \( N \). If \( / \) represents integer division and \( \backslash \) represents an integer remainder operation, then the result is that \( \left( (E + 1) \times S \right) \backslash N \) processors have to calculate \( \left( (E + 1) \times S \right) / N + 1 \) elements and the rest have to calculate just \( \left( (E + 1) \times S \right) / N \) elements. This is also performed for the \( M \times S \) magnetic data. The distribution of work is described in figure 58. The maximum idle time of any processor is now comparable to the time required to calculate a single \( L_{ij}^e \) or \( L_{ij}^m \) value. This is much smaller than when coarse granularity is used: the fine granularity improves the efficiency of the algorithm and hence increases its speed. This strategy is therefore the method of choice for generating the gain matrix in parallel.

It is worth considering how the other types of parallelism could be used to generate the gain matrix. The elementary computations used to generate the gain matrix are the electric and magnetic forward problems. These are executed many times on different data. With function parallelism, the actual forward problem calculations are
Figure 58. Computation time, $t$, as a function of processor number, $p$, for an $N$-processor parallel system generating the gain matrix using fine granularity geometric parallelism. The gain matrix has $(E + 1)$ electric rows and $S$ columns, giving a total of $(E + 1) \times S$ electric elements. The data with which to calculate these elements are distributed as evenly as possible across the $N$ processors, disregarding any row or column structure in the gain matrix. Some processors have to generate one more element than others if $(E + 1) \times S$ is not an exact multiple of $N$. The same occurs for the $M \times S$ magnetic elements, but since the processors which have to calculate an extra magnetic element require only a small amount of extra time to do this, no attempt is made to avoid calculating the extra magnetic elements on processors which are already calculating an extra electric element.

decomposed into parallel processes. Although each element of the gain matrix is calculated in sequence, each individual forward problem is executed in parallel and is therefore fast. For instance, the electric forward problem consists of a summation of the first 100 terms of an infinite series. The sum could be divided into smaller sums, which are executed concurrently and added together at the end. A slight drawback, however, is that the series uses spherical harmonics, $Y_{nma}(\theta, \phi)$, which are best calculated using a recurrence relation (see equations 36). Calculation of $Y_{n+1,ma}$ requires knowledge of $Y_{n-1,ma}$ and $Y_{nma}$. In a sequential summation of increasing order terms, $Y_{n-1,ma}$ and $Y_{nma}$ are naturally available from previous iterates, but in a parallel situation, the lower order terms are being dealt with by another process, so that the required spherical harmonics are unavailable. This is typical of the type of problem that occurs when decomposing a sequential algorithm into a parallel one. In this particular case, all the spherical harmonics would have to be calculated beforehand by one process, or each individual process would have to start by performing a recurrence in order to obtain the two required initial values. For example, a process calculating terms $n = 30$ to $n = 40$ would have to use recurrence to obtain $Y_{30ma}$ and $Y_{30ma}$. Alternatively, the forward problem
could be further decomposed, so that each process calculates only certain parts of its allocated terms and another parallel process evaluates the spherical harmonics.

The magnetic forward problem consists of a numerical integration over the coils of a flux gradiometer. Twenty-four flux densities have to be calculated on the surface of each coil. For a second order gradiometer, a total of 72 flux densities are required. Hence, a parallel strategy for the magnetic forward problem using function parallelism might carry out the flux density calculations in parallel. A further process may be needed to unify these parallel calculations.

As discussed in an earlier section, the principal difficulty with function parallelism is that it is not very scalable. The above technique would work satisfactorily on a few processors, but would become inefficient if the number of nodes were increased. In the case of the electric forward problem, the inefficiency would begin when more than 100 nodes were used, since any additional nodes would just be idle. The magnetic forward problem would also suffer from the same effect, since there are 72 basic calculations regardless of the number of nodes in the system. Because of this poor scalability, function parallelism is not optimal for generating the gain matrix.

If a pipeline strategy is used, the forward problems are again divided into parallel processes. This time, however, instead of repeating the entire parallel forward problem for each element of the gain matrix, the data for each element in turn are passed along the chain of processes and the forward problem is successively computed as they go. This method is preferable to function parallelism as the spherical harmonics required by the electric forward problem can be evaluated easily. A given forward problem actually executes sequentially as the data pass along the processes. Hence, the last values of $Y_{n \lambda \alpha}$ to be produced by one process can be taken along the pipeline with the data to the next process, where they are used to continue the recurrence. However, pipelining again suffers from poor scalability; for the electric forward problem, the maximum number of processes is fixed by the number of terms evaluated in the infinite series, while for the magnetic forward problem, the maximum number of processes is determined by the number of points in the numerical integration. Hence, when generating the gain matrix, neither function parallelism nor pipelining have the natural simplicity and scalability of data parallelism.

Solution of the Linear System  Singular value decomposition is applied to the gain matrix only once (step 3), but it is inherently sequential and dividing it into parallel processes is therefore difficult. At present, no parallel algorithm is available, so it must be performed sequentially within a single process. The regularisation and backsubstitution computations (steps 4-7) must be carried out for each time instant, $t$, for which a data vector, $\mathbf{f}(t)$, is to be analysed. Typically, from a data collection time (epoch) of 511 ms, there are 512 data vectors. Hence, solving the linear system involves performing steps 4-7 of the algorithm on each of 512 data items. Much of the preceding discussion regarding the optimal choice of parallel technique applies here also. Again,
data parallelism is ideal because it is both simple and scalable. The 512 data sets are divided into $N$ parts as equally as possible and distributed over the $N$ processors, such that each processor deals with approximately $512/N$ data vectors. Function parallelism is not worth considering for solving the linear system as each of steps 4-7 requires that the previous step has been completed, which is not the case when the steps are operating in parallel. Pipelining is more suitable, as the steps of the computation can form parallel processes and the various items of data can then be passed through the steps sequentially. However, as the steps require considerably different computational effort, it is difficult to divide them into equal-sized processes, which are required if the pipeline strategy is to be efficient. In addition, there are a fixed number of steps in the algorithm, and hence a fixed number of parallel processes: the method is not scalable. Hence, data parallelism is the optimal technique for solving the linear system.

Process Topology It has been shown above that the optimal parallel inversion algorithm uses fine granularity geometric parallelism to generate the gain matrix. Singular value decomposition is then performed sequentially within a single process, and data parallelism is used to solve the linear system for a variety of measurement sets. This algorithm has been implemented using a master-slave process topology (figure 59). The master process obtains data from the user via the host computer and performs any necessary preliminary calculations, such as transforming the coordinates of the sensors into the occipital spherical coordinate system. The master process then initiates and coordinates transfer of data to the slave processes for generation of the gain matrix. An array representing matrix $\mathbf{L}$ is located within the master process and the slave processes return blocks of elements with which to fill it (step 1). The trivial generation of the combined data vectors, $\mathbf{f}(t)$, is also performed by the master process (step 2). The complete gain matrix is then decomposed by the master process (step 3). Once this has been done, the set of data vectors is divided as evenly as possible and some vectors are sent to each of the slave processes. The matrices $\mathbf{U}$, $\Lambda$ and $\mathbf{V}$, resulting from the singular value decomposition, are also sent to each slave process. Each slave calculates a number of solutions, $\mathbf{q}(t)$, and goodnesses-of-fit, $g(t)$, and returns them to the master in a block (steps 4-7). Once the master has received the solutions and goodnesses-of-fit from all of the slaves, it enters a menu loop and sends the first results to the graphics process; further results are sent when the user requests. The graphics process is subdivided into two parallel processes, the first of which communicates with the master process and the second of which is the TTGS graphics server. The master process sends a solution to the first process and requests that it be displayed. The first process then interprets the request and sends a series of elementary commands, such as `draw.line`, `plot.point` etc., to the graphics server so that the solution is displayed.

Each slave process contains two buffers for the transfer of data along the pipeline channels. There are also three calculation processes within each slave process. The first two of these compute $L_{g}^{c}$ and $L_{g}^{m}$ elements of the gain matrix, respectively, while the
Figure 59. Topology of the master-slave process configuration used to implement the fine granularity data parallelism algorithm on a linear pipeline of transputers. H: host computer; M: master process; S: slave process; LS: last slave process containing no transmission channels and thus ending the process pipeline; G: graphics process. Figures in brackets denote the transputers onto which the processes are mapped: the slave processes are mapped onto the transputers on a one-to-one basis, but the master process and the first slave process are both mapped onto the root transputer so as to reduce its idle time.
third performs backsubstitutions (figure 60). The buffers operate at high priority to ensure the free flow of data along the pipeline, while the calculation processes run at normal priority. Note that the master process and the first slave process are both mapped onto the root transputer. This is because the master process is virtually idle once the slave processes are in operation. Including the first slave process on the root transputer ensures that the root transputer is never idle and thus that efficiency is maximal.

If the source space is to be rotated to fit an individual brain, the coordinates of the dipoles are placed into temporary storage and rotated about the y-axis by an angle corresponding to 20mm at the scalp radius of 81mm; i.e. a rotation of -20/81 radians about the y-axis is applied. The dipole coordinates are successively rotated in intervals of 2mm at 81mm radius, until the angle of rotation reaches +20/81 radians. For each source space location, the linear system is generated in parallel, as previously described. However, this method uses the goodness-of-fit at one latency to determine the optimal source space location, so only one backsubstitution has to be performed once the linear system has been generated. This backsubstitution is done sequentially by the master process. Once the source space has been rotated through the complete range of angles, the optimal source space position is identified as that having the maximum goodness-of-fit. The linear system is then generated and solved for the optimal source space location, but this time, solutions are produced in parallel for the complete range of latencies.

![Diagram](image)

Figure 60. Topology of a slave process. B: buffer for transmission and selection of data; E: electric forward problem calculation process; M: magnetic forward problem calculation process; S: solution calculation process. Channels OA and OB are outward pipeline channels, channels RA and RB are return pipeline channels. Channels OB and RB are omitted from the last slave process, which has correspondingly simpler buffers.
Performance Characteristics

**Scalability** A scalable parallel program can be made to run efficiently on any number of nodes, with only trivial code changes. This is important for the program to be portable between different parallel machines. The neuro-electromagnetic inversion program has no upper limit of scalability, but its memory requirements dictate that it cannot be run on less than four transputers. This is because as the number of processors is reduced, the quantity of data to be dealt with by each slave process increases, and eventually the 1Mbyte of memory on each transputer module is insufficient. The processes dealing with the backsubstitutions are a particular cause of this problem. These processes must hold the matrices \( U, A \) and \( V \) from the singular value decomposition, as well as a subset of the (typically 512) measurement vectors. If the backsubstitutions are removed from the slave processes and the solutions calculated sequentially within the master process, the program can run on any number of transputers, but with a speed which is inferior to that of the fully parallel program.

**Computation Time** The simulations of the previous chapter were used to test the speed of the parallel program. The simulated electric and magnetic distributions of figure 31, calculated over the occipital montage of figure 30, formed the basis of the measurement set. The 20 electric potentials (relative to the inion) and 20 magnetic flux gradients were replicated 512 times to simulate 512 samples taken at 1ms intervals. These data were then inverted using the 24-dipole cruciform model (figure 27).

A sequential inversion program, running on the root transputer, required 294.2s (approximately 5 minutes) to generate, decompose and solve the linear system for the simulation described. Generation of the linear system (steps 1 and 2) required 282.4s, singular value decomposition (step 3) required 0.7s, and solution of the linear system (steps 4 to 7 repeated 512 times) required 11.1s. Meanwhile, the parallel inversion program, running on the full system of 22 transputers, took 16.0s to perform the same task. Generation of the linear system took 13.1s, singular value decomposition took 0.7s, and solution of the linear system took 2.2s.

**Speedup** The speedup of a parallel program, running on \( p \) processors, is defined as:

\[
S(p) = \frac{T(1)}{T(p)}
\]  

(149)

where \( T(1) \) is the time taken by a sequential program to perform a given task and \( T(p) \) is the time taken by the parallel program to complete the same task (Golub and Van Loan 1989). A perfectly parallelised program is said to have linear speedup; in this case, \( S(p) = p \). The majority of programs, however, have \( S(p) < p \). It is possible for a parallel program to have superlinear speedup, when \( S(p) > p \), but this is rare. The
speedup of the parallel inversion program is given in figure 61; results are given for both
the fully parallel program and for the more scalable program which solves the linear
system sequentially. The fully parallel program, running on the 22-processor system,
has a speedup of 18.3. The speedup curves for the generation and solution stages of this
program are shown in figure 62.

*Efficiency* The efficiency of a parallel program running on \( p \) processors is defined as the
speedup per processor (Golub and Van Loan 1989). That is:

\[
E(p) = \frac{S(p)}{p} = \frac{T(1)}{pT(p)}.
\]  

(150)

This is a measure of how well a sequential program is redistributed over a concurrent
network. For example, a program with linear speedup has an efficiency of 100\%. The
efficiency of the parallel inversion program is given in figure 63, along with that of the
more scalable program which backsubstitutes sequentially. Running on the 22
transputers, the fully parallel program has an efficiency of 0.83. Figure 64 shows the
efficiency of the generation and solution stages of this program.

*Discussion* The drop in efficiency of the program with increasing number of processors
can be explained by communication overheads. Data must be transmitted from the
master process to the relevant slave processes and time is taken by the processors to do
this. The speedup is therefore less than linear and the efficiency is less than 100\%. In
the case of the linear pipeline of processors used in this work, addition of an extra
processor onto the end of the pipeline causes that all the processors must be delayed in
order to pass information along the pipeline from the master process to the last slave
process. The global processor time occupied by such data communication increases as
an arithmetic progression as the pipeline becomes longer, and hence the efficiency of the
program drops. This is particularly prevalent in the backsubstitution stages of the
program (steps 4-7), where the communications are extensive but the calculations
relatively fast (see figures 62 and 64). If the program were scaled up onto a large
transputer network, the speedup of the complete program would plateau and ultimately
drop, eventually invalidating the use of transputers. A more efficient processor topology
is the answer to this problem. If each slave process were connected directly to the
master process, communication time would increase in proportion to the number of
transputers and efficiency would remain constant. However, since a transputer only has
four links and a transputer motherboard only has eight edge connectors, this is
impractical and the modest efficiency of a pipeline or ring topology must suffice.

A second limit to the speedup is the distribution of data among the processors.
In the test described, there are \( 20 \times 24 = 480 \) \( \mathcal{I}_\gamma \) values and \( 20 \times 24 = 480 \) \( \mathcal{I}_\gamma^n \) values to
Figure 61. The speedup of the parallel inversion program running on a linear pipeline of transputers. The lower limit of network size is governed by the scalability of the parallel program. The solid line shows the performance of the optimal parallel program; the broken line shows the performance of a more scalable program which performs the backsubstitution process sequentially on the root transputer. The feint line indicates the optimal performance achieved by a perfectly parallelised program (linear speedup).

Figure 62. The speedup of algorithm steps 1-2 (solid line), and algorithm steps 4-7 (broken line). Linear speedup is indicated by the feint line. The matrix generation process is more efficient (closer to linear) than the solution process as the ratio of communication time to calculation time is lower. The speedup of the singular value decomposition (step 3) is unity by definition, as it is always executed sequentially.
Figure 63. The efficiency of the parallel inversion program running on a linear pipeline of transputers. The solid line shows the performance of the optimal parallel program; the broken line shows the performance of a more scalable program which performs the backsubstitution process sequentially on the root transputer. The feint line indicates 100% efficiency (linear speedup).

Figure 64. The efficiency of algorithm steps 1-2 (solid line), and algorithm steps 4-7 (broken line). 100% efficiency (linear speedup) is indicated by the feint line. The matrix generation process is more efficient than the solution process as the ratio of communication time to calculation time is lower.
be calculated. Thus, if the number of processors were increased beyond 480, some processors would become idle, speedup would plateau and efficiency would drop.

The final limit to the speedup is the proportion of inherently sequential code within the program. If \( T_s \) is the time required by a single processor to perform the computations which cannot be parallelised, and \( T_p \) is the time spent sequentially performing those which can, then:

\[
T(1) = T_s + T_p, \tag{151}
\]

and:

\[
T(p) = T_s + \frac{T_p}{p}, \tag{152}
\]

assuming zero communication time and 100% efficiency for those computations which can be parallelised. Then:

\[
S(p) = \frac{T(1)}{T(p)} = \frac{T_s + T_p}{T_s + \frac{T_p}{p}}, \tag{153}
\]

whence:

\[
S_{\text{max}} = \lim_{p \to \infty} S(p) = 1 + \frac{T_p}{T_s}, \tag{154}
\]

which is known as Amdahl's law. In the case of the parallel inversion algorithm, \( T_s \) corresponds roughly to the time required for the singular value decomposition, which is 0.7s for the test data described. \( T_p \) is the remaining 293.5s of the sequential computation time. Thus, the maximum achievable speedup is approximately 440.

It should be noted that both of these two latter considerations are somewhat hypothetical, as in practice, if the number of processors were increased, the communication inefficiency would diminish the speedup before its theoretical maximum were reached. The limited maximum speedup implies that even if a large parallel computer were available, the parallel program could not be used to obtain neuro-electromagnetic inversions at the data acquisition rate, which is usually about 1kHz. Solutions would have to be produced in around 1ms, necessitating a speedup of around \( 10^3 \), which is well beyond the practically achievable figure.
Minimum Norm Estimation

Although the linear inverse procedure has been discussed for the case of an anatomically constrained source model, it is also possible to use it to calculate conventional unconstrained minimum norm estimates. If this is required, the source space is simply defined as the two- or three-dimensional region in which the primary current density could possibly be non-zero. This source region is discretised into voxels and three orthogonal current dipoles are situated in each voxel. Collectively, these dipoles then represent the three orthogonal components of primary current density at each point in the source space. The inverse problem is solved by calculating the magnitudes of the dipoles. Clearly, this situation is exactly the same as when the source space is constrained, except that rather more dipoles are required to define the source space.

Numerical Algorithm To calculate a minimum norm estimate, the linear system is generated as for a constrained analysis. If there are $S/3$ voxels in the source space, giving $S$ current dipoles, matrix $L$ has dimensions of $E + M$ by $S$. Typically, for voxels of $10\text{mm} \times 10\text{mm} \times 10\text{mm}$, $S$ is around 1000. $L_{ij}$ now represents the electric potential at electrode $i$, relative to a reference site, due to the first orthogonal unit dipole of the first voxel. Similarly, $L_{j\parallel}^\varepsilon$ is the potential resulting from the second component of the first voxel, $L_{j\perp}^\varepsilon$ the potential resulting from the first component of the second voxel, and so on for all of the voxels. Values $L_j^m$ are similarly defined. As for the anatomically constrained source space, each $L_{ij}^\varepsilon$ or $L_j^m$ value represents the component of the $i$th electrode or magnetometer lead field in the direction of dipole $j$. Using this definition, it can be seen that the lead field of the $i$th electrode at the $k$th voxel is given by $(L_{i,k}^\varepsilon, L_{i,k-1}^\varepsilon, L_{i,k-2}^\varepsilon)^\top$, and the lead field of the $i$th magnetometer at the $k$th voxel is given by $(L_{i,k}^m, L_{i,k-1}^m, L_{i,k-2}^m)^\top$, i.e. each set of three $L_{ij}^\varepsilon$ or $L_j^m$ values provides the three components of the lead field at a particular voxel.

Having generated the linear system, the next step is to decompose matrix $L$. In principle, this can be achieved exactly as before, but in practice, the previous method is unsuitable, for two reasons. The first is that $L$ now has dimensions of about $40 \times 1000$. Singular value decomposition of such a large matrix would be slow. Secondly, and more significantly, the singular value decomposition algorithm given by Press et al. (1989) requires that the matrix to be decomposed has at least as many rows as columns. In the present case, this means that $L$ would have to be augmented with zeros to a dimension of $1000 \times 1000$. The resulting matrix - consisting almost entirely of zeros - would (a) require too much memory to be of any use, and (b) would require a very long time for decomposition. Straightforward decomposition of matrix $L$ is therefore inappropriate and it is necessary to diverge from the anatomically constrained approach.
When anatomical constraints are used, it is possible for there to be either more or less measurements than dipole sites; i.e. \( E + M > S \) or \( E + M < S \). Solutions \( \tilde{q}(t) = L^* f(t) \) are sought by singular value decomposition and backsubstitution. However, in the case of unconstrained minimum norm estimation, it is guaranteed that matrix \( L \) will have more columns than rows; i.e. \( E + M < S \). Hence, it is possible to make use of an alternative definition of the pseudoinverse of \( L \) (Ben-Israel and Greville 1974; Barnett 1990; Hämäläinen and Ilmoniemi 1984; Wang et al. 1992):

\[
L^* = L^T (L L^T)^+. \tag{155}
\]

If the total number of independent measurements, \( E + M \), is denoted by \( R \), then \( L L^T \) is an \( R \times R \) matrix. Since \( R \) is typically about 40, the pseudoinverse of \( LL^T \) can be computed readily using singular value decomposition:

\[
LL^T = U \Lambda V^T, \tag{156}
\]

where \( U \) and \( V \) are orthogonal matrices and \( \Lambda \) is a diagonal matrix containing the singular values of \( LL^T \). All three of these matrices have dimensions of \( R \times R \). There is no need for \( LL^T \) to be augmented beforehand, as it is already square. However, since the elements of \( LL^T \) can be large enough to cause floating overflow during decomposition, they are normalised beforehand. The element with the largest modulus is found, and all elements are divided by this modulus. After decomposition, all the elements of \( \Lambda \) are multiplied by the modulus, to denormalise. The pseudoinverse of \( \Lambda \) is then defined as:

\[
\lambda^*_j(t) = \frac{1}{\lambda_j}, \quad \lambda_j > \frac{\lambda_1}{\gamma^2(t)}, \\
\lambda^*_j(t) = 0 \quad \text{otherwise,} \quad j = 1...R, \ t = 1...T, \tag{157}
\]

where \( \gamma(t) \) is the maximum signal-to-noise ratio of the data for time \( t \), given by the element of \( f(t) \) which has the largest modulus. This is identical to the previous procedure, except that the cut-off value in \( \Lambda \), beyond which the elements of \( \Lambda^* \) are set to zero, is chosen according to the square of the maximum signal-to-noise ratio. This is purely a consequence of inverting \( LL^T \) as opposed to \( L \) and results in an identical solution to that produced by evaluating \( L^* \) directly. The regularised pseudoinverse of \( LL^T \) is then given by:

\[
(LL^T)^* = V \Lambda^* U^T, \tag{158}
\]

whence the minimum norm solution is constructed as:
\[ q^*(t) = L^T V \Lambda^*(t) U^T f(t). \]  \hspace{1cm} (159)

Hence, at the expense of several matrix multiplications, the computationally intensive process of singular value decomposition can be limited to an \( R \times R \) matrix, as opposed to an \( R \times S \) or \( S \times S \) matrix.

**Parallel Computation Strategy** The above method is only a minor modification of the previous method, and contains the same three essential steps: matrix generation, singular value decomposition, and backsubstitution. As before, a master-slave process topology is used to carry out these tasks. Matrix \( L \) is generated by dividing it into \( N \) parts and calculating each part within one of the \( N \) slave processes. The results of the forward problem calculations are returned to the master process, which collates the individual portions of matrix \( L \) and groups them into one complete matrix.

The master process calculates \( LL^T \) and decomposes it. There are \( T \) data vectors, \( f(t) \), to be inverted and these are divided into \( N \) sets and distributed to the \( N \) solution processes. Matrices \( L, V, \Lambda \) and \( U \) are also transmitted to each solution process, so that equation 159 can be evaluated for the data vectors. As before, the solution processes calculate \( \Lambda^* \) for each time instant separately, as the signal-to-noise ratio varies with time. The solutions are returned to the master process, which groups them together into one solution matrix.

The distribution of logical processes onto the physical transputers is identical to that for anatomically constrained distributed-source analysis, except that the root transputer contains only the master process, instead of the master and a slave process. This is because the master process requires more memory than previously, so that the slave process cannot be accommodated. Unfortunately, this means that the root transputer is idle while the slave processes are calculating forward problems or solutions.

**Scalability** The algorithm described can run on 18 transputers or more. The lower limit to the number of nodes on which it can be run is governed by the memory capacity of the transputer modules. Running the code on fewer transputers dictates that more data has to be handled by each one, and eventually the remaining transputers run out of memory. Scalability can obviously be improved by increasing the external memory on the transputer modules.

**Computation Time** The computation time of the parallel minimum norm estimation program was determined by inverting the test data described earlier. The simulated data consisted of 20 electric potentials and 20 magnetic flux gradients, as recorded by an occipital montage. The data were replicated 500 times to simulate 500 samples taken at 1ms intervals. The source space consisted of a quasi-cubic region, bounded by the spherical coordinates \( r = 25 \text{mm}, \ r = 65 \text{mm}, \ \theta = 20/65 \text{rads}, \ \theta = 110/65 \text{rads}, \ \phi = -45/65 \text{rads} \).
rads, $\phi = 45/65$ rads. There were a total of 500 grid points, with a spacing of 10mm at $r = 65$mm, and a correspondingly smaller spacing at smaller radii. The radii changed in steps of 10mm, giving five spherical surfaces. Within each surface, $\theta$ and $\phi$ were both increased in steps of 10/65 rads. The three orthogonal dipoles at a given grid point were oriented in the directions of the spherical basis vectors at that point. The source space gave a total of 1500 discrete current dipoles.

The parallel program, running on 22 transputers, took 898.0s (about 15 minutes) to invert the test data. Generating the linear system took 847.7s (around 14 minutes) and solving it took 50.3s. It was not possible to calculate the speedup, as this necessitated a sequential version of the code running on one transputer. With the parallel program requiring at least 18 transputers, a single-processor version of the code was clearly not feasible. However, the majority of the computation time was taken up with calculating forward problems, and since the ratio of communication to computation for this was identical to that for the anatomically constrained algorithm, it was suggested that the speedup and efficiency would be similar for both programs.

**Equivalent Current Dipole Estimation**

The position coordinates of an equivalent current dipole are non-linear with respect to the computed electric potential and magnetic flux density, and the equivalent current dipole therefore has to be located by non-linear parameter estimation. This type of estimation technique is inherently sequential, as it involves an iterative process in which the dipole location is successively adjusted until the dipole accurately accounts for the measured data. At each stage of the process, the new location of the current dipole is based upon a goodness-of-fit measure obtained at several previous locations. Clearly, this method cannot be executed in parallel, as each step cannot proceed until the previous step has been completed.

For a given dipole location, the computed electric potentials and magnetic fields are linear functions of the components of the dipole moment. Consequently, the dipole moment can be obtained from the measured data by a linear estimation technique (see equation 61 of chapter 4). If the dipole moment is treated as the vector sum of three orthogonal dipole moments, the problem is that of determining the magnitudes of three collocated current dipoles of fixed location and orientation. This situation is a simplified form of the anatomically constrained model. The dipole moments are found by generating and inverting a matrix - a process which can be distributed across a network of processors and executed in parallel.

Now consider the complete procedure for estimating an equivalent current dipole, which involves the determination of both the position and moment vectors. The non-linear position estimation is sequential, and therefore cannot be parallelised. However, at each stage of the iteration, the dipole moment must be found, and this
procedure can be efficiently executed in parallel. Thus, the natural strategy for parallelising an equivalent current dipole estimation process is to perform a sequential non-linear estimation routine to find the dipole location, but to speed this up by finding the dipole moment in parallel at each stage of the iteration. This is, in fact, the normal (and only) way of "parallelising" a non-linear parameter estimation process: the iteration continues to run sequentially, but the evaluation of the cost function is speeded up by calculating it in parallel. In the equivalent current dipole case, the cost function is a goodness-of-fit measure, which requires the dipole moment to be evaluated.

A parallel program for equivalent current dipole fitting has been implemented on the transputer array and the details of this program are now given.

Numerical and Computational Algorithm  The equivalent current dipole program is based on Powell's algorithm (Powell 1964), as given by Press et al. (1989). This algorithm is used to determine the dipole position parameters which minimise the combined chi-square for the electric and magnetic measurements. The Pascal algorithm of Press et al. (1989) has been translated into Occam 2 for execution on the transputers. Although the Occam code is primarily a direct, line-by-line translation of the Pascal code, the parameter-passing mechanism has been altered. The Pascal routine is designed to call a user-defined function which evaluates the quantity to be minimised. Since an Occam function cannot communicate with other processes, which is required for the dipole moment and chi-square to be evaluated in parallel, the parameter-passing mechanism has had to be revised. In the parallel algorithm the minimisation routine is run in parallel with a second process which coordinates the calculation of chi-square values. The Powell process requests and receives chi-square values from the second process via a pair of data channels (figure 65). This arrangement is the parallel equivalent of parameter passing. The minimisation process and its partner process (hereafter referred to as the server process) are both located on the root transputer. Although they are defined as parallel processes, they do not actually work in parallel, as the minimisation routine cannot proceed until a chi-square value is returned to it, and the server process does not operate unless a chi-square is requested. Hence, the two processes are mutually exclusive in terms of the times at which they become operative.

The minimisation procedure sequentially performs six dipole localisations, each one starting with a different initial dipole location. The initial dipole locations are evenly distributed throughout the occipital lobes of the brain so as to cover areas 17, 18 and 19. The starting locations are given in table 9. The purpose of the repeated localisation is to overcome the possibility that the minimisation procedure could become "trapped" in a local minimum of the chi-square function, thus returning a false result. If only one localisation is performed, the resulting location may correspond to that of a local minimum in the chi-square function. Repeating the localisation from a variety of different starting points ensures that the overall minimum in the chi-square function is found by at least one of the attempts.
Figure 65. Process topology of the parallel equivalent current dipole program. H: host computer; RCS: request chi-square; SCS: supply chi-square; GLS: generate linear system; SLS: solve linear system; CCS: calculate chi-square; EFP: electric forward problem; MFP: magnetic forward problem. Figures in brackets indicate the transputers onto which the processes are mapped.
<table>
<thead>
<tr>
<th>Starting Location</th>
<th>$r$ (mm)</th>
<th>$\theta$ (degs.)</th>
<th>$\phi$ (degs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>45</td>
<td>-30</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>45</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>45</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>90</td>
<td>-30</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>90</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>90</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 9. Equivalent current dipole starting locations.

During the course of a particular localisation, the minimisation process sends a dipole location to the server process and requests the chi-square value for the dipole in that location. The server then generates the linear system describing the external electric potentials, $e_i$ ($i = 1 \ldots E$), and the magnetic flux gradients, $m_i$ ($i = 1 \ldots M$), in terms of the dipole moment vector, $q$. The dipole moment is considered to be the vector sum of three orthogonal dipoles located at the given dipole site, directed along the $x$-, $y$- and $z$-axes, respectively. The moments of the three dipoles are denoted $q_x$, $q_y$, and $q_z$. Then, the total of $R$ external measurements are given by:

$$
\begin{pmatrix}
  e_i \\
  \sigma_e \\
  \vdots \\
  e_e \\
  \sigma_e \\
  \vdots \\
  m_i \\
  \sigma_m \\
  \vdots \\
  m_m \\
  \sigma_m
\end{pmatrix}
= 
\begin{pmatrix}
  I_{Rx}^e & I_{Ry}^e & I_{Rz}^e \\
  \sigma_e & \sigma_e & \sigma_e \\
  \vdots & \vdots & \vdots \\
  I_{Rx}^e & I_{Ry}^e & I_{Rz}^e \\
  \sigma_e & \sigma_e & \sigma_e \\
  \vdots & \vdots & \vdots \\
  I_{Rx}^m & I_{Ry}^m & I_{Rz}^m \\
  \sigma_m & \sigma_m & \sigma_m \\
  \vdots & \vdots & \vdots \\
  I_{Rx}^m & I_{Ry}^m & I_{Rz}^m \\
  \sigma_m & \sigma_m & \sigma_m
\end{pmatrix}
\cdot
\begin{pmatrix}
  q_x \\
  q_y \\
  q_z
\end{pmatrix}
$$

(160a)

or:

$$
\mathbf{f} = \mathbf{Lq}.
$$

(160b)

The terms $I_{Rx}^e$, $I_{Ry}^e$ and $I_{Rz}^e$ ($i = 1 \ldots E$) represent the electric potentials produced at the $i^{th}$ electrode, relative to the reference potential, by dipole moments $q = (1, 0, 0)^T$, $(0, 1, 0)^T$ and $(0, 0, 1)^T$, respectively. These dipole moments produce magnetic flux gradients of $I_{Rx}^m$, $I_{Ry}^m$ and $I_{Rz}^m$ at the $i^{th}$ magnetometer site. $\sigma_e$ and $\sigma_m$ are the electric and magnetic standard deviations, respectively. The elements of matrix $\mathbf{L}$ are calculated by the linear
network of slave processes, which are mapped on a one-to-one basis onto all the transputers. The fine granularity geometric parallelism method is used for the calculations.

The required dipole moment, \( \tilde{\mathbf{q}} \), is that which minimises the chi-square. This solution is found by premultiplying the data vector by the pseudoinverse of \( \mathbf{L} \):

\[
\tilde{\mathbf{q}} = \mathbf{L}^\dagger \mathbf{f}.
\]  \hfill (161)

The pseudoinverse, \( \mathbf{L}^\dagger \), could be obtained by decomposing \( \mathbf{L} \) by singular value decomposition. However, it is more economical to calculate:

\[
\mathbf{L}^\dagger = (\mathbf{L}^\top \mathbf{L})^{-1} \mathbf{L}^\top,
\]  \hfill (162)

since \( \mathbf{L} \) is an \( R \times 3 \) matrix, whereas \( \mathbf{L}^\top \mathbf{L} \) is only a \( 3 \times 3 \) matrix. Hence, the server process computes \( \mathbf{L}^\top \mathbf{L} \) and decomposes it by singular value decomposition:

\[
\mathbf{L}^\top \mathbf{L} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{V}^\top,
\]  \hfill (163)

where \( \mathbf{U} \) and \( \mathbf{V} \) are \( 3 \times 3 \) orthogonal matrices and \( \boldsymbol{\Lambda} \) is a \( 3 \times 3 \) diagonal matrix containing the singular values of \( \mathbf{L}^\top \mathbf{L} \). As with minimum norm estimation, matrix \( \mathbf{L}^\top \mathbf{L} \) is normalised before decomposition, and \( \boldsymbol{\Lambda} \) is denormalised immediately after. The truncated matrix, \( \Lambda^* \), is calculated such that \( \lambda_j^* = 1/\lambda_j \) when \( \lambda_j > 10^{-5} \lambda_1 \), and \( \lambda_j^* = 0 \) otherwise. This serves to regularise \( (\mathbf{L}^\top \mathbf{L})^\dagger \) just sufficiently to overcome any roundoff error which may occur during the decomposition. The chi-square estimate of the dipole moment is calculated as:

\[
\tilde{\mathbf{q}} = \mathbf{V} \Lambda^* \mathbf{U}^\top \mathbf{L}^\top \mathbf{f}.
\]  \hfill (164)

Note that this method naturally zeros the radial component of the dipole moment when only magnetic data are used. It is therefore a good choice of method for use in the present situation, since either electric data or magnetic data or both can be analysed by the one algorithm.

The chi-square value for the equivalent current dipole can be obtained from the matrices of equation 160:

\[
\chi^2 = \sum_{i=1}^{R} \left( f_i - \sum_{j=1}^{3} L_{ij} \tilde{q}_j \right)^2.
\]  \hfill (165)

This value is returned to the minimisation process. The decomposition of \( \mathbf{L}^\top \mathbf{L} \), the backsubstitution, and the evaluation of the chi-square are all performed sequentially by
the server process as these are relatively trivial operations, compared to the forward problem calculations required to generate matrix $L$.

On some occasions, the minimisation procedure causes the dipole to "wander" out of the spherical conductivity model. This situation is detected when the minimisation process requests the chi-square for a dipole whose position lies outside of the white matter sphere. The server process does not attempt to calculate the moment of such a dipole, but instead forces the dipole to return to the innermost sphere by returning an artificially high chi-square, equal to $10^{20}$ times the dipole radius.

When the Powell procedure has located the optimal location of the dipole, it terminates and repeats the process for a different starting location. Termination occurs when the fractional reduction in the chi-square between two successive iterates falls below 0.1. This criterion is satisfied only when the dipole lies within a millimetre of the optimal location. Before the Powell process terminates, it must signal to the server process that it has terminated, so that the server process can store the most recent dipole location, dipole moment and chi-square as those of the equivalent current dipole. At this point, the server process compares the chi-square for the equivalent current dipole with a running "best" chi-square: if the new chi-square is lower than the previous lowest value, then the new equivalent current dipole parameters are stored as optimal, but if the new chi-square is equal to or higher than the previous lowest value, the new dipole parameters are rejected. The server process thus updates the details of the optimal dipole and then repeats its service to the minimisation process. At the end of the six dipole localisations, the server process contains the coordinates and chi-square for the optimal equivalent current dipole. Hence, after the sixth localisation, the server calculates the goodness-of-fit for this optimal dipole and terminates, while the minimisation process also terminates.

**Scalability** Since the linear system involved in this method is considerably smaller than that in either of the previous two methods, the quantity of memory required for the calculations is much less. Hence, the program can be run on any number of transputers, from one upward. This is aided by the fact that the slave processes only calculate forward problems, and not solutions, thus not requiring memory for matrix storage.

**Computation Time** The previous 20 electric potentials and 20 magnetic flux gradients were applied to the equivalent current dipole program. The dipole was started at the first location of table 9. A single dipole localisation took 89.8 minutes when the parallel code was run on the root transputer. When the code was run on all 22 transputers, this time was reduced to 263.0s (about 4½ minutes). Note that these were times for one localisation only, as opposed to those for six localisations from different starting points. The time for six localisations on the 22 transputers would have required approximately $6 \times 263\text{s} = 1578\text{s}$ (approximately 26 minutes). The exact localisation time was, of
course, a function of the dipole starting and finishing positions. Various localisations, from different starting points, required 266.5s, 244.0s and 529.5s, respectively.

**Speedup** Assuming that the root transputer takes the same time to execute the parallel program as it does to execute a sequential version of the algorithm (which is a reasonable assumption as a transputer performs both sequential and parallel operations serially and communications within one transputer take negligible time), then the speedup of the program is 20.5. Note that this is independent of dipole starting point or number of repeated localisations.

**Efficiency** The efficiency of the algorithm is 0.93. Again, this is independent of the dipole starting point and number of localisations. The speedup and efficiency of this algorithm are higher than those of the previous algorithms; this is because the slave processes only calculate forward problems, which have a lower communication/computation ratio than the solution calculations. Hence, a lower proportion of the total computation time is spent communicating.

**Conclusion**

This chapter clearly shows that a transputer network can be used to speed up anatomically constrained distributed-source analysis, minimum norm estimation and equivalent current dipole fitting. The starting point for efficient parallel solution of the inverse problem is a linear description of the measurements in terms of the underlying sources. This type of mathematical formulation is directly analogous to the general spatio-temporal formulation proposed by Mosher et al. (1992), in which current dipole sources are constrained in both location and orientation, constrained in location only, or are completely unconstrained (see chapter 4). This chapter has shown that a single linear system, or a modification of it, can be used to solve the inverse problem for the case of discrete dipoles which are (a) fixed in location and orientation - anatomically constrained analysis, (b) fixed in location but able to rotate - minimum norm estimation, and (c) allowed to move and rotate - equivalent current dipole fitting. This mathematical framework allows an efficient computational strategy to be developed using data parallelism, even when the problem is conventionally classed as non-linear, as with equivalent current dipole fitting.

In the case of anatomically constrained distributed-source analysis, the computation time required for inversion of a typical set of data can be reduced by a factor of about 18 using a parallel system of 22 transputers. In absolute terms, this reduction in time is from 5 minutes to around 16 seconds. This is obviously a useful achievement, but since 5 minutes is not a particularly long computational time to begin with, one might well ask whether the reduction in time is worth the cost and effort of
using a parallel system. The answer is that when the parallel program is used repetitively, either in research or for clinical purposes, the reduction in processing time becomes very noticeable to the operator. Furthermore, if the source space is to be rotated to match the head of an individual subject, a total of 21 inversions have to be performed. This requires about 1½ hours using sequential methods, but only takes about 5½ minutes in parallel. Hence, under these conditions, parallel techniques become indispensable.

Minimum norm estimation also becomes much more convenient when a parallel computer is available. In this case, a typical data set could require about 4½ hours for sequential inversion, but only requires 15 minutes in parallel. Again, parallel computing is well worth the cost and programming effort. The same situation occurs with equivalent current dipole fitting, which requires about 9 hours \((i.e. \ 6 \times 90 \text{ minutes for 6 localisations})\) when performed sequentially, but only about 26 minutes \((i.e. \ 6 \times 263s)\) when performed concurrently. A modest transputer network, such as that used in this thesis, is therefore a very valuable tool for fast solution of the neuro-electromagnetic inverse problem.

Unfortunately, real time inverse solutions, which are displayed virtually instantaneously and updated at the same rate as the electric and magnetic fields are sampled, are still unfeasible, even with parallel computing technology. This is firstly because reducing the computation time from 100s to 1ms would require 100,000 processors, which is somewhat large! Secondly, it has been shown that the efficiency of such a system would be low, so that even if the number of processors were available, their speedup would be considerably less than linear and the desired processing time would not be achieved.

One area in which parallel computing has a great deal of potential is in the boundary element method. Here, even forward problems require several hours. An electric forward problem calculation involves generating and solving a linear system of equations:

\[
\mathbf{v} = \mathbf{g} + \mathbf{A} \mathbf{v},
\]

where \(v_i\) is the potential at the \(i\)th discretisation point of the triangulated conductivity model, \(g_i\) is the potential that would occur at the \(i\)th discretisation point if the medium were infinite and homogeneous, and \(A_{ij}\) is proportional to the solid angle subtended by triangle \(j\) at discretisation point \(i\) (see equation 19 of chapter 3). The elements of \(\mathbf{g}\) and \(\mathbf{A}\) can be calculated simply, but \(\mathbf{A}\) can have dimensions of as large as 1000 \(\times\) 1000, so the operation is not trivial. The procedure could easily be decomposed into parallel processes using geometric parallelism and executed on a transputer array for maximum speed. Equations 166 can either be solved iteratively using the Gauss-Seidel iteration scheme or the Jacobi iteration scheme, for which parallel algorithms are available, or solved directly by computing:
\[ \mathbf{v} = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{g}, \tag{167} \]

which, being a straightforward matrix inversion method, can also be carried out in parallel (Golub and Van Loan 1989).

The magnetic forward problem is calculated by generating the equations:

\[ \mathbf{e} = \mathbf{e}^\infty + \tilde{D}\mathbf{v}, \tag{168} \]

where \( e_i \) is the \( i \)th external magnetic flux gradient, \( e_i^\infty \) is the flux gradient that would occur at the \( i \)th magnetometer site if the medium were infinite and homogeneous, and \( \tilde{D} \) is a summation over the coils of the gradiometer (see equations 25). This linear system can be generated effectively by geometric parallelism, and the evaluation of the vector \( \mathbf{e} \) is then trivial.

If the boundary element method were used with the linear inverse procedure described in this chapter, then solving the above equations for \( \mathbf{v} \) and \( \mathbf{e} \) would provide one column of the gain matrix, \( \mathbf{L} \). The entire gain matrix could therefore be generated by sequentially repeating the parallel boundary element method. Alternatively, the boundary element method could be performed sequentially on each node of a transputer network, so that each node calculated one column of the gain matrix. The linear system would then be solved as described earlier in the chapter. Hence, parallel computing could be used to advantage in the boundary element method, whether this were used for the forward problem in its own right or for the forward aspect of the inverse problem.
7. Source Analyses of Pattern Onset and Pattern Reversal Visual Evoked Magnetic Responses

Introduction

The previous chapters have been concerned with the development of techniques for determining the cortical origins of visual evoked electric and magnetic responses. Three methods of analysis have been established: equivalent current dipole fitting, unconstrained minimum norm estimation and anatomically constrained distributed-source analysis. This chapter and the following chapter discuss the application of these techniques to real neuro-electromagnetic records.

In this chapter, attention is confined to evoked magnetic responses, as opposed to electric or combined electromagnetic responses. The choice of a single modality is to enable the performance of the algorithms to be readily assessed before the additional problems associated with analysing electric and magnetic data simultaneously are encountered. Magnetic responses are selected primarily because a wide variety of data have been painstakingly collected prior to this study (Armstrong et al. 1991a; Harding et al. 1991; Degg et al. 1992), but also because single dipole analyses have shown that current sources can be determined more accurately from magnetic records than from electric records (Stok 1987; Ishiyama and Kanai 1992). Hence, it is natural to concentrate initially on evoked magnetic responses.

The reasons for choosing pattern onset and pattern reversal visual evoked magnetic responses are twofold. Firstly, these responses are both of significant amplitude, so that the signal-to-noise ratio is sufficiently elevated, which is necessary for accurate source analysis. Secondly, the sources of these responses are controversial: the conclusions of previous studies are by no means unanimous. Thus, a comprehensive source analysis is judged to be of value.

The first part of this chapter contains a select review of pattern onset and pattern reversal visual evoked electric and magnetic responses. For completeness, both electric and magnetic modalities are considered, although treatment is restricted to (1) achromatic stimuli, and (2) normal responses. Specific emphasis is given to the
effect of stimulus parameters on the responses. Previous source analyses are then discussed, before the author's work is described in detail. Some of the work in this chapter is outlined in Bedford et al. (1993).

The Pattern Onset Visual Evoked Potential

The visual evoked potential elicited by the abrupt onset of a contrast-modulated pattern consists of three major peaks. These peaks are referred to as CI, CII and CIII. The first peak, CI, is positive and has a latency of 65-80ms. The second peak, CII, is negative and generally much larger in amplitude than CI, having a latency of 90-110ms. The third peak, CIII, has a latency of around 160ms and has a low amplitude, comparable to that of CI (figure 66; Spekreijse et al. 1973). In addition, an earlier negative peak of latency around 60ms, designated CO, has been reported (Kulikowski 1977; Lesèvre and Joseph 1979; Drasdo 1980).

The shape of a recorded waveform is influenced significantly by the temporal properties of the stimulus (Jackson and Barber 1980). Components CI and CIII remain approximately constant in amplitude regardless of interstimulus interval, but CII shows a dramatic adaptive effect as the interstimulus interval is increased from 0.1s to 9.0s. For an interstimulus interval of 0.1s, CII is virtually absent, but it increases in amplitude to around 15µV as the interstimulus interval rises to 9.0s. The most dramatic rise in amplitude occurs for interstimulus intervals up to 2.0s (figure 67; Jackson and Barber 1980). The interpretation of this phenomenon is that at short interstimulus intervals, one pattern presentation influences the next. This finding is in accord with the observations of James and Jeffreys (1975), which indicate that CII and CIII can be attenuated by quite brief pre-exposures of the stimulus pattern, but CI is unaffected.

Figure 66. Example of a binocular pattern onset visual evoked potential from a single subject. The response was elicited by the onset and offset of a 3° checkerboard having 10' checks of 20% contrast, and was recorded occipitally relative to a vertex reference. The response clearly shows the onset peaks, CI, CII and CIII, and the offset peak. Positive upwards. After Spekreijse et al. (1973).
Figure 67. Variation of the pattern onset visual evoked potential with interstimulus interval. The potentials were elicited by the onset of a 20° checkerboard of 30' checks, persisting for 0.5s, and were recorded from 25mm above the inion on the midline, relative to a mid-frontal reference. Positive upwards. After Jackson and Barber (1980).

Moreover, the amplitude of CII and CIII can be greatly reduced by the presence of steady outlines which approximately match the outlines of the stimulus pattern, whereas C1 persists.

The duration for which the pattern is presented also affects the recorded waveform. This is because the offset of the pattern also evokes a characteristic response, dominated by a single positive peak which is equivalent to C1 (James and Jeffreys 1975). If the pattern is presented for only a few milliseconds, the offset response confounds the later components of the onset response, and a new evoked potential configuration results. However, a presentation time of around 25ms produces an evoked response of maximal amplitude (Jeffreys 1977) and has therefore been used by Jeffreys and Axford (1972a,b) for studies of the pattern onset response.

To distinguish it from a luminance onset or flash stimulus, a pattern onset stimulus is temporally isoluminant. That is, the contrast-modulated pattern evolves in such a way that the overall luminance of the stimulus remains constant. The precise form of the pattern can significantly affect the evoked potentials. Spekreijse et al. (1973) used a bar mirror to generate first vertical and then horizontal achromatic square-wave gratings. Both of these stimuli produced similar evoked potentials (figure 68). However, when the mirrors were combined to produce a square lattice similar to a
Figure 68. Effect of stimulus configuration on the pattern onset visual evoked potential. The stimuli were generated by superimposing two bar mirrors and the responses were recorded from the midline above the inion, relative to a vertex reference. Note that the offset response increases in amplitude with decreasing angle of mirror intersection, but the onset response has a different form with each stimulus. Positive upwards. After Spekreijse et al. (1973).

checkerboard, the response was markedly different. Changing the intersections of the lattice to a more acute angle had the effect of reducing the onset response while dramatically increasing the amplitude of the offset response (figure 68). Hence, it is evident that choice of stimulus pattern strongly influences the form and amplitude of the evoked potentials.

With regard to the individual peaks of the onset response, CI appears to be relatively unaffected by the stimulus shape (James and Jeffreys 1975). It is not very sensitive to the sharpness of the stimulus pattern and is not attenuated by pre-exposure or by the presence of temporally invariant outlines near to and matching the edges of the pattern elements. It therefore appears to be the result of contrast-specific processes which do not readily adapt to prolonged stimulation (James and Jeffreys 1975). In contrast, CII and CIII are mainly sensitive to discontinuous contours. This can be neatly illustrated by replacing a checkerboard with a grid pattern whose lines correspond to the edges of the checks. The CII response to the grid pattern is virtually identical to that from the conventional checkerboard, regardless of how the contrast onset is achieves, i.e. whether the grid lines are dark or bright and whether or not there is an overall luminance change (Spekreijse 1980). CII is more sensitive to the sharpness of the
stimulus pattern than CI, since it can be halved in amplitude by miscorrection of about \( \frac{1}{4} \text{D} \) (Spekreijse 1980). Blurring the stimulus also increases the latency of CII, a blur of +5D delaying it by as much as 72ms for a check size of 14' (Harding and Wright 1986). Hence, CII appears to be generated by contour-specific neurones, such as hypercomplex cells (James and Jeffreys 1975).

The most commonly used pattern onset stimulus is a checkerboard of square checks. In this case, the check size influences the latency, amplitude and form of the recorded potentials. Small checks of side 5' favour a large offset response while producing only a small onset response. Checks of around 15' produce a large onset response as well as a large offset response, and large checks of about 80' elicit mainly onset components, but these are reduced in amplitude and less well developed in form than for small checks (figure 69; Spekreijse et al. 1973). The latency of CI decreases from around 80ms to 71ms as check size is increased from 13' to 56', while the latency of CII decreases from about 120ms to 100ms for the same increase in check size. Large checks can also decrease the latency of CIII by up to about 10ms (Wright et al. 1985).

Figure 69. Effect of check size on the pattern onset visual evoked potential. The responses were recorded from the midline above the inion, relative to a vertex reference. Positive upwards. After Spekreijse et al. (1973).
For a given spatial frequency, the pattern onset evoked potential is influenced by the luminance and contrast of the presented pattern (Spekreijse et al. 1973). For a given luminance, increasing the contrast has the effect of increasing the amplitude of the response, until a saturation level is reached, beyond which further contrast increase has no effect. The contrast saturation level is dependent on such properties as pattern type and check size and most notably on luminance. If the luminance of the stimulus is high, then the contrast may saturate at as low as 10%. Contrast also alters the form of the evoked potential (Spekreijse et al. 1973).

The Pattern Onset Visual Evoked Magnetic Response

One of the earliest studies of the evoked magnetic response to the abrupt onset of a patterned stimulus was that of Brenner et al. (1975). They used a vertical bar grating which was presented for 50ms and then turned off for 50ms to produce a 10Hz flicker. A 10Hz steady-state response was recorded, having maximal amplitude on the midline between 50 and 80mm above the inion. The lateral distribution of the evoked magnetic field was measured 76mm above the inion. Moving towards the left side of the head, the response remained almost constant in amplitude until about 60mm left of the midline, when it abruptly reversed in polarity. From this, it was evident that the response consisted of an outwardly directed magnetic field near the midline and an inwardly directed field at the left side of the head.

Degg et al. (1992) studied the transient pattern onset response. A 7° × 6° achromatic stimulus of check size 38′ × 27′ was used to obtain full-field and half-field responses at 20 points around the occipital surface of the head. The waveforms consisted of three peaks which were analogous to the CI, CII and CIII peaks of the corresponding visual evoked potential, and were therefore termed C1m, CIIm and CIIIm (figure 70). For the stimulus used, the CIIIm peak was much greater in amplitude than either C1m or CIIIm. A left half-field CIIIm response consisted of an outward-going field in the upper part of the montage, some 60-80mm above the inion, and an inward-going field at the height of the inion. This topography was consistent with the hypothesis that the field was generated by a medial current dipole situated in the hemisphere contralateral to the stimulated half-field and directed contralaterally. A right half-field CIIIm response was the reverse of the left half-field response, consisting of an ingoing field about 60mm above the inion, and an outgoing field around the inion itself. This was also consistent with a contralaterally-directed medial dipole. The full-field response was somewhat variable between subjects, but in each subject, was approximately equal to the sum of the left and right half-field responses. In a subsequent study (Degg et al. 1993), it was shown that the precise morphology, latency, amplitude and spatial topography of the CIIIm peak was dependent upon the exact latency and duration of the two half-field responses. At some latencies, the two half-
Figure 70. Example of a binocular pattern onset visual evoked magnetic response from a single subject. The stimulus was a 7° × 6° checkerboard having 38′ × 27′ checks of contrast 66%, and the response was recorded from 60mm above the inion. The onset peaks, C1m, CIIIm and CIIIIm, and the offset peak are shown. Outgoing field upwards. After Degg et al. (1992).

Field responses tended to cancel, and at others, they summated, resulting in a full-field CIIIm whose peak latency was different to that of either half-field CIIIm, and which therefore reflected the current source of neither.

The Pattern Reversal Visual Evoked Potential

A pattern reversal stimulus is a contrast-modulated pattern containing bright and dark regions whose luminances alternate in counterphase. In the case of an achromatic checkerboard - the most common pattern reversal stimulus - the effect is that the black checks periodically exchange position with the white checks, giving an illusion of motion.

The evoked response to the abrupt reversal of a checkerboard more closely resembles a pattern offset response than a pattern onset response (figure 71; Kriss and Halliday 1980). An occipitally recorded waveform consists of negative peaks at latencies of around 75 and 135ms, and positive peaks at latencies of about 60, 100 and 180ms (Halliday 1982). The negative peaks are denoted N75 and N135, while the positive peaks are denoted P60, P100 and P180. However, some of these peaks are insignificant in amplitude and are often not observed, even in normal subjects. Peaks N75, P100 and N135 generally form a triphasic complex, although even amongst these more prominent peaks, N75 and N135 can be ill-defined. For a screen size of 8.7° and a check size of 25.8′, P100 has a mean latency of 102.3ms and mean amplitude of 10.1μV (Shahrokhi et al. 1978), although these parameters vary depending on the precise characteristics of the stimulus.
Figure 71. Example of a monocular pattern reversal visual evoked potential from a single subject, showing the major positive peak, P100, and the negative peaks, N75 and N135. The response was recorded from Oz, relative to Cz. Positive downwards. After Chiappa (1990).

The response to full-field stimulation shows maximum deflection at a mid-occipital electrode when a mid-frontal reference is used. The response spreads laterally however, and can also be recorded, with considerably reduced amplitude, from 50-100mm either side of the midline. The potential is usually symmetrically distributed between the left and right sides (Blumhardt and Halliday 1979).

A left or right half-field stimulus subtending a large visual angle (e.g. 16°) yields a triphasic NPN complex over the midline, but the response is markedly asymmetric (Halliday 1982). Recording from a transverse electrode montage extending from 100mm to the left of the midline to 100mm to the right, it is found that the NPN complex extends from the midline to the side of the head ipsilateral to the stimulated half-field, but there is a smaller triphasic complex of the same latency but opposite polarity on the contralateral side. This PNP complex has maximal amplitude at the most lateral electrode, situated 100mm out from the midline. At an electrode 50mm to the contralateral side, the response is transitional between the NPN and PNP waveforms (Blumhardt and Halliday 1979). For large field sizes, it can be shown that the full-field response is merely a summation of the left and right half-field responses, and that the central maximum deflection of the full-field response arises from the summation of the two half-field NPN complexes. The rapid lateral decrement which occurs with a full-field stimulus can be explained by the contralateral PNP complex of one half-field response cancelling the ipsilateral NPN complex of the other (Blumhardt et al. 1977).

Historically, the large deflection observed ipsilateral to the stimulated half-field when using a large stimulus was paradoxical (Barrett et al. 1976). From knowledge of the retinotopic representation of the visual field in striate cortex, it was believed that the evoked response was generated in cortex contralateral to the stimulated half-field, and that the P100 should therefore be maximal in amplitude contralaterally. Instead, the response was observed ipsilaterally, and contralateral electrodes revealed only the low-
amplitude PNP complex. The paradox was resolved when it was suggested that the
P100 was generated by current dipoles on the medial and posteromedial surfaces of the
visual cortex contralateral to the stimulated half-field; the dipoles being oriented
perpendicular to these surfaces, with current flowing towards the pial membrane. Such
an arrangement of current sources would produce a positivity on the ipsilateral side, an
absence of response just to the contralateral side of the midline, and a weak negativity
contralaterally, as observed experimentally (Barrett et al. 1976). This was subsequently
proven by Clement et al. (1985) using a source derivation procedure based on the
Laplacian of the evoked potential distribution. Harding et al. (1980) found that if the
stimulus size was reduced in steps from 14° to 2.5°, the location of the largest NPN
complex shifted from ipsilateral to slightly contralateral. This was because the smaller
stimuli preferentially activated the cortex representing the central visual field, located
posteromedially and on the occipital convexity itself. Consequently, as the stimulus size
was reduced, the current dipole associated with the cortical response shifted posteriorly
and rotated so that its moment was directed dorsally rather than medially, with the result
that the NPN complex shifted contralaterally (Barrett et al. 1976; Harding et al. 1980;
Halliday 1982).

The summation effect, whereby the two half-field responses sum to equal the
full-field response, also occurs within each respective half-field. By introducing a
central scotoma into a half-field stimulus, and varying the radius of the scotoma, the
distributions of the ipsilateral NPN and contralateral PNP complexes are altered
significantly (Blumhardt et al. 1978), indicating that the half-field response is a sum of
central and peripheral responses whose individual lateralisations are different to that of
the overall half-field response. As the radius of the scotoma is increased, the amplitude
of the ipsilateral P100 peak falls linearly, suggesting that both central and peripheral
responses contribute equally to this wave (Blumhardt et al. 1978; Haimovic and Pedley
1982). However, peripheral vision seems to contribute more than central vision to the
contralateral negative peak at 100ms, since this peak is initially enhanced by a central
scotoma, while it is attenuated, without attenuation of the ipsilateral P100 wave, by
occlusion of the paramacular portion of the half-field stimulus (Blumhardt et al. 1978;
Haimovic and Pedley 1982).

Asselman et al. (1975) found that P100 amplitudes were reduced by 50% when
the central 5° of an 18° field were occluded, and reduced by 80% when the central 10°
were occluded. Yiannikas and Walsh (1983) found that at least 80% of the P100
response arose from the central 8°. However, the peripheral 8 to 32° also produced a
significant contribution to the amplitude of the P100 peak. A plot of the percentage of
the maximum P100 amplitude generated by each degree of the visual field showed that
the central 4° of the stimulus produced by far the greatest proportion of the response.
Furthermore, the plot closely resembled a plot of cortical magnification over the same
visual angles, illustrating a direct correspondence between the area of cortex associated
with each degree of visual field and the corresponding contribution of that area to the evoked response.

As with the pattern onset visual evoked potential, the pattern reversal response is influenced significantly by the temporal properties of the stimulus. For an interstimulus interval of greater than 200ms, the response is transient, as described above. However, if the stimulus reversal rate is increased to 8 or 10Hz, the response becomes a continuous train of waves having a sinusoidal appearance (Chiappa 1990). This is a steady-state evoked potential. Each peak of the waveform occurs at some latency after the stimulus that produces it. This latency or phase lag varies with stimulus frequency and is not equal to the latency of the transient evoked potential (Milner et al. 1974). Moreover, if the phase lag of the response exceeds 180°, then it is interpreted as a much smaller lag; for example, a phase lag of 190° is interpreted as a lag of 10°: it becomes difficult to interpret which wave is produced by each stimulus. The amplitude of the steady-state evoked potential can be used to produce plots of amplitude versus stimulation frequency (Milner et al. 1974; Regan et al. 1976; Spekreijse et al. 1978). However, due to the relative difficulty of using steady-state evoked potentials, transient evoked potentials are much more common and our attention is confined to them hereafter.

The check size of the stimulus influences the amplitude of the transient pattern reversal response, small checks producing a slightly larger response than large checks. Asselman et al. (1975) used a checkerboard stimulus of 18° diameter and compared 30 and 57' checks, obtaining larger amplitudes with the smaller checks. Similarly, Erwin (1981) found that the positive peak had highest amplitude when the checks subtended an angle of between 15 and 30'. Meredith and Celesia (1982) studied different stimulus and check sizes in varying parts of the visual field, finding that small stimuli and small checks were optimal at the fovea, while larger stimuli and checks were optimal at the periphery. Yiannikas and Walsh (1983) found that the response was maximal in amplitude with 27' checks when the central 4° of the field were stimulated. When greater areas of the visual field were stimulated, up to a maximum of 32°, 55' checks produced the response of largest amplitude. Using a high-contrast 15° screen, Chiappa (Chiappa et al. 1985; Chiappa 1990) found that a check size of 17' produced a mean P100 amplitude of 8.34μV, with a standard deviation of 3.71μV (60 subjects), whereas 70' checks produced a mean P100 amplitude of 7.95μV, with a standard deviation of 3.64μV (58 subjects).

Both stimulus size and check size alter the latency of the P100. In the study of Yiannikas and Walsh (1983), the peak latency of the P100 was around 97-98ms for all check sizes when a 32° stimulus was used. When the stimulus size was reduced to 2°, an increase in latency was observed, the increase being dependent on the check size. Large checks of 110' caused the latency to rise to around 120ms, whereas small checks of 27' caused it to rise to only about 102ms. Thus, for a given field size, increasing the check size has the effect of slightly increasing the latency of the P100.
The mean luminance of a pattern reversal stimulus has a marked effect on the latency and amplitude of the response. The P100 latency increases as pattern luminance is decreased (Cant et al. 1978; Sokol 1980); the latency increase is approximately 15ms per log unit reduction in luminance for a stimulus diameter of 32°, check size of 50', and initial intensity of 60cd (Halliday et al. 1973). For example, reducing the luminance by a factor of 100 (two log units) increases the P100 latency by about 30ms. There is also approximately a 15% reduction in P100 amplitude per log unit intensity reduction (Halliday et al. 1973).

Reducing the contrast of an achromatic checkerboard pattern also increases the latency and reduces the amplitude of the P100 (Wright and Johnston 1982). However, this effect saturates at contrasts of about 20 to 40% (Halliday and McDonald 1981). In the majority of cases, contrast is much higher than this, so a contrast increment or decrement produces little change in P100 latency and amplitude. Hence, provided a contrast of at least 60-70% is used, the effect of contrast variation can be neglected.

It should be noted that the type of stimulator can affect the latency of the pattern reversal response. This is because the time to reverse the pattern has a definite influence on the P100 latency. Specifically, P100 latency increases linearly by 0.6ms per millisecond as reversal time increases from 3 to 50ms (Trojaborg and Petersen 1979). Several methods of stimulation are in common use; these are (1) back-projection onto a translucent screen via a rotating mirror, (2) use of a video monitor, and (3) using an array of light emitting diodes, either naked, behind coloured squares, or projecting into optical fibres. Mirror or light emitting diode stimuli usually reverse faster than video stimuli, as a video has a refresh rate of 50-100Hz and therefore requires 10-20ms to completely reverse the pattern, the top left of the screen being reversed immediately after triggering and the bottom right being reversed last. Hence, a video stimulus produces a response with a longer latency than that of the response from a light emitting diode device (Stockard et al. 1979). The differences in the times at which the various parts of a video screen are reversed can cause latency differences between the responses to different parts of the screen. Although a mirror device moves the pattern rather than truly reversing the intensities of the elements, this has been shown to have no consistent effect on the resulting response (Lowitzsch et al. 1980). This is presumably because the angular velocity of the moving pattern is so high that it produces an on-off response (Chiappa 1990).

Numerous patient factors, such as age, gender and visual acuity, affect the latency of the pattern reversal response. The precise variation of P100 latency with age is controversial but it is clear that (1) there is a latency increase with increasing age, and (2) smaller check sizes produce greater differences in latency between young and old subjects (Sokol et al. 1981; Wright et al. 1985; Celesia et al. 1987). For example, Wright et al. (1985) found using 13' checks that the latency of P100 was 107ms in 10-19 year olds, rising to 118ms in 70-79 year olds. With 56' checks, P100 latency rose only from 109ms to 111ms.
The effect of gender on the P100 is also controversial, but females may have a slightly shorter P100 latency than males. For example, Stockard et al. (1979) studied 100 normal age-matched male and female subjects of 13-67 years and found the female mean P100 latency to be 2.7ms shorter than that of the males. However, Shearer and Dustman (1980) and La Marche et al. (1986) found no gender differences. Mean P100 amplitude is usually greater in females than in males (Allison et al. 1984; La Marche et al. 1986; Celesia et al. 1987).

A loss in visual acuity causes an increase in P100 latency which depends on the check size. For instance, Collins et al. (1979) used 12' checks and found that reducing visual acuity to 20/200 with a +2D lens caused a 20ms increase in P100 latency, whereas Halliday et al. (1973) used 50' checks and found no latency change with visual acuity in patients recovering from optic neuritis. Sokol and Moskowitz (1981) found that P100 latency was shortest when the stimulus pattern was least blurred, and that the latency increased more when 12' checks were defocused than when 48' checks were defocused. This was confirmed by Harding and Wright (1986), who found that a +3D blur introduced a mean delay of 5ms into the P100 response to 56' checks, whereas the same blur produced a larger 20ms delay when 13' checks were used.

The Pattern Reversal Visual Evoked Magnetic Response

Many of the early studies of the pattern reversal evoked magnetic response concentrated on the steady-state response obtained by the use of a high reversal rate. For example, Brenner et al. (1981) used a vertical sinusoidal grating stimulus with a spatial frequency of 5 cycles/degree and a reversal rate of 13Hz (fundamental frequency 6.5Hz). Accordingly, a 13Hz steady-state response was observed. Stimulation of the left half-field produced a pair of field extrema about 20mm to the right of the midline, the one extremum being at the level of the inion, with the other being about 60mm above it and having the opposite polarity. These field extrema both oscillated approximately sinusoidally, retaining their 180° phase difference. The field pattern could be explained by a horizontal current dipole in the right hemisphere. A symmetrical result was obtained by stimulating the right half-field: the field extrema were located to the left of the midline and could be accounted for in terms of a horizontal dipole in the left hemisphere. To a first approximation, the full-field response represented the sum of the left and right half-field responses.

Maclin et al. (1983) carried out a similar experiment to show that stimulating the peripheral part of the visual field elicited a response deeper below the scalp than did stimulating the macular region. The sinusoidal grating was presented to the central 1.6° of the right half-field, and to the semi-annular regions extending from 1.6 to 3.3° and from 3.3 to 5.0° in the right half-field. Each stimulus yielded a pair of extrema in the left hemisphere, with the one extremum located approximately vertically above the other.
The two extrema were located symmetrically about a point some 20mm above the level of the inion and 20-30mm to the left of it. However, the separation of the extrema increased as the stimulus became more peripheral. For each stimulus, the current source was assumed to be a dipole in a spherical head, and the depth of the dipole below the scalp was 12mm for the central stimulus, 22mm for the paramacular stimulus, and 33mm for the peripheral stimulus. Hence, provided that the current dipole was a reasonable source model, this experiment indicated that the peripheral retina projected to a deeper region of the brain, in accordance with models of cortical projection (Drasdo 1980).

The steady-state pattern reversal evoked magnetic response was also studied extensively by Okada et al. (1982). A sinusoidal grating of mean luminance 10cdm\(^{-2}\) was presented to the right half-field. The amplitude of one of the field extrema was measured as a function of the spatial frequency of the stimulus. It was found that the amplitude of the evoked magnetic response was largest for a grating of 1-4 cycles/degree, with spatial frequencies of 0.5-1.0 and 4-6 cycles/degree producing a smaller response. However, the spatial frequency for which the response was maximal shifted from 3 cycles/degree at a reversal rate of 7Hz to 1 cycle/degree at a reversal rate of 23Hz, suggesting that the response was tuned to higher spatial frequencies at low temporal frequencies and vice versa (Okada 1983b).

Williamson et al. (1978) found that the spatial frequency of a sinusoidal grating also influenced the latency of the steady-state response. When the spatial frequency was increased from 0.2 to 10 cycles/degree, the latency increased from about 80 to 200ms. This was in accord with previous psychophysical data (Breitmeyer 1975) which showed that the reaction time to brief onset of a sinusoidal grating also increased in approximately the same manner.

The temporal frequency of the stimulus has also been shown to affect the phase lag of the steady-state pattern reversal response. Williamson et al. (1978) showed that phase lag increased linearly with reversal rate for rates of 5-20Hz. A greater phase lag was introduced by a stimulus of 8 cycles/degree than by a stimulus of 1 cycle/degree. A comparison of the steady-state evoked potential and evoked magnetic field (Okada et al. 1982) showed that in general, the phase lags of both modalities were equal for a given set of stimulus parameters.

The effect of stimulus luminance on the pattern reversal visual evoked magnetic response was also measured by Okada (1983b). A sinusoidal grating of 1 cycle/degree was presented to the right half-field at a reversal rate of 13Hz and mean luminances of 50, 5 and 0.5cdm\(^{-2}\). For a given contrast level of 10%, the amplitude of the response at one of the spatial field extrema was greater for the 50cdm\(^{-2}\) stimulus than for the 5cdm\(^{-2}\) stimulus, which, in turn, produced a larger response than the 0.5cdm\(^{-2}\) stimulus. Hence, the stimulus luminance appeared to influence the amplitude of the steady-state response.

Similarly, Okada et al. (1982) found that contrast influenced the steady-state response. Increasing the contrast from 0.5% upwards caused the evoked response to rise above the noise level then to increase in amplitude until a saturation level was
reached. For a reversal rate of 13Hz, the amplitude rose linearly with the logarithm of the contrast until the saturation point. This finding was consistent with that of Halliday and McDonald (1981) for the case of the evoked potential, discussed earlier.

The above studies, using high-frequency pattern reversal stimulation, served to show the direct correspondence in many cases between the electric and magnetic responses. The reader's attention is now directed to the low-frequency or transient pattern reversal evoked magnetic response. This was studied by Armstrong et al. (1991a) using a checkerboard of size 16° × 14° and a check size of 70°, with the checks reversing at a rate of 2Hz. The transient response was found to consist of a large outgoing field over the midline between 90 and 120ms, which, by analogy with the major positive component in the pattern reversal evoked potential, was denoted P100m. This peak was sometimes preceded and followed by weak ingoing fields of latency 75-96ms and 125-155ms respectively (figure 72).

Harding et al. (1991) studied the spatial topography of the P100m response. It was found that the response to half-field stimulation was generally confined to the side of the head contralateral to the half-field stimulated, although some ipsilateral activity was also observed. The P100m had maximal amplitude when large 70° checks were used. A left half-field response consisted of an ingoing field in the upper part of the recording montage, and an outgoing field at around the elevation of the inion. A right half-field response was the opposite of this, consisting of an outgoing field above an ingoing field. These responses were consistent with a current source in the contralateral hemisphere, directed medially. For large 70° checks, the full-field response was approximately equal to the sum of the left and right half-field responses. However, check sizes of 34 and 22° produced full-field responses which were considerably variable between subjects and which could not be explained in terms of summated half-field responses.

![Figure 72. Example of a binocular pattern reversal visual evoked magnetic response from a single subject, showing the P100m peak. The stimulus was a 16° × 14° checkerboard, having 70° checks of contrast 76%, and the response was recorded from 60mm above the inion. Outgoing field upwards. After Armstrong et al. (1991a).]
The P100m peak of the pattern reversal response was found by Armstrong et al. (1991a,b) to show greater age-dependence than the corresponding P100 peak of the evoked potential. The latency of the P100m was found to increase from around 105ms in an age-group of 10-20 years to around 108ms in an age-group of 50-60 years, thereafter increasing more rapidly to around 153ms in an age-group of 80-90 years (Armstrong et al. 1991a). This was in contrast to the P100 peak of the pattern reversal evoked potential, which rose in latency only from about 106ms to about 109ms over an age range of 10 years to 80 years (Wright et al. 1985). Meanwhile, the amplitude of the P100m peak in the magnetic response dropped steadily with age, from about 300fT in subjects of 10-20 years, to about 150fT in subjects of 80-90 years (Armstrong et al. 1991b). These changes could have been the result of ageing processes in the crystalline lens, the retina, optic nerve or visual cortex, or a combination of all of these.

Source Analyses

The first major attempt to deduce the cortical origin of the evoked response to patterned stimulation was probably that of Jeffreys and Axford (1972a,b). These workers carried out a qualitative analysis of the pattern onset evoked potential using the cruciform model of primary visual cortex. (For a description of the model, see figure 5 of chapter 2.) They reasoned that if the cortical response were confined to striate cortex, then each stimulated octant of the visual field should produce a response in the appropriate part of the cruciform model. It was assumed that the cortical response was dipolar, with the dipole oriented perpendicular to the surface of the active cortex. Under these conditions, it was reasoned, a left half-field stimulus should produce a response in the right half of the cruciform model. This response would consist of four current dipoles, two of which were normal to the right medial surfaces and two of which were normal to the surfaces of the right calcarine fissure (figure 73). Moreover, if all regions of striate cortex contained primary current which flowed either towards or away from the cortical surface, then the two dipoles on the right calcarine fissure would cancel, leaving only the medial sources to generate the external electric potentials. Together, the medial dipoles would constitute one single equivalent current dipole, located on the midline; this would be directed towards or away from the left hemisphere, depending on whether the cortical current flowed towards or away from the pial surface. By the same argument, it was reasoned that a right half-field stimulus might produce a medial current dipole which was directed in the opposite direction, while upper and lower half-field stimuli might produce medial dipoles which were directed superiorly or inferiorly. In all cases, the theoretical distribution of external potentials could be quantitatively deduced from observations of a current dipole within a simple spherical conductivity model (Shaw and Roth 1955). Thus, Jeffreys and Axford predicted what
the distribution of the evoked potentials would be if the cortical response were striate in origin.

The CI component of the pattern onset visual evoked potential was found to exactly match the scalp distributions that were predicted for the case of primary current directed towards the white matter. Furthermore, quadrant and octant stimuli produced potential distributions which were in accord with those predicted by the cruciform model. Hence, it was concluded by Jeffreys and Axford (1972a) that CI was generated by striate cortex. However, the CII component did not have the same scalp topography as CI, and could not be explained using the cruciform model. From this, it was deduced that CII was generated by a different cortical region to CI, and that it must therefore be extrastriate in origin (Jeffreys and Axford 1972a).

In a further study (Jeffreys and Axford 1972b), it was hypothesised that the CII component originated from the ventral and dorsal surfaces of visual cortex. It was assumed that extrastriate visual cortex was retinotopically mapped, so that the upper half-field was represented on the ventral surface, and the lower half-field was represented on the dorsal surface. (From figure 4, it can be seen that these regions correspond approximately to Brodmann area 19, possibly including area 18.) As for CI, predicted distributions of external potential were calculated for the case of primary current flowing towards the white matter, and in the majority of cases, the experimentally recorded potentials were in agreement with these. The scalp distribution of CII was thus consistent with the peak having an extrastriate origin.

Although the experiments of Jeffreys and Axford were undoubtedly valuable, several authors seemed to disagree with their conclusions. For example, Lesèvre and Joseph (1979) carried out a similar study of pattern onset evoked potentials and made rather different deductions. They observed an early negative wave of 60ms latency - the CO or N60 wave - which persisted even when the fovea was selectively stimulated.
Because of this persistence, the wave fitted the hypothesis that it was generated on the convexity of the occipital lobe in the region of area 17 representing the fovea. Moreover, this wave failed to change in polarity or topography when the stimulated half-field was changed, which was also consistent with such a hypothesis.

Lesèvre and Joseph (1979) recorded a similar CI response to Jeffreys and Axford (1972a,b). However, they pointed out that while the response could be interpreted in terms of the cruciform model, it could also be explained by retinotopy within area 19. An origin within area 19 was, in fact, more likely, as the dipoles resulting from upper and lower half-field stimulation were too far apart to be contained within striate cortex. Furthermore, the recorded data showed that the response to lower half-field stimulation contained well-differentiated peaks, but the response to upper half-field stimulation did not, a finding which was consistent with the lower half-field source being closer to the scalp than the upper half-field source. This also fitted the suggestion that CI was generated in area 19.

The scalp topography of the CII component did not change in orientation between left and right half-field stimulation, or between upper and lower half-field stimulation, which ruled out Jeffreys and Axford’s interpretation of CII in terms of area 19. The absence of phase reversal when changing from upper to lower half-field stimulation also ruled out the possibility of a response within the calcarine fissure. It was suggested that area 18 was the most likely generator of CII and CIII.

Richer et al. (1983) recorded visual evoked magnetic responses to pattern onset stimuli. The two most significant components appeared at latencies of 120 and 180 ms, corresponding to the CIIIm and CIIIIm. Both components reversed polarity as a function of the stimulated half-field. Simple equivalent current dipole fitting was performed by locating a dipole midway between the two field extrema and estimating its depth from their separation. This showed that both left and right half-field responses were produced by a medial current dipole which was directed contraterally to the stimulated half-field. The source appeared to be common to both the CIIIm and CIIIIm components. However, no conclusions were made as to the visual area responsible for the current source.

The above studies all used simple, fairly qualitative methods of determining the cortical sources. A far more sophisticated study was that of Maier et al. (1987), which used principal components analysis to determine the locations and timecourses of the current sources responsible for visual evoked potentials elicited by various stimuli. The study implicitly assumed that throughout the duration of the evoked response, there were a number of simultaneously active dipolar sources of fixed location and orientation but variable timecourse. This approach explored the possibility that more than one source could contribute to a given waveform peak. An isotropic three-sphere model was used to model the conductivities of the brain, skull and scalp. Further details of the principal components analysis method are given in chapter four.
A pattern onset response was obtained by presenting a central or peripheral half-field checkerboard of contrast 20% for 300-500ms each second. The central stimulus consisted of a semicircular pattern whose radius subtended 2°, while the peripheral stimulus was a half-annulus whose radius extended from 2 to 4°. Each of these stimuli yielded two principal components. The first component had a potential distribution which was similar to that of the single component produced by a flicker stimulus, which was known to originate in striate cortex. Hence, the first principal component was attributed to striate cortex. The timecourse of this component consisted of a single negative peak, occurring between 120 and 150ms. Meanwhile, the second principal component had a timecourse which consisted of a positive peak at 80-90ms, followed by a negative peak which occurred 30ms later than the peak of the first component. The maximum of the potential distribution for the second principal component shifted upward by about 30mm when changing from central to peripheral stimulation. This component was attributed to a current dipole in area 18 or 19, its reversal in polarity signifying a reversal in the direction of the current dipole. From the timecourses of the two principal components, it was deduced that CI was generated by an extrastriate current source, while CII was both striate and extrastriate in origin, with the current flow responsible for CI reversing its direction and contributing to the CII-CIII complex.

Pattern offset also produced a similar cortical response. Two principal components were again observed, the dipoles for which were located similarly to those for pattern onset. The temporal variations of the components were also comparable to those for pattern onset, but reversed in polarity. The main difference between pattern onset and pattern offset responses was that with a pattern offset stimulus, the striate component had a larger amplitude than the extrastriate component.

A pattern reversal stimulus, of contrast 40% and alternating once every 300ms, also produced two principal components. These components had similar spatial distributions to those for pattern onset stimuli, with the striate component being larger in amplitude than the extrastriate component. Both components had a biphasic timecourse, the duration of which suggested that P100 could contain contributions from both striate and extrastriate cortex.

Stok (1986) came to similar conclusions. He found that visual evoked electric and magnetic responses to pattern onset could be described by two equivalent current dipoles which were active one after the other. The first dipole was located laterally, while the second was located near to the midline and/or above the first dipole. Between them, these two dipoles accounted for the response during the period from 60 to 150ms after stimulus presentation. However, the exact latencies at which they were active varied considerably between subjects. It was suggested that the first dipole was associated with area 18 or 19, while the second was located in area 17, approximately in agreement with the proposition of Maier et al. that CI was generated by an extrastriate current source and CII by a combination of both striate and extrastriate sources.
A number of researchers have stimulated quadrants and octants of the visual field in order to examine cortical retinotopy. These experiments are considered in detail in the next chapter, but their inclusion here is warranted because they help to show the origins of the various features of electric and magnetic responses to patterned stimulation. Darcey et al. (1980a,b) carried out an extensive study of pattern onset visual evoked potentials and performed equivalent current dipole analyses for half-field and quadrant stimulation. Analyses were carried out at a latency of 90ms, corresponding approximately to the latency of CI, and the equivalent current dipoles were found to be located in striate cortex. The dipole orientations were consistent with a cruciform model which was modified slightly to include the representation of the central visual field on posterolateral cortex. Upper and lower half-field responses could be explained using a two-dipole model, while the responses to left and right half-field and quadrant stimuli could be adequately described by a single current dipole; in both cases, the model gave an excellent fit to the measured data and had a fairly small confidence region. This study therefore indicated that the source of CI was located in primary visual cortex and could be modelled by one current dipole if the stimulus was restricted to one half-field only, and by two current dipoles if it stimulated both half-fields.

A very similar experiment was carried out by Butler et al. (1987). They stimulated the eight octants of the visual field using a pattern onset stimulus which excluded the central 4°. Half-field and quadrant stimuli were also used to stimulate the central 2° of the visual field. The evoked potentials were analysed using an equivalent current dipole model. It was found that the dipole locations were consistent with CI being generated by striate cortex, but incompatible with alternative arrangements of sources in extrastriate cortex. The dipole orientations indicated that primary current was directed from the pial surface towards the white matter. However, foveal stimuli elicited potentials which were positive at posterior electrodes and consistent with an outwardly-directed current dipole, a finding which could not be explained using a simple model of the occipital pole. Hence, the conclusions of this study were uncertain.

In contrast with the above experiments, which suggested that striate cortex was responsible for CI, Ossenblok and Spekreijse (1991) reported that both CI and CIII were generated by extrastriate cortex and therefore showed no correspondence with the cruciform model. Their stimulus used large 24’ or 48’ checks so as to minimise any edge-specific striate component in the response. The evoked potentials were analysed using principal components analysis and two components were found to account for the measured signals, the first dominating the CI peak and the second being maximal at the latency of CIII. The current dipoles associated with the components were both sufficiently lateral to be recognised as lying in extrastriate cortex. The dipole responsible for the first principal component responded to changes in retinal stimulus location (i.e. quadrant or octant stimulated) and was therefore deduced to lie within area 18. Conversely, the second dipole showed no retinotopic behaviour and was thus
concluded to lie in area 19, which does not have a single retinotopic representation of the visual field (see chapter two). Hence, this experiment conflicted with the conventional view of CI as being striate in origin, but was in accord with the previous study of Maier et al. (1987), which also deduced CI to be extrastriate.

Ahlfors et al. (1992) examined retinotopy using visual evoked magnetic fields. They stimulated the central 1.8° of the visual field using a foveal pattern onset octant stimulus, and the region from 1.8 to 3.8° using an annular octant stimulus. Evoked magnetic responses were recorded using a 24-channel magnetometer, and the CI peak was analysed using two-dimensional minimum norm estimates and single equivalent current dipoles. Although this was a comprehensive study, the results did not show any particular trend, other than that cortical current was contralateral to the stimulated octant. No conclusions were made as to which visual area the magnetic responses originated from.

By now, it will doubtless be clear to the reader that the cortical origins of the pattern onset and pattern reversal visual evoked electric and magnetic responses are still very uncertain. In general, the evidence seems to suggest that the CI peak of the pattern onset evoked potential is generated in striate cortex, while CII is extrastriate in origin. However, experiments such as that of Lesèvre and Joseph (1979) and Maier et al. (1987) conflict with this opinion. Less attention has been paid to the pattern reversal response, probably because of its smaller amplitude, which makes it more difficult to analyse reliably. In view of the considerable controversy which exists with regard to the cortical origin of pattern onset and pattern reversal evoked electric and magnetic responses, the author has chosen to use pattern onset and pattern reversal magnetic responses for trials of the source analysis methods presented in the last two chapters.

Methods

Note None of the data collection described subsequently in this chapter has been performed by the author of this thesis. The recording procedure and evoked magnetic fields have been published previously (Armstrong et al. 1991a; Harding et al. 1991; Degg et al. 1992, 1993), but are included here for completeness. Waveform peak identification has also been carried out previously (Degg et al. 1993). Contour plots appearing later in this chapter are the responsibility of the author, but these are included merely so that the reader can view the data from which the source analyses are derived; they are not intended as an original presentation of the data.

Left half-field, full-field and right half-field pattern onset and pattern reversal stimuli were presented to ten normal subjects, aged 23 to 41 years. All subjects were without ophthalmological problems and had corrected visual acuities of better than 6/6.

Pattern onset stimuli were presented on a BioLogic VM4515 monitor at 1Hz using a Medelec OS5 stimulation controller. Upright rectangular checks of 66%
contrast were displayed for 200ms within a field whose width was 7°20' and whose height was 5°40'. Check size was 38' x 27' unless otherwise stated. At all times, the mean luminance of the screen was 74cdm⁻²; during half-field stimulation, one half of the screen remained uniform in luminance, and for the 800ms between each checkerboard presentation, the screen also remained uniformly illuminated.

Pattern reversal stimulation was achieved by back-projecting a checkerboard pattern onto a translucent screen and shifting it at a fundamental frequency of 1Hz (reversal rate 2Hz) using a rotating mirror. The screen had a circular field whose radius subtended 7° at the eye of the viewer. Check size was 70', as this had been shown to produce the largest pattern reversal response (Harding et al. 1991). The mean luminance of the screen was 1050cdm⁻² and contrast was 76%.

Evoked magnetic responses were recorded during daytime in a semi-darkened, unshielded room using a BTI 601 single-channel magnetometer. This device was a symmetric second-order gradiometer with baseline 50mm. The pick-up coil was located 12mm from the outer surface of the dewar tail, and the diameter of all the coils was 20mm. The system had an intrinsic noise level of 16.6fT/√Hz at 5Hz, and a total noise level, including environmental noise, of 30-40fT/√Hz. Responses were recorded from a 20-point occipital montage containing four rows of five magnetometer sites. The inion coincided with the middle point of the bottom row. The magnetometer sites were separated by 10% of the half-circumference of each subject's head, this separation being between 28 and 31mm in the ten subjects considered. A recording was made from each point of the montage in turn, with the magnetometer located approximately normally to the scalp. Recordings at a sample of locations were repeated to ensure that the response was consistent. A more closely spaced montage was also used in some instances, containing six rows of seven points, with a horizontal and vertical spacing of 20mm. The inion coincided with the middle point of the second row from the bottom. The magnetometer was operated at a sensitivity of 28fWb/V (89pT/V) and the signals were bandpass filtered between 0.3 and 30Hz, with a roll-off of 24dB/octave. The signals were also passed through a 50Hz comb filter to remove mains noise which was present in the unshielded environment. After attenuation by a factor of 100, 50 epochs from each magnetometer site were averaged on a BioLogic Traveler EEG machine incorporating artefact rejection. The output of the EEG machine (in μV) was converted using the calibration factor for the magnetometer into units of flux, i.e. atto-Weber (10⁻¹⁸Wb or aWb). Note that the gradiometer strictly measured a flux difference (units aWb), but conversion into mean flux density (units fT) could be achieved simply by dividing the flux difference by the area of the pickup coil. (Results are subsequently given in units of aWb, but for ease of comparison, certain figures are also given in fT.)

The CIIm peak of the pattern onset response and the P100m peak of the pattern reversal response were selected for source analyses, as they were the most significant features in the respective records, thus having the highest signal-to-noise ratio. Where a specific peak was to be analysed, the precise latency of the peak was determined using
the method of Lehmann and Skrandies (1980). This method used the total variance between signals at two successive time instants to define the stability of the response. The records were analysed at latencies where temporal signal variance was minimal, such latencies representing stable periods in the brain's activity. Group averages were used in some cases; these were found by averaging the response at the peak latency in each respective individual, rather than by simply averaging all responses at the same latency. This was to ensure that the same event was identified in all contributors to the group average response. (The above procedure was carried out by Harding et al. (1991) and Degg et al. (1992, 1993), rather than by the author.)

Source analyses were carried out (by the author) using equivalent current dipole fitting, minimum norm estimation and anatomically constrained distributed-source analysis, as described in chapters five and six. The occipital region of the scalp was assumed to be a sphere of radius 81mm, the centre of which defined the origin of coordinates. The x-axis then passed through the inion, the z-axis through the midline posterior to the vertex, and the y-axis through the right side of the head, mutually perpendicular to the x- and y-axes. The radii and conductivities of the various layers of the head were taken to be those of table 3.

It was necessary to define the locations and orientations of the magnetometer sites. This was done using an azimuth-elevation coordinate system in which a point in the recording montage was specified by (1) the shortest length of arc between the site and the half-circumference line, and (2) by the length of arc joining the site to the midline, such that the arc lay in a plane parallel to the half-circumference line (figure 74). This coordinate system complemented the method of locating the montage points on a subject's head, which was done by (1) locating a given row of the montage a certain distance up the midline from the inion, and (2) locating a given point a certain distance to one side of the midline at the appropriate elevation. The magnetometer was assumed to be oriented radially to the conductivity model.

The montage coordinate system did not coincide with the occipital spherical coordinate system used for the source analyses, being rotated from it. Locating the half-circumference line on a stereotaxic study (Talairach et al. 1967) showed that the two coordinate systems had a rotational discrepancy of 20° in the sagittal plane. Hence, a simple coordinate transformation was applied to the montage coordinates to convert them into the occipital spherical coordinate system, in which they could be used for source analysis (see figure 74).

Figure 74 (overleaf). Gross layout of the occipital spherical coordinate system in relation to principal anatomical features and the recording montage. Recording sites were defined by (1) their elevation above the inion and (2) their sideways displacement from the midline. N: nasion; I: inion; AC: anterior commissure; PC: posterior commissure; O: origin of coordinates; PO: parieto-occipital sulcus; C: calcarine sulcus. Derived from Talairach et al. (1967).
The source space for the equivalent current dipole fitting procedure was simply the innermost sphere of the conductivity model. For the anatomically constrained analyses, the source space of table 3 and figure 28 was used. The source region used for minimum norm estimation consisted of five truncated spherical shells, at radii of 25, 35, 45, 55 and 65mm. Each spherical shell contained a $10 \times 10$ grid of discrete dipoles and extended vertically from $\theta = 20/65$ radians to $\theta = 110/65$ radians at intervals of $10/65$ radians, and horizontally from $\phi = -45/65$ radians to $\phi = +45/65$ radians, also at intervals of $10/65$ radians. In other words, the grid points were spaced 10mm apart in the outermost shell of the source space, and correspondingly closer together in the less eccentric shells.

When displaying a minimum norm estimate, the current dipole moment at a given grid point was divided by the volume of its associated voxel. This was because the estimation method described in the last chapter intrinsically returned a measure of dipole moment, whereas cortical activity was proportional to current density. For an example of the difference between these two quantities, assume that there is a certain level of activity in the cortex, but at two depths. The deep dipoles in the source space register a lower dipole moment than the superficial dipoles, because they are more closely spaced and therefore represent a smaller region of cortex each. Hence, the estimates were scaled so as to represent current density. Note that the volume of one voxel was $1000\text{mm}^3$ at the most superficial layer of the source space, decreasing as $r^2$ in the less eccentric layers.

All methods required the standard deviation of the noise to be estimated. For both pattern onset and pattern reversal responses, this was done by examining recorded waveforms at latencies remote from evoked responses. It was assumed that the largest noise peaks lay two standard deviations from the mean signal level, so the peak-to-peak noise level was divided by four to give a measure of the standard deviation. The value obtained was about 40fT (13aWb). This procedure was repeated using records obtained without a subject under the magnetometer and similar results were produced. In the case of a group average response from $n$ subjects, the standard deviation of the noise was assumed to be $13/\sqrt{n}$ aWb.

Model goodness-of-fit was calculated using the chi-square and incomplete gamma function as described in chapter five. A goodness-of-fit of 10% or higher indicated that the source model was a very good fit to the measured data, a goodness-of-fit between 10% and 10-2% was deemed acceptable, and a value below 10-2% indicated a rather unlikely source model. These criteria were applied to all of the inverse methods.
Results

*Group Average Pattern Onset Responses*  Figure 75 is a sample of the magnetic response from a single subject, produced by pattern onset stimulation. The most prominent peak, the CIIm, has a latency of 117ms. The spatial distributions of the group average CIIm responses to left half-field, full-field and right half-field stimulation are shown in the upper half of figure 76. In this figure and those to follow, positivity denotes an outgoing field, negativity an ingoing field. The left half-field response shows an outgoing field in the upper part of the montage, over the midline, and a broad ingoing field at the level of the inion. The right half-field response consists of a strong ingoing field some 60mm up from the inion and 30mm to the left of the midline. The right side of the montage contains a very weak outgoing field. The full-field response is much weaker than either the left or right half-field response, but consists of a slight outgoing field in the upper right part of the montage and an ingoing field of comparable magnitude in the bottom right area of the montage. The low magnitude of the full-field response is most probably because it is produced by both left and right half-field mechanisms simultaneously, whose respective magnetic fields cancel each other.

Figure 75. Left half-field pattern onset visual evoked magnetic response recorded from a single subject. CIIm is seen clearly on waves 8, 9 and 10. The montage is numbered from left to right; uppermost row first. Outgoing field upwards.
Figure 76. Group average left half-field (LHF), full-field (FF) and right half-field (RHF) magnetic responses to (above) pattern onset and (below) pattern reversal stimulation. Number of subjects: pattern onset left half-field: 5; full-field: 10; right half-field: 10; pattern reversal left half-field: 5; full-field: 8; right half-field: 5. Contours are marked in aWb.
Group Average Pattern Reversal Responses Figure 77 is a sample of the pattern reversal response from a single subject, showing the P100m peak at a latency of around 110ms. The group average P100m responses to left half-field, full-field and right half-field stimulation are given in the lower half of figure 76. The left half-field response shows an ingoing field about 60-90mm above the inion and slightly to the right of the midline, with a broad outgoing field at the level of the inion. The right half-field response shows an outgoing field about 60mm above the inion, almost over the midline. The full-field response consists of an outgoing field in the upper left of the montage, with weak ingoing fields in the bottom left and top right of the montage. Again, the low amplitude of the full-field response is probably because it is produced by the same neural populations as those responsible for the left and right half-field responses, and the fields of these neural populations act so as to cancel each other. Note that the pattern reversal responses have similar distributions - but opposite polarity - to the pattern onset responses, with the exception of the full-field response, which, due to its low amplitude, is probably influenced by noise (standard deviation approximately 5aWb).

![Graph](image)

Figure 77. Right half-field pattern reversal visual evoked magnetic response recorded from a single subject. P100m is seen clearly on waves 8, 9 and 10. The montage is numbered from left to right; uppermost row first. Outgoing field upwards. After Harding et al. (1991).
Source Analyses of the Group Average Pattern Onset Responses  Equivalent current dipole fits for the left half-field, full-field and right half-field CIIIm responses are given in figure 78. The reader should refer to figure 28 for the relationship between the brain and the coordinate axes. The left half-field analysis shows a dipole situated on the medial surface of the right hemisphere near to the calcarine fissure, directed contralaterally. This is compatible with the spatial distribution of the left half-field response (figure 76). The equivalent current dipole for the right half-field response is located in the left hemisphere, and is also directed contralaterally to the stimulated half-field. Its z-coordinate is some 6mm less than that of the dipole for left half-field stimulation, but the two dipoles are otherwise symmetrical. The full-field current dipole is much lower in magnitude than either the left or right half-field current source, which is expected as the full-field response has low magnitude. Note that the primary current underling the full-field response is probably the union of the currents underlying the left and right half-field responses, and a large current dipole should therefore result. However, since the magnetic fields produced by the left and right half-field primary currents cancel, an insignificant dipole results: the inverse problem cannot determine current sources when they do not produce an external magnetic field.

The goodnesses-of-fit for the left half-field and full-field dipoles are acceptable, indicating that a single equivalent current dipole is an adequate current source model. However, the dipole for the right half-field response has a poor goodness-of-fit, showing that a single dipole is inadequate to describe the data: multiple sources must be present and an alternative source model, which is capable of reflecting this, must be used.

A minimum norm estimate of the right half-field group average pattern onset response is shown in figure 79. Each grid shows the current density in one of the layers of the source space, and in each case, the x-axis lies between the two middle grid points of the second row from the bottom. The grid points are spaced at 10mm intervals in the layer of 65mm radius, and are proportionately closer together in the less eccentric layers. Note from figure 28 and table 3 that the calcarine fissure lies some 15-20mm above the x-axis. Bearing this in mind, it can be seen from figure 79 that the right half-field response is generated by a region of current density in the left hemisphere, near to the calcarine fissure, oriented towards the left of the head. Some downwardly-oriented current density also appears in the centre of the grids for radii 55 and 65mm. This additional primary current, constituting a second current source, may explain why the single dipole model was inadequate to describe the evoked field. However, minimum norm estimates are known to suffer from spurious "whirls" of current which surround the true current source, and the medial current may be of this nature, so its significance cannot be relied upon. Nevertheless, the goodness-of-fit of the estimate is 97.2%, indicating that the current distribution corresponds closely to the external magnetic fields.

From figure 79, it can be seen that most of the primary current is confined to the superficial layers of the source space. This is a consequence of the magnetometer lead
Figure 78. Equivalent current dipoles for the left half-field (LHF), full-field (FF) and right half-field (RHF) group average pattern onset CIIm peak.
LHF: x = 39mm, y = 7mm, z = 13mm, q = 21.0nAm, goodness-of-fit = 6.8%.
FF: x = 36mm, y = 27mm, z = 23mm, q = 3.8nAm, goodness-of-fit = 46.4%.
RHF: x = 47mm, y = -17mm, z = 7mm, q = 15.2nAm, goodness-of-fit = 3.3\times10^{-9}%. 
fields being greater in magnitude at the surface of the spherical conductivity model than deep within it, as discussed in chapters four and five. Although some have used probability weighting to correct for the effect, there is no physiological justification for this, and, furthermore, no weighted minimum norm estimation technique can successfully image every single possible configuration of current, because of the non-uniqueness. In view of these difficulties, the author prefers not to use probability weighting, but to be aware of the limitations of the method. The resulting estimate is the most probable current distribution which is compatible with the data.

Minimum norm estimates for the left half-field, full-field and right half-field responses are shown in figure 80. In this figure and those following, only the most superficial layer of the source space is shown, as the other layers always show a similar pattern of primary current density, but with a lower magnitude. The estimates for left half-field, full-field and right half-field stimulation are similar to the single dipole results, with the exception that they are less focal than the single dipole fits (and that they are more superficial).

Finally, figure 81 shows the anatomically constrained distributed-source analyses for the group average CIIm responses. The left half-field analysis shows primary current in the dorsal, medial and ventral cortices of the right hemisphere, directed inwards towards the white matter. Primary current also appears on the medial surface of the left hemisphere, but this is most likely to be the result of "smoothing" caused by regularisation. An insignificant current source also appears on the left calcarine fissure; this is probably spurious current as observed in the simulations of chapter five. The right half-field analysis shows primary current in the left hemisphere, directed towards the white matter. Primary current also appears on the superior surface of the left calcarine fissure, directed towards the pial surface. In the right hemisphere, there are low levels of spurious current, except for inwardly-directed current on the dorsal surface and inferiorly-directed current on both the superior and inferior surfaces of the calcarine fissure. The full-field analysis, plotted on the same scale as the half-field analyses, shows no primary current whatsoever: the magnetic field distribution is so low in amplitude, relative to the half-field distributions, that the corresponding primary current is negligible.

It has already been remarked that no inverse procedure can correctly determine primary current sources when they are arranged in such a way that their magnetic fields cancel. If the magnetic field has low amplitude, then the primary current is also deduced to be small in magnitude. This is the reason why none of the three analysis methods can produce a significant result for full-field stimulation. In this case, the true source distribution probably involves current sources in both hemispheres, but the magnetic

Figure 79 (previous page). Minimum norm estimate for the group average CIIm peak of the response to right half-field pattern onset stimulation. Goodness-of-fit: 97.2%. Distances quoted are the radii of the layers of the source space.
Figure 80. Minimum norm estimates for the left half-field (LHF), full-field (FF) and right half-field (RHF) group average pattern onset ClIrm peak. Goodness-of-fit: LHF 70.7%; FF 80.3%; RHF 97.2%. Only the 65mm radius layer of the source space is shown.
Figure 81. Anatomically constrained analyses of the left half-field (LHF), full-field (FF) and right half-field (RHF) group average pattern onset CIIIm peak. Goodness-of-fit: LHF 74.7%; FF 20.4%; RHF 81.9%.
fields of these sources almost cancel, and the analysis methods then determine only the current source associated with the residual magnetic field. This is a problem with all neuro-electromagnetic inversion algorithms.

The effect also manifests itself with regard to the calcarine fissures in the anatomically constrained analyses. In the left half-field analysis, it is impossible to say whether the right calcarine fissure contains primary current or not, as symmetrical currents on the upper and lower surfaces would produce cancelling magnetic fields and zero current would be reconstructed. Hence, the apparent absence of current should be treated cautiously, as it may be misleading. In the right half-field analysis, current can be seen explicitly on both left and right calcarine fissures. In this case, the reconstructed current is almost certainly the vector sum of two oppositely oriented current dipoles, one on the superior surface of the fissure and the other on the inferior surface. Because of the regularisation effect, the current may be reconstructed on either surface, or both together, calling for careful interpretation, as was discussed in chapter five.

Consideration is now given to the origins of the pattern onset response. The equivalent current dipoles indicate that the response is centred on striate cortex. It could be that the response is indeed striate in origin, with the current in the calcarine fissure producing cancelling fields, leaving the medial part of area 17 as the response generator. This medial region would produce primary current in the form of a horizontal dipole, as observed. However, it is also possible that the current dipoles represent complex distributions of primary current which just happen to surround striate cortex; in other words, the response is extrastriate, but the dipole lies in striate cortex. For example, the response could be in area 18, which lies above and below area 17; then the equivalent current dipole would lie in area 17. In view of the goodness-of-fit of the dipole for the right half-field response, it seems likely that this latter suggestion is the case: the response involves a complex distribution of primary current which is centred around primary visual cortex. However, a contribution from area 17 cannot be ruled out.

The minimum norm estimates provide little further information. This is typical of such estimates - they tend to be simple and diffuse, providing little more information than a one- or two-dipole fit. (For another example, see Ahlfors et al. 1992.) The anatomically constrained solutions, however, reveal rather more. The left half-field analysis suggests that primary current is located at least in areas 18 and 19 of the contralateral hemisphere. The right half-field response confirms this, and also shows current in area 17, suggesting that areas 17, 18 and 19 are all involved in generating the CIIm peak of the pattern onset response. Could the large extent of the primary current be the result of a "smoothing" or "scattering" effect of the model? It is very unlikely, since simulations in chapter five showed that if the sources were confined to one quadrant of visual cortex, for example, then the model reconstructed them correctly from the group average response, with the exception that smearing across immediately
adjacent fissures was observed. Hence, there is no reason to believe that the analysis method reconstructs a distributed response when it is really localised.

Some ipsilateral current also appears in the right half-field analysis, this may be genuine, or it may be that group averaging a number of responses from montages which differed slightly in size caused the original magnetic field distribution to be smeared slightly, giving the solution a lateral spread. However, the ipsilateral current is probably genuine, as later analyses will show that an almost identical current distribution is reconstructed from the data of a single subject. The downwardly-oriented sources on the calcarine fissures probably correspond to the downwardly-directed current density on the right half-field minimum norm estimate. In all the analyses, primary current appears to be directed from the pial surface towards the white matter. Hence, it is deduced that the C1Im peak is produced by cortical current in areas 17, 18 and 19, with the current flowing towards the white matter.

Source Analyses of the Group Average Pattern Reversal Responses  Figures 82, 83 and 84 give equivalent current dipoles, minimum norm estimates and anatomically constrained analyses for the group average P100m peak of the left half-field, full-field and right half-field pattern reversal responses. Left half-field stimulation produces a current dipole in the right hemisphere and vice versa, but the direction of the current dipoles does not show any particular trend. Full-field stimulation produces an insignificant current dipole, as previously. The goodness-of-fit of the left half-field analysis is unacceptable; those of the full-field and right half-field analyses are not very good, but adequate. While both the left and right half-field dipoles are situated near to striate cortex, they are both somewhat lateral to represent primary current confined exclusively to area 17, and, in view of the poor goodnesses-of-fit, it seems likely that multiple sources are involved in generating the magnetic response.

The minimum norm estimates support this suggestion, showing numerous current sources. The analysis of the left half-field response shows three main areas of current density, the first situated in the upper left region of the source space, the second just below and to the right of it, approximately centrally in the source space grid, and the third in the lower right hand corner of the grid. However, these sources do not seem to bear any relationship to the anatomy of the visual cortex. The right half-field analysis contains primary current density just to the right of the medial plane, some 40mm above the x-axis, and a second region of primary current in the lower left of the source space grid. Apart from the greater magnitude of the left half-field estimate, the left and right half-field results are approximately symmetrical. The goodnesses-of-fit of the estimates are acceptable.

The anatomically constrained analyses suggest that the left and right half-field responses are generated in areas 17, 18 and 19 contralateral to the stimulated half-field, by current flowing from the white matter towards the pial surface, the only exception being a current source on the left dorsal surface upon left half-field stimulation, and
Figure 82. Equivalent current dipoles for the left half-field (LHF), full-field (FF) and right half-field (RHF) group average pattern reversal P100m peak.
LHF: $x = 31$mm, $y = 27$mm, $z = 10$mm, $q = 19.4$nAm, goodness-of-fit = $4.4 \times 10^{-12}\%$.
FF: $x = 22$mm, $y = -6$mm, $z = 58$mm, $q = 4.9$nAm, goodness-of-fit = $0.1\%$.
RHF: $x = 25$mm, $y = -34$mm, $z = 25$mm, $q = 12.9$nAm, goodness-of-fit = $9.2 \times 10^{-3}\%$. 

252
Figure 83. Minimum norm estimates of the left half-field (LHF), full-field (FF) and right half-field (RHF) group average pattern reversal P100m peak. Goodness-of-fit: LHF $1.0 \times 10^{-2} \%$; FF 16.9\%; RHF 3.7\%.
Figure 84. Anatomically constrained analyses of the left half-field (LHF), full-field (FF) and right half-field (RHF) group average pattern reversal P100m peak. Goodnesses-of-fit: LHF $3.3 \times 10^{-7}\%$, FF $2.9 \times 10^{-8}\%$, RHF $0.4\%$. 

$\Rightarrow 7.2\text{nAm}$. 

otherwise low levels of spurious ipsilateral activity. Left half-field stimulation produces current sources within the right calcarine fissure, but no current sources are reconstructed within the left calcarine fissure upon right half-field stimulation. It is assumed that this is due to complete field cancellation in the right half-field case, but not the left. From the left half-field analysis, it is deduced that a half-field response involves the contralateral calcarine fissure. However, of the three anatomically constrained analyses, only the right half-field case should be treated with confidence, as this is the only one with a reasonable goodness-of-fit to the data. The conclusion from this analysis is that the P100m peak of the half-field pattern reversal response is generated by contralateral current sources in areas 18, 19 and possibly area 17, with the current flowing towards the pial surface. The pattern reversal P100m peak therefore appears to be generated in the same areas as the pattern onset CIIIm peak, but with the current flowing in the opposite direction.

*Pattern Onset CIIIm Peak in a Single Subject* Pattern onset responses were recorded over a 42-point montage in a single subject, and the topography of the CIIIm peak is shown in the upper part of figure 85 for the case of full-field and right half-field stimulation. The full-field topography is complex, probably due to the presence of noise (estimated standard deviation 13aWb). The distribution consists of an outgoing field in the lower left part of the montage, and an ingoing field in the upper right. The right half-field response has greater magnitude and is therefore less affected by noise. It contains outgoing fields in the lower left corner and right side of the montage, and an ingoing field about 40mm up from the inion and 40mm to the left of the midline.

The equivalent current dipoles are shown in figure 86. The dipole for full-field stimulation is located in striate cortex close to the medial plane, and is directed superiority and to the left, as would be expected from the topographic field map. The goodness-of-fit is acceptable at 2.5%. The dipole for right half-field stimulation is situated above striate cortex in the right hemisphere. It has a much lower dipole moment than the dipole for full-field stimulation, despite the greater amplitude of the topographic map, and a poor goodness-of-fit. The low dipole moment and inadequate goodness-of-fit almost certainly occur because the topographic map contains three field extrema, rather than two of opposite polarity which signify a dipolar current distribution. A single dipole is therefore not a suitable source model in this case.

The minimum norm estimates (figure 87) are more suitable for showing the multiple sources underlying the field maps. The full-field analysis shows two regions of primary current density, the first centred on the x-axis beneath the inion, and the second extending into the upper left part of the source space. These currents are located approximately below the line of inflexion (the zero contour line) of the field map, as would be expected from the right-hand grip rule. The right half-field analysis also shows two regions of current density, the one in the lower left region of the source space and the other just to the right of the medial plane, about 30mm above the inion. These two
Figure 85. Full-field (FF) and right half-field (RHF) magnetic responses to (above) pattern onset and (below) pattern reversal stimulation, recorded from a single subject using a 42-point montage. Contours are marked in nWb.

Current flows are also situated beneath the two lines of inflexion of the field topography. This analysis is a good example of the ability of minimum norm estimation to show multiple sources: the field map is clearly the result of at least two dipolar sources, and these are resolved in the resulting estimate. The goodness-of-fit of both full-field and half-field analyses are fairly good.

Although the minimum norm estimates are fully compatible with the data, it is difficult to relate them to the anatomy of the occipital lobe. For example, full-field stimulation produces current in the upper left region of the source space; this current is located in the dorsal cortex of the left hemisphere. However, the current is tangential (radial primary current cannot be detected by magnetic recording alone), whereas physiologically, primary current flowing normal to the dorsal cortical surface would be predominantly radial. A similar paradox occurs with right half-field stimulation, where the minimum norm estimate shows tangential primary current in the lateral region of the left hemisphere. Again, the minimum norm estimate is unrealistic, as primary current in
Figure 86. Equivalent current dipoles for the full-field (FF) and right half-field (RHF) pattern onset CI1m peak recorded from a single subject using a 42-point montage.

FF: $x = 32\text{mm}, y = -3\text{mm}, z = 15\text{mm}, q = 30.5\text{nAm}$, goodness-of-fit = 2.5%.

RHF: $x = 50\text{mm}, y = 14\text{mm}, z = 36\text{mm}, q = 7.8\text{nAm}$, goodness-of-fit = $2.0 \times 10^{-9}\%$. 
Figure 87. Minimum norm estimates for the full-field (FF) and right half-field (RHF) pattern onset CIIIm peak recorded from a single subject using a 42-point montage. Goodness-of-fit: FF 10.6%, RHF 26.2%.
this part of the cortex would be predominantly radial - not tangential as predicted by the estimate.

The anatomically constrained analyses (figure 88) provide a more physiologically plausible interpretation of the data, as the primary current is constrained to flow normally to the cortex. The full-field magnetic response is shown to be the result of current sources in areas 17 and 18 bilaterally. No dorsal, ventral or lateral current is observed at all. The right half-field magnetic response is the result of primary current in areas 17, 18 and 19, with almost as much activity in the right hemisphere as the left. This is perhaps surprising for a right half-field response, as one might expect the current to be contralateral to the stimulated half-field. However, the original topographic map, from which the current distribution is derived, also has a bilateral nature, causing the current distribution to follow similarly. The current is directed from the pial surface towards the white matter, except in the calcarine fissures, where it is impossible to interpret which surface the current is on. The goodness-of-fit of the full-field and right half-field analyses are both acceptable. Note the similarity of the right half-field analysis with the corresponding group average analysis shown in figure 81.

The conclusions of this study are thus similar to those of the group average study. Although the original topographic map for right half-field stimulation is bilateral, giving unexpected lateralisation in the three types of source analysis, it is clear that the CIIm response is generated by at least two sources. The anatomically constrained analyses show that the magnetic response is generated in areas 17, 18 and 19 by current flowing inwards towards the white matter.

*Pattern Reversal P100m Peak in a Single Subject* Full-field and right half-field pattern reversal responses were recorded from a single subject using a 42-point montage and the spatial distribution of the P100m peak is shown in the lower part of figure 85. The full-field topography consists of four field extrema: an ingoing and outgoing field in the upper right of the montage and an ingoing and outgoing field in the lower left of the montage. In each case, the outgoing field is stronger than the ingoing field, suggesting that the two pairs of extrema may not simply be due to two current dipoles. The topography of the right half-field response has a low amplitude and is therefore somewhat noisy. The only notable features are an outgoing field about 40mm up from the inion and 20mm to the left of the midline, and another outgoing field in the lower right of the montage.

The equivalent current dipoles for full-field and right half-field stimulation (figure 89) are both situated in the left hemisphere, the former at about the level of the calcarine fissure and the latter near to the ventral surface of the brain. The full-field dipole is located between the two leftmost extrema of the field map, although it is somewhat rotated from the orientation required to produce these two extrema. The right half-field dipole is located beneath the lower left part of the recording montage, having a direction compatible with the outgoing fields in the upper and right areas of the montage. The
Figure 88. Anatomically constrained analyses of the full-field (FF) and right half-field (RHF) pattern onset CIIm peak recorded from a single subject using a 42-point montage. Goodness-of-fit: FF 0.2%; RHF 7.2%.
Figure 89. Equivalent current dipoles for the full-field (FF) and right half-field (RHF) pattern reversal P100m peak recorded from a single subject using a 42-point montage.

FF: $x = 32\text{mm}$, $y = -33\text{mm}$, $z = 20\text{mm}$, $q = 10.9\text{nAm}$, goodness-of-fit = 0.7%.

RHF: $x = 32\text{mm}$, $y = -15\text{mm}$, $z = 1\text{mm}$, $q = 11.0\text{nAm}$, goodness-of-fit = 85.4%.
goodnesses-of-fit of the two dipoles are acceptable, the fit for the right half-field dipole being very good.

The minimum norm estimate for full-field stimulation (figure 90) shows a region of significant primary current centred about 10mm above the inion and 30mm to the left of it. Other primary currents appear centrally within the source space grid and also to the extreme left of it. It is uncertain whether these two smaller current flows are "whirls" associated with the main current flow, or whether they represent true current sources. Given the locations of the extrema in the field map, it is likely that the medial current is legitimate, although the lateral current probably is not. The minimum norm estimate for right half-field stimulation contains a variety of current sources, the two most prominent being in the lower left and right parts of the source space grid. As with the estimates for pattern onset responses, it is difficult to interpret the lateral current flow in terms of the visual cortex: the orientation of the current is incompatible with its location.

The anatomically constrained analyses are shown in figure 91. The full-field analysis shows a single region of primary current on the ventral surface of the brain. This is compatible with the equivalent current dipole and the current density in the lower region of the minimum norm estimate. The ventral activity is sufficient to account for the data with a goodness-of-fit of 7.6%, which is fairly good. The right half-field response has a low magnitude, in accordance with the recorded magnetic field. It is not clear whether the response is predominantly contralateral or not, the only indication that the left hemisphere may contribute more to the response than the right is that there is significant current in the left calcarine fissure, but not the right. It is also difficult to say whether the current is directed towards or away from the pial surface.

In conclusion, it appears that the recorded data for full-field stimulation can be accounted for by a region of primary current in the left hemisphere, directed inferomedially. This current is probably generated by the ventral surface of the brain, such that it flows from the white matter towards the pial surface. The right half-field response can be accounted for with a goodness-of-fit of 85.4% by a surface-going current dipole on the ventral surface of the left hemisphere, but distributed-source analyses show the current to be more widely spread throughout areas 17, 18 and 19. It is not clear whether the cortical response is entirely contralateral to the stimulated half-field, and current direction relative to the cortical surface is also uncertain. However, this study does confirm the involvement of extrastriate cortex in the generation of the pattern reversal P100m peak.
Figure 90. Minimum norm estimates of the full-field (FF) and right half-field (RHF) pattern reversal P100m peak recorded from a single subject using a 42-point montage.
Goodness-of-fit: FF 3.8%; RHF 92.1%.
Figure 91. Anatomically constrained analyses of the full-field (FF) and right half-field (RHF) pattern reversal P100m peak recorded from a single subject using a 42-point montage. Note that the scale is different for each diagram. Goodnesses-of-fit: FF 7.6%; RHF 41.3%.
Sequential Topography of the Pattern Onset Visual Evoked Magnetic Response  
Full-field pattern onset responses were recorded over a 20-point montage in two subjects and the spatial distribution of the magnetic field was plotted at 10ms intervals (figures 92 and 93). The check size of the stimulus was 9′ × 6′. The experiment was then repeated in subject 1 using large checks of size 77′ × 55′ (figure 94). (This was carried out by Degg et al. (1992, 1993) and the magnetic responses described below are by the kind permission of those concerned.)

In subject 1 (figure 92), the response from 80-100ms consists of an ingoing field at the left of the montage and an outgoing field on the right. The field pattern suggests that the underlying current source may be a vertically-oriented current dipole near to striate cortex. From 100ms onwards, the topography changes to one involving typically two or three extrema in various stages of formation. In this subject, CIIm occurs at 110ms.

In subject 2 (figure 93), the response from 80-100ms also consists of an ingoing field on the left and an outgoing field on the right, but the outgoing field does not become very significant until 100ms, when it contributes to a dipolar pattern centred immediately over the left calcarine fissure. Again, from 100ms onward, the field is divided into two or three extrema which evolve as time progresses. CIIm occurs at 120ms.

Returning to subject 1, the response to large 77′ × 55′ checks (figure 94) can be seen to be of lower amplitude than the response of either subject to small checks. Furthermore, specific field extrema are less commonly observed. From 80-100ms, a single ingoing field is observed about 60mm above the inion; from 110ms onward it spreads into a more distributed field topography which persists, in varying degrees of evolution, for the remainder of the response. In this case, CIIm is at 120ms.

Sequential Source Analyses of the Pattern Onset Visual Evoked Magnetic Response

Figure 95 shows the equivalent current dipoles for the sequence of maps shown in figure 92, the response of subject 1 to 9′ × 6′ checks. Dipole fits were discarded if the dipoles were located with x < 0, (i.e. if they were in the anterior half of the spherical model), or if the eccentricity was less than several millimetres, giving an unreasonably large dipole moment. All other current dipoles are shown, regardless of goodness-of-fit. For this reason, the results should be treated with caution, as there is no guarantee that each

Figure 92 (overleaf). Sequential topography of the magnetic response to onset of 9′ × 7′ checks, recorded from subject 1. Contours are marked in aWb.

Figure 93 (two pages ahead). Sequential topography of the magnetic response to onset of 9′ × 7′ checks, recorded from subject 2. Contours are marked in aWb.

Figure 94 (three pages ahead). Sequential topography of the magnetic response to onset of 77′ × 55′ checks, recorded from subject 1. Contours are marked in aWb.
dipole is an adequate model for its respective magnetic field topography. Nevertheless, it is interesting to observe that the majority of dipoles are clustered around the calcarine fissure and biased slightly towards the left hemisphere. If the full-field response is considered to be the sum of the left and right half-field responses, the lateral bias may suggest that the full-field response is dominated by the constituent right half-field response. Dipoles 8 (160ms) and 9 (170ms) are both located on the ventral surface of the brain.

Figure 96 shows the dipole plot for the response of subject 2 to 9' × 6' checks. The locations of the current dipoles are not dissimilar to the locations found in subject 1, the dipoles generally being slightly superior to the calcarine fissure and biased towards the left hemisphere, although the latter effect is not as marked as in subject 1. Dipoles 6 (140ms) and 7 (150ms) are located near to the ventral surface of the brain.

Figure 97 gives the equivalent current dipoles for the response of subject 1 to 77' × 55' checks. The dipoles are located in the vicinity of and slightly superior to the calcarine fissure, dipoles 5, 6, 7 and 8 (130-160ms) being exactly coincident with it. Dipole 0 (80ms) is apparently spurious, being located in the cerebellum! It is interesting to compare the dipole fits for the responses of subject 1 to small 9' × 6' checks and large 77' × 55' checks (figures 95 and 97). Figure 95 shows the dipoles clustered fairly eccentrically, while figure 97 shows them spread evenly across both deep and superficial regions of the visual cortex. This seems to support previous evidence (Maclin et al. 1983) that small checks evoke a more superficial response than large checks. It is also consistent with the hypothesis that small checks preferentially stimulate the parvocellular pathway, which is associated with central vision and therefore projects to the convexity of the occipital pole (Drasdo 1980), while large checks preferentially stimulate the magnocellular pathway, which dominates peripheral vision and projects to the anterior part of visual cortex (Meredith and Celesia 1982; Novak et al. 1988).

The usefulness of equivalent current dipole fitting as a means of carrying out a sequential source analysis is limited by the goodness-of-fit. At some time instants, the single dipole is an adequate source model, and at others it is not. Distributed-source analyses do not suffer from such a problem: they can be used at each time instant in a sequence to produce a solution with an acceptable goodness-of-fit. This is a distinct advantage in the present situation. Hence, the remainder of the sequential analyses are performed using distributed-source methods.

Sequential minimum norm estimates of the response of subject 1 to 9' × 7' checks are shown in figures 98 and 99 (figure 98 amplifies the low-amplitude estimates at 70, 80 and 90ms). At 70 and 80ms, the estimates show a region of lateral current density on the left of the source space, directed medially, and a region of medial current density, directed inferiorly. The strength of the current is higher at 80ms than at 70ms. At 90ms, the current distribution alters slightly, so that a single area of high current density dominates. This current is located near to the calcarine fissure, about 30mm above the x-axis and 15mm to the left of the medial plane. The vertical orientation of
Figure 95. Locations of the equivalent current dipoles for the response of subject 1 to $9' \times 7'$ checks. Latencies run from 80ms (0) to 190ms (11) in 10ms intervals.
Figure 96. Locations of the equivalent current dipoles for the response of subject 2 to 9' x 7' checks. Latencies run from 80ms (0) to 190ms (11) in 10ms intervals.
Figure 97. Locations of the equivalent current dipoles for the response of subject 1 to 77' x 55' checks. Latencies run from 80ms (0) to 190ms (11) in 10ms intervals.
\[ \rightarrow 1 \times 10^{-4} \text{Am}^{-2}. \]

Figure 98. Minimum norm estimates for the response of subject 1 to $9' \times 6'$ checks: early latencies.
the primary current suggests that it could be a product of activity within the calcarine fissure. Note that CIIm occurs at 80-90ms and these results may therefore indicate that CIIm is generated by striate cortex within the calcarine fissure.

At 120ms, several current regions are apparent, some of which are medial and some of which are lateral. However, the sources in the left hemisphere are difficult to interpret, as they form a "whirl". Part of the "whirl" is undoubtedly genuine, but the remainder of the loop is artefact caused by lack of anatomical constraint. As noted earlier, lateral tangential sources are difficult to relate to the underlying anatomy, which raises questions as to their validity. Nevertheless, comparing the estimate with the magnetic field distribution from which it is derived, and using the right-hand grip rule, it can be seen that the estimate approximately accounts for the recorded data.

At 140ms, there are three distinct current sources, all of which are directed in a superior-inferior direction. The medial sources can be accounted for by activity within the left and right calcarine fissures, but the origins of the lateral sources are less certain. Again, the source distribution qualitatively explains the magnetic data. At 180-190ms, a similar source distribution arises, the current having the opposite orientation and not quite such a high magnitude. Interestingly, periods of stable magnetic field are observed in this subject at around 141 and 188ms. Correspondingly, two similar but oppositely oriented current distributions of significant amplitude characterise these latencies. The direct reversal of the current suggests that the same cortical regions may be involved, but that current flow is in the opposite direction.

The sequence of magnetic fields was analysed using anatomically constrained distributed-source estimation. In order to ensure that the brain and source space were correctly co-registered, the source space was rotated as described in chapter five. Maximum goodness-of-fit was obtained when the source space was in its default, unrotated, position (figure 100), so no rotation was used for the subsequent analyses (figures 101 and 102). At latencies of 70-90ms, the primary current is seen to be building up on the left calcarine fissure and on the cuneal gyrus of one or other of the hemispheres. The direction of current flow is in accord with that of the minimum norm estimates, and the current source on the calcarine fissure at 90ms exactly corresponds to the current density observed on the minimum norm estimate at 90ms. At 110ms, the latency of CIIm, primary current is distributed throughout the calcarine fissures, the cuneal gyri and the dorsal surfaces of the brain. The arrangement of current has features in common with an analysis of a right half-field CIIm response (see figures 81 and 88). Ten milliseconds later, the current distribution is mainly confined to the right hemisphere, with current directed from the pial membrane towards the white matter. The distribution resembles a left half-field response (see figure 81). It is likely that the

Figure 99 (previous two pages). Minimum norm estimates for the complete response of subject 1 to 9' x 6' checks.
Figure 100. The goodness-of-fit of an anatomically constrained solution as a function of axis rotation for the response of subject 1 to 9' x 6' checks at a latency of 80ms. The peak at approximately zero rotation indicates that the standard source space is applicable to this subject. Hence, in the sequence following, no source space rotation is applied.

cortical response to full-field stimulation is the sum of independent left and right half-field responses, one having maximum amplitude at 110ms and the other at 120ms. The externally-observed CIIIm peak occurs when the combined cortical response produces a magnetic field of maximum amplitude. This is consistent with the observation of Degg et al. (1993) that the full-field CIIIm peak is the sum of the two half-field peaks, neither of which necessarily occurs at the latency of the full-field peak.

From 120 to 150ms, a dominant ventral source in the right hemisphere builds up and then declines, while numerous other sources on the dorsal and medial surfaces and calcarine fissures also evolve. At 190ms, the source distribution consists almost entirely of current sources in the right hemisphere, directed from the white matter towards the pial surface. These sources extend at least throughout areas 18 and 19, possibly including area 17. The distribution is almost identical to that at 120ms, except for the opposite current direction. This is a similar effect to that observed in the minimum norm estimates for 140 and 190ms. Since CIIIIm has a latency of 190ms, it is likely that this peak is generated in areas 18 and 19, and possibly area 17, by current directed away from the white matter. In terms of cortical current, CIIIIm appears to be very similar to CIIIm, apart from the reversal in the current direction.

Minimum norm estimates of the response of subject 2 to 9' x 6' checks are shown in figures 103 and 104. At the early latencies, from 70 to 80ms, the primary current appears to be medial and directed from left to right. At 90ms, around the latency of CIIIm, the current appears to be located in the left hemisphere and to be separated into two parts, one of which is directed medially, and the other of which is directed inferiorly. Some current also appears in the right hemisphere, although with all the current, it is difficult to know what parts are legitimate and what parts are "whirls"
Figure 101. Anatomically constrained analyses for the response of subject 1 to 9 × 6 checks early latencies.
Figure 102. Anatomically constrained analyses for the complete response of subject 1 to 9' x 6' checks. ⇒ 10.9 nAm.
$\rightarrow 1 \times 10^{-4} \text{Am}^{-2}$.

Figure 103. Minimum norm estimates for the response of subject 2 to 9' x 6' checks: early latencies.
due to regularisation. The remaining estimates are unremarkable, except for lateral current in the right hemisphere at 110-120ms, around the latency of CIIm. At 160-170 and 190ms, superiorly-directed current appears near to the ventral surface of the brain, the current at 170ms being in the right hemisphere, and that at 190ms being in the left.

The anatomicallly constrained analyses for subject 2 are shown in figure 105 and 106. As with subject 1, no source space rotation was required. At 70-90ms, the response is seen to be located in areas 17 and 18, with most of the response at 90ms being on the cuneal gyrus. The analyses are consistent with the corresponding minimum norm estimates, the primary current flow being from left to right, with some flow in the superior-inferior direction on the left calcarine fissure. From 110ms onward, the magnetic response is generated almost entirely by sources on the ventral surface. This is also consistent with the minimum norm estimates, which showed superiorly-directed current near to the ventral surface at 160-170 and 190ms. The explanation for this ventral current is that the currents on the medial and calcarine surfaces produce magnetic fields which cancel each other, whereas the ventral currents do not. Hence, only the ventral current is observed by the source analyses. This argument also applies to dorsal current, but since dorsal sources are not observed, it would seem that the dorsal surface is not as active as the ventral surface. In addition, the dorsal surface is more radial than the ventral surface, which would promote the latter in analyses of evoked magnetic responses.

The final study examines the evoked magnetic response of subject 1 to large 77' x 55' checks. Minimum norm estimates of the magnetic response at latencies of 70-80ms (figures 107 and 108) show both bilateral and medial current. The medial current density probably represents inferiorly-directed sources on one or both of the calcarine fissures, while the lateral current is probably artefact. This can be deduced from the fact that there is no artefactual medial current at the top of the source space, which there would be if either of the two lateral currents were genuine. At 90ms, the current density is centred on the left calcarine fissure, oriented upwards. The later estimates are difficult to interpret, due to regularisation artefacts, but they generally consist of horizontally-oriented bilateral currents and inferiorly-oriented medial currents. Since the lateral currents are physiologically unreasonable, as discussed earlier, no further attempt is made to explain these sources; anatomically constrained analyses provide more plausible solutions.

The anatomically constrained solutions (figures 109 and 110) show that at latencies of 70-80ms, the primary current is situated on the right calcarine fissure and is directed downward. The current flow is consistent with that of the minimum norm estimates for the same latencies. At 90ms, both the left calcarine fissure and the cuneal

Figure 104 (previous two pages). Minimum norm estimates for the complete response of subject 2 to 9' x 6' checks.
Figure 105. Anatomically constrained analyses for the response of subject 2 to 9' × 6' checks: early latencies.
Figure 106. Anatomically constrained analyses for the complete response of subject 2 to $9' \times 6'$ checks. ➔ 8.6nAm.
Figure 107. Minimum norm estimates for the response of subject 1 to 77' x 55' checks: early latencies.
gyrus are involved in generating the magnetic response. At 120ms, the latency of CIIIm, the dorsal, ventral and cuneal surfaces are active, with most primary current being ventral. At higher latencies, similar current distributions are observed, but the cuneal source predominates. It would seem that activity previously observed on the ventral surface has now been transferred to the cuneal gyrus, presumably because the magnocellular pathway is now being stimulated, rather than the parvocellular pathway for small checks. At 170-190ms, the cuneal source has the opposite orientation to that at 120-130ms, again suggesting that CIIIm is generated by the same areas as CIIIm, but by current of reversed polarity. Note particularly the similarity, apart from reversal of the current direction, between the analyses at 130 and 170ms.

Conclusions

*Cortical Origins of the Pattern Onset Visual Evoked Magnetic Response* The sequential studies demonstrate that at latencies of 70-90ms, during which time CIIIm develops, the cortical response is located in the calcarine fissures and on the cuneal and lingual gyri. CIIIm can therefore be said to be generated by area 17 and possibly by area 18. The only question remains as to whether area 18 is involved or not. If area 17 extends significantly onto the medial surfaces, then the current could be located entirely within area 17. On the other hand, area 17 is sometimes confined to the calcarine fissure, extending only a few millimetres onto the cuneal and lingual gyri (Brindley 1972; Talairach and Tournoux 1988; Gebhard et al. 1993). There is thus insufficient primary visual cortex on the medial surface to generate such large medial currents as are observed. Hence, the CIIIm peak appears to be generated by areas 17 and 18 (areas V1 and V2). This conclusion is partly consistent with that of Jeffreys and Axford (1972a), Darcey et al. (1980a,b) and Butler et al. (1987) (area 17), but inconsistent with that of Lesèvre and Joseph (1979) (area 19) and Maier et al. (1987) (area 18 or 19). However, in comparing the conclusions of this study of evoked magnetic responses with those of evoked potential studies, it must be remembered that the population of neurones which produces a magnetic response may not be the same as the one which produces an evoked potential. In other words, each modality may represent the activity of a distinct neural population, and the source analyses will reflect this.

Both the group average and individual studies have shown the CIIIm waveform peak to be produced by a complex arrangement of sources throughout areas 17, 18 and 19. Since the CIIIm peak has the largest amplitude of the three waveform peaks, it is natural for it to be generated by the greatest area of visual cortex. Hence, the finding that CIIIm

Figure 108 (previous two pages). Minimum norm estimates for the complete response of subject 1 to 77' x 55' checks.
Figure 109. Anatomically constrained analyses for the response of subject 1 to 77 × 55 checks: early latencies.
Figure 110. Anatomically constrained analyses for the complete response of subject 1 to 77' x 55' checks. → 4.8 nAm.
is generated by three Brodmann visual areas as opposed to one or two for CII, is physiologically reasonable. The analyses have shown that the primary current underlying CII is directed from the pial surface towards the white matter. This is consistent with the hypothesis that the current is the product of excitatory postsynaptic potentials in the large pyramidal cells of layer 5.

When using a full-field stimulus, the medial and calcarine sources produce fields which cancel, so that there is no external evidence for the sources and they are not reconstructed by a source analysis. In this case, the only current sources which produce significant external fields are those on the dorsal and ventral surfaces, with the ventral contribution being greatest. The reason for the predominance of the ventral cortex is probably because it is located in such a way that current sources which are normal to its surface are tangential to the spherical conductivity model, whereas those on the dorsal surface are approximately radial. On the other hand, the bias towards the ventral cortex may be a real one, perhaps suggesting a greater involvement of VP than dV3.

The minimum norm estimates sometimes show significant tangential currents in a lateral location. There are two possible explanations for this: either the current is located within minor sulci, oriented normally to the cortical surfaces, so that it is tangential to the conductivity model, or the current is merely incorrect as a result of applying insufficient prior information to the inverse procedure. If the primary current were genuinely located in lateral sulci, there would be every reason to believe that the currents on opposite walls of the sulci would produce cancelling fields, since this occurs even in the much more extensive medial and calcarine regions involving several distinct visual areas. Hence, it is unlikely that the evoked magnetic fields are the result of lateral tangential currents. The reconstructed lateral currents must therefore be the result of minimum norm estimation producing oversimplified solutions. When anatomical constraints are applied, so that lateral current must be radial, a more complex source distribution confined to less lateral regions is the result. In general, the anatomically constrained analyses show no primary current on the occipital pole or on the lateral surfaces. This is almost certainly due to the fact that primary currents in these regions are predominantly radial, producing no external magnetic field. Consequently, no current is reconstructed, even though it possibly exists. This is a limitation of analysing evoked magnetic responses.

The proposed locations of the cortical generators of CII are partly consistent with previous hypotheses relating to the electrical counterpart, CII. For example, Jeffreys and Axford (1972b) realised that CII could be produced by the dorsal and ventral cortices, although their simple approach did not predict the additional sources on the medial and calcarine surfaces. Lesèvre and Joseph (1979) suggested that area 18 was the most likely source of CII, which is also partly in accord with the result of this chapter inasmuch as area 18 seems to contribute to the response. The extensive cortical activity found in this chapter also supports the conclusion of Maier et al. (1987) that both striate and extrastriate sources are involved in producing CII. However, as with
CIIm, it needs to be borne in mind that CII and CIIm are not necessarily produced by the same neural population.

The sequential analyses show that extrastriate cortex is mainly responsible for the evoked magnetic response beyond latencies of about 150ms, although a striate contribution cannot be ruled out because the absence of reconstructed sources of significant amplitude on the calcarine fissures may be due to field cancellation effects rather than genuine absence of sources. The response from 150 to 190ms seems to vary between subjects and between stimuli in a given subject. However, there is evidence that the cortical response underlying CIIm may be very similar to that responsible for CIIm, but with current flow reversed. This similarity was observed by Lesèvre and Joseph (1979), who suggested that area 18 was responsible for both CII and CIII peaks of the evoked potential. Richer et al. (1983) also noted that a single current source appeared to be common to both the CIIm and CIII peaks of the evoked magnetic response. Maier et al. (1987) deduced that CIII was predominantly extrastriate in origin, and Ossenblok and Spekreijse (1991) further concluded that it was produced within area 19. The results of the present study confirm this suggestion, as the ventral region of area 19 seems to contribute most significantly to the response to small checks at latencies beyond about 150ms.

Cortical Origins of the Pattern Reversal Visual Evoked Magnetic Response The analyses of group average and single subject responses to pattern reversal stimulation show that the P100m peak is generated by areas 17, 18 and 19 simultaneously. The principal surfaces involved in the response are the dorsal, ventral, medial and calcarine surfaces. Virtually no primary current is observed on the occipital pole or in lateral cortex for reasons already discussed. As with pattern onset stimulation, the medial and calcarine surfaces produce cancelling fields under some conditions, notably with full-field stimulation, leading to an absence of external magnetic field and reconstructed current. The primary current is generally directed from the white matter towards the pial surface. The proposed involvement of both striate and extrastriate areas is in agreement with the principal components analyses of Maier et al. (1987).

Effectiveness of the Source Analysis Techniques Equivalent current dipole fitting is robust and straightforward to use, but is of little value if the goodness-of-fit indicates that the single current dipole is an inadequate model of the source. When this occurs, a multiple- or distributed-source model must be used. In a sequence of analyses at different latencies, there are invariably several instants at which the single dipole model is inadequate, so its utility is limited.

Minimum norm estimation overcomes the difficulties associated with a single current dipole, as it is capable of imaging multiple current sources. However, due to regularisation, the solutions suffer from artefactual "whirls", which make image interpretation difficult. Furthermore, the solution tends to be oversimplified due to lack
of specific spatial constraint of any form; often, an equivalent current dipole is reconstructed from data generated by a complex source distribution. Because of this oversimplification, a current source sometimes appears in a physiologically implausible region of the brain. Even when a source is reconstructed correctly, it is difficult to visualise where the source is relative to the brain, unless the result is superimposed on an anatomical image of the brain.

Anatomically constrained distributed-source estimation overcomes these difficulties in a natural and effective way. It allows multiple sources to be reconstructed and can therefore be used for virtually all magnetic field topographies, except when the goodness-of-fit is occasionally low. Hence, it is of value when a sequence of field maps are to be analysed. Because anatomical constraints are used, the number of unknown parameters are reduced and the oversimplification observed in minimum norm estimates is overcome. Consequently, source distributions can be imaged with considerable detail. Since the source model is based explicitly on the anatomical shape of the brain, the results can also be displayed relative to a brain shape, allowing easy interpretation. The potential problem of using a standard source space in conjunction with a variety of subjects does not seem to present a problem: the analyses of group average responses recorded over a 20-point montage gave very similar source distributions to analyses of responses recorded from a single subject using a 42-point montage. In one case, the group average analysis was virtually identical to the corresponding single-subject analysis (c.f. figures 81 and 88, right half-field response).

The main difficulties which have not been resolved in this chapter are: (1) evoked magnetic responses are produced by tangential current dipoles only and therefore provide no information about purely radial sources such as those likely to be found at the occipital convexity, and (2) when using half-field and full-field stimuli, some areas of visual cortex produce negligible external magnetic fields and primary current in these areas cannot therefore be imaged successfully. In view of these difficulties, the following chapter uses both electric and magnetic responses to octant stimuli, ensuring that radial current sources are observed and that opposite surfaces of major fissures are simultaneously active as little as possible. This enables a more comprehensive image of cortical current to be realised.
8. Cortical Retinotopy Examined Using Visual Evoked Electric and Magnetic Responses

Introduction

One of the fundamental features of visual cortex is that it is retinotopically mapped. That is, a stimulus in a particular part of the visual field produces a response in a specific region of visual cortex. In human, retinotopy is known to exist in area V1 (Holmes 1945; Spalding 1952; Teuber et al. 1960), but in monkey, retinotopy is also known to be present in areas V2 (Zeki 1969a; Allman and Kaas 1971, 1974; Van Essen et al. 1979; Gattass et al. 1981; Tootell et al. 1983) and V3 (Cragg 1969; Zeki 1969a, 1978b; Van Essen et al. 1979). Because of the ordered mapping of visual field to cortical location, a stimulus in a particular portion of visual field elicits electric and magnetic responses which originate in the corresponding region of cortex. The analysis methods developed and tested in the previous chapters are a means of determining this region. However, for this approach to be entirely successful, it is necessary to stimulate only one visual area: if two or more visual areas are stimulated there will be two or more retinotopically ordered regions of cortex which are simultaneously active, making the mapping of each respective area more difficult to determine.

Several studies of retinotopic mapping have already been conducted by various researchers; these were considered briefly in the last chapter, but are now discussed in greater detail. Darcey et al. (1980a,b) stimulated left, right, upper and lower half-fields and all four quadrants of the visual field using a chromatic pattern onset stimulus. The stimulus consisted of a red checkerboard which contained 18' checks and subtended 10°. The pattern was presented for 30ms and a uniform field was then presented for the following 420-620ms. The short stimulus duration was similar to that of Jeffreys and Axford (1972a,b). Evoked potentials were recorded using 40 electrodes which covered the posterior half of the head, and the CI peak in the response was identified using global field power. The evoked potentials were analysed using one or two equivalent current dipoles. At a latency of around 90ms, corresponding to the latency of CI, the dipoles were generally located fairly near to the midline, having fairly small confidence regions and high goodnesses-of-fit. Stimuli which were presented to either the left or
right half of the visual field produced one contralateral current source, while stimuli which involved both halves produced bilateral sources. In other words, stimuli which projected to one hemisphere gave one source, while those which projected to both hemispheres gave two. The dipole orientations were exactly as would be expected from the cruciform model, and because of this accurate correspondence, the evoked potentials were assumed to be generated by area V1. Cancellation effects, similar to those noted in the previous chapter, were apparent in some cases. For example, in analyses of left and right half-field responses, there was no evidence of activity within the calcarine fissure, despite its presence in analyses of upper and lower half-field and quarter-field responses. It was deduced that the neurones on the upper and lower surfaces of the calcarine fissure produced cancelling electric potentials, which left only the medial current sources to produce the evoked response. Additive effects were also noted, the various quarter-field responses adding to give the half-field responses, and so on. It was thus concluded that the cortical projections of the different areas of the visual field were independent at the gross level of quadrant stimulation. Probably the most important observation of this study, however, was that the orientation of a current dipole representing the cortical response to quarter-field stimulation was dependent on the relative areas of intracalcarine and medial cortex within area V1. Although the cruciform model assumes that the area of medial cortex within V1 is equal to the area of intracalcarine cortex, there is evidence that the involvement of medial cortex in V1 is considerably variable between subjects and is in some cases negligible (Polyak 1957; Brindley 1972; Stensaas et al. 1974; Clarke and Miklossy 1990). If area V1 is confined almost exclusively to the calcarine fissure, a dipole representing a quarter-field response would have a vertical rather than oblique orientation when viewed coronally.

A similar study was conducted by Butler et al. (1987). They used an achromatic pattern onset stimulus which was presented for 170ms and had an interstimulus interval of between 600 and 1200ms. The stimulus was divided into peripheral octants which extended from 4 to 15° within the visual field. Foveal half-field stimuli, subtending approximately 1.6°, and foveal quadrant stimuli, subtending 2°, were also used. The check size was 50' and contrast was 90%. Evoked potentials were recorded from 13 posterior electrodes and the CI peak (90-110ms) was analysed using a single equivalent current dipole.

The peripheral stimuli produced current sources which were generally consistent with the cruciform model, assuming primary current was directed from the pial surface towards the white matter. However, there was one notable exception: whereas octants next to the vertical meridian were predicted by the cruciform model to cause a response on the medial surface of the cortex, thereby producing horizontal current dipoles, no horizontally oriented current dipoles were observed. Instead, current dipoles elicited by the octants next to the vertical meridian were almost vertically oriented. This suggested that the medial cortex made only a minor contribution to the evoked response, which was instead produced by the calcarine fissure.
Stimuli in the upper and lower foveal regions of the visual field produced sources in the inferior and superior parts of the occipital lobe, respectively, in agreement with the inverted representation of the visual field in occipital cortex. Although stimulating the two left foveal quadrants produced a pair of current dipoles which were mutually perpendicular, stimulating the right quadrants produced no such orientation difference. This indicated that the cruciform model was inadequate when foveal stimuli were used. The direction of the current flow with respect to the cortical lamina was also difficult to explain. The current dipoles were directed outwardly towards the scalp, suggesting that primary current was directed from the white matter towards the pial surface. This was paradoxical since the peripheral stimuli produced current of the opposite orientation. In order to resolve this paradox, the possibility was considered that CI was generated in extrastriate cortex. However, it was reasoned that if this were true, the locations of the dipoles for upper and lower quadrants would be further apart. An alternative suggestion was that the calcarine fissure was intricately folded within the occipital pole, thereby giving rise to outwardly-directed currents.

The two studies described so far have mainly used single current dipoles, which, given the complexity of the cortical response to pattern onset stimulation seen in the last chapter, are not likely to be sufficient. Ossenblok and Spekreijse (1991) used a spatio-temporal model consisting of several dipoles of fixed location and orientation, the number of dipoles being determined from the dimensionality of the recorded evoked potentials. The potentials were elicited using half-field, quarter-field and octant stimuli whose radii subtended 4°. The stimuli appeared for 300ms every 800ms. Check size was 24' or 48' and contrast was 80%. Evoked responses were recorded using 24 scalp electrodes of 40mm spacing, and the data were analysed using principal components analysis. The maximum time window chosen for the analysis was from 58 to 204ms after stimulus onset.

The large stimuli and high contrast checks elicited responses which contained two major positive peaks. The first peak, at a latency of around 80ms, corresponded to CI, and the second, at around 140ms, corresponded to CIII. Principal components analyses showed that two sources underlay these peaks. The first source had a triphasic timecourse and was mainly responsible for CI. Meanwhile, the second source had a biphasic timecourse and contributed mainly to CIII. The location of the early source varied depending on the location of the stimulus within the visual field. With left and right half-field stimulation, the early source was contralateral and oriented mainly radially. The late source had about the same position, but was predominantly tangential. Both early and late sources were located at an azimuthal angle (relative to the inion) of greater than about 15°, indicating an extrastriate origin.

Stimulation of lower quadrants of the visual field yielded an early source which was more lateral than that for half-field stimulation, but which maintained a radial orientation. Meanwhile, stimulation of the upper quadrants produced a tangential early source. This suggested that the early source was located in extrastriate cortex on the
dorsal surface when the lower quadrant was stimulated and on the ventral surface when
the upper quadrant was stimulated. The position and orientation of the late source were
the same for both upper and lower quadrant stimulation. This indicated that the late
source was located in one of the relatively small functional areas within area 19, the
small extent of the visual area and its large receptive fields causing that the same current
source was observed for all of the stimuli.

Octant stimuli presented within the lower half-field produced current sources
which were predominantly radial, while those presented in the upper half-field produced
tangential current sources. Contrary to the results of Butler et al. (1987), sources
produced by stimuli either side of the horizontal meridian were not opposite in direction,
but sources activated by stimuli either side of the vertical meridian were. Stimulating
octants in the upper half-field produced early current sources which were deeper than
those for octants in the lower half-field. These findings supported the results of
quadrant stimulation, suggesting that the early source, which was responsible for CI,
was located in area 18 on the dorsal and ventral surfaces of the occipital lobe.

Ahlfors et al. (1992) examined cortical retinotopy by analysing pattern onset
visual evoked magnetic responses using equivalent current dipoles and minimum norm
estimation. They stimulated the eight octants of the visual field using a sectorial foveal
stimulus whose radius subtended 1.8°, and an annular parafoveal stimulus, which
extended from 1.8 to 3.8°. The checks at the centre of the stimulus had a radial width
of 12', but this increased to 36' towards the periphery of the stimulus so as to evoke the
largest possible magnetic response. The contrast of the checks was 100%, and the
stimulus was presented for 250ms during a random interstimulus interval of between
850 and 1250ms. A 24-channel magnetometer was used to record the evoked magnetic
fields. The magnetometer measured the gradient of the radial magnetic field relative to
two mutually perpendicular tangents to the scalp, i.e. it measured $\partial B_r / \partial \theta$ and $\partial B_r / \partial \phi$.
In general, two prominent peaks were recorded, having respective latencies of 75-100ms
and 115-165ms in one subject and 105-135ms and 170-200ms in another. The two field
gradients at each location peaked at different latencies, e.g. 80 and 100ms, indicating
that more than one source was responsible for the magnetic response.

Minimum norm estimates were calculated for latencies earlier than 100ms using a
two-dimensional curvilinear source space which covered part of a spherical surface.
This surface was located at a radius of 60mm, 30mm below the surface of the scalp.
Although this approach produced satisfactory current estimates, the two-dimensional
source space amounted to a systematic error, since the cortical current was known to be
distributed throughout a three-dimensional volume. Nevertheless, using a two-
dimensional source space was an effective method of overcoming problems of depth
bias. The minimum norm estimates were compared with single equivalent current
dipoles, and the two different techniques gave comparable results. The site of maximum
primary current was generally contralateral to the stimulus location, but apart from this,
the orientation of the primary current flow was inconsistent with the cruciform model.
Neither was any consistent difference observed between the estimates for stimulation within the upper and lower or left and right half-fields. This may have been due to analysing all responses at a fixed latency, rather than identifying a specific peak, or due to several cortical regions being simultaneously active. Hence, although this study was a sophisticated examination of visual evoked magnetic responses, it was somewhat inconclusive.

Most recently, Aine et al. (1993) reported a study of cortical retinotopy using spatio-temporal analyses of evoked magnetic responses. The pattern onset stimuli consisted of circular targets which were displayed at eccentricities of between 3° and 12° in the right half-field. The stimuli were visible for 266ms and the repetition rate was random, with a mean rate of 1Hz. The diameter of each target varied between 0.4° and 1.0° according to the eccentricity at which it was presented, being adjusted according to cortical magnification. Evoked magnetic responses were recorded using a seven-channel magnetometer which was positioned at several locations for each stimulus. A five-dipole spatio-temporal source model (see chapter four) was used to analyse the responses between selected latencies (e.g. 80-120ms or 80-150ms). Three to five sources were identified in each subject and plotted on appropriate magnetic resonance scans. These sources were located in V1, V2, left occipito-temporal cortex and left medial parieto-occipital cortex. An ipsilateral source was also observed. The V1 and V2 sources were examined in most detail. It was found that the V1 source became progressively deeper as the stimulus target was positioned more peripherally. The V2 source was closest to the V1 source when the stimulus was close to the vertical meridian, and furthest from V1 when the stimulus was near to the horizontal meridian. Sources produced by stimuli in the lower right quadrant of the visual field were more dorsal when the stimulus was presented near to the horizontal meridian than when it was presented near to the vertical meridian. These observations confirmed the retinotopic organisation of area V1 and the inverted, mirror-image organisation of area V2.

As was mentioned at the end of the previous chapter, visual evoked magnetic responses yield no information about radial current sources, and a current estimate based on magnetic data alone is therefore an incomplete image of the cortical response. Meanwhile, visual evoked potentials reflect both radial and tangential sources, but yield less precise current estimates than magnetic data. Chapter five showed that optimal current images are obtained when both electric and magnetic responses are used simultaneously. Hence, this chapter describes a study of cortical retinotopy based on source analyses of both electric and magnetic data simultaneously. This chapter is an advancement of the previous one in two respects. Firstly, it uses both electric and magnetic responses, and secondly, it uses smaller stimuli, which are designed to reduce cancellation effects, thereby allowing interfissural currents to be observed. As with previous chapters, equivalent current dipole fitting, minimum norm estimation and anatomically constrained distributed-source analysis form the basis of the study.
Methods

Subjects Three male subjects (IM, KS and PF) and three female subjects (AW, FF and CS) volunteered to participate in the experiment. The ages of these subjects ranged from 23 to 37 years. All subjects had a corrected visual acuity of better than 6/6.

Stimulus Each octant of the visual field was stimulated in turn using a sectorial pattern onset stimulus whose radius subtended 8° at the eye of the viewer (figure 111). The stimulus was generated using a Cambridge Research Systems Visual Stimulus Generator board, which was installed into a 386 personal computer. The graphics board was programmed using Borland C++ in conjunction with the function libraries supplied by Cambridge Research Systems. An achromatic checkerboard pattern was generated on an Eizo Flexscan T560i video monitor. The checkerboard had 60’ checks of contrast 94%, which were present for 200ms and absent for the following 800ms. At all times, the mean luminance of the screen was 25cdm⁻². Sectorial stimulation was achieved by excluding the unwanted part of the checkerboard with a black mask. Fixation was facilitated by a small green spot which subtended 15’ and was located at the apex of the stimulus sector, on the mask. Note from figure 111 that the stimulus sector had an included angle of only 30°; this was to ensure that the stimulus remained within the appropriate octant of the visual field despite minor rotations caused by misalignment of

![Diagram](image)

Figure 111. Details of the sectorial stimulus, showing the position of octants 1-8 within the visual field. F: fixation spot.
the subject's head. The large, high-contrast checks were chosen to elicit CI and CLm peaks of high amplitude. Previous studies such as those of Darcey et al. (1980a,b) and Butler et al. (1987) indicated that CI was most likely to arise from a retinotopically mapped visual area. Similarly, the results of the previous chapter showed CLm to be the product of area V1 and possibly V2. Since the purpose of sectorial stimulation was to examine cortical retinotopy, it was therefore appropriate to enhance and analyse CI and CLm.

For both electric recording (evoked potentials) and magnetic recording (evoked magnetic responses), the subject viewed the stimulus from within a magnetically shielded room. Although magnetic shielding was not strictly a requirement of the electrical recording procedure, the subjects were seated in exactly the same position for both types of recording in order to ensure that the stimulus conditions were identical. The stimulus monitor was located just outside a $350 \times 350$mm aperture in the magnetic shielding. The subject was seated on a stool and leaned forward to rest their forehead on a purpose-built viewing table, so that they were looking downwards into a $45^\circ$ mirror reflecting an image of the monitor. The total distance from the monitor to the eye was 1.5m. Vertical inversion of the stimulus was taken into account when considering the octant of the visual field stimulated by each stimulus sector.

Evoked Potential Recording Each of the eight octants of the visual field was stimulated in turn and evoked potentials were recorded from a 20-point occipital montage (figure 112). The horizontal spacing of the electrodes was 10% of the half-circumference of the head and the vertical spacing was 10% of the nasion-inion distance. A reference electrode was situated at Fz of the international 10-20 system, and a ground electrode was attached near the vertex. The locations of the 20 electrodes were recorded relative to a head-centred coordinate system based on the nasion and left and right pre-auricular points. The origin of this coordinate system was located midway between the left and right pre-auricular points, and the x-axis projected from the origin through the nasion. The y-axis then passed through the left pre-auricular point and the z-axis was directed upwards towards the vertex. The locations of the electrodes were measured using a Polhemus 3Space Isotрак. This device consisted of a magnetic field source, which was fixed in position, and a magnetic field detector probe, whose position relative to the field source could be calculated from the detected signal. By locating the probe on the nasion and the two pre-auricular points, and then on each of the electrodes in turn, the Cartesian coordinates of the electrodes could be found relative to the chosen coordinate system. The accuracy of this method was found by simple geometric tests to be ±3mm.

The subject was instructed to fixate on the fixation spot and a BioLogic Brain Atlas EEG machine was used to average 50 evoked potentials. The signals were bandpass filtered between 0.3 and 30Hz, and a 50Hz notch filter was also used. Artefactual epochs were rejected. The Visual Stimulus Generator produced a positive-edge signal each time the stimulus pattern was presented, and this signal was used to
Figure 112. Occipital electrode montage, showing the relationship of the electrodes to the 10-20 electrode sites. The horizontal spacing of the electrodes is 10% of the half-circumference of the head, and the vertical spacing is 10% of the nasion-vertex-inion distance.

trigger the EEG machine. After a short break, a second 50 epochs were recorded, and the two sets of 50 epochs were subsequently off-line averaged using the facility provided on the Brain Atlas.

Each average of 100 epochs was obtained from the EEG machine in ASCII format and converted to a binary format suitable for use in the analysis procedure. Since the EEG machine was known to have a filter delay of 8ms, the latencies of the signals were reduced by 8ms. This was necessary to ensure that the electric and magnetic records were correctly synchronised. The reference was also transferred from Fz to the inion by subtracting the potential at the inion from the potential at all of the other electrodes in the montage.

*Evoked Magnetic Response Recording* The subject was seated in the magnetically shielded room as described previously, resting their forehead on the viewing table and looking downwards. This meant that the occipital region of the head faced upwards in a suitably accessible position for recording. The magnetometer was positioned vertically and the base of the dewar was adjusted so as to touch (but not rest upon) the back of the head. The dewar was centred on Oz, although due to its size, it was not possible to locate it accurately. The Polhemus 3Space Isotrac system was therefore used to
determine the exact position and orientation of each of the magnetometer channels. The magnetic field source was attached to the dewar at a predetermined location and the field detector probe was positioned on the nasion and two pre-auricular points in turn. Computerised geometrical calculations then yielded the location and orientation of the magnetometer channels, relative to the head-centred coordinate system.

The magnetometer itself was a system of 19 second-order gradiometers, designed and constructed by the Russian Academy of Sciences (Harding et al. 1994). The 19 channels were arranged in a hexagonal array over a flat plane (figure 113). The baseline of the individual gradiometers was approximately 50mm, although the individual turns of the gradiometer coils were separated slightly so as to reduce the self-inductance (table 10). The pickup coils were located about 14mm from the external surface of the dewar base. Note that because the gradiometers were arranged over a flat plane, the outer channels of the array were located about 20mm from the scalp and had a significant tangential component in their orientation. The white noise level of the gradiometers was approximately 30fT/√Hz, and the peak-to-peak amplitude of 50Hz environmental noise was around 20pT.

In addition to the 19 gradiometer channels, the Russian system contained three vector magnetometers, measuring the magnetic flux (as opposed to its gradient) in three mutually orthogonal directions. The vector magnetometers were located approximately 100mm further up the dewar than the gradiometer pickup coils, and because of their

![Figure 113. Gradiometer montage, showing its approximate position relative to the 10-20 electrode sites.](image-url)
<table>
<thead>
<tr>
<th>Coil offset, relative to the pickup coil (mm)</th>
<th>Coil diameter (mm)</th>
<th>No. of turns</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>15.0</td>
<td>1</td>
</tr>
<tr>
<td>1.5</td>
<td>15.0</td>
<td>1</td>
</tr>
<tr>
<td>51.0</td>
<td>15.0</td>
<td>−1</td>
</tr>
<tr>
<td>56.0</td>
<td>15.0</td>
<td>−1</td>
</tr>
<tr>
<td>61.0</td>
<td>15.0</td>
<td>−1</td>
</tr>
<tr>
<td>66.0</td>
<td>15.0</td>
<td>−1</td>
</tr>
<tr>
<td>115.5</td>
<td>15.0</td>
<td>1</td>
</tr>
<tr>
<td>117.0</td>
<td>15.0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 10. Details of the gradiometer coils in the 19-channel magnetometer.

large separation from the head, they merely measured the three components of the environmental noise in the magnetically shielded room.

The analogue output of the magnetometer system was low-pass filtered between 0 and 106Hz, and then converted to digital form at a frequency of 1kHz using a 32-bit analogue-to-digital converter. Recording and signal processing were then performed on a 486 personal computer. For each stimulus, the subject was instructed to concentrate on the fixation spot and 50 epochs were recorded. After a short break, a further 50 epochs were recorded. The records were stored for off-line filtering and averaging. The first stage in this process consisted of active filtering. This procedure subtracted a multiple of each vector magnetometer signal from the gradiometer signals. Since the vector magnetometers measured the environmental noise in the magnetically shielded room, subtracting their output from the gradiometer outputs reduced the noise in the recorded signals. The resulting signals were then digitally filtered between 0.3 and 30Hz. Finally, the 100 epochs were averaged. At the time when the recording was carried out, some of the 19 gradiometers were malfunctioning, so the associated data, characterised by excessive noise levels, were rejected. Depending on the subject, this left between 13 and 17 channels of data for analysis. The low-pass analogue filters introduced a delay of 4ms into the arrival of the signals, so 4ms were subtracted from the latencies to maintain synchronisation with the evoked potentials. The output of the data acquisition program (in fT) was converted into units of flux (atto-Weber; 10⁻¹⁸Wb; aWb) by multiplying the signal amplitudes by the area of the pickup coil. This conversion was necessary to make the data compatible with the source analysis procedure.

The complete experimental protocol was as follows:

1. The subject was seated in the magnetically shielded room and the dewar was positioned. The subject then kept his/her head as still as possible so that the gradiometer coordinates remained constant throughout the recording.
2. The door of the magnetically shielded room was closed and the bias voltage of the magnetometer was adjusted so that the modulated output of the SQUID lay within the dynamic range of the SQUID amplifier.

3. Two sets of 50 epochs were recorded for each of four sectorial stimuli.

4. The door of the magnetically shielded room was opened, the Polhemus source was attached to the dewar, and the coordinates of the magnetometer recorded. The subject was then allowed to move freely for several minutes.

5. The above procedure was repeated for the remaining four octants of the visual field.

The total procedure required about 40 minutes.

**Peak Identification** Peaks in the evoked potential waveforms were identified using global field power. This was defined simply as the sum of the squares of the evoked potentials at a given latency. Peaks in a millisecond-by-millisecond plot of global field power identified times of global electrical activity. Since electric global field power was reference-dependent, it was calculated using Fz as reference, before the reference was transferred to the inion.

Peaks in the evoked magnetic responses were identified using magnetic global field power. This was defined as the sum of the squares of the magnetometer outputs at a given time instant. Times of global magnetic activity were determined by plotting magnetic global field power at a range of latencies and observing the maxima.

The above methods were adequate when electric or magnetic measurements were analysed separately. However, a peak in the electric global field power did not necessarily occur at the same latency as the corresponding peak in the magnetic global field power. Moreover, neither the electric peak nor the magnetic peak necessarily represented a maximum in source strength: a strong source might produce only weak fields due to its location and orientation, and *vice versa* (*c.f.* Ossenblok and Spekreijse 1991). It was therefore necessary to find a single method of peak identification which (1) was global to both electric and magnetic responses, and (2) which measured source strength rather than field strength. A suitable method was devised and denoted "global source power". This technique involved carrying out a minimum norm estimate at each latency in the evoked response, using both electric and magnetic data (details are given in the following subsection). At each latency, the total power dissipated by the primary current was found by summing the squares of the estimated dipole moments over the entire source space grid. The resulting quantity was equal to the square of the Euclidean norm of the solution vector, which was minimised by the estimation procedure. Hence, the "global source power" was proportional to the power dissipated...
by the "smallest" current distribution which could account for both electric and magnetic responses. Because global source power was minimised by the estimation procedure, false power peaks were naturally suppressed. The technique provided a means of unifying both electric and magnetic responses, and also provided an indication of activity in the source domain rather than in the external signal domain, which was important as the experiment was conducted to determine source features. Source analyses were carried out for latencies at which global source power was maximal. A similar method, using a two-dimensional source space, was applied to visual evoked magnetic responses by Ioannides et al. (1989).

The pattern onset stimuli evoked a variety of peaks in the plots of electric and magnetic global field power and global source power. The earliest peak, corresponding to CI/CIm, had the largest amplitude due to the large stimulus checks. This peak was selected for source analysis as it was most likely to represent a response in area 17.

Source Analysis  The recorded data were analysed using equivalent current dipole fitting, minimum norm estimation and anatomically constrained distributed-source analysis, as described previously. In order to do this, it was necessary to convert the electrode and magnetometer coordinates from the head-centred coordinate system used for the recording into the occipital coordinate system used for the analyses. The centre of the occipital region of the head was found by fitting a sphere to the electrode coordinates, using a least-squares algorithm. A coordinate transformation was then applied to the electrode and magnetometer sites, corresponding to the following operations: (1) movement of the head-centred origin to the occipital origin, (2) a rotation about the z-axis of approximately 180°, required to locate the inion in the xz-plane, and (3) a rotation about the new y-axis sufficient to cause the new x-axis to pass through the inion. The resulting coordinate system was that required for source analysis. Note that the occipital sphere-fitting procedure also conveniently defined the radius of each subject's head. Since the electrode coordinates related to the top surfaces of the electrodes, the electrode thickness (2mm) was subtracted from the radius of the fitted sphere to give the scalp radius. The compartments of the conductivity model were assumed to have thicknesses equal to those of table 1, but the radii were fixed by the scalp radius.

The source region for equivalent current dipole fitting was simply the innermost sphere of the conductivity model, the radius of which was defined by subtracting the thicknesses of the scalp, skull, cerebrospinal fluid and grey matter from the radius of the head. Minimum norm estimation was carried out using a source space which was scaled to suit each individual's head. The angular spacing of the discrete dipoles was the same as in the previous chapter, i.e. θ ranged from 20/65 radians to 110/65 radians and φ ranged from -45/65 radians to +45/65 radians. However, the radii of the shells were scaled according to the ratio of the innermost sphere radius to the standard radius of 66mm. For example, subject AW had an external head radius of 79mm, which, after
subtracting thicknesses of 6mm for the scalp, 5mm for the skull, 2mm for the cerebrospinal fluid and 2mm for the grey matter, gave an innermost sphere radius of 64mm. Hence, the standard radii of the source space shells, 25, 35, 45, 55 and 65mm, were all multiplied by 64/66. This scaling ensured that the volume of the source space matched the volume of each subject's brain.

Similar scaling was also applied to the anatomically constrained source space. In this case, the Cartesian coordinates of the discrete dipoles were all scaled according to the ratio of the innermost sphere radius to the standard radius. For example, in subject AW, the position coordinates given in table 3 were all multiplied by 64/66. This operation was equivalent to using a proportional coordinate system such as that described by Talairach and Tournoux (1988), and ensured that the model brain matched each real brain as closely as possible.

As previously, it was necessary to estimate the standard deviation of the noise in the electric and magnetic records. The standard deviation of each record was found by making two measurements of the maximum peak-to-peak signal amplitude at latencies remote from those of the evoked response. The two measurements were made on any of the recording channels; the noise level was assumed to be equal for all of the electrodes or gradiometers. In the electric case, the measurements were made using the inion-referenced data. The peak-to-peak amplitudes were divided by four, on the assumption that the extreme noise fluctuations lay two standard deviations from the mean signal level. This procedure was repeated for the eight electric and eight magnetic records obtained from each subject. Since the eight stimuli yielded estimates of noise standard deviation which were almost constant (highest and lowest estimates differed at most by a factor of two), the mean standard deviation was found and this was used for all eight records. By this method, the electric standard deviation was found to be between 0.4 and 0.8μV for the six subjects, and the magnetic standard deviation was found to be between 14 and 22nWb (80-120fT).

The equivalent current dipoles for subject FF were superimposed on magnetic resonance images of her head. At the time of magnetic resonance scanning, vitamin capsules were taped to the nasion and pre-auricular points so that these landmarks could be located on the images. The head-centred coordinate system was then constructed from these landmarks, as described earlier. The dipole coordinates were transformed from the occipital coordinate system to the head-centred coordinate system and plotted on relevant sagittal and coronal sections.

Results

Signals Figure 114 is a sample of the evoked potential elicited by sectorial stimulation. The onset response is biphasic, consisting of a positive deflection, CI, at around 90ms, and a negative deflection of comparable amplitude, CIII, at around 150ms. A positive
offset peak also occurs at around 330ms, although this is generally of lower amplitude than either of the onset peaks. Figure 115 shows the corresponding evoked magnetic response. The waveforms vary considerably between recording channels, but two substantial peaks are present: CIm, at a latency of about 90ms (most clearly visible on channel 1), and CIIIm, at around 140ms (channel 18). Due to the noise level, it is not possible to conclusively identify an offset peak, although a deflection occurs at about 330ms on channel 1, which could be such a response.

Figure 116 compares electric responses obtained from a single recording channel for all eight sectorial stimuli. Although the waveforms are somewhat variable, a positive peak and a later negative peak are both visible on most of the traces. The peaks generally occur at around 110 and 170ms respectively, and correspond to CI and CIII. The variability of the waveforms between stimuli is probably due to the change in the stimulus shape, but the signals are also affected by noise. Magnetic responses to all stimuli are shown in figure 117. The waveforms are clearest for octants 3, 4 and 5, showing a negative deflection (ingoing field), CIm, at around 90ms, and a positive deflection (outgoing field), CIIIm, at around 140ms. Note that these waveforms are significantly affected by noise. This is because the stimulus is relatively small, thus producing a small response, and because the total noise level of the magnetometer (standard deviation approximately 16aWb or 100fT) is fairly high.

Global Field Power  Figure 118 shows the electric global field power for the response of a single subject to all eight stimuli. The plots are highly variable, but each plot shows one or more of three distinct peaks. The first peak, at around 100ms, corresponds to CI, the second, at around 170ms, corresponds to CIII, and the third, at around 330ms, is the offset peak. The CI peak occurs with stimuli 1, 3, 4, 5, 6 and 8, being absent only with stimuli 2 and 7, while the other peaks are less consistent. The magnetic global field power (figure 119) shows similar trends: the CIm peak appears in most of the responses at around 90ms, while the later peaks are often absent and/or variable in latency. The magnetic global field power of a second subject (figure 120) illustrates clearly the persistence of the initial peak and its fairly consistent latency of around 80ms.

Global Source Power  Figure 121 is an example of the global source power in the response of one subject to all eight stimuli. It can be seen that global source power is negligible until a latency of about 60ms, when it increases to produce the most prominent peak, which corresponds to CI/CIm. (Hereafter, peaks will simply be referred to using the electric terminology; both electric and magnetic modalities are implied.) The first peak is present for stimuli 2, 4, 5, 6 and 7, has a reduced amplitude and slightly increased latency for stimulus 3, and is absent for stimuli 1 and 8. Absence of a clear response from octants 1 and 8 is common among the six subjects: it appears that these octants do not evoke particularly large electric and magnetic responses, consistent with the conclusion of Lesèvre and Joseph (1979) and Ossenblok
Figure 115. Visual evoked magnetic responses produced in subject AW by stimulus 4. Horizontal axis: \(-4 \rightarrow 507\) ms; vertical axis: \(-193 \rightarrow 193\) aWb. Outgoing field upwards.
Figure 116. Visual evoked electric responses to the eight stimuli, recorded from channel 16 on subject AW. Horizontal axis: -8 → 503ms; vertical axis: -6 → 6μV approx. Fz reference. Positive upwards.
Figure 117. Visual evoked magnetic responses to the eight stimuli, recorded from channel 5 on subject AW. Horizontal axis: 0 → 507ms; vertical axis: -70 → 70aWb approx. Outward field upwards.
Figure 118. Electric global field power for the response of subject AW to the eight stimuli. Horizontal axis: -8 → 503 ms; vertical axis: 0 → $1.6 \times 10^{-11}$ V$^2$ approx.
Figure 119. Magnetic global field power for the response of subject AW to the eight stimuli. Horizontal axis: -4 → 507ms; vertical axis: 0 → 2.5×10^{-33}Wb² approx.
Figure 120. Magnetic global field power for the response of subject FF to the eight stimuli. Horizontal axis: $-4 \rightarrow 507\text{ms}$; vertical axis: $0 \rightarrow 1.2 \times 10^{-3}\text{Wb}^2$ approx.
Figure 121. Global source power as a function of latency for the response of subject AW to the eight stimuli. Vertical axis marked in units of $A^2 m^2$. 

316
Spekreijse (1991) that the lower visual field produces a larger response than the upper.

The later peaks are variable in latency and amplitude. For instance, stimulus 2 produces three onset peaks, at around 80, 130 and 190ms, but stimulus 4 produces just two peaks, at 80 and 140ms respectively. Stimuli 2-6 all produce an offset peak at around 320ms. The relative consistency of the CI peak, compared to the other peaks, is one reason why CI is used for source analyses.

Discontinuities or "spikes" are observed at several latencies, notably with stimuli 3 and 5. The global source power plots are produced using regularised minimum norm estimation, and the discontinuities are a product of the regularisation process. For given signal and noise levels, a certain number of inverted singular values, \( \lambda_i^j (r) \), are non-zero (see equation 157). As the signal level changes with time, so the number of non-zero \( \lambda_i^j \) values increases or decreases discretely. Whilst this does not dramatically affect the spatial distribution of the primary current, it does influence the Euclidean norm of the solution. This is seen as a "step" in the global source power. If the maximum signal happens to fluctuate around the value required to change the number of non-zero singular values, then some oscillation in the global source power can occur. This problem could be overcome by using the same number of singular values throughout the entire latency range of the global source power plot. However, if the signal-to-noise ratio determining the cut-off of the singular value spectrum were chosen to be that of one of the response peaks, then the response peaks would be analysed correctly, but periods of noise in-between would be analysed as though the noise were a genuine response, and the global source power would therefore vary randomly during these periods, obscuring the true peaks. Conversely, if the signal-to-noise ratio were chosen to be that of the noise, the global source power would be zero during the noise periods, but also zero at the latency of an evoked response. Hence, it is essential that the singular value spectrum is truncated separately at each latency, so the discontinuities in the global source power cannot be overcome.

Figure 122 compares the electric global field power, magnetic global field power and global source power produced in a single subject by stimulation of one octant. Although all three plots show similar CI, CIII and offset peaks, the precise latencies of these peaks differ by up to 20ms. Note particularly that the first electric peak is wider than the corresponding magnetic peak, and that all the electric peaks are delayed by about 15ms relative to the magnetic peaks (despite correction for filtration differences between the two recording systems). The latency differences between corresponding peaks in the electric global field power and magnetic global field power clearly illustrate the need to identify the true latency of events within the brain, a requirement which is fulfilled by the global source power. The CI peak of the global source power corresponds closely to the first peak in the magnetic global field power, whereas the CIII peak of the global source power coincides with neither of the global field power peaks. The offset response is only distinguishable in the electric global field power, and the offset peak in the global source power therefore occurs at the same latency.
Figure 122. Electric global field power (E), magnetic global field power (M), and global source power (S) for the response of subject AW to stimulus 4.

*Equivalent Current Dipoles for the Six Subjects* For each subject and stimulus, the latency of CI was identified using global source power. An equivalent current dipole was fitted to the electric and magnetic measurements at this latency. Figures 123-128 show the results of these dipole fits for all subjects, and table 11 gives the corresponding numeric data. Absence of graphic or numeric data indicates that a clear CI peak could not be identified in the global source power. As described in chapters five and seven, a goodness-of-fit of $10^{-20}$ or greater indicates that a dipole is acceptable, while a goodness-of-fit of less than this shows that the single dipole model is inadequate. Dipole moments of greater than 200nAm are physiologically implausible.

Before the equivalent current dipoles can be interpreted in terms of the occipital anatomy, it is necessary to define a working hypothesis which specifies what the expected results might be. This hypothesis contains the following elements:

1. We begin by assuming that the response is generated in V1, which is retinotopically mapped according to the cruciform model.

2. Stimulation of octants 2, 3, 6 and 7 should then yield sources entirely within the calcarine fissures.

3. Stimulation of octants 1 and 8 should produce sources on the lingual gyri.
Figure 123. Equivalent current dipoles for the response of subject IM to the eight stimuli. The axes are those of figure 28 and the circles represent the outer surface of the scalp. The coronal view is from the posterior. Note that the locations of the nose and ears are schematic only.
Figure 124. Equivalent current dipoles for the response of subject KS to the eight stimuli.
Figure 125. Equivalent current dipoles for the response of subject PF to the eight stimuli.
Figure 126. Equivalent current dipoles for the response of subject AW to the eight stimuli.
Figure 127. Equivalent current dipoles for the response of subject FF to the eight stimuli.
Figure 128. Equivalent current dipoles for the response of subject CS to the eight stimuli.
<table>
<thead>
<tr>
<th>Subj.</th>
<th>Oct.</th>
<th>$x$ (mm)</th>
<th>$y$ (mm)</th>
<th>$z$ (mm)</th>
<th>$q_x$ (nAm)</th>
<th>$q_y$ (nAm)</th>
<th>$q_z$ (nAm)</th>
<th>$q$ (nAm)</th>
<th>Fit (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>55</td>
<td>-14</td>
<td>22</td>
<td>-8.9</td>
<td>-17.8</td>
<td>-13.2</td>
<td>23.9</td>
<td>58.2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>51</td>
<td>-20</td>
<td>30</td>
<td>-4.1</td>
<td>-14.3</td>
<td>4.2</td>
<td>15.5</td>
<td>89.5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>61</td>
<td>-12</td>
<td>32</td>
<td>60.7</td>
<td>-27.8</td>
<td>35.3</td>
<td>75.5</td>
<td>91.6</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>59</td>
<td>15</td>
<td>36</td>
<td>-20.1</td>
<td>-7.3</td>
<td>1.0</td>
<td>21.4</td>
<td>95.9</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>56</td>
<td>19</td>
<td>29</td>
<td>30.9</td>
<td>20.5</td>
<td>34.6</td>
<td>50.7</td>
<td>41.7</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>54</td>
<td>12</td>
<td>25</td>
<td>-53.7</td>
<td>-9.7</td>
<td>-40.4</td>
<td>67.9</td>
<td>22.3</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>53</td>
<td>23</td>
<td>32</td>
<td>4.0</td>
<td>7.7</td>
<td>15.7</td>
<td>17.9</td>
<td>74.0</td>
<td></td>
</tr>
<tr>
<td>KS</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>79</td>
<td>-24</td>
<td>6</td>
<td>-6.8</td>
<td>-7.7</td>
<td>-1.3</td>
<td>10.4</td>
<td>75.1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>78</td>
<td>-22</td>
<td>19</td>
<td>-32.1</td>
<td>0.4</td>
<td>-18.8</td>
<td>37.2</td>
<td>30.4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>80</td>
<td>-15</td>
<td>-16</td>
<td>-44.4</td>
<td>3.1</td>
<td>10.0</td>
<td>45.6</td>
<td>10^2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-1</td>
<td>18</td>
<td>-3</td>
<td>-62.7</td>
<td>280.5</td>
<td>-60.3</td>
<td>293.7</td>
<td>10^4</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>71</td>
<td>42</td>
<td>10</td>
<td>61.2</td>
<td>56.7</td>
<td>9.7</td>
<td>84.0</td>
<td>9.5</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>57</td>
<td>49</td>
<td>8</td>
<td>77.0</td>
<td>74.5</td>
<td>8.4</td>
<td>107.5</td>
<td>99.1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>66</td>
<td>37</td>
<td>12</td>
<td>106.4</td>
<td>68.0</td>
<td>30.4</td>
<td>129.9</td>
<td>78.5</td>
<td></td>
</tr>
<tr>
<td>PF</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>48</td>
<td>-11</td>
<td>35</td>
<td>-21.5</td>
<td>9.0</td>
<td>-23.9</td>
<td>33.4</td>
<td>89.4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>44</td>
<td>-17</td>
<td>43</td>
<td>5.5</td>
<td>3.1</td>
<td>-2.7</td>
<td>6.9</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>55</td>
<td>-12</td>
<td>30</td>
<td>54.4</td>
<td>-15.6</td>
<td>36.6</td>
<td>67.4</td>
<td>10^2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>-64</td>
<td>23.0</td>
<td>3.8</td>
<td>542.0</td>
<td>542.5</td>
<td>94.7</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>42</td>
<td>12</td>
<td>23</td>
<td>13.8</td>
<td>8.1</td>
<td>24.1</td>
<td>28.9</td>
<td>39.0</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AW</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>-3</td>
<td>-2</td>
<td>-63.5</td>
<td>9.2</td>
<td>-14.0</td>
<td>65.7</td>
<td>10^2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>62</td>
<td>-12</td>
<td>9</td>
<td>2.5</td>
<td>1.6</td>
<td>11.2</td>
<td>11.6</td>
<td>10^4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>63</td>
<td>-10</td>
<td>-3</td>
<td>106.7</td>
<td>-29.9</td>
<td>7.7</td>
<td>111.1</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>59</td>
<td>4</td>
<td>0</td>
<td>-32.2</td>
<td>7.7</td>
<td>11.5</td>
<td>35.0</td>
<td>10^4</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>51</td>
<td>17</td>
<td>-5</td>
<td>120.1</td>
<td>54.9</td>
<td>-2.8</td>
<td>132.1</td>
<td>13.9</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>61</td>
<td>6</td>
<td>-3</td>
<td>-48.5</td>
<td>-3.4</td>
<td>-10.7</td>
<td>49.8</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>48</td>
<td>-1</td>
<td>2</td>
<td>1.2</td>
<td>-5.8</td>
<td>10.1</td>
<td>11.7</td>
<td>88.7</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>41</td>
<td>-5</td>
<td>12</td>
<td>-77.9</td>
<td>3.7</td>
<td>-36.9</td>
<td>86.3</td>
<td>97.5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>44</td>
<td>-9</td>
<td>16</td>
<td>-6.1</td>
<td>-14.9</td>
<td>7.1</td>
<td>17.6</td>
<td>64.1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>43</td>
<td>-24</td>
<td>21</td>
<td>31.7</td>
<td>-31.4</td>
<td>20.3</td>
<td>49.0</td>
<td>32.2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>7</td>
<td>19</td>
<td>-127.0</td>
<td>-4.2</td>
<td>-59.6</td>
<td>140.4</td>
<td>10^2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>30</td>
<td>14</td>
<td>11</td>
<td>-20.0</td>
<td>4.4</td>
<td>-24.1</td>
<td>31.6</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>42</td>
<td>24</td>
<td>-2</td>
<td>133.4</td>
<td>91.5</td>
<td>-11.9</td>
<td>162.2</td>
<td>10^2</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>18</td>
<td>16</td>
<td>9.6</td>
<td>21.8</td>
<td>6.1</td>
<td>24.6</td>
<td>59.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>42</td>
<td>43</td>
<td>36</td>
<td>-7.2</td>
<td>-17.2</td>
<td>-21.4</td>
<td>28.4</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>65</td>
<td>12</td>
<td>22</td>
<td>-16.0</td>
<td>-2.8</td>
<td>-20.9</td>
<td>26.5</td>
<td>10^3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>70</td>
<td>2</td>
<td>5</td>
<td>-11.2</td>
<td>-8.4</td>
<td>-2.2</td>
<td>14.2</td>
<td>10^2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>67</td>
<td>-2</td>
<td>19</td>
<td>32.1</td>
<td>-11.7</td>
<td>18.5</td>
<td>38.9</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>57</td>
<td>20</td>
<td>35</td>
<td>36.8</td>
<td>22.1</td>
<td>33.5</td>
<td>54.5</td>
<td>41.0</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>59</td>
<td>29</td>
<td>17</td>
<td>68.9</td>
<td>44.5</td>
<td>25.4</td>
<td>85.9</td>
<td>10^4</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>62</td>
<td>20</td>
<td>24</td>
<td>-37.3</td>
<td>-21.1</td>
<td>-23.3</td>
<td>48.8</td>
<td>18.6</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>7</td>
<td>27</td>
<td>-16.6</td>
<td>18.7</td>
<td>-19.3</td>
<td>31.6</td>
<td>79.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 11. Equivalent current dipole parameters for the six subjects and eight stimuli.
4. Stimulation of octants 4 and 5 should produce sources on the cuneal gyri.

5. Depending on how far V1 extends onto the medial surfaces, stimulation of octants 1 and 4 and/or octants 5 and 8 may yield sources in the calcarine fissures rather than on the medial surfaces.

6. Area V2 may also be involved in generating the response. This will accentuate the contribution of the lingual gyrus with octants 1, 2, 7 and 8, and the cuneal gyrus with octants 3, 4, 5 and 6.

7. Current may flow towards or away from the pial surface.

8. Area V3 may also contribute to the response, although areas V4 and V5 are unlikely to contribute as the stimulus does not contain colour or motion attributes.

This hypothesis is now discussed in terms of the equivalent current dipoles for subject FF (figure 127). This subject is selected because all eight stimuli yield a Cl peak in the global source power, enabling a current dipole to be determined for each stimulus. The goodnesses-of-fit of the dipoles are also consistently high (except with stimulation of octants 5 and 7, where the goodnesses-of-fit are acceptable, but rather poor). The reader should refer to figure 28 for the approximate locations of the various surfaces of visual cortex.

Octant 1. This stimulus is expected to produce a laterally oriented current dipole on the left lingual gyrus. The fitted dipole is in the correct location, but the superior component of its orientation is difficult to interpret. Either the primary current is flowing away from the pial membrane on the lingual gyrus and towards the pial membrane on the inferior surface of the calcarine sulcus, or the current is confined to the lingual gyrus, but extends onto the ventral aspect, flowing towards the white matter. If the latter is true, then V2 must be involved.

Octant 2. According to the basic cruciform model, this stimulus is expected to produce a current source on the inferior surface of the left calcarine sulcus. The current dipole is located correctly and directed anteriorly. The direction of current flow suggests that CI is produced by current flowing towards the white matter.

Octant 3. A current source located on the superior surface of the left calcarine sulcus is expected. Current is expected to flow superiorly or inferiorly. The observed dipole location is in accord with this prediction, but the dipole is predominantly lateral in orientation. This strongly suggests that V2 on the cuneal gyrus is involved, current in both V1 and V2 being directed towards the white matter.

Octant 4. This stimulus is expected to produce primary current mainly on the left cuneal gyrus. However, the lateral location and superior direction of the observed
current dipole suggest that the superior surface of the calcarine sulcus is still significantly involved. Current is flowing towards the white matter. Note that there is now also a significant posterior component to the direction of the current dipole, which can be explained by the gradient of the calcarine fissure relative to the x-axis.

Octant 5. From here onwards, the cortical response is expected to be in the right hemisphere. Stimulating octant 5 should produce current on the right cuneal gyrus. The observed dipole lies in the correct location, but is directed inferiorly and anteriorly. This can only be explained by proposing that the superior surface of the calcarine sulcus is significantly involved in producing the current. However, if the primary current flows towards the white matter, then the dipole orientation is opposite to that expected. The anterior current flow is also somewhat large compared to the gradient of the calcarine fissure relative to the x-axis. Note that the goodness-of-fit of the dipole is only just acceptable: there are probably several closely-spaced sources which are not particularly well represented by a single current dipole.

Octant 6. This is expected to produce current on the superior surface of the right calcarine sulcus. The location and orientation of the observed current dipole are exactly in accord with this, except that primary current now appears to be flowing away from the white matter. This apparent reversal of current direction is similar to that discussed by Butler et al. (1987) for the case of a foveal stimulus.

Octant 7. The current dipole is predicted to lie on the inferior surface of the right calcarine sulcus, such that it is directed inferiorly. Instead, it is located about 15mm below the fissure and is directed laterally and posteriorly. The lateral orientation could be explained by the action of current sources within V2 on the right lingual gyrus. However, the posterior orientation is difficult to interpret. Note that the dipole goodness-of-fit is only just acceptable, so several closely-spaced sources are possibly present.

Octant 8. This stimulus is expected to produce a current source on the right lingual gyrus. The observed current dipole is exactly consistent with this, the current being directed towards the white matter.

Although these results relate to one subject only, similar trends can be observed in the results for the other subjects. In general, the cortical response is contralateral to the stimulus octant, but there are numerous discrepancies with the current sources predicted by the simple cruciform model. The discrepancies suggest that, besides area V1, area V2 is also involved in generating CI/Clm, and that current may not always be directed from the pial surface towards the white matter.

These deductions are supported by the magnetic resonance images for subject FF (figure 129). Owing to the poor signal-to-noise ratio of the recorded data, there is a considerable degree of scatter in the locations of the equivalent current dipoles. However, almost all of the dipoles (with the exception of dipole 7) lie in the vicinity of the calcarine sulcus, within areas V1 and V2. The sagittal section shows that the dipoles generally lie slightly superior to the calcarine sulcus, although this may be due to errors.
Figure 129. Magnetic resonance images of the visual cortex of subject FF, showing the locations of the dipoles for the eight sectorial stimuli. (a) Paramedial section through the left hemisphere, and (b) coronal section, viewed from the front.
in aligning the measurement frame of reference with the imaging frame of reference. Most notably, the coronal section shows that the dipoles for stimuli 1, 2, 3 and 4 are situated in the left hemisphere, while the dipoles for stimuli 5, 6, 7 and 8 are situated in the right hemisphere, i.e. the cortical response is contralateral to the stimulus location.

**Detailed Analyses for a Single Subject** The equivalent current dipoles conveniently summarise the cortical responses of all the subjects. Detailed results of more comprehensive analyses are now presented for a single typical subject, IM. Firstly, note the topography of the evoked potentials (figure 130) and evoked magnetic fields (figure 131) elicited in this subject by the sectorial stimuli. (Stimulus 1 failed to produce a significant peak in the global source power, so no distribution is shown for this octant.) The electric distributions (figure 130) exhibit either a maximum and a minimum, such as would be expected from a tangential current dipole (e.g. octant 2), or a single extremum, such as that produced by a radial dipole (e.g. octant 8). The magnetic distributions (figure 131) are generally dipolar; in some cases the extrema are asymmetrical (e.g. octant 8), but this could simply be due to the attitude of the dewar. Rather than attempt to interpret the electric and magnetic data by eye, we now proceed with the detailed analyses, resorting, where necessary, to the external data to show that the source analyses are a true reflection of the evoked electric and magnetic responses.

**Minimum Norm Estimates for a Single Subject** Figures 132-134 are minimum norm estimates of the response of subject IM to stimulus 5. All three estimates were computed at the latency of CI as identified by combined electric/magnetic global source power (82ms). However, figure 132 was obtained using evoked potentials only, figure 133 was obtained using evoked magnetic responses only, and figure 134 was obtained using both electric and magnetic responses. Figure 132 is typical of an estimate obtained from electric data: primary current is mainly deposited in the deepest layer of the source space, reflecting the bias of the lead fields towards the deep part of the volume conductor. The estimate obtained from magnetic data (figure 133) is also typical: primary current is deposited mainly in the most superficial layer of the source space. Note that although the magnetic estimate appears to be less distributed than the electric estimate, this is partly due to the wider spacing of the grid nodes at 73mm radius than at 28mm radius. (Grid points are spaced by 11mm at 73mm radius, but by only 4mm at 28mm radius.) Note also that the superficial bias of the magnetic estimate is not as strong as the deep bias of the electric estimate - some current appears in all of the source space layers with the former.

The minimum norm estimate constructed from both electric and magnetic data (figure 134) has features in common with the two separate estimates. The primary current in the deepest layer of the source space is similar to that observed in the electric estimate, except that the current has now shifted superiorly whilst retaining its overall orientation. The current in the most superficial layer of the source space resembles the
Figure 130. Visual evoked potentials produced in subject IM by the eight stimuli. (No clear peak occurs in the global source power for stimulus 1.) Colour scale marked in units of μV. Fz reference.
Figure 131. Visual evoked magnetic responses produced in subject IM by the eight stimuli. (No clear peak occurs in the global source power for stimulus 1.) Colour scale marked in units of aWb. Positivity denotes an outgoing field.
Figure 132. Minimum norm estimate of the response of subject IM to stimulus 5, obtained using electric data only. The quoted figures are the radii of the source space shells. \( \rightarrow \) \( 3 \times 10^{-4} \text{Am}^2 \). Goodness-of-fit: 69.5\%.
Figure 133. Minimum norm estimate of the response of subject IM to stimulus 5, obtained using magnetic data only. The quoted figures are the radii of the source space shells. → $4 \times 10^4$Am$^2$. Goodness-of-fit: 91.3%.
Figure 134. Minimum norm estimate of the response of subject IM to stimulus 5, obtained using both electric and magnetic data. The quoted figures are the radii of the source space shells. $\rightarrow 3 \times 10^{-4}$ Am$^2$. Goodness-of-fit: 5.4%. 
current that appears in the magnetic estimate. Hence, the combined minimum norm estimate is approximately the sum of the separate estimates. In this respect, using both electric and magnetic data to produce an unconstrained minimum norm estimate is not entirely successful, as ideally, the two modalities should provide complementary information about one current source, whereas in practice, two separate sources are reconstructed, one which accounts for the electric data and the other which accounts for the magnetic data. Although these two sources completely account for the measured data, it is unlikely that there is always one deep source and one superficial source! It was mentioned in chapter four that a magnetic estimate, for instance, fully accounts for the measurements but always predicts a superficial current distribution, which is unlikely in practice, and the combined estimates suffer from exactly the same problem. Thus, in the context of unconstrained minimum norm estimation, combined electric and magnetic measurements do not overcome the problem of depth bias. When using anatomically constrained estimation, electric and magnetic measurements can be combined to reduce depth bias and to improve the accuracy with which a source feature is reconstructed, but in unconstrained minimum norm estimation, the same does not hold because of the lack of constraint and the large number of unknowns. However, provided that they are interpreted carefully, combined estimates can provide some information about the current sources. For example, the deep and superficial primary currents in figure 134 both have the same orientation, showing that the current sources are directed superiorly.

The reader may consider it anomalous that unconstrained minimum norm estimation has been used to calculate global source power when the estimates do not give an entirely accurate representation of the primary current. The paradox is resolved by recalling that the minimum norm estimate is the most probable current distribution of all the distributions which successfully account for the electric and magnetic measurements. Thus, although the minimum norm estimate is somewhat implausible, it is the most likely solution, given that there is no prior information with regard to the number or geometry of the current sources. In this respect, the minimum norm estimate is the current distribution which should be used to deduce the global source power. As mentioned earlier, a further reason for using the minimum norm estimate is that it is the solution having the lowest Euclidean norm of all the possible solutions to the inverse problem. This prevents false peaks from appearing in the global source power. Any other estimate - such as an artificially weighted estimate - would have a higher norm, thus producing peaks in the global source power for which there was no evidence in the measured data.

Figure 135 shows minimum norm estimates of the response of subject IM to each of the eight stimuli. Since the maximum current density occurs in the least eccentric layer of the source space, only this layer is shown. These estimates are therefore predominantly due to the evoked potentials. Due to their distributed nature, it is difficult to make any conclusive deductions from them. It is not even clear that the current sources are contralateral to the stimulus locations. Whenever the primary
Figure 135. Minimum norm estimates for the response of subject IM to the eight stimuli, obtained using both electric and magnetic data. The radius of the source space shell is 28mm. (The first stimulus does not produce a clear peak in the global source power.)
current is reconstructed as a simple distribution with only one orientation, such as with stimuli 2 and 8, the current direction bears little resemblance to that predicted by the cruciform model. Hence, unconstrained minimum norm estimates using both electric and magnetic data are dismissed as being rather unhelpful.

Anatomically Constrained Analyses for a Single Subject Before using the scaled average source space for subject IM, the goodness-of-fit of a solution was tested as a function of source space rotation (figure 136). Since the goodness-of-fit was maximal for normal source space orientation, no rotation was applied to the source space in the subsequent analyses.

Anatomically constrained analyses of the electric and magnetic responses of subject IM to the eight stimuli are shown in figure 137. In general, the analyses show primary current in the calcarine fissures and on the medial surfaces, although some current appears on the ventral surfaces, notably with stimulus 4. Stimuli presented within the left half-field (octants 5-8) yield a cortical response mainly in the right calcarine fissure, but stimuli presented within the right half-field (octants 1-4) do not appear to activate the calcarine fissure much at all, and the response is less lateralised. Examining the results in detail, the following features are observed:

Octant 1. No clear peak was observed in the global source power, so no analysis was performed.

Octant 2. From the cruciform model, a source on the inferior surface of the left calcarine fissure is expected. No such source is observed, but current appears on the cuneal and lingual gyri. Assuming that this current is in the left hemisphere, then the response appears to be generated by V2. Current is probably directed from the pial surface towards the white matter.

![Figure 136](image)

Figure 136. Goodness-of-fit as a function of source space orientation for the response of subject IM to stimulus 5.
Figure 137. Anatomically constrained analyses for the response of subject IM to the eight stimuli. (No clear peak occurs in the global source power for the first octant.)

⇒ 5nAm approx.
Octant 3. A current source on the superior surface of the left calcarine sulcus is expected. This is present in the analysis, but a source on the cuneal gyrus is also present, suggesting that V2 is involved in the cortical response. Sources are also observed in the right calcarine sulcus and elsewhere.

Octant 4. A current source on the left cuneal gyrus is expected. Correspondingly, current sources are observed on the left and right cuneal gyri; it is assumed that these sources are associated with the left cuneus but are reconstructed across both gyri due to regularisation effects. Sources of high magnitude are present on both the left and right ventral surfaces. These sources are most likely to be associated with the portion of V2 which passes beneath the occipital pole. For this stimulus, the electric potential distribution consists of a broad positivity extending over the entire montage (see figure 130). Current dipoles, lying beneath the lower edge of the montage, in the ventral cortices, would explain this distribution. Applying the right-hand rule to the magnetic field distribution (figure 131), it is found that the underlying source is medial and directed towards the left side of the head, exactly in accord with the cuneal sources reconstructed by the anatomically constrained analysis. Thus, it would seem that the electric and magnetic measurements each reflect a different source population.

Octant 5. The cruciform model predicts a current source on the right cuneal gyrus. Instead, a current source appears within the right calcarine fissure. This source is consistent with primary current on the superior surface of the sulcus, directed towards the white matter. In this case, it seems that V1 is primarily responsible for the evoked electric and magnetic responses. Presumably V1 ends at the lips of the calcarine sulcus, so that octant 5 is represented within the fissure.

Octant 6. A current source is expected on the superior surface of the right calcarine sulcus. The observed current sources in the right calcarine sulcus are compatible with this suggestion, so long as regularisation is taken into account. Current sources are also observed on the cuneal gyri, compatible with the involvement of V2. Current appears to be directed towards the white matter.

Octant 7. According to the cruciform model, current should appear on the inferior surface of the right calcarine sulcus. Current density does appear here, oriented in the opposite direction to that for octant 6. The orientation of the current suggests that it has shifted from the upper to the lower surface of the fissure, but is still directed towards the white matter.

Octant 8. A current source is expected on the right lingual gyrus. A diminutive source does appear here, but most of the current density lies in the right calcarine sulcus, directed superiorly. This supports the earlier suggestion that V1 is confined mainly to the calcarine sulcus, so that octants 5 and 8 are represented within it. However, if the current is located on the inferior surface of the calcarine sulcus, then it is directed away from the white matter, which conflicts with the preceding anatomically constrained
analyses, but is consistent with the corresponding equivalent current dipole (see figure 123).

Hence, the anatomically constrained analyses show the same trends as the equivalent current dipoles: current does not necessarily follow the cruciform model and is not always directed in the same direction with respect to the cortical laminae. Most notably, current sometimes appears on the medial surface when only the calcarine fissure is expected to be active, and at other times, the calcarine fissure is active when the response is predicted to occur on the medial gyri. This strongly suggests that V1 is restricted mainly to the calcarine sulcus but that V2 contributes to the recorded response.

**Sequential Analyses for a Single Subject** Although this chapter is primarily concerned with the first major peak of the pattern onset response, it is interesting to compare the complete temporal response with that observed in chapter seven. Figures 138 and 139 show the sequential topography of the electric and magnetic evoked responses of subject IM to stimulation of octant 5, and figure 140 gives the corresponding series of anatomically constrained analyses. It is instructive to compare the electric topography, magnetic topography and source distribution at each latency; particularly interesting features are summarised below:

80ms. This is close to the latency of CI. The electric distribution shows an upper positivity and a lower negativity. The magnetic distribution shows an outward-going field on the left and an inward-going field on the right. Both of these distributions suggest superiorly directed primary current below the centre of the montage. This current duly appears in the right calcarine sulcus on the anatomically constrained analysis at 80ms.

120ms. Both the electric and magnetic distributions have reversed.
Correspondingly, primary current appears in the right calcarine fissure, having the opposite direction to at 80ms.

150ms. This is approximately the latency of CIII. The electric distribution has rotated so that there is a positivity in the lower left region of the montage and a negativity in the upper right. The magnetic distribution shows an ingoing field in the centre and an outgoing field in the lower right. These topographies suggest a current dipole directed inferiorly and to the left. The anatomically constrained analysis shows current sources on the right calcarine sulcus and on the cuneal and lingual gyri; the vector sum of these sources equates to such a dipole.

Figure 138 (overleaf). The visual evoked potential elicited in subject IM by stimulus 5. The latencies are in ms and the colour scale is marked in $\mu$V.

Figure 139 (two pages ahead). The visual evoked magnetic response elicited in subject IM by stimulus 5. The latencies are in ms and the colour scale is marked in aWb.
Figure 140. Anatomically constrained analyses for the response of subject IM to stimulus 5. ⇒ 5.3nAm.
Figures 141, 142 and 143 show the electric, magnetic and cortical responses of subject IM to stimulation of octant 6. Again, the reader may wish to examine the three distributions occurring at various latencies. Overall, the source analyses show two distinct source distributions: the CI distribution, centred in the right calcarine sulcus and lasting from about 60 to 100ms, and the CIII distribution, which is bilateral and lasts from about 130 to 170ms. The responses at representative latencies are briefly described:

80ms. The electric distribution consists of a positivity in the upper part of the montage and a weak negativity in the lower part. The magnetic distribution consists of a broad outgoing field on the left and a focal ingoing field on the right. Both of these topographies suggest that the current source is directed superiorly, although the asymmetry of the maxima and minima indicate that the source distribution is unlikely to be a simple current dipole. The source analysis bears this out: a dominant source is located in the right calcarine fissure, directed superiorly, but current density is also found on the cuneal gyri and the right ventral surface.

150ms. The electric distribution is entirely negative, having greater amplitude in the right side of the montage. The magnetic distribution contains an ingoing field at the centre of the montage and an outgoing field to the right. The electric distribution possibly suggests a current source directed towards the left side of the head, while the magnetic distribution indicates a current source directed inferiorly and to the left. In accordance with this, the source distribution shows current density in both calcarine fissures, directed inferiorly, and current density on the medial surface, directed towards the left.

Figure 141 (overleaf). The visual evoked potential elicited in subject IM by stimulus 6. The latencies are in ms and the colour scale is marked in μV.

Figure 142 (two pages ahead). The visual evoked magnetic response elicited in subject IM by stimulus 6. The latencies are in ms and the colour scale is marked in aWb.
Figure 143. Anatomically constrained analyses for the response of subject IM to stimulus 6. ➔ 7.5nAm.
Conclusions

This chapter has examined the cortical current which gives rise to the CI peak of the pattern onset evoked potential and the CIm peak of the pattern onset evoked magnetic field, with a view to observing retinotopy in the early visual areas, particularly in area V1. The previous chapter and other studies (Darcey et al. 1980a,b; Butler et al. 1987) showed that the CI and CIm peaks were most likely to reflect primary current within area V1, so these peaks were preferentially enhanced using an achromatic checkerboard of large checks. Both electric and magnetic responses were used so as to image the cortical current as comprehensively as possible.

The equivalent current dipoles show the cortical current underlying CI and CIm to be situated near to striate cortex. The goodnesses-of-fit of the dipoles are generally high, indicating that the current source is fairly localised, but it is difficult to reconcile the equivalent current dipoles with the cruciform model. Deviations of the dipole orientation from that predicted by the cruciform model can be explained by three proposals: (1) the medial extent of striate cortex is variable between subjects, so that in some cases, V1 may be confined almost exclusively to the calcarine sulcus; (2) primary current is not confined to V1, but is situated in V2 also; and (3) although primary current is generally directed from the pial surface towards the white matter, it can also flow in the opposite direction. The equivalent current dipoles are contralateral to the stimulated octant, with very few exceptions.

It should be noted that the conclusions with regard to the direction of current flow depend upon the deduced location of the current dipole being correct. For example, a superiorly directed current dipole in the calcarine fissure could represent primary current on the superior surface, directed towards the white matter, or primary current on the inferior surface, directed away from the white matter. Deductions of current direction in this chapter are based upon the assumption that the dipoles originate approximately on the surfaces predicted by the cruciform model; e.g. in the above example, if the model predicts current on the superior surface of the calcarine fissure, then the observed dipole is assumed to represent current on the superior surface of the fissure, directed towards the white matter.

These results agree with previous studies of retinotopy. Ahlfors et al. (1992) reported that equivalent current dipoles and minimum norm estimates for the CIm peak of the pattern onset evoked magnetic response showed little correspondence with the cruciform model, other than that the cortical response was contralateral to the stimulus. Butler et al. (1987) reported limited correspondence using evoked potentials, but the dipole orientations were not always as predicted. The most favourable results were produced by Darcey et al. (1980a,b), who showed that the cruciform model provided an accurate prediction of the evoked potentials to quadrantic stimuli. However, they used a red chromatic stimulus which almost certainly produced a different cortical response to that elicited by the achromatic stimulus of the present study. The finding that V2
contributes to CI and Clm is consistent with previous evidence (Ossenblok and Spekreijse 1991) that CI is produced by two source components, one of which is located in area 18 and the other of which is to be found in a higher visual area. Butler et al. (1987) observed that current flowed towards the white matter with paramacular stimulation, but then inexplicably reversed in direction with foveal stimulation. This type of current reversal has been found in the present study.

The anatomically constrained distributed-source analyses support the conclusions from the equivalent current dipoles. The cortical current underlying CI and Clm is generally found to be located in the calcarine sulci and on the medial surfaces. The active surfaces show little correspondence with those predicted by the cruciform model. It is deduced that V2 must be involved as well as V1, since medial currents sometimes persist even for stimuli which the cruciform model predicts should produce a response well within the calcarine sulcus. In general, primary current is directed from the pial surface towards the white matter, but there are several exceptions, when current density is approximately in the expected location, but is apparently directed away from the white matter. Note, however, that in the cases where current sources are spread over two parallel surfaces due to regularisation, deductions of the current direction are dependent upon an accurate assessment of which surface of the two is active, as discussed for the equivalent current dipoles.

Stimuli presented within the left half-field produce responses which are predominantly confined to the right hemisphere, but stimulating octants within the right half-field fails to produce any significant lateralisation. The coordinates of the equivalent current dipoles show greater magnitude for stimuli 5-8 than for stimuli 1-4, supporting this observation. A similar effect was also noticed in the previous chapter, although the reason for the difference in lateralisation is unclear.

In several instances, ventral current sources are reconstructed by the anatomically constrained analyses. It is unlikely that these current sources are merely artefacts of the analysis method, as their presence is warranted by the data. They could be associated with VP, although this seems to be a very advanced visual area to be involved so soon in the response. The most likely explanation of the ventral sources is that they are generated in the portion of V2 which lies below the occipital convexity.

The sequential analyses for a single subject show that CI and Clm are produced mainly by a striate current source, and that CIII and CIIIIm are generated by current flowing in the opposite direction and spreading out onto the medial surface. Assuming that the persistent current density on the right ventral surface is associated with area 18, then area 19 does not seem to be involved.

Effectiveness of the Source Analysis Methods Fitting equivalent current dipoles to the evoked electric and magnetic responses has proved to be a robust means of summarising both types of data. Because of the relatively small stimulus size, a single dipole is
generally an adequate model of the current source. Hence, the equivalent current dipoles are fairly useful.

Minimum norm estimation is less helpful. The lead fields of the electric and magnetic detectors fail to overlap sufficiently in depth, so that current is reconstructed in the deepest and most superficial layers of the source space, with little current density in-between. This could be overcome by reducing the range of radii which the source space covers, so that the lead fields of both modalities overlap. However, a depth range of about 40 mm is necessary to accommodate the whole of visual cortex and thus avoid systematic error. The deep current density, attributable to the electric measurements, is more strongly biased in depth than the superficial current density, which is attributable to the magnetic measurements. The deep current is also slightly more diffuse than the superficial current. In the same way that both electric and magnetic responses are generally distributed in accord with a common current dipole, so the current orientation is common to both superficial and deep parts of the source space.

Because of the separation of the electric and magnetic parts of the minimum norm estimates, and the diffuseness of the deep current, the estimates are of little use as images in their own right. However, the estimates are useful for determining the global source power. The minimum norm criterion ensures that all available information with regard to the source is taken into account, while unjustified peaks are not constructed. The anatomically constrained distributed-source analysis method has performed successfully using both electric and magnetic measurements simultaneously. In this chapter, there are typically 19 electric and 15 magnetic measurements, giving a total of 34 measurements, compared to 42 unknowns. Thus, although the linear system is underdetermined, there are sufficient data to avoid the problems associated with minimum norm estimation. Although the rotation method indicated that the brain of one subject was significantly different to the scaled average brain, there were otherwise no source space alignment problems. The sequential analyses show the solutions to be stable with time, despite the presence of substantial noise in the data. Since each current dipole is associated explicitly with a specific region of cortex, the results are easy to interpret in terms of the anatomy of the visual cortex.
9. Conclusions

Neuro-electromagnetism is a wide-ranging multidisciplinary field, involving aspects of physics, engineering, mathematics, neurophysiology and other disciplines. To date, there has been little coherence between the technical disciplines, such as mathematics, and the clinical disciplines, such as neurophysiology. This thesis has attempted to bridge this gap, using physics and mathematics in conjunction with anatomy and physiology to seek a more satisfactory method of solving the neuro-electromagnetic inverse problem, and to provide new insights into the cortical origins of commonly recorded evoked electric and magnetic responses.

In the areas of anatomy and physiology, use has been made of the current knowledge that visual cortex is comprised of a number of distinct functional areas, each of which has its own characteristics and functional speciality. V1 and V2 are functional segregators concerned with separating the signals related to form, colour and motion. V3 is concerned with form analysis, V4 is concerned with colour in conjunction with form, and V5 is concerned with motion. All of these visual cortical areas are arranged in laminae, with the large pyramidal cells of layer 5 and many of the smaller pyramidal cells oriented such that their apical dendrite is perpendicular to the cortical laminae. Since there are many large pyramidal cells of a common orientation in each small region of visual cortex, these cells seem most likely to be the source of measured external signals. Postsynaptic potentials are much more likely to be the source of neuro-electromagnetic signals than action potentials, since the former give rise to a single extended current dipole whereas the latter produce two opposed current dipoles whose electric and magnetic fields almost cancel. An assembly of coherent current sources in adjacent neurones constitutes a current dipole.

Meanwhile, physics and mathematics shows that in an anisotropic concentric spherical volume conductor, a radial current dipole produces an electric potential distribution on the scalp, but the magnetic field is zero everywhere outside of the conductor. However, a tangential current dipole produces both an electric field and a magnetic field. A deep current dipole yields a large electric field, whereas a superficial current dipole yields a large magnetic field. The inverse problem is non-unique and its solution therefore requires that prior information is included with the measured electric
and magnetic fields. This prior information usually takes the form of (1) the assumption that there is a single dipolar current source (equivalent current dipole fitting), or (2) that there is a distribution of current dipoles, but the probability density function of current dipole moment is Gaussian (minimum norm estimation), although other forms of assumption have also been discussed.

Since primary current is restricted to the cortex, such that it flows along the apical dendrites of the pyramidal cells towards or away from the cortical surface, this current can be modelled as a distribution of discrete current dipoles which are arrayed throughout visual cortex, oriented normally to the pial membrane. This approach introduces anatomical prior information into the inverse problem. Analysing both electric and magnetic evoked responses simultaneously in conjunction with this constrained source model eliminates silent sources and increases the number of measurements, enabling an overdetermined or almost overdetermined linear system to be obtained. This increases the accuracy of the solution. Whereas conventional minimum norm estimation tends to reconstruct deep current sources from electric responses and superficial current sources from magnetic responses, anatomically constrained analyses using both electric and magnetic data yield solutions whose depth is more realistic. If the shape and location of the source model deviates from that of the true brain, the solutions are slightly less accurate but still acceptable. If there is a gross rotation of the source model relative to the real brain, the solution shifts unacceptably, but by rotating the source model through an angle which maximises the goodness-of-fit, the solution can be recovered adequately. A standard anatomical source model can be used to obtain accurate source images from group average electric and magnetic responses.

Anatomically constrained distributed-source analysis involves generating and solving a linear system of equations. This can be efficiently performed using data parallelism on a linear pipeline of transputers. The speedup of the parallel program is close to linear for up to several tens of processors. Parallel computing ensures that computation times for simultaneous analyses of electric and magnetic responses are kept below about five minutes. Moreover, the parallel linear inversion algorithm can be adapted simply to perform minimum norm estimation and equivalent current dipole fitting. In this case, parallel processing reduces the computation time from as long as nine hours to less than half an hour. This increase in speed would be extremely advantageous if the boundary element method were used.

Comparing methods, equivalent current dipoles provide a useful interpretation of evoked electric and magnetic fields, provided that the goodness-of-fit is acceptable, but they must otherwise be discarded, which is often inconvenient. Unconstrained minimum norm estimates are difficult to interpret as they are oversimplified, too diffuse, and suffer from artefactual "whirls". Furthermore, the solution is reconstructed in the deep part of the source space if electric data are used, and in the superficial part of the source space if magnetic data are used. Unfortunately, if both electric and magnetic responses are analysed simultaneously, the solution does not improve in accuracy, but instead shows
deep sources due to the electric measurements and superficial sources due to the magnetic measurements, when in reality, one intermediate source is probably more appropriate.

Current source images produced by anatomically constrained distributed-source analysis show considerably more detail than minimum norm estimates. Because of the reduced number of unknown parameters, the solutions are not diffuse. The solutions do not suffer from spurious whirls, but current spreads across both of the opposing surfaces of a fissure: the method cannot distinguish which surface is active, so current appears on both. If both electric and magnetic data are analysed simultaneously, the reconstructed current has approximately the correct depth profile. Because the anatomical model relates explicitly to the visual cortex, it is easy to interpret the locations of the reconstructed current sources.

By applying the techniques to real data, it has been shown that the ClIm peak of the pattern onset visual evoked magnetic response appears to be generated in area 17 and possibly area 18. CIIm is produced by sources distributed throughout areas 17, 18 and 19, with primary current directed from the pial surface towards the white matter. CIIIm appears to be produced by a similar source distribution to CIIm, but with current direction reversed. Meanwhile, the P100m peak of the pattern reversal visual evoked magnetic response is generated by areas 17, 18 and 19 simultaneously; primary current flows from the white matter towards the pial surface.

The onset of an achromatic checkerboard of large checks evokes an electromagnetic response which, at the latency of the first major primary current flow (approximately the latency of CI/ClIm), can be attributed to areas 17 and 18. Primary current sometimes flows towards the white matter and sometimes away from it. Because of these attributes, the cortical response to a sectorial stimulus does not obey the simple cruciform model, other than that the response is contralateral to the stimulus. Nevertheless, neuro-electromagnetic imaging is able to identify changes in the location of the cortical response to a luminance contrast stimulus as the position of the stimulus in the visual field is varied.

There is a promising future for anatomically constrained distributed-source analysis. The most obvious advance is to use magnetic resonance scans to define the source model for each individual subject. This would involve using an automated region-growing or triangulation procedure to define the surface of the visual cortex (or any other desired region of cortex). Once the surface were defined, it would be possible to automatically locate a series of discrete current dipoles throughout the source region so that they were perpendicular to the cortical laminae. This type of procedure has already been carried out by Fuchs et al. (1993; see the introduction to chapter five). The advantage of this method is that it reduces the systematic error between the true sources and the source model. However, this approach also has its own associated problems. For example, although efficient algorithms are available for automatically defining the gyri, such algorithms are not yet able to define deep sulci such as the
calcarine sulcus. In addition, the method cannot remove systematic error entirely as it becomes necessary to co-register the coordinate system used for the magnetic resonance scan with that used for the neuro-electromagnetic measurements. An error of $\pm 2$mm in the locations of predefined bony landmarks such as the nasion and pre-auricular points can cause mis-registration of the source model with the cortex by up to about 5mm. Hence, the studies of chapter five, where the true location of the source model is found using the solution goodness-of-fit, are directly relevant to this situation. Parallel processing could also be used to derive a source space using magnetic resonance images, as well as being used to generate and solve the resulting linear system.

Using a magnetic resonance scan to define the source model is ideal if such a scan is available. But magnetic resonance scanning is costly and it is therefore desirable to carry out neuro-electromagnetic imaging without the prior acquisition of a magnetic resonance scan. Indeed, in future, when functional magnetic resonance imaging becomes more commonplace, clinical centres having access to a magnetic resonance scanner may simply use this to produce both functional and anatomical images. However, the functional images produced by this procedure would have low temporal resolution, so neuro-electromagnetic imaging would still have an important place in the range of functional localisation tools available to the surgeon or clinician. Hence, there is a considerable argument in favour of retaining neuro-electromagnetic imaging as a cost-effective standalone procedure with high temporal resolution, and the use of an average model of the brain is therefore an important area for research. One possible future direction would be to construct the average source model using a series of small spherical source regions in which a current dipole of fixed or variable orientation would be located. Each dipole would be free to move within the confines of its own sphere, so that if the true brain were slightly different from the average brain, the dipole locations could move accordingly. This would, however, increase the number of unknown parameters in the source model.

There are many different evoked responses which could be analysed using the anatomically constrained distributed-source analysis method. One particularly interesting and important response is the pattern onset response to sectorial stimulation, as used in chapter eight. Evidence seems to be accumulating that an achromatic stimulus evokes a cortical response in areas V1 and V2 which does not closely obey the cruciform model (e.g. Butler et al. 1987; Ahlfors et al. 1992, chapter eight), while a chromatic stimulus evokes a response in V1 which consequently can be modelled successfully with the cruciform model (e.g. Darcey et al. 1980a, b). Hence, a particularly interesting application of the techniques developed in this thesis would be to compare the cortical responses to achromatic and chromatic sectorial pattern onset stimuli. This study may reveal some significant differences in the way the visual cortex handles achromatic and chromatic information, further underlining the importance of neuro-electromagnetic imaging in functional localisation.
References


Brindley GS (1972) The variability of the human striate cortex. *J. Physiol.** 225**: 1P-3P.


Clark WELG (1941) The laminar organisation and cell content of the lateral geniculate body in the monkey. *J. Anat.* **75**: 419-433.


365


Lowitzsch K, Rudolph HD, Trincker D and Muller E (1980) Flash and pattern-reversal visual evoked responses in retrobulbar neuritis and controls: a comparison of conventional and TV stimulation techniques. In: Electroencephalography and

Lund JS (1973) Organization of neurons in the visual cortex, area 17, of the monkey (Macaca mulatta). *J. comp. Neurol.* **147**: 455-496.


Wong-Riley M (1978) Reciprocal connections between striate and prestriate cortex in squirrel monkey as demonstrated by combined peroxidase histochemistry and autoradiography. *Brain Res.* **147**: 159-164.


