Field Geophysics

The Geological Field Guide Series

Barnes, J.W. and Lisle, R.J. (2004) *Basic Geological Mapping*, 4th edn. ISBN: 978-0-470-84986-6, 5th edn publishing (2011). ISBN: 978-0-470-68634-8

Fry, N. (1991) *The Field Description of Metamorphic Rocks*. ISBN: 978-0-471-93221-5

McClay, K.R. (1991) *The Mapping of Geological Structures*. ISBN: 978-0-471-93243-7

Milsom, J. and Eriksen, A. (2011) *Field Geophysics*, 4th edn. ISBN: 978-0-470-74984-5

Tucker, M.E. (2011) *Sedimentary Rocks in the Field*, 4th edn. ISBN: 978-0-470-68916-5

Field Geophysics

FOURTH EDITION

John Milsom
Gladestry Associates

Asger Eriksen
Zetica Ltd

CONTENTS

PREFACE TO THE FIRST EDITION

The purpose of this book is to help anyone involved in small-scale geophysical surveys. It is not a textbook in the traditional sense, in that it is designed for use in the field and concerns itself with practical matters – with theory taking second place. Where theory determines field practice, it is stated, not developed or justified. For example, no attempt is made to explain why four-electrode resistivity works where two-electrode surveys do not.

The book does not deal with marine, airborne or downhole geophysics, nor with deep seismic reflection work. In part this is dictated by the space available, but also by the fact that such surveys are usually carried out by quite large field crews, at least some of whom, it is to be hoped, are both experienced and willing to spread the benefit of that experience more widely.

Where appropriate, some attention is given to jargon. A field observer needs not only to know what to do but also the right words to use, and right in this context means the words which will be understood by others in the same line of business, if not by the compilers of standard dictionaries.

A word of apology is necessary. The field observer is sometimes referred to as 'he'. This is unfortunately realistic, as 'she' is still all too rare, but is not intended to indicate that 'she' is either unknown or unwelcome in the geophysical world. It is hoped that all geophysical field workers, whether male or female and whether geophysicists, geologists or unspecialized field hands, will find something useful in this book.

Finally, a word of thanks. Paul Hayston of BP Minerals and Tim Langdale-Smith of Terronics read early drafts of the text and made numerous invaluable suggestions. To them, to Janet Baker, who drew many of the sketches, and to the companies which provided data and illustrations, I am extremely grateful.

PREFACE TO THE SECOND EDITION

Since the first edition of this book was published in 1989, there have been some changes in the world of field geophysics, not least in its frequent appearance in television coverage of archaeological 'digs'. In this work, and in surveys of contaminated ground and landfill sites (the archaeological treasure houses of the future), very large numbers of readings are taken at very small spacings and writing down the results could absorb a major part of the entire time in the field. Automatic data logging has therefore become much more important and is being make ever easier as personal computers become smaller and more powerful. New field techniques have been developed and image processing methods are now routinely used to handle the large volumes of data. Comments made in the first edition on the need to record information about the survey area as well as geophysical data have equal, and perhaps even more, force in these instances, but it is obviously usually not practical or appropriate to make individual notes relating to individual readings.

The increase in the number of geophysical surveys directed at the very shallow subsurface (1–5 m) has also led to the increasing use of noncontacting (electromagnetic) methods of conductivity mapping. Moreover, the increased computing power now at every geophysicist's disposal has introduced inversion methods into the interpretation of conventional direct current resistivity soundings and has required corresponding modifications to field operations. It is hoped that these changes are adequately covered in this new edition. a further development has been the much wider availability of ground penetrating radar systems and a recent and fairly rapid fall in their cost. A chapter has been added to cover this relatively new method.

Much else has remained unchanged, and advances in airborne techniques have actually inhibited research into improving ground-based instrumentation for mineral exploration. Automatic and self-levelling gravity meters are becoming more widely available, but are still fairly uncommon. Magnetometers more sensitive than the conventional proton precession or fluxgate instruments are widely advertised, but in most circumstances provide more precision than can be readily used, except in the measurement of field gradients.

VLF methods are enjoying something of a revival in exploration for fracture aquifers in basement rocks, and the importance of ease of use is being recognized by manufacturers. Instruments for induced polarization

and time-domain electromagnetic surveys also continue to be improved, but their basic principles remain unchanged. More use is being made of reflected seismic waves, partly because of the formerly undreamed of processing power now available in portable field seismographs, but refraction still dominates seismic studies of the shallow subsurface.

Inevitably, not all the methods currently in use could be covered in the space available. Seismo-electrical methods, in which the source pulses are mechanical and the signal pulses are electrical, are beginning to make their presence felt and may demand a place in textbooks in the future. Few case histories have yet been published. Magnetotelluric methods have a much longer history and continue to be developed, in conjunction with developments in the use of controlled (CSAMT) rather than natural sources, but many general purpose geophysicists will go through their entire careers without being involved in one such survey.

Despite the considerable rewriting, and the slight increase in size (for which I am immensely grateful to the new publishers), the aim of the book remains the same. Like its predecessor it is not a textbook in the conventional sense, but aims to provide practical information and assistance to anyone engaged in small-scale surveys on the ground. In helping me towards this objective, I am grateful particularly to Paul Hayston (RTZ) for introducing me to mineral exploration in a new and exciting area, to Asgeir Eriksen of Geophysical Services International (UK) for keeping me in touch with the realities of engineering and ground-water geophysics, and to my students for reminding me every year of where the worst problems lie. I am also grateful to all those who have given their permission for illustrations to be reproduced (including my daughter, Kate, whose view of field geophysics is shown in Fig. 5.1), and most especially to my wife, Pam, for retyping the original text and for putting up with this all over again.

<div align="right">John Milsom</div>

PREFACE TO THE THIRD EDITION

In the decade and a half since the preparation of the first edition of this handbook there have been few fundamental changes in the methods used in small-scale ground geophysical surveys. There have, however, been radical changes in instrumentation, and far-reaching developments in applications.

The use of geophysics in mineral exploration has declined, both in absolute terms (along with the world-wide decline in the mining industry itself), and relative to other uses. What is loosely termed environmental, engineering or industrial geophysics has taken up much of the slack. Sadly, the search for unexploded ordnance (UXO) is also assuming ever-increasing importance as more and more parts of the world become littered with the detritus of military training and military operations (the much more lethal search for landmines which, unlike UXO, are deliberately designed to escape detection, also uses geophysical methods but is emphatically *not* covered in this book).

Archaeological usage is also increasing, although still inhibited in many cases by the relatively high cost of the equipment.

In instrumentation, the automation of reading and data storage, which was only just becoming significant in the late 1980s, has proceeded apace. Virtually all the new instruments coming on to the market incorporate data loggers and many include devices (such as automatic levelling) to make operations quicker and easier. This, and the fact that virtually every field crew now goes into the field equipped with at least one laptop PC, has had two main, and contrasting, consequences. On the one hand, the need for specialist skills in the field personnel actually operating the instruments has been reduced, and this is leading to a general decline in the quality of field notes. On the other hand, much more can now be done in the field by way of processing and data display, and even interpretation. The change is exemplified by ground radar units, which provide users with visual (even though distorted) pictures of the subsurface while the survey is actually under way. Interestingly, the trend towards instruments that provide effectively continuous coverage as they are dragged or carried along lines has led to the emergence in ground surveys of errors that have long plagued airborne surveys but have now been largely eliminated there. Comments made in the first edition on the need to record information about the survey area as well as geophysical data have equal, and perhaps even more, force in

these instances, but it is obviously usually neither practical nor appropriate to make individual notes relating to individual readings.

The increase in the number of geophysical surveys directed at the very shallow subsurface (1–5 m) has also led to the increasing use of electromagnetic methods of conductivity mapping and the development of non-contacting electrical methods which use capacitative rather than inductive coupling. A chapter section has been added to cover this latter, relatively new, method. Other new sections deal with GPS navigation, which has become immensely more useful to geophysicists since the removal of 'selective availability' and with audio-magnetotellurics (AMT), largely considered in the context of controlled sources (CSAMT) that mimic the natural signals but provide greater consistency.

There has also been a slight change in the notes and bibliography. Providing references to individual papers is a problem in a book of this size, and I have actually reduced the number of such references, confining myself to older papers containing some fundamental discussion, and to papers that are the sources of illustrations used. I have also eliminated the section on manufacturers' literature, not because this literature is any less voluminous or important, but because it is now largely available through the Internet. A number of key URLs are therefore given.

Despite the considerable rewriting, and the slight increase in size (for which I am again immensely grateful to the publishers), the aim of the book remains unchanged. Like its predecessors, it is not a textbook in the conventional sense, but aims to provide practical information and assistance to anyone engaged in small-scale surveys on the ground. In helping me towards achieving this objective, I am grateful particularly to Chris Leech of Geomatrix for involving me in some of his training and demonstration surveys, to Asgeir Eriksen of Geophysical Services International (UK) for keeping me in touch with the realities of engineering and groundwater geophysics, and to my students for incisive and uninhibited criticisms of earlier editions. I am also grateful to all those who have given their permission for illustrations to be reproduced (including my daughter, Kate, whose view of field geophysics is shown in Figure 5.1), and most especially to my wife, Pam, for exhaustive (and exhausting) proofreading and for putting up with this for a third time.

PREFACE TO THE FOURTH EDITION

Becoming a co-author of this established handbook on Field Geophysics has been a fascinating exercise, since the changes in emphasis over the years have reflected my own experience. In my company, we moved our main focus from mineral exploration to engineering, environmental and archaeological geophysics in 1993, and have found the wide range of applications in the built environment to be challenging and remarkably satisfying. New routine uses of applied geophysics (e.g. in ballast scanning of railway trackbed) are continually being brought to market. As this new edition goes to press, large productivity gains are being achieved in data collection by using towed array systems, and the potential of remotely monitored, fixed installation geophysical systems to measure changes in material properties is being demonstrated. These are portents of further exciting innovations and developments. There has never been a better time to be involved in applied geophysics.

With the increasing use of geophysical data to provide evidence of change in existing infrastructure and to reduce the in-ground risk when developing new structures, there has come an increased necessity for high levels of professionalism in the collection, management and reporting of geophysical data. The benefits of the growth in data volumes provided by modern data collection systems can only be realised if strict fieldwork procedures are followed. It is no simple task to monitor the quality of data from multi-instrument platforms whilst ensuring both that the positional control meets design specifications and that everything is being kept dry in the pouring rain. Today's field geophysicist has to be a client-friendly, weather-impervious and patient manager of electronics systems, with an eye for detail.

In this edition, new sections on geophysical survey design, procedures, data quality control and limits of detection have been added to emphasise the importance of the fieldwork stage in delivering reliable information to end users. Updates to the resistivity and ground-penetrating radar sections have been made to reflect recent developments. A new chapter on surface wave seismics has been added, partly to increase awareness of this now-active area.

One reason for the decline in the use of small-scale geophysics in mineral exploration has been that the increasing accuracy in three dimensions of GPS positioning, and the parallel reduction in minimum reading times for some instruments, has given airborne data a previously unobtainable quality.

Where small-scale geophysics has expanded, in archaeology, site investigation, other forms of engineering and hydrological investigations and the search for unexploded ordnance (UXO), GPS has also had an impact. Field crews have not only been gifted with a whole range of new tools that will (supposedly) make their lives easier, but have been pointed towards whole new libraries of things that they need to know. In recognition of this fact, a new Chapter 15 is included, to deal with mapping problems and with GPS.

Reluctantly, we have omitted almost all of the sections that, in previous editions, dealt with the geophysical uses of military Very Low Frequency (VLF) radio-wave transmissions. It seems that the military no longer need them. Many transmitters have already been decommissioned, and in many parts of the world it is now impossible to receive adequate signals from even one source, let alone the two, at widely different azimuths, that are required for coverage to be satisfactory. The VLF band is therefore discussed only in the context of broad-band natural and controlled-source magnetotellurics.

Technology has also produced some changes in the book itself. "Field Geophysics" is designed to be taken into the field and used there. As long as the required information is available elsewhere, there therefore seems little point in including a bibliography in the printed edition that would, necessarily, be severely restricted by the space available, and would, inevitably, be listing material unlikely to be available in the field. It is now possible to provide readers with an associated on-line bibliography that is much more comprehensive, and much more searchable, than could ever be possible in print, and this is the route that we have taken. The result will be found on www.wiley.com/go/milsom/geophysics4e. Even there, we have abandoned the referencing of other websites, since these so frequently change or vanish. An internet search engine is all that the reader will need to locate manufacturers' manuals and applications sheets, to find out about the latest IGRFs or IGFs, or to source SRTM and ASTER topographic grids.

It has been a privilege and a great pleasure to work as a second author with John Milsom on this edition of the handbook.

Asger Eriksen

1
INTRODUCTION

1.1   What Geophysics Measures

Applied or exploration geophysics can be defined as mapping the subsurface through the remote measurement of its physical properties. The discipline dates back to ancient times but only since the advent of modern-day instrumentation has its use become widespread. The development of geophysical techniques and equipment during the early to middle parts of the twentieth century was driven by oil and mineral exploration, for targets that could be several kilometres deep. Many of the instruments used today in archaeological, environmental and engineering surveys owe their development to this kind of geophysics, but have been adapted to investigations of the near-surface, in the range of 0.5–100 m.

The success of any geophysical method relies on there being a measurable contrast between the physical properties of the target and the surrounding medium. The properties utilised are, typically, density, elasticity, magnetic susceptibility, electrical conductivity and radioactivity (Table 1.1). Whether a physical contrast is in practice measurable is inextricably linked to the physics of the problem, the design of the geophysical survey and the selection of suitable equipment. Not all equipment is fit for purpose. Often a combination of methods provides the best means of solving a complex problem, and sometimes a target that does not provide a measurable physical contrast can be detected indirectly by its association with conditions or materials that do. One of the aims of this handbook is to give the field observer an appreciation of the notional detectability of targets and the influence of burial setting, survey design, equipment selection and operating procedures on actual detectability.

1.2   Fields

Although there are many different types of geophysical measurement, small-scale surveys all tend to be rather similar and involve similar, and sometimes ambiguous, jargon. For example, the word *base* has three different common meanings, and *stacked* and *field* have two each.

Measurements in geophysical surveys are made *in the field* but, unfortunately, many are also *of* fields. Field theory is fundamental to gravity,
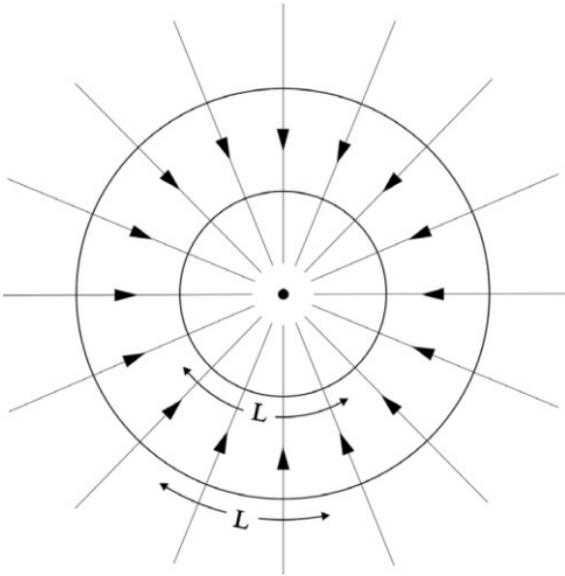
Figure 1.1 *Lines of force from an infinite line source (viewed end on). The distance between the lines increases linearly with distance from the source so that an arc of the inner circle of length L is cut by four lines but an arc of the same length on the outer circle, with double the radius, is cut by only two.*

magnetic and electromagnetic (EM) work, and even particle fluxes and seismic wavefronts can be described in terms of radiation fields. Sometimes ambiguity is unimportant, and sometimes both meanings are appropriate (and intended), but there are occasions when it is necessary to make clear distinctions. In particular, the term *field reading* is nearly always used to identify readings made *in* the field, i.e. not at a base station.

Physical fields can be illustrated by lines of force that show the field direction at any point (Figure 1.1). Intensity can also be indicated, by using more closely spaced lines for strong fields, but it is difficult to do this quantitatively where three-dimensional situations are being illustrated on two-dimensional media.

In Table 1.1 there is a broad division into *passive* and *active* methods. Passive methods use naturally occurring fields (such as the Earth's magnetic field), over which the observer has no control, and detect variations caused by geology or man-made objects. Interpretation is usually non-unique, relying a great deal on the experience of the interpreter. Active methods involve generating signals in order to induce a measurable response associated with

Table 1.1 *Common geophysical techniques*

Technique	Passive/ active	Physical property utilised	Source/signal
Magnetics	Passive	Magnetic susceptibility/ remanence	Earth's magnetic field
Gravity	Passive	Density	Earth's gravitational field
Continuous Wave and Time-Domain Electromagnetics (EM)	Active/ passive	Electrical conductivity/ resistivity	Hz/kHz band electromagnetic waves
Resistivity Imaging/ Sounding	Active	Electrical resistivity	DC electric current
Induced Polarisation	Active	Electrical resistivity/ complex resistivity and chargeability	Pulsed electric current
Self potential (SP)	Passive	Redox and electrokinetic	Redox, streaming and diffusion potentials
Seismic Refraction and Reflection/ Sonic	Active/ passive	Density/elasticity	Explosives, weight drops, vibrations, earthquakes, sonic transducers
Radiometrics	Active/ passive	Radioactivity	Natural or artificial radioactive sources
Ground Penetrating Radar (GPR)	Active	Dielectric properties (permittivity)	Pulsed or stepped frequency microwave EM (50–2000 MHz)
Wireline Logging	Active/ passive	Various	Various
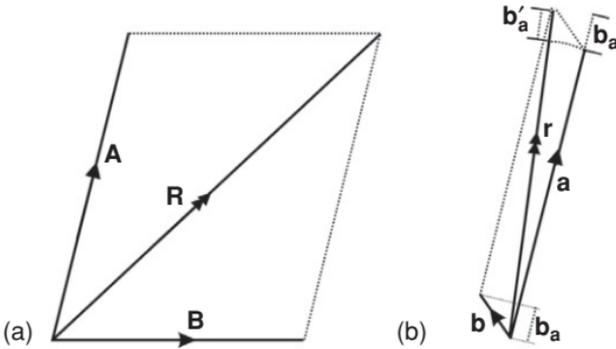
Figure 1.2 *Vector addition by the parallelogram rule. Fields in (a) that are represented in magnitude and direction by the vectors **A** and **B** combine to give the resultant **R**. In (b), the resultant **r** of the large field **a** and the small field **b** is approximately equal in length to the sum of **a** and the component **b**$_a$ of **b** in the direction of **a**. The angular difference in direction between **a** and **r** is small and therefore the component **b**$'_a$ in the direction of **r** is almost identical to **b**$_a$.*

a target. The observer can control the level of energy input to the ground and also measure variations in energy transmissibility over distance and time. Interpretation of this type of data can be more quantitative. Depth discrimination is often better than with passive methods, but ease of interpretation is not guaranteed.

1.2.1 Vector addition

When combining fields from different sources, vector addition (Figure 1.2) must be used. In passive methods, knowledge of the principles of vector addition is needed to understand the ways in which measurements of local anomalies are affected by regional backgrounds. In active methods, a local anomaly (*secondary field*) is often superimposed on a *primary field* produced by a transmitter. In either case, if the local field is much the weaker of the two (in practice, less than one-tenth the strength of the primary or background field), then the measurement will, to a first approximation, be made in the direction of the stronger field and only the component of the anomaly in that direction will be measured (Figure 1.2b). The slight difference in direction between the resultant and the background or primary field is usually ignored in such cases.

If the two fields are similar in strength, there will be no simple relationship between the magnitude of the anomalous field and the magnitude of

the observed anomaly. However, variations in any given *component* of the secondary field can be measured by taking all measurements in a single direction and assuming that the component of the background or primary field in that direction is constant over the survey area. Measurements of vertical rather than total field are sometimes preferred in magnetic and electromagnetic surveys for this reason.

The fields due to multiple sources are not necessarily equal to the vector sums of the fields that would have existed had those sources been present in isolation. A strong magnetic field from one body can affect the magnetisation in another, or even in itself (*demagnetisation effect*), and the interactions between fields, conductors and currents in electrical and electromagnetic surveys can be very complicated.

1.2.2 The inverse-square law

An inverse-square law attenuation of signal strength occurs in most branches of applied geophysics. It is at its simplest in gravity work, where the field due to a point mass is inversely proportional to the square of the distance from the mass, and the constant of proportionality (the *gravitational constant G*) is invariant. Magnetic fields also obey an inverse-square law, and the fact that, in principle, their strength varies with the permeability of the medium is irrelevant in most geophysical work, where measurements are made in either air or water. More important is the fact, which significantly modifies the simple inverse-square law decrease in field strength, that magnetic sources are essentially bipolar (Section 1.2.5).

Electric current flowing from an isolated point-electrode embedded in a continuous homogeneous ground provides a physical illustration of the significance of the inverse-square law. All of the current radiating from the electrode must cross any closed surface that surrounds it. If this surface is a sphere concentric with the electrode, the same fraction of the total current will cross each unit area on the surface of the sphere. The current *per unit area* will therefore be inversely proportional to the *total* surface area, which is in turn proportional to the square of the radius. Current flow in the real Earth is, of course, drastically modified by conductivity variations.

One problem inherent in the inverse-square law control of so many of the fields important in geophysics is *ambiguity*, i.e. the fact that a set of measurements made over a single surface can, in principle, be produced by an infinite number of possible source distributions. Most of these will be geologically impossible, but enough usually remain to render non-geophysical information essential to most interpretations. Figure 1.3 shows two spherical bodies, each with its centre at 5.5 m depth. One, an air void, has a radius of 2.25 m and zero density, whereas the other, a zone of weathered chalk, has a radius of 5 m and a density of $1.9 \, \text{Mg m}^{-3}$. The surrounding rock is
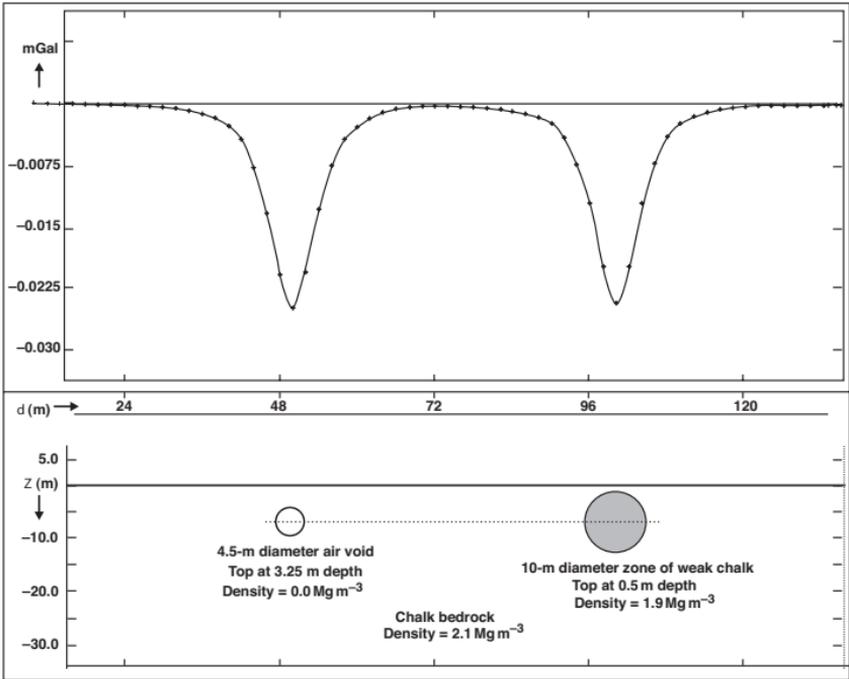
Figure 1.3 *Ambiguity in potential field interpretation. The two very different sources produce almost identical gravity anomalies.*

modelled with the density of $2.1\,\mathrm{Mg\,m^{-3}}$ typical of more competent chalk. The gravitational attraction of each sphere can be calculated assuming the mass deficit is concentrated at its centre. The two anomalies are almost identical, and a follow-on intrusive investigation of each, or a survey using a corroborative geophysical method such as electrical resistivity tomography (Section 6.5) would be required to resolve the ambiguity. Even non-identical anomalies may, of course, differ by amounts so small that they cannot be distinguished in field data.

Ambiguity worries interpreters more than it does the observers in the field, but its existence does emphasise the importance of those observers including in their field notes anything that might possibly contribute to a better understanding of the data that they collect.

1.2.3 Two-dimensional sources

Rates of decrease in field strengths depend on source shapes as well as on the inverse-square law. Infinitely long sources of constant cross-section are
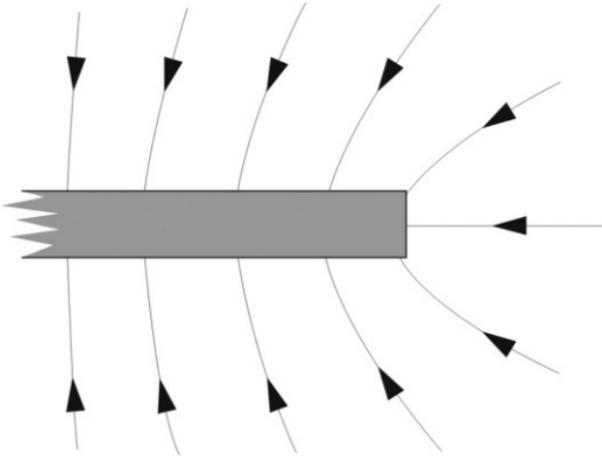
Figure 1.4 *Lines of force from a semi-infinite slab. The lines diverge appreciably only near the edge of the slab, implying that elsewhere the field strength will decrease negligibly with distance.*

termed *two-dimensional (2D)* and are often used in computer modelling to approximate bodies of large strike extent. If the source 'point' in Figure 1.1 represents an infinite line-source seen end-on rather than an actual point, the area of the enclosing (cylindrical) surface is proportional to its radius. The argument applied in the previous section to a point source then leads to the conclusion that the field strength for a line-source will be inversely proportional to distance and not to its square. It follows that, in 2D situations, lines of force drawn on pieces of paper can indicate field intensity (by their separation) as well as direction.

1.2.4 One-dimensional sources

The lines of force or radiation intensity from a source consisting of a homogeneous layer of constant thickness diverge only near its edges (Figure 1.4). The *Bouguer plate* of gravity reductions (Section 2.5.1) and the radioactive source with *2π geometry* (Section 4.3.4) are examples of infinitely extended layer sources, for which field strengths are independent of distance. This condition is approximately achieved if a detector is only a short distance above an extended source and a long way from its edges.

1.2.5 Dipoles

A dipole consists of equal-strength positive and negative point sources a very small distance apart. Its *moment* is equal to the pole strength multiplied
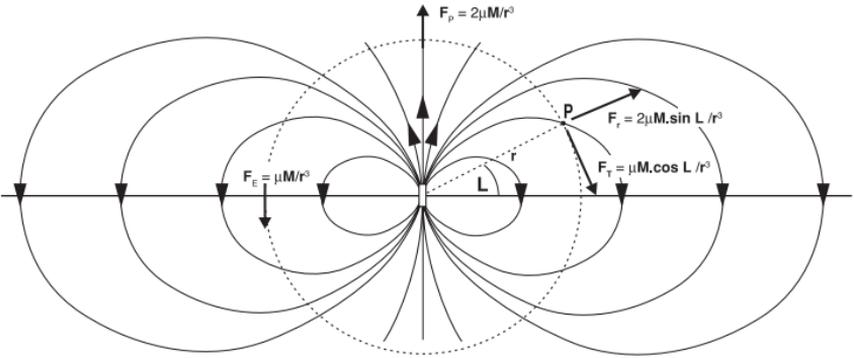
Figure 1.5 *The dipole field. The plane through the dipole at right angles to its axis is known as the equatorial plane, and the angle,* **L***, between this plane and the line joining the centre of the dipole to any point P is sometimes referred to as the latitude of P. The fields shown, at distances r from the dipole centre, are for a dipole with strength (moment)* **M** *(see Section 3.1.1). The values for the radial and tangential fields at P follow from the fact that* **M** *is a vector and can therefore be resolved according to the parallelogram law. The symbol μ is used for the proportionality constant where magnetic fields are concerned (Chapter 3).*

by the separation distance. Field strength decreases as the inverse cube of distance, and both strength and direction change with 'latitude' (Figure 1.5). The intensity of the field at a point on a dipole 'equator' is only half the intensity at a point the same distance away on the dipole axis, and in the opposite direction.

Magnetisation is fundamentally dipolar, and electric currents circulating in small loops are dipolar sources of magnetic field. Many radar antennas are dipolar, and in some electrical surveys the electrodes are set out in approximately dipole pairs.

1.2.6 Exponential decay

Radioactive particle fluxes and seismic and electromagnetic waves are subject to absorption as well as geometrical attenuation, and the energy crossing closed surfaces is less than the energy emitted by the sources they enclose. In homogeneous media, the percentage loss experienced by a plane wave is determined by the path length and the *attenuation constant*. The absolute loss is proportional also to the signal strength. A similar *exponential* law (Figure 1.6), governed by a *decay constant*, determines the rate of loss of mass by a radioactive substance.
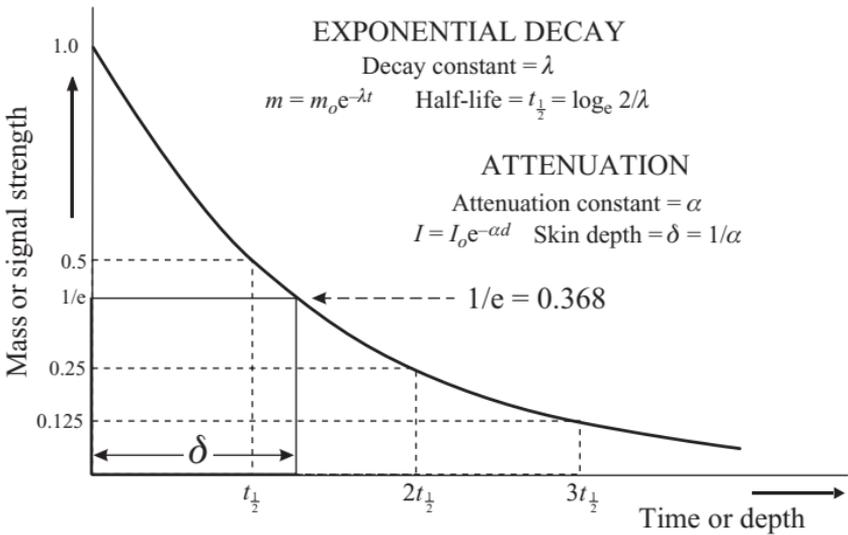
Figure 1.6 *The exponential law, illustrating the parameters used to characterise radioactive decay and radio wave attenuation.*

Attenuation rates are alternatively characterised by *skin-depths*, which are the reciprocals of attenuation constants. For each skin depth travelled, the signal strength decreases to $1/e$ of its original value, where e ($= 2.718$) is the base of natural logarithms. Radioactive decay rates are normally described in terms of the *half-lives,* equal to $\log_e 2$ ($= 0.693$) divided by the decay constant. During each half-life period, one half of the material present at its start is lost.

1.3 Geophysical Survey Design

1.3.1 Will geophysics work?

Geophysical techniques cannot be applied indiscriminately. Knowledge of the material properties likely to be associated with a target (and its burial setting) is essential to choosing the correct method(s) and interpreting the results obtained.

Armed with such knowledge, the geophysicist can assess feasibility and, where possible, select a geophysical method to meet the survey objectives. Table 1.2 lists some of the more important physical properties, for some of the commoner rocks and minerals. Inevitably, the values given are no more than broad generalisations, but the table does at least indicate some of the circumstances in which large contrasts in physical properties might be expected, or at least be hoped for.

Table 1.2 Important physical properties of common rocks and ore minerals

Material	Density $Mg\,m^{-3}$	Susceptibility $SI \times 10^6$	Resistivity Ohm-m	Conductivity $mS\,m^{-1}$
Air	0	0	8	0
Ice	0.9	−9	100 000–8	0–0.01
Fresh water	1	0	1 000 000	0.001
Seawater	1.03	0	0.2	5000
Topsoil	1.2–1.8	0.1–10	50–100	10–20
Coal	1.2–1.5	0–1000	500–2000	2–0.5
Dry sand	1.4–1.65	30–1000	1000–5000	1–0.02
Wet sand	1.95–2.05	30–1000	500–5000	0.2–2
Gravel	1.5–1.8	20–5000	100–1000	1–10
Clay	1.5–2.2	10–500	1–100	10–1000
Weathered bedrock	1.8–2.2	10–10 000	100–1000	1–10
Salt	2.1–2.4	−10	10–10 000 000	0.01–1
Shale	2.1–2.7	0–500	10–1000	1–100
Siltstone	2.1–2.6	10–1000	10–10 000	0.1–100
Sandstone	2.15–2.65	20–3000	200–8000	0.125–5
Chalk	1.9–2.1	0–1000	50–200	5–20
Limestone	2.6–2.7	10–1000	500–10 000	0.1–2
Slate	2.6–2.8	0–2000	500–500 000	0.002–2
Graphitic schist	2.5–2.7	10–1000	10–500	2–100
Quartzite	2.6–2.7	−15	500–800 000	0.00125–2
Gneiss	2.6–2.9	0–3000	100–1 000 000	0.001–10
Greenstone	2.7–3.1	500–10 000	500–200 000	0.005–2
Serpentinite	2.5–2.6	2000–100 000	10–10 000	0.1–100
Granulite	2.7–2.9	100–5000	500–1 000 000	0.001–2
Granite	2.5–2.7	20–5000	200–1 000 000	0.001–5
Rhyolite	2.5–2.7	100–5000	1000–1 000 000	0.001–1
Basalt	2.7–3.1	500–100 000	200–100 000	0.01–5
Dolerite	2.8–3.1	500–100 000	100–100 000	0.01–10
Gabbro	2.7–3.3	100–10 000	1000–1 000 000	0.001–1
Peridotite	3.1–3.4	10–10 000	100–100 000	0.01–10
Pyrite	4.9–5.0	100–5000	0.01–100	10–1 000 000
Pyrrhotite	4.4–4.7	1000–50 000	0.001–0.01	1 000 000–10 000 000
Sphalerite	3.8–4.2	10–100	1000–1 000 000	0.001–1
Galena	7.3–7.7	10–500	0.001–100	10–10 000 000
Chalcopyrite	4.1–4.3	100–5000	0.005–0.1	10 000–200 000
Chromite	4.5–4.7	750–50 000	0.1–1000	1–10 000
Hematite	5.0–5.1	100–1000	0.01–1 000 000	0.001–100 000
Magnetite	5.1–5.3	10 000–10 000 000	0.01–1000	0.001–1
Cassiterite	7.0–7.2	10–500	0.001–10 000	0.1–10 000 000

The design and implementation of a geophysical survey requires careful consideration of the following main factors:

(a) *Target discrimination*

The nature and degree of the contrast in physical properties between a target and its surroundings is of primary importance in the feasibility assessment and choice of techniques. However, information may be limited or non-existent, and in these cases the geophysicist should recommend a trial survey or the application of multiple techniques. Trials are recommended wherever the assumptions made in designing the survey are suspect. Usually a day is all that is required to determine whether the chosen methods can detect the presence of a target in actual field conditions. This is an often neglected stage in the execution of a geophysical survey but is one that could save much geophysicist's pride and client's money were it more routinely used.

Once it has been decided, on the basis of observation, modelling and/or experience, what the geophysical response of a buried target is likely to be, the sensitivity of the equipment and the distribution of the survey stations needed to meet the survey objectives can be specified.

(b) *Detection distance*

In addition to the composition of the target and its surroundings, geophysical methods are sensitive to the relationship between target size and detection distance. In general, the greater the depth of the target, the larger its volume and/or cross-sectional area must be for it to be detectable.

(c) *Survey resolution*

The choice of sampling interval (frequency or spacing of sampling points) is critical to the success of a survey and its cost-effectiveness. The appropriate interval is dictated by the geophysical 'footprint' of the target, which may be tens of centimetres for small-diameter shallow pipes, a few metres for narrow fault zones, and kilometres for ore bodies at depth. An anomaly must be adequately sampled to meet the survey objectives. Although it is almost equally important that resources are not wasted in collecting more data than are required, it has to be remembered that under-sampling can produce completely fictitious anomalies (Figure 1.7).

In some cases, particularly on brownfield sites, surface obstructions can prevent the collection of regularly spaced data. The obstructions may be removable, but unless their impact on the survey outcome is fully understood by the field observer, they may not be dealt with at the appropriate time.
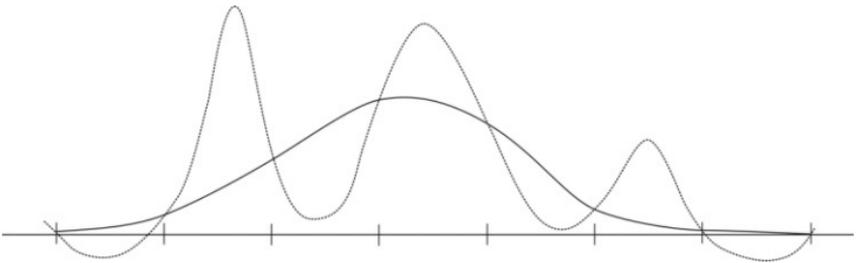
Figure 1.7 *Aliasing. The dashed curve shows a magnetic profile as it should have been recorded, the solid line shows the spurious anomaly that would be deduced by using only data from the widely spaced reading points indicated by vertical lines on the distance axis. Aliasing can occur in time as well as in space, if time-varying signals are sampled too infrequently.*

(d) *Site conditions*

The suitability of a site for collecting good quality geophysical data is often overlooked in survey design. The issues affecting data quality that could be of concern are often specific to the method or methods being proposed. For example, signal degradation may occur or geophysical 'noise' may be introduced in electromagnetic and magnetic surveys by the presence of surface metallic structures and overhead power lines. In microgravity or seismic surveys, noise may result from traffic movements or wind and waves. If the noise exceeds the amplitude of the anomaly due to the target and cannot be successfully removed, the target will not be detectable. The best way to assess the likely influence of site conditions is to visit the site at the design stage and/or carry out a trial survey.

Field observers should be fully briefed on the objectives of the survey and mindful of the design aspects, so that departures of the field conditions from any assumptions made can be reported in good time, allowing the design to be modified where possible. They should immediately report any unexpected conditions, and any geological information provided by drillers to which the geophysicist who designed the survey may not have been privy. They may also obtain useful information relating to previous land-use in conversations with the client or casual passers-by, and this also should be passed on.

1.3.2 Preparing for a survey

The design of a regional or even a local geophysical survey can be greatly assisted by using the geographic data now freely available on the internet.

Google Earth is familiar not only to geophysicists but to almost everyone who has internet access. Only freely available satellite imagery and aerial photographs are used, and quality and geo-registration accuracy vary with location. Images can be saved as .jpg files and it is possible, before so doing, to superimpose survey area outlines or survey grids using standard .kml (ASCII) or .kmz (binary) files. Area dimensions can be quickly estimated, and the area to be surveyed (and the parts that may be unsurveyable because of access restrictions) can be discussed and agreed with the client. The images also provide a practical basis for planning access along routes through farmers' fields. 'Forewarned, forearmed; to be prepared is half the victory' [Miguel de Cervantes Saavedra, seventeenth-century Spanish writer].

Internet-available elevation grids are less widely known, but can be equally useful. The Satellite Radar Tracking Mission (SRTM) used a satellite-mounted synthetic aperture radar interferometer to obtain data during a period of 11 days in February 2000. The targeted landmass extended from 56°S to 60°N, and within this region (containing about 80% of the Earth's land surface) elevation estimates were obtained once for at least 99.96%, twice for at least 94.59% and three or more times for about 50%. The data are now available as one-degree square 'tiles', with a 3 arc-second cell-size (equivalent to about 90 m at the Equator) globally (SRTM3) and an optional 1 arc-second (30-m) cell size in the USA (SRTM1). The Version 2 processed data set was replaced in 2009 by an improved (although usually imperceptibly so) Version 2.1.

The SRTM data as distributed suffered from data gaps in areas of steep topographic gradients. It was inevitable that, with a swathe width of about 225 km and a satellite altitude of 233 km, there would be areas that could not be imaged by a side-looking system. These disadvantages have, to a considerable extent, been overcome in the ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) data obtained with Japanese instrumentation mounted from December 1999 onwards on a US Terra spacecraft. Coverage was also wider, from 83°S to 83°N. As with SRTM, ASTER data are distributed in one-degree 'tiles' but with a worldwide 1 arc-second (∼ 30-m) cell size. Elevation data are provided in GeoTIFF format, and each data file is accompanied by a quality (QA) file that indicates data reliability, pixel by pixel.

The ASTER instrument operated stereoscopically in the near infra-red, and could therefore be affected by cloud cover. In most cases this problem was solved by the high degree of redundancy (since the mission lasted much longer than the 11 days of SRTM), but in some cases SRTM data have had to be used for infill. In a few areas where SRTM coverage did not exist, 'bad' pixels remain and are flagged by a −9999 value.

The amount of data collected by ASTER is truly enormous, and analysis and verification is a continuing process.

1.3.3 Procedures
All surveys require adherence to some form of procedure, and the field crew should ensure that this is agreed with the geophysicist before commencing fieldwork. Common aspects include, but are not limited to: daily checks on equipment functionality and sensitivity (sometimes with target seeding, depending on the target); survey station layout (to a specified accuracy); survey grid referencing (to previously agreed mapped features); frequency and nature of data quality and repeatability checks; frequency of data archiving; maintenance and format of decipherable field logbooks; and recording of all client communications. Assumption is the mother of all miscommunication between the office and the field, and a formal record of the agreed procedures is worth its weight in gold.

1.3.4 Metadata
Automation in geophysical work proceeds apace, and is giving increased importance to a distinction that, while always present, was sometimes not even recognised when all information was stored in field notebooks. These notebooks contained not only the numerical values displayed on whatever instruments were being used, but also positional and logistical data and other vital information, such as the observer's name. The term *metadata* is now widely used for this largely non-numeric information. Modern data loggers vary widely in the extent to which metadata can be entered into them, but none, so far, have reached a level of sophistication that would allow notebooks to be dispensed with altogether.

1.4 Geophysical Fieldwork
Geophysical instruments vary greatly in size and complexity but all are used to make physical measurements, of the sort commonly made in laboratories, at temporary sites under sometimes hostile conditions. They should be economical in power use, portable, rugged, reliable and simple. These criteria are satisfied to variable extents by the commercial equipment currently available.

1.4.1 Choosing geophysical instruments
It seems that few instrument designers have ever tried to use their own products for long periods in the field, since operator comfort seldom seems to have been considered. Moreover, although many real improvements have been made in the last 50 years, design features have been introduced during the same period, for no obvious reasons, that have actually made fieldwork

more difficult. The foldable proton magnetometer staff, discussed below, is a case in point.

If different instruments can, in principle, do the same job to the same standards, practical considerations become paramount. Some of these are listed below:

Serviceability: Is the manual comprehensive and comprehensible? Is a breakdown likely to be repairable in the field? Are there facilities for repairing major failures in the country of use or would the instrument have to be sent overseas, risking long delays *en route* and in customs? Reliability is vital, but some manufacturers seem to use their customers to evaluate prototypes.

Power supplies: If dry batteries are used, are they of types that are easy to replace or will they be impossible to find outside major cities? If rechargeable batteries are used, how heavy are they, and will they be acceptable for airline transportation? In either case, how long will they keep the instruments working at the temperatures expected in the field? Battery life is reduced in cold climates, and the reduction can be dramatic if the battery is used to keep the instrument at a constant temperature, since not only is the available power reduced but the demands made are increased.

Data displays: Are these clearly legible under all circumstances? A torch is needed to read some displays in poor light, and others are almost invisible in bright sunlight. Large displays are needed if continuous traces or profiles are to be shown, but can exhaust batteries very quickly.

Hard copy: If hard-copy records can be produced directly from the field instrument, are they of adequate quality? Are they truly permanent, or will they become illegible if they get wet or are abraded?

Comfort: Is prolonged use likely to cripple the operator? Some instruments are designed to be suspended on a strap passing across the back of the neck. This is tiring under any circumstances and can cause actual medical problems if the instrument has to be levelled by bracing it against the strap. Passing the strap over one shoulder and under the other arm may reduce the strain, but not all instruments are easy to operate when carried in this way.

Convenience: If the instrument is placed on the ground, will it stand upright? Is the cable then long enough to reach the sensor in its normal operating position? If the sensor is mounted on a tripod or pole, is this strong enough? The traditional magnetometer pole, made up of sections that screwed together and ended in a spike that could be stuck into soft ground, has now been largely replaced by unspiked hinged rods that are more awkward to stow away, much more fragile (the hinges can twist and break), can only be used if fully extended and must be supported at all times.

Fieldworthiness: Are the control knobs and connectors protected from accidental impact? Is the casing truly waterproof? Does protection from damp grass depend on the instrument being set down in a certain way? Are there depressions on the console where water will collect and then inevitably seep inside?

Automation: Computer control has been introduced into almost all the instruments in current production. Switches have almost vanished, and every instruction has to be entered via a keypad. This has reduced the problems that used to be caused by electrical 'spikes' generated by switches but, because the settings are usually not permanently visible, unsuitable values may be repeatedly used in error. Moreover, simple operations have sometimes been made unduly complicated by the need to access nested menus. Some instruments do not allow readings to be taken until line and station numbers have been entered, and in extreme cases may demand to know the distance to the next station and even to the next line!

The computer revolution has produced real advances in field geophysics, but has its drawbacks. Most notably, the ability to store data digitally within data loggers has discouraged the making of notes on field conditions where these, however important, do not fall within a restricted range of options. This problem is further discussed in Section 1.7.

1.4.2 Cables

Almost all geophysical work involves cables, which may be short, linking instruments to sensors or batteries, or hundreds of metres long. Electrical induction between cables (electromagnetic coupling, also known as *cross-talk*) can be a serious source of noise.

Efficiency in cable-handling is an absolute necessity. Long cables always tend to become tangled, often because of well-intentioned attempts to make neat coils using hand and elbow. Figures of eight are better than simple loops, but even so it takes an expert to construct a coil from which cable can be run freely once it has been removed from the arm. On the other hand, a seemingly chaotic pile of wire spread loosely on the ground can be quite trouble-free. The basic rule is that cable must be fed on and off such piles in opposite directions; that is, the last bit of cable fed on must be the first to be pulled off. Any attempts to pull cable from the bottom will almost certainly end in disaster.

Cable piles are also unlikely to cause the permanent kinks that are often features of neat and tidy coils and that may have to be removed by allowing the cable to hang freely and untwist naturally. Places where this is possible with 100-metre lengths are rare.

Cable piles can be made portable by dumping cables into open boxes, and on many seismic surveys the shot-firers carried their firing lines in this way in old gelignite boxes. Ideally, however, if cables are to be carried from place to place, they should be wound on properly designed drums. Even then, problems can occur. If a cable is being unwound by pulling on its free end, the drum will not stop simply because the pull stops, and a free-running drum is an effective, but untidy, knitting machine.

A drum carried as a back-pack should have an efficient brake and should be reversible so that it can be carried across the chest and be wound from a standing position. Some drums sold with geophysical instruments combine total impracticality with inordinate expense and are inferior to garden-centre or even home-made versions.

Geophysical cables exert an almost hypnotic influence on livestock, and cattle have been known to desert lush pastures in favour of midnight treks through hedges and across ditches in search of them. Not only can a survey be delayed but a valuable animal may be killed by chewing on a live conductor. Constant vigilance is essential.

1.4.3 Connections

Crocodile clips are usually adequate for electrical connections between single conductors. Heavy plugs must be used for multi-conductor connections and are usually the weakest links in the entire field system. They should be placed on the ground very gently and as seldom as possible and, if they do not have screw-on caps, be protected with plastic bags or 'clingfilm'. They must be shielded from grit as well as moisture. Faults are often caused by dirt, which increases the wear on the contacts in socket plugs, which are almost impossible to clean.

Plugs should be clamped to their cables, since any strain will otherwise be borne by the weak soldered connections to the individual pins. Inevitably, cables are flexed repeatedly just beyond the clamps, and wires may break within their insulated sleeves at these points. Any break there, or a broken or *dry* joint inside the plug, means work with a soldering iron. This is never easy when connector pins are clotted with old solder, and is especially difficult if many wires crowd into a single plug.

Problems with plugs can be minimised by ensuring that, when moving, they are always carried, never dragged along the ground. Two hands should always be used, one holding the cable to take the strain of any sudden pull, the other to support the plug itself. The rate at which cable is reeled-in should never exceed a comfortable walking pace, and special care is needed when the last few metres are being wound on to a drum. Drums should be fitted with clips or sockets where the plugs can be secured when not in use.

Figure 1.8 The geophysical cape in action. Electronics and observer are both dry, with only the sensor bottle exposed to the elements. The observer can retreat still further, to view the display.

1.4.4 Geophysics in the rain

Geophysicists huddled over their instruments are sitting targets for rain, hail, snow and dust, as well as mosquitoes, snakes and dogs. Their most useful piece of field clothing is often a large waterproof cape, which they can not only wrap around themselves but into which they can retreat, along with their instruments, to continue work (Figure 1.8).

Electrical methods that rely on direct or close contact with the ground generally do not work in the rain, and heavy rain can be a source of seismic noise. Other types of survey can continue, since most geophysical instruments are supposed to be waterproof and some actually are. However, unless dry weather can be guaranteed, field crews should be plentifully supplied with plastic bags and sheeting to protect instruments, and paper towels for drying them. Large transparent plastic bags can often be used to enclose instruments completely while they are being used, but even then condensation may create new conductive paths, leading to drift and erratic behaviour.

Silica gel within instruments can absorb minor amounts of moisture but cannot cope with large volumes, and a portable hair-drier held at the base camp may be invaluable.

1.4.5 A geophysical toolkit

Regardless of the specific type of geophysical instruments involved, similar tools are likely to be needed. A field toolkit should include the following:

Long-nose pliers (the longer and thinner the better)
Slot-head screwdrivers (one very fine, one normal)
Phillips screwdriver
Allen keys (metric and imperial)
Scalpels (light, expendable types are best)
Wire cutters/strippers
Electrical contact cleaner (spray)
Fine-point 12-V soldering iron
Solder and 'Solder-sucker'
Multimeter (mainly for continuity and battery checks, so small size and durability are more important than high sensitivity)
Torch/flashlight (either a type that will stand unsupported and double as a table lamp or a 'head torch')
Hand lens
Insulating tape, preferably self-amalgamating
Strong epoxy glue/'super-glue'
Silicone grease
Waterproof sealing compound
Spare insulated and bare wire, and connectors
Spare insulating sleeving
Kitchen cloths and paper towels
Plastic bags and 'clingfilm'

A comprehensive first-aid kit is equally important, and a legal necessity in many countries.

1.5 Geophysical Data

Geophysical readings may be of true *point data* but may also be obtained using *arrays* where sources are separated from detectors and where values are determined *between* rather than *at* points. In most such cases, readings will be affected by array orientation. Precise field notes are always important but especially if arrays are involved, since reading points must then be defined and array orientations must be recorded.

If transmitters, receivers and/or electrodes are laid out in straight lines and the whole array can be reversed without changing the reading, the mid-point should be considered as the reading point. Special notations are needed for asymmetric arrays, and the increased probability of positioning error is in itself a reason for avoiding asymmetry. Great care must be taken in recording the positions of sources and detectors in seismic work.

1.5.1 Station numbering

Station numbering should be logical and consistent. Where data are collected along traverses, numbers should define positions in relation to the traverse grid. Infilling between traverse stations 3 and 4 with stations $3^1/_4$, $3^1/_2$ and $3^3/_4$ is clumsy and may create typing problems, whereas defining as 325E a station halfway between stations 300E and 350E, which are 50 metres apart, is easy and unambiguous. The fashion for labelling such a station 300+25E has no discernible advantages and uses a plus sign that may be needed, with digital field systems or in subsequent processing, to stand for N or E. It is good practice to define the grid origin in such a way that S or W stations do not occur, and this may be essential with data loggers that cannot cope with either negatives or directions.

Stations scattered randomly through an area are best numbered sequentially, as read. Positions can be recorded in the field by pricking through the field maps or air-photos and labelling the reverse sides. Estimating co-ordinates from maps in the field may seem desirable but mistakes are easily made and valuable time is lost. Station coordinates are now often obtained from GPS receivers (see Section 15.2), but differential or RTK (*real-time kinetic*) GPS may be needed to provide enough accuracy in detailed surveys.

If several observers are involved in a single survey, numbers can easily be accidentally duplicated. All field books and sheets should record the name of the observer. The interpreter or data processor will need to know who to look for when things go wrong.

1.5.2 Recording results

Geophysical results are primarily numerical and must be recorded even more carefully than the qualitative observations of field geology. Words, although sometimes difficult to read, can usually be deciphered eventually, but a set of numbers may be wholly illegible or, even worse, may be misread. The need for extra care has to be reconciled with the fact that geophysical observers are usually in more of a hurry than are geologists, since their work may involve instruments that are subject to drift, draw power from batteries at frightening speed or are on hire at high daily rates.

Numbers may, of course, not only be misread but also miswritten. The circumstances under which data are recorded in the field are varied but

seldom ideal. Observers are usually either too hot, too cold, too wet or too thirsty. They may, under such conditions, delete correct results and replace them with incorrect ones, and data recorded on geophysical field sheets should therefore never be erased. Corrections should be made by crossing out the incorrect items, preserving their legibility, and writing the corrected values alongside. Something may then be salvaged even if the correction is wrong. Precise reporting standards must be enforced and strict routines must be followed if errors are to be minimised. Reading the instrument twice at each occupation of a station, and recording both values, reduces the incidence of major errors.

Loss of geophysical data tends to be final. Some of the qualitative observations in a geological notebook might be remembered and re-recorded, but not strings of numbers. Copies are therefore essential and should be made in the field, using duplicating sheets or carbon paper, or by transcribing the results each evening. Whichever method is used, originals and duplicates must be separated immediately and stored separately thereafter. Duplication is useless if copies are stored, and lost, together with the originals. This, of course, applies equally to data stored in data loggers incorporated in, or linked to, field instruments. Such data should be downloaded, checked and backed-up each evening.

Digital data loggers can greatly simplify field operations but are often poorly adapted to storing non-numeric metadata. This design feature ignores the fact that observers are uniquely placed to note and comment on a multitude of topographic, geological, man-made (*cultural*) and climatic factors that may affect the geophysical results. If they fail to do so, the data they have gathered may be interpreted incorrectly. If data loggers are not being used, comments should normally be recorded in the notebooks, alongside the readings concerned. If they are being used, adequate supplementary positional data must be stored elsewhere. In archaeological and site investigation surveys, where large numbers of readings are taken in very small areas, annotated sketches are always useful and may be essential. Sketch maps should be made wherever the distances of survey points or lines from features in the environment are important. Field observers also have a responsibility to pass on to their geological or geophysical colleagues information of interest about places that only they may visit. Where these would be useful, they should be prepared to record dips and strikes, and perhaps to return with rock samples.

1.5.3 Accuracy, sensitivity, precision

Accuracy must be distinguished from sensitivity. A modern gravity meter, for example, may be sensitive to field changes of 1 microGal but an equivalent level of accuracy will be achieved only if readings are carefully made and

drift and tidal corrections are correctly applied. Accuracy is thus limited, but not determined, by instrument sensitivity. Precision, which is concerned only with the numerical presentation of results (e.g. the number of decimal places used), should always be appropriate to accuracy (see Example 1.1). Not only does superfluous precision waste time but false conclusions may be drawn from a high implied accuracy.

Example 1.1

Gravity reading = 858.3 scale units
Calibration constant = 0.10245 mGal per scale division (see Section 2.2.6)
Converted reading = 87.932835 mGal
But reading accuracy is only 0.01 mGal (approximately), and therefore:
Converted reading = 87.93 mGal

(Note that five decimal place precision is needed in the calibration constant, because 858.3 multiplied by 0.00001 is equal to almost 0.01 mGal)

Geophysical measurements can sometimes be made with more accuracy than the interpreters need or can use. However, the highest possible accuracy should always be sought, as later advances may allow the data to be analysed more effectively.

1.5.4 Drift

A geophysical instrument will usually not record the same result if read repeatedly at the same place. *Drift* may be due to changes in background field but can also be caused by changes in the instrument itself. Drift correction is often the essential first stage in data analysis, and is usually based on repeat readings at *base stations* (Section 1.6).

Drift is often related to temperature and is unlikely to be linear between two readings taken in the relative cool at the beginning and end of a day if temperatures are 10 or 20 degrees higher at noon. Survey *loops* may therefore have to be limited to periods of only 1 or 2 hours.

Changes in background field are sometimes treated as drift but in most cases the variations can either be monitored directly (as in magnetics) or calculated (as in gravity). Where such alternatives exist, it is preferable they be used, since poor instrument performance may otherwise be overlooked. Drift calculations should be made whilst the field crew is still in the survey area, so that readings can be repeated if the drift-corrected results appear suspect.

1.5.5 Repeatability

Repeat data are vital for checking whether an instrument is performing to specification. Ideally, a repeat survey line should be completed on every survey grid before moving to the next grid. For linear transects or meandering surveys, a minimum of 5% of repeat data is required. Repeat line-data achieve two things – they confirm that the instrument is responding consistently and they also provide a measure of the positioning accuracy. Where geophysical anomalies are small, it may be prudent to collect more than one repeat line per survey grid, because of low signal-to-noise ratios. In gravity surveys requiring microGal resolution, it may be necessary to reoccupy two or more stations in each loop. Repeatability requirements should be discussed and agreed with the client before a survey begins.

1.5.6 Detection limits

To a geophysicist, *signal* is the object of the survey and *noise* is anything else that is measured but is considered to contain no useful information. Using geophysics to locate a target is in some ways analogous to receiving a mobile phone message. If the ratio of signal to noise is high (good 'reception'), a target may be found at close to the theoretical limits of detection. If the signal is weak it may not be possible to distinguish enough of the 'conversation' to make it understandable, or the 'connection' may be lost completely. 'Made' ground often contains material that interferes with the geophysical signal, so that the signal-to-noise ratio may be low even though the signal is strong. It may then not be possible to distinguish the target.

One observer's signal may be another's noise. The magnetic effect of a buried pipe is a nuisance when interpreting magnetic data in geological terms but may be invaluable to a site developer. Much geophysical field practice is dictated by the need to improve signal-to-noise ratios. In many cases, as in magnetic surveys, variations in a background field are a source of noise and must be precisely monitored.

1.5.7 Variance and standard deviation

The statistics of random noise are important in seismic, ground radar, radiometric and induced polarisation (IP) surveys. Adding together N statistically-long random series, each of average amplitude A, produces a random series with amplitude $A \times \sqrt{N}$. Since N identical signals of average amplitude A treated in this way produce a signal of amplitude $A \times N$, adding together (*stacking*) N signals containing some random noise should improve signal-to-noise ratios by a factor of $\sqrt{N}$.

Random variations may have a *normal* or *Gaussian* distribution, producing a bell-shaped probability curve. A normal distribution can be characterised by a *mean* (equal to the sum of all the values divided by the total
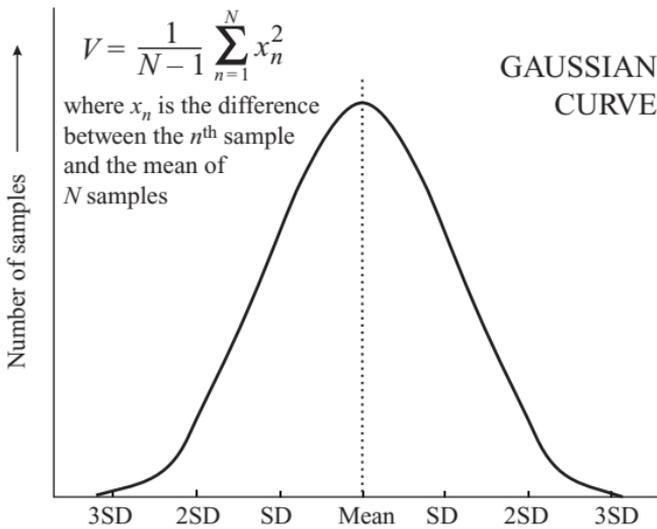
Figure 1.9 *Gaussian distribution. The curve is symmetric, and approximately two-thirds of the area beneath it lies within one standard deviation (SD) of the mean.* V = *variance.*

number of values) and a *variance* (*V*, defined in Figure 1.9) or its square root, the *standard deviation* (SD). About two-thirds of the readings in a normal distribution lie within 1 SD of the mean, and less than 0.3% differ from it by more than 3 SDs. The SD is popular with contractors when quoting survey reliability, since a small value can efficiently conceal several major errors. However, it is rare, in many types of geophysical survey, for enough field data to be obtained for statistical methods to be validly applied, and distributions are often assumed to be normal when they cannot be shown to be so.

Gaussian and more sophisticated statistical summaries of data (both background and target-related) are recommended for unexploded ordnance (UXO) surveys, where confidence is essential, to quantify the detection assurance level (the distance from a sensor within which a target of a certain size can be detected with 100% confidence). This measure will vary from site to site, and within a site, depending on the variable composition of made ground or geology, as well as on target size.

1.5.8 Anomalies

Only rarely is a single geophysical observation significant. Usually, many readings are needed, and regional background levels must be determined,

before interpretation can begin. Interpreters tend to concentrate on *anomalies* – that is, on differences from a constant or smoothly varying background. Anomalies take many forms. A massive sulphide deposit containing pyrrhotite would be dense, magnetic and electrically conductive (Table 1.2). Typical anomaly profiles recorded over such a body by various types of geophysical survey are shown in Figure 1.10. A wide variety of possible contour patterns correspond to these differently shaped profiles.

Background fields also vary and may, at different scales, be regarded as anomalous. A 'mineralisation' gravity anomaly, for example, might lie on a broader high due to a mass of basic rock. Separation of regionals from residuals is an important part of geophysical data processing, and even in the field it may be necessary to estimate background so that the significance of local anomalies can be assessed. On profiles, background fields estimated by eye may be more reliable than those obtained using a computer, because of the virtual impossibility of writing a computer program that will produce a background field that is not influenced by the anomalous values (Figure 1.11). Computer methods are, however, essential when deriving backgrounds from data gathered over areas rather than along single lines.

The existence of an anomaly indicates a difference between the real world and some simple model, and in gravity work the terms *free-air anomaly*, *Bouguer anomaly* and *isostatic anomaly* are commonly used to denote derived quantities that represent differences from gross Earth models. These so-called 'anomalies' are sometimes almost constant within a small survey area – that is, the area is not anomalous! Use of terms such as Bouguer *gravity* (rather than Bouguer anomaly) avoids this confusion.

1.5.9   Wavelengths and half-widths

Geophysical anomalies in profile often resemble transient waves but vary in space rather than time. In describing them the terms *frequency* and *frequency content* are often loosely used, although *wavenumber* (the number of complete waves in unit distance) is pedantically correct. *Wavelength* may be quite properly used of a spatially varying quantity, but where geophysical anomalies are concerned the use is imprecise, since an anomaly described as having a single 'wavelength' would be resolved by Fourier analysis into a number of components with different wavelengths.

A more easily estimated quantity is the *half-width*, which is equal to half the distance between the points at which the amplitude has fallen to half the anomaly maximum (cf. Figure 1.10a). This is roughly equal to a quarter of the wavelength of the dominant sinusoidal component, but has the advantage of being directly measurable on field data. Wavelengths and half-widths are important because they are related to the depths of sources. Other things being equal, the deeper the source, the broader the anomaly.
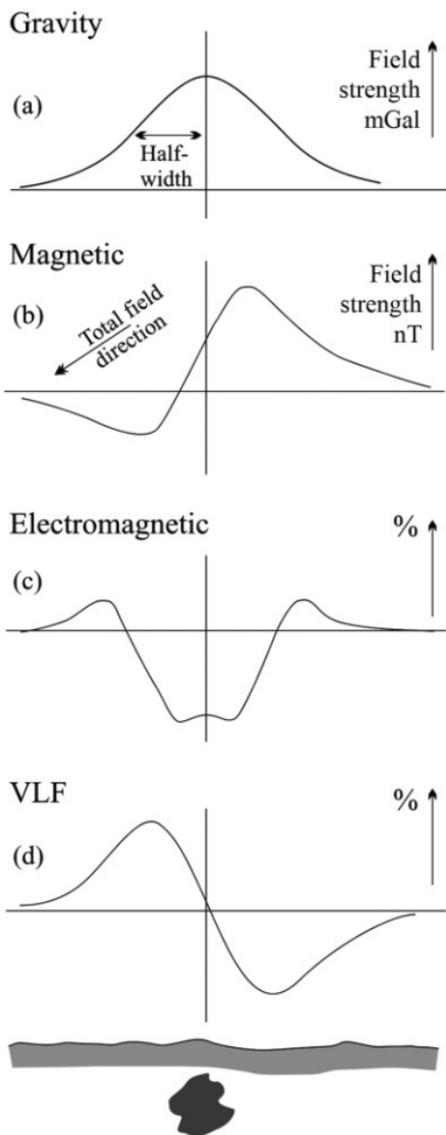
Figure 1.10 *Geophysical profiles across a pyrrhotite-bearing sulphide mass. The amplitude of the gravity anomaly (a) might be a few tenths of a milliGal, and of the magnetic anomaly (b) a few hundred nanotesla (nT). The electromagnetic anomalies are for a two-coil co-planar system (c) and a dip-angle system (d). Neither of these is likely to have an amplitude of more than about 20%.*
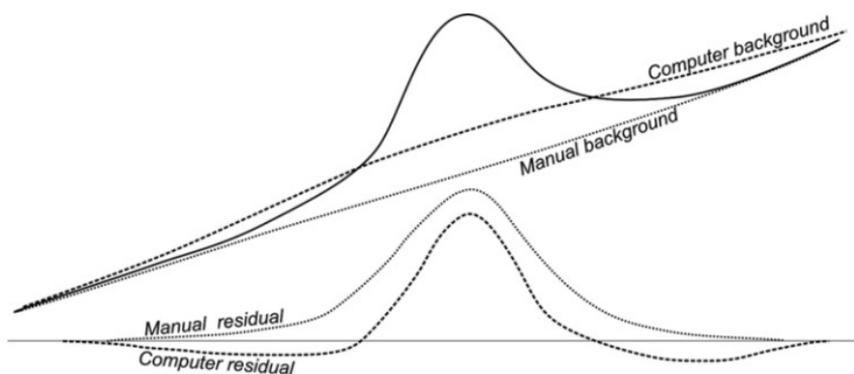
Figure 1.11 *Computer and manual residuals. The background field drawn by eye recognises the separation between regional and local anomaly, and the corresponding residual anomaly is probably a good approximation to the actual effect of the local source. The computer-drawn background field is biased by the presence of the local anomaly, and the corresponding residual anomaly is therefore flanked by troughs.*

1.5.10 Presentation of results

The results of surveys along traverse lines can be presented in profile form, as in Figure 1.10. It is usually possible to plot profiles in the field, or at least each evening, as work progresses, and such plots are vital for quality control. Most field crews now carry laptop computers, which can reduce the work involved, and many modern instruments and data loggers will display profiles in real time as work proceeds.

A traverse line plotted on a topographic map can be used as the baseline for a geophysical profile. This type of presentation is particularly helpful in identifying anomalies due to man-made features, since correlations with features such as roads and field boundaries are obvious. If profiles along a number of parallel traverses are plotted in this way on a single map they are said to be *stacked*, a word otherwise used for the addition of multiple data sets to form a single output set (see Section 1.5.7).

Contour maps used to be drawn in the field only if the strike of some feature had to be defined quickly so that infill work could be planned, but the routine use of laptop computers has vastly reduced the work involved. Information is, however, lost in contouring because it is not generally possible to choose a contour interval that faithfully records all the features of the original data. Also, contour lines are drawn in the areas between traverses, where there are no data, and inevitably introduce a form of noise. Examination of contour patterns is not, therefore, the complete answer to

Figure 1.12 Image-processed magnetic data over an archaeological site. (Reproduced by permission of Professor Irwin Scollar.)

field quality control. Contoured cross-sections (*pseudo-sections*) are used to display the results of some types of electrical survey.

In engineering site surveys, pollution monitoring and archaeology, the objects of interest are generally close to the surface and their positions in plan are usually much more important than their depths. They are, moreover, likely to be small and to produce anomalies detectable only over very small areas. Data have therefore to be collected on very closely spaced grids and can often be presented most effectively if background-adjusted values are used to determine the colour or grey-scale shades of picture elements (*pixels*) that can be manipulated by image-processing techniques. Interpretation then relies on pattern recognition and a single pixel value is seldom critically important. Noise is filtered by eye, patterns such as those in Figure 1.12 being easily recognised as due to human activity.

It can also be revealing to overlay contoured results on a Google Earth or other image. Many tools are available for doing this, ranging from full Geographic Information Systems (GIS) to simpler packages such as Global Mapper. Some also allow the transparency of overlaid pixel-based images to be adjusted so that features on the ground can be correlated with patterns in the geophysical data. This can be a powerful interpretation tool, provided, of course, that the ground features imaged were actually there at the time of the survey. It is also a valuable way of showing results to clients.

1.6   Bases and Base Networks

Bases or *base stations* are important in gravity and magnetic surveys, and in some electrical and radiometric work. They may be:

1. *Drift bases* – Repeat stations that mark the starts and ends of sequences of readings and are used to control drift.
2. *Reference bases* – Points where the value of the field being measured has already been established.
3. *Diurnal bases* – Points where regular measurements of background are made whilst field readings are taken elsewhere.

A single base may fulfil more than one of these functions. The reliability of a survey, and the ease with which later work can be tied to it, will often depend on the quality of the base stations. Base-station requirements for individual geophysical methods are considered in the appropriate chapters, but procedures common to more than one type of survey are discussed below.

1.6.1   Base station principles

There is no absolute reason why any of the three types of base should coincide, but surveys tend to be simpler and fewer errors are made if every *drift base* is also a *reference base*. If, as is usually the case, there are too few existing reference points for this to be done efficiently, the first step in a survey should be to establish an adequate base network.

It is not essential that the *diurnal base* be part of this network and, because two instruments cannot occupy exactly the same point at the same time, it may actually be inconvenient for it to be so. However, if a diurnal monitor has to be used, work will normally start each day by setting it up and end with its removal. It is good practice to read the field instruments at a drift base at or near the monitor position on these occasions, noting any differences between the simultaneous readings of the base and field instruments.

1.6.2   ABAB ties

Bases are normally linked together using ABAB ties (Figure 1.13). A reading is made at Base A and the instrument is then taken as quickly as possible to Base B. Repeat readings are then made at A and again at B. The times between readings should be short so that drift, and sometimes also background variation, can be assumed linear. The second reading at B may also be the first in a similar set linking B to a Base C, in a process known as *forward-looping*.

Each set of four readings provides two estimates of the difference in field strength between the two bases, and if these do not agree within the limits of instrument accuracy (± 1 nT in Figure 1.13), further ties should be made.
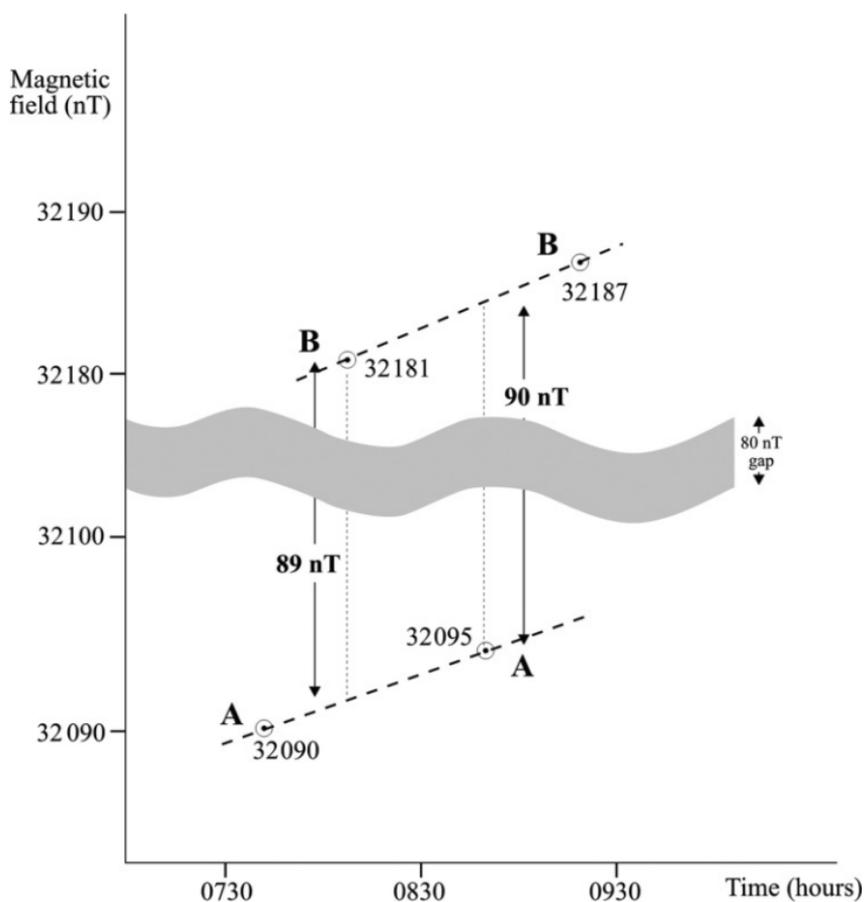
Figure 1.13 *ABAB tie between bases in a magnetic survey with a 1-nT instrument. The estimated difference between the two stations would be 89 nT. Note that the plotting scale should be appropriate to instrument sensitivity and that it may be necessary to 'remove' some of the range of the graph to allow points to be plotted with sufficient precision.*

Differences should be calculated in the field so that any necessary extra links can be added immediately.

1.6.3 Base networks

Most modern geophysical instruments are accurate and quite easy to read, so that the error in any ABAB estimate of the difference in value between two points should be small. Even so, the final value obtained at the end of an
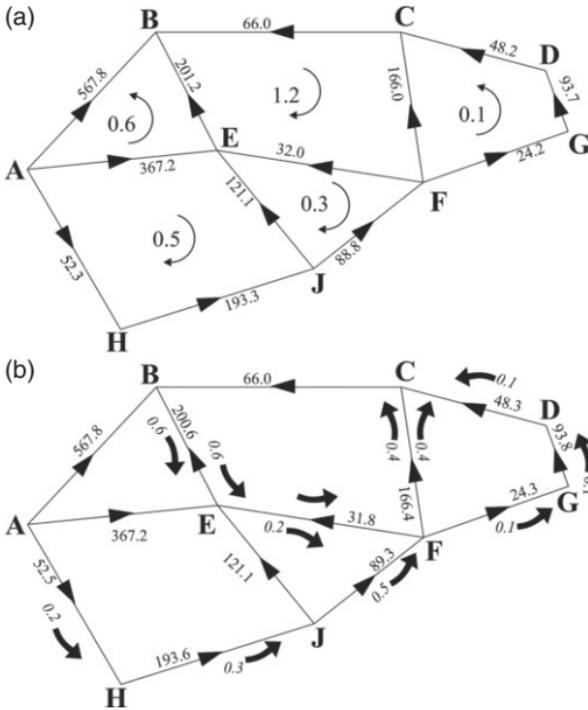
Figure 1.14 *Network adjustment. (a) The 1.2-unit misclosure in loop BCFE suggests a large error in either the 'unsupported' link BC or in BE, the only link shared with another loop with a large misclosure. (b) Adjustments made on the assumption that BC was checked and found to be correct but that no other checks could be made.*

extended series of links could include quite large accumulated errors. The integrity of a system of bases can be assured if they form part of a network in which each base is linked to at least two others. *Misclosures* are calculated by summing differences around each loop, with due regard to sign, and are then reduced to zero by making the smallest possible adjustments to individual differences. The network in Figure 1.14 is sufficiently simple to be adjusted by inspection. A more complicated network could be adjusted by computer, using least-squares or other criteria, but this is not generally necessary in small-scale surveys.

1.6.4 Selecting base stations
It is important that bases be adequately described and, where possible, permanently marked, so that extensions or infills can be linked to previous work

by exact reoccupations. Concrete or steel markers can be quickly destroyed, either deliberately or accidentally, and it is usually better to describe station locations in terms of existing features that are likely to be permanent. In any survey area there will be such points that are distinctive because of the presence of man-made or natural features. Written descriptions and sketches are the best way to preserve the information for the future. Sketches, such as those shown in Figure 2.7, are usually better than photographs, because they can emphasise salient points.

Permanence can be a problem, and maintaining gravity bases at international airports is almost impossible because building work is almost always underway (and, these days, because attempting to read a geophysical instrument anywhere near an airport is likely to trigger a security alert). Geodetic survey markers are usually reliable but may be in isolated and exposed locations. Statues, memorials and historic or religious buildings often provide sites that are not only quiet and permanent but also offer some shelter from sun, wind and rain.

1.7 Real-Time Profiling
During the past 20 years, automation of the geophysical equipment used in small-scale surveys has progressed from a rarity to a fact of life. Although many of the older types of instrument are still in use, and giving valuable service, they now compete with variants containing the sort of computer power employed, 40 years ago, to put a man on the Moon.

1.7.1 Data loggers
The integration of data loggers into geophysical instruments has its drawbacks. At least one manufacturer proudly boasted 'no notebook', even though the instrument in question was equipped with only a numerical key pad so that there was no way of entering text comments (*metadata*) into the (more than ample) memory. Other automated instruments have data displays that are so small and so poorly positioned that the possibility that the observer might actually want to look at, and even think about, the observations as they are being collected has clearly not been considered. Unfortunately, pessimism in this respect is often justified, partly because of the speed with which readings, even when essentially discontinuous, can now be taken and logged. Quality control thus often depends on the subsequent playback and display of whole sets of data, and it is absolutely essential that this is done at least once every day. As Oscar Wilde might have said (had he opted for a career in field geophysics), to spend a few hours recording rubbish might be accounted a misfortune. To spend anything more than a day doing so looks suspiciously like carelessness.

Automatic data loggers, whether 'built-in' or attached, are essential rather than optional if instruments are dragged, pushed or carried along a traverse to provide virtually continuous readings. In many cases, all that is required of the operators is that they press a key to initiate the reading process, walk along the traverse at constant speed and press the key again when the traverse is completed. Additional keystrokes should be used to 'mark' the passing of intermediate survey points on lines more than about 20 m long, but even this can be made unnecessary by integrating a DGPS unit (see Section 15.2) into the system. Many instruments can now record GPS data and can be synchronised using the GPS signal as a common time reference, enabling on-the-move recording to almost 1-metre positional accuracy with relatively cheap systems and without significant loss of data quality (Figure 1.15). Apart from the obvious productivity benefits of lines being traversed more quickly and survey grids being set out in significantly less time, the permanent record of where the instrument has actually measured data is valuable for quality control.

Figure 1.15 *Magnetometer coupled to a differential GPS navigation system for continuous profiling. Unless allowance is made in processing for the offset between the GPS and magnetic sensors, anomaly locations will be incorrectly plotted (photo courtesy of Geometrics).*

Figure 1.16 *Geonics EM-31 mounted on a quad bike. Induced currents will flow in the vehicle as well as in the ground, but should be reasonably constant.*

1.7.2 Vehicle-mounted systems

The increasing use of vehicle-mounted systems in medium- to large-scale geophysical surveys is a particularly welcome trend for those who (like the authors) have been worn down by a lifetime of walking lines carrying instruments. The system shown in Figure 1.16 is a good example of life being made very much easier. It was used to record ground conductivity data to delineate a fault, using a Geonics EM31-Mk2 and a DGPS system with EGNOS capability (see Section 15.2) to 2 m spatial accuracy, in less than one-third of the time it would have taken on foot.

Most continuously recording geophysical instruments can be mounted in this way, achieving significant cost benefits in open areas more than 5 hectares in area, if these are to be covered by lines more than 2 m apart. The agricultural quad-bike is the vehicle of choice. The main precaution required is regular checking of satellite coverage, and care must also be taken to travel at speeds compatible with the station interval needed to map the target. It is all too tempting to try to squeeze in a few extra lines by opening the throttle.

Figure 1.17 *Multiple systems mounted on a purpose-built wooden sledge.*

1.7.3 Towed systems

Putting a sensor on a survey vehicle is undesirable if the vehicle is likely to be a source of noise. It may also be difficult to mount all the equipment needed for a multi-system survey on the vehicle and still leave room for the driver. Towing the instruments behind the vehicle on a purpose-built sledge then becomes a better option. The towed system in Figure 1.17 was used to record combined ground conductivity, natural gamma and multiple total field magnetometer data to simultaneously map shallow geological deposits, cross-cutting pipelines, archaeological features and buried pits along a proposed linear route. A DGPS system with EGNOS availability (see Section 15.2) was used to record locations to approximately 2-m spatial accuracy.

Multi-instrument platforms in their most advanced form have been developed in the USA to improve the efficiency of scanning firing ranges and battlefields with multiple magnetometers and time domain EM systems synchronised so that they do not affect each other. The systems utilise not only real-time GPS control but are integrated with inertial navigation units to provide accurate dead reckoning navigation in the event of poor GPS signal. These platforms offer huge cost savings compared with separate surveys or

surveys on foot. The cost of negotiating land access can also be significantly reduced, because only a single visit is needed.

Designing towed systems can be challenging. Some signal sources will interfere with unrelated sensors if located too close to them without transmission and data-capture synchronisation, so the sensor layout must be carefully planned. Systems such as that shown in Figure 1.17 work well in reasonably flat terrain and in dry conditions. Add topography and wet weather, and the need to monitor the state of multiple instruments whilst tracking a survey grid barely visible on the screen of a tablet PC, and the days of trudging along lines on foot can seem like a lost paradise.

Handling the increase in data volumes and ensuring the accurate synchronisation of multiple datasets is also non-trivial, requiring a rigid set of procedures involving daily tests of repeatability and sensitivity and of the influence of the towing vehicle. It will also, inevitably, be necessary to deal with dropouts in one or more data channels. The process is not for the faint-hearted, and commercial surveys demand previous experience in running multi-instrument platforms and careful and detailed planning.

1.7.4 Errors in continuously recorded data

One consequence of semi-continuous operation has been the appearance in ground surveys of the sorts of errors that were once common in airborne surveys but have now been almost eliminated by improved compensation methods and GPS navigation. These were broadly divided into *parallax* errors, *heading* errors, ground clearance/coupling errors and errors due to speed variations.

With the carried system shown in Figure 1.15, parallax errors can occur because the magnetic sensor is about a metre ahead of the GPS sensor. Similar errors can occur in surveys where positions are recorded by keystrokes on a data logger. If the key is depressed when the operator, rather than the sensor, passes a survey peg, all readings will be displaced from their true positions. If, as is normal practice, alternate lines on a grid are traversed in opposite directions, a *herringbone* pattern (Figure 1.18) can be imposed on a linear anomaly, with the position of the peak fluctuating backwards and forwards according to the direction in which the operator was walking.

False anomalies can be produced in airborne surveys if ground clearance is allowed to vary, and similar effects can now be observed in ground surveys. Keeping the sensor shown in Figure 1.15 at a constant height above the ground is not easy (although a light flexible 'spacer' hanging from it can help). On level ground there tends to be a rhythmic effect associated with the operator's motion, and this can sometimes appear on contour maps as 'striping' at right angles to the traverse, as minor peaks and troughs on adjacent lines are linked to each other during contouring. On slopes there
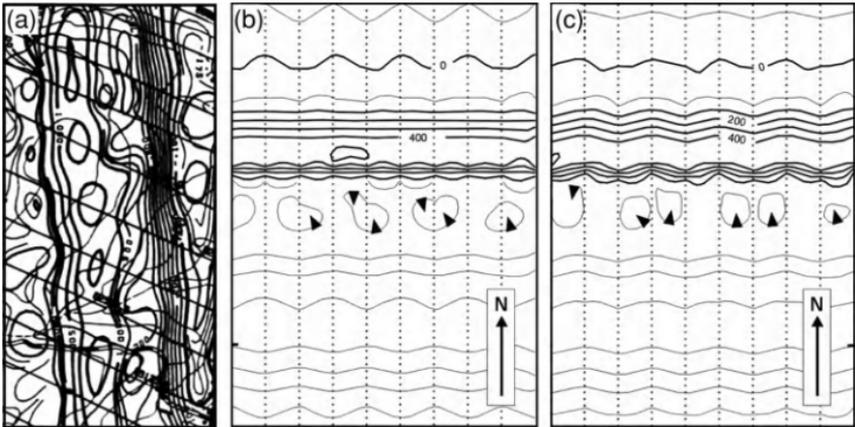
Figure 1.18 *Distortion in automated contouring of linear anomalies. (a) Introduction of closures in the peak of a linear aeromagnetic anomaly caused by the contouring program seeking (as most do) to equalise gradients in all directions. A similar effect can be seen in the 'bubbling' of the very closely spaced contours on the south side of the anomaly in (b). In neither case are the features required by the actual data, which exist only along the traverse lines indicated by lines in (a) and by discrete points in (b). (b) 'Herringbone' pattern due to a consistent difference in background levels on lines measured in opposite directions (see discussion in text). The effect is barely visible on the large main anomaly (thick contours at 100-nT intervals) but very obvious in the low-gradient areas where contours are at 10-nT intervals. (c) 'Herringbone' pattern due to parallax error. In this case there is a consistent offset between contour 'cuts' along lines recorded in opposite directions, regardless of anomaly magnitude.*

will inevitably be a tendency for a sensor carried in front of the observer to be closer to the ground when going uphill than when going downhill. How this effect will appear on the final maps will vary with the nature of the terrain, but in an area with constant slope there will be a tendency for background levels to be different on parallel lines traversed in opposite directions. This can produce herringbone effects on individual contour lines in low gradient areas (Figure 1.18).

Heading errors occurred in airborne (especially aeromagnetic) surveys because the effect of the aircraft on the sensor depended on aircraft orientation. A similar effect can occur in a ground magnetic survey if the observer is carrying any iron or steel material. The induced magnetisation in these objects will vary according to the facing direction, producing effects similar to those described above as being produced by constant slopes.

Before the introduction of GPS navigation, flight path recovery in airborne surveys relied on interpolation between points identified photographically. Necessarily, ground speed was assumed constant between these points, and anomalies were displaced if this was not the case. Similar effects can now be seen in data-logged ground surveys. Common reasons for slight displacements of anomalies are that the observer either presses the key to start recording at the start of the traverse, and then starts walking or, at the end of the traverse, stops walking and only then presses the key to stop recording. These effects can be avoided by insisting that observers begin walking before the start of the traverse and continue walking until the end point has been safely passed. If, however, speed changes are due to rugged ground, the most that can be done is to increase the number of 'marker' points.

2
GRAVITY METHOD

Differences in rock density produce small changes in the Earth's gravity field that can be measured using portable instruments known as gravity meters or gravimeters.

2.1 Physical Basis of the Gravity Method

The attraction between two point masses m_1 and m_2 is given by Newton's Law:

$$F = G\, m_1.m_2/r^2$$

The gravitational constant, G, has a value of 6.67×10^{-11} N m^2 kg^{-2}. Gravity fields, measured in the SI system in newtons per kilogram, are equivalent to accelerations, for which the numerically identical and more widely used unit is the m s^{-2}. This is inconveniently large for geophysical work and is divided by a million to produce the more practical 'gravity unit' (μN kg^{-1}, μm s^{-2} or 'g.u.'). In principle, the g.u. should long ago have replaced the c.g.s. 'milliGal' (mGal), which is equal to 10 g.u., but the older unit obstinately refuses to die. Because it is used in virtually all equipment manuals and most publications, it is also used in this chapter.

2.1.1 Gravity field of the Earth

The Earth's gravity field is approximately equal to that of a sphere with the same average radius and total mass, but increases slightly (by about 0.5%, or 5000 mGal) from the Equator to the poles. The rate of change with latitude is zero at the poles and Equator, and reaches a maximum of about 0.8 mGal per kilometre north or south at 45° latitude (Figure 2.1). The relationship, in mGal, between 'normal' sea-level gravity and latitude (λ) is described by the 1967 *International Gravity Formula (IGF67)*:

$$g_{\text{norm}} = 978\,031.85 + 5162.927 \sin^2 \lambda + 22.95 \sin^4 \lambda$$

The theoretical sea-level gravity at the Equator is thus $978\,031.85$ mGal. This formula replaced an earlier, 1930, version with slightly different
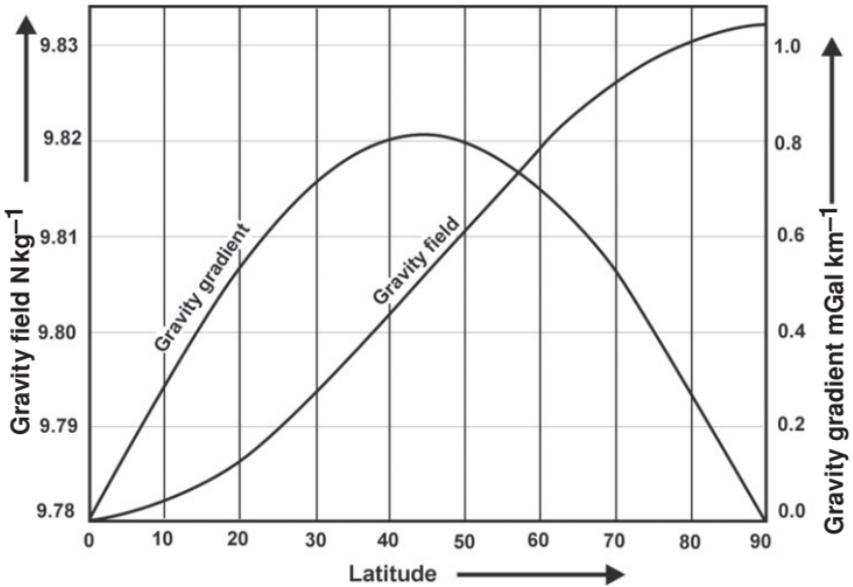
Figure 2.1 *Variation in theoretical sea-level gravity field and in the corresponding north-south rate of change with latitude. There is no east-west change in the theoretical field.*

constants, including an equatorial sea-level gravity of 978 049 mGal. The change of formula became necessary once it had been realised that the absolute gravity values at the base stations of the then current 'Potsdam' system were in error by some 16 mGal, and that correcting this error should incorporate the improved knowledge of the shape of the Earth. The network of international base stations compatible with IGF67 is known as IGSN71. It is still all too common to find surveys in which the 1930 formula has been applied to data referenced to IGSN71 bases or the IGF67 formula has been applied to Potsdam values, leading to errors of sometimes more than 16 mGal in latitude-corrected values.

In recent years, a more complicated formula compatible with the 1980 Geodetic Reference System (GRS80) has come into use. This is:

$$g_{norm} = \frac{978\,032.677\,15\,(1 + 0.001\,931\,851\,353\,\sin^2 \lambda)}{(1 - 0.006\,694\,380\,022\,9\,\sin^2 \lambda)^{1/2}}$$

The difference, of about 0.8 mGal, between this and IGF67 is largely accounted for by an additional, elevation dependent, correction for the mass

of the atmosphere, given by:

$$\delta g = 0.874 - 0.000\,099\,h + 0.000\,000\,003\,56\,h^2 \text{ mGal}$$

Provided this correction is included (and it is all too often overlooked), the actual differences in theoretical gravity between the 1967 and 1980 formulae are usually smaller than the errors in absolute gravity values at individual gravity stations. Since no changes in base station values are required, the changeover has been widely regarded as not urgent and is proceeding only slowly. The reluctance of many organisations to move to the new standard is increased by the fact that further refinements are being proposed almost yearly, although with negligible implications for practical gravity processing.

Gravity measurements are useful because subsurface changes can produce measurable deviations from the theoretical field. A major sedimentary basin can reduce the gravity field by more than 100 mGal, while targets such as massive ore bodies may produce anomalies of the order of a milliGal. The effects of caves and artificial cavities such as mine workings, even when very close to the surface, are usually even smaller. Gravity differences may therefore have to be measured to accuracies of at least 0.01 mGal (approximately one-hundred-millionth of the Earth's field), and this is the sensitivity of most manual instruments. Automatic meters and the so-called 'micro-gravity meters' have readout precisions of a microGal (μGal), but even their manufacturers do not claim them to be consistently accurate to better than about 3 μGal.

Topographic effects may be much larger. Elevation alone produces a gravity difference of nearly 2000 mGal between sea level and the summit of Mt Everest.

2.1.2   Rock density

The SI unit of density is the $kg\,m^{-3}$, but the $Mg\,m^{-3}$ is widely used because the values are, numerically, the same as those in the c.g.s. system, in which water has unit density. The density ranges for some common materials are shown in the first column of Table 1.2. Most crustal rocks have densities of between 2.0 and $2.9\,Mg\,m^{-3}$. A density of $2.67\,Mg\,m^{-3}$ was adopted as standard for the upper crust in the early days of gravity work, and is still widely used in modelling and in calculating elevation corrections for standardised gravity maps.

2.2   Gravity Meters

For the past 70 years the vast majority of gravity measurements on land have been made using meters with unstable (*astatic*) spring systems, and this
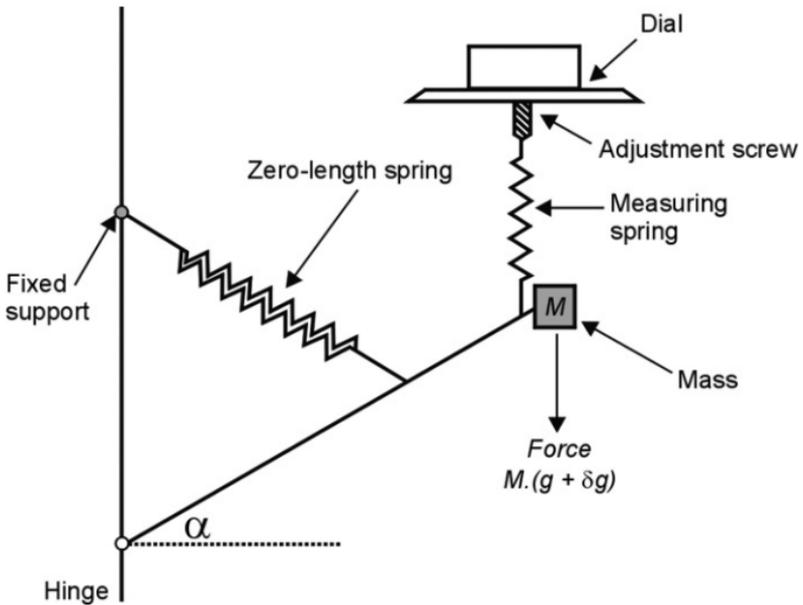
Figure 2.2 *A simple astatic gravity meter. The tension in the zero-length spring is proportional to its length and, with the geometry shown, it provides a moment that will support a mass* M *in some selected field* g, *regardless of the value of the angle* α. *Measurements are made by rotating the dial, which raises or lowers the measuring spring to provide a force* M.δg *to return the mass to a standard position.*

seems likely to remain the case for the foreseeable future. Gravity surveys are complicated by the fact that such meters measure gravity differences, not absolute field strengths.

2.2.1 Astatic spring systems

Astatic systems use *zero-length* main springs, in which tension is proportional to actual length. With the geometry shown in Figure 2.2 and for one particular value of gravity field, the main spring will support the balance arm in any position. In stronger fields, a much weaker auxiliary spring can be used to support the increase in weight, which will be equal to the product of the total mass and the increase in gravity field. To use an expression common in other geophysical methods, the zero-length spring *backs-off* a constant weight so that the measuring spring can respond to small changes in gravity field. All modern commercial gravity meters use this principle, although with additional complications. For example, LaCoste meters have

no auxiliary springs and measurements are made by displacing the support point of the main spring.

Because spring systems are mechanical, they are subject to drift. Temperature changes affect elasticity and cause short-term drift despite the use of various compensation devices. There is also longer-term extensional *creep* of springs under continual tension. Repeated readings at base stations are required to monitor drift and to allow corrections to be made.

Gravity meters remained essentially unchanged for almost the entire second half of the twentieth century, but major moves were made in its last decade towards automating readings and reducing the need for skilled operators. LaCoste G and D meters were fitted with automatic readouts, and the Scintrex CG-3 also pioneered automated tilt correction. The LaCoste meter was then completely redesigned as the fully automatic Graviton-EG, in which actual levelling, rather than merely the levelling correction, was automated. Inevitably, data loggers were also added to allow direct download to laptop PCs. The merger in 2001 between LaCoste and Scintrex effectively eliminated competition in the commercial manufacture of land gravity meters, and the CG-5 is the only instrument now being actively marketed. However, the durability (and, to some extent, the high cost) of gravity meters ensures that manual instruments will be around for many years to come.

2.2.2 Quartz astatic meters

Manual meters with spring systems made of fused quartz enclosed in vacuum chambers (to maximise thermal insulation) dominated land gravity surveying between about 1930 and 1970. Although now obsolete, they are still occasionally encountered and are recognisable by their resemblance to (rather large) thermos flasks. They are characterised by high drift rates, vulnerability to shock (because the spring systems cannot be clamped for transit) and limited ranges. It took a very experienced and conscientious observer to maintain the theoretical reading accuracy of 0.01 mGal. Quartz spring/vacuum chamber systems now again dominate the market, in the form of the Scintrex CG-5 automatic meter.

2.2.3 Steel astatic meters

LaCoste-Romberg meters, which are the only manually operated meters still in common use, have steel spring systems that can be clamped for transit. When clamped, these meters are reputedly able to survive any shocks that do not fracture the outer casing. There are two main types. The G (geodetic) meter has a single long measuring screw that is used to give readings worldwide without resetting, whereas the D-meter, used for 'microgravity' surveys, sacrifices this advantage in the interests of greater reading

precision. Calibration factors vary slightly over G-meter ranges, being tabulated by the manufacturer for 100-mGal intervals.

Because steel is a good conductor of heat, LaCoste G and D meters cannot be adequately insulated and thermostatic control is essential. Their weight, of about 5 kg, is effectively doubled by the rechargeable battery required by the heater. A battery charger is needed in the field, since a single charge lasts only one or two days, depending on thermostat setting and external temperature. Batteries must be changed before they are exhausted because the drift for two or three hours after reaching operating temperature is so high that the instrument is unusable. Drift may then be very low, in which case it can be extrapolated linearly across intervals during which the meter has been off-heat. However, the main form of drift is by discontinuous *tares*, of perhaps as much as 1 mGal, which can occur at any time. An instrument that suffers tares more than about once a month requires servicing.

Even quite inexperienced observers have little difficulty in achieving accuracies of 0.01 mGal with LaCoste meters, particularly if an electronic readout has been installed.

2.2.4 Setting up a manual gravity meter

Manual gravity meters are normally read on concave dishes supported by three short stubs, which should be pressed firmly but not too deeply into the ground. The undersurface of the dish must not touch the ground since a fourth support point allows 'rocking' back and forth. Thick grass under the dish may have to be removed before a reading can be taken. Extension legs may be attached to the stubs but readings will then take longer, the dish itself may have to be levelled (some incorporate a bull's-eye bubble) and the height of the meter above the ground will have to be measured.

The meters themselves rest on three adjustable, screw-threaded feet and are levelled using two horizontal spirit-levels (Figure 2.3), initially by being moved around the dish until both level bubbles are 'floating'. The temptation to hurry this stage and use the foot-screws almost immediately should be resisted.

Usually one of the levels (probably the *cross-level*, at right angles to the plane of movement of the balance arm) is set parallel to a line joining two of the feet. Adjustments to the third foot then scarcely affect this level. The quickest method of levelling is to centre the cross-level bubble, using one or both of the two foot-screws that control it, and then use the third screw to set the *long-level*. Experienced observers often use two screws simultaneously but the ability to do this efficiently comes only with practice.

Once the meter is level, the reading is obtained by rotating a calibrated dial to bring a pointer linked to the spring system to a fixed point on a graduated scale viewed through a telescopic eye-piece. The alignment is
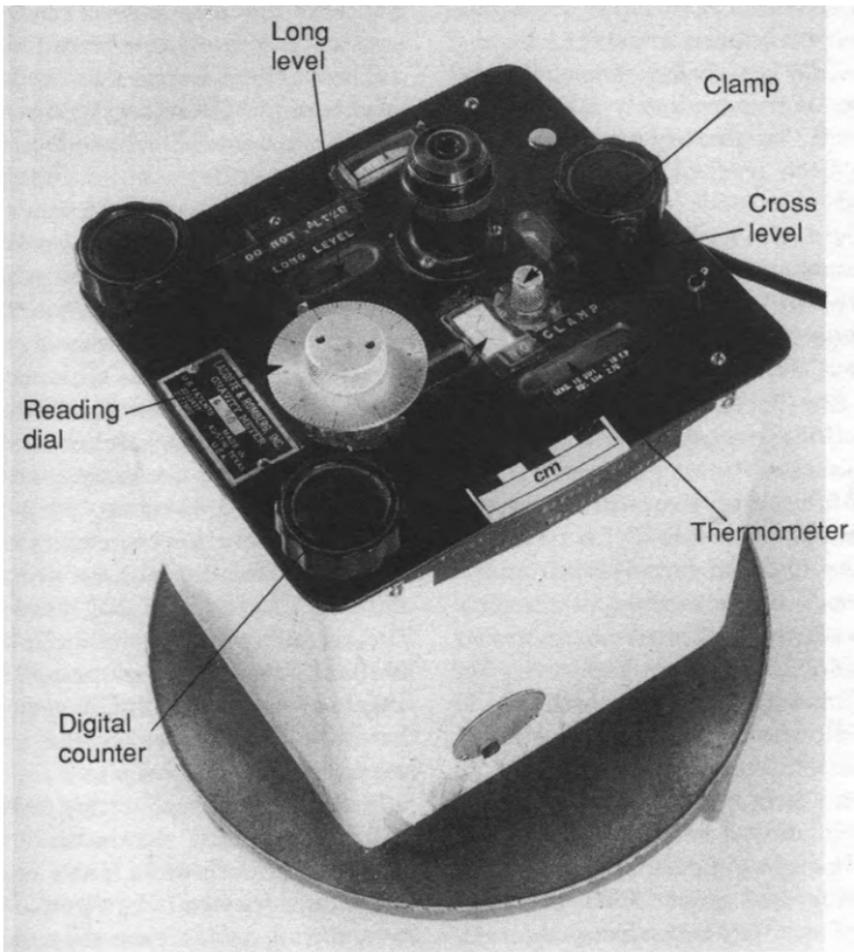
Figure 2.3 *Controls of the LaCoste 'G' meter. Note the two level bubbles at right angles, the clamp and the aluminium reading dial. The digital counter is behind the small window between the clamp and the dial. The thermometer, viewed through a window in front of the clamp, monitors internal temperature and must show the pre-set operating temperature if the instrument is to be usable.*

rather subjective, and all readings in a single loop should be made by the same observer so that this subjectivity is eliminated when the base reading is subtracted. The subjective element is much reduced in instruments fitted with electronic repeaters.

Other precautions have to be taken. It is vital that the level bubbles are checked whilst the dial is being adjusted, and especially immediately after a supposedly satisfactory reading has been taken, since natural surfaces can subside slowly under the weight of an observer, off-levelling the meter. On snow or ice, the levels have to be adjusted almost continuously as the dish melts its way down, unless it is insulated from the surface by being placed on a small piece of plywood.

All mechanical measuring systems suffer from *whiplash*, and two readings will differ, even if taken within seconds of each other, if the final adjustments are made by opposite rotations of the reading dial. The only remedy is total consistency in the direction of the final adjustment.

Severe continuous vibration, as from nearby machinery or the roots of trees moved by the wind, can make reading difficult and may even displace the reading point. Earthquakes can make the pointer swing slowly from side to side across the field of view and survey work must be stopped until the disturbance is over. Fortunately, this effect is rare in most parts of the world, although very large earthquakes can affect gravity meters at distances of more than 10 000 km.

2.2.5 Meter checks

A series of checks should be made each day before beginning routine survey work. The meter should first be *shaken down* by tapping the dish gently with a pencil between readings until a constant value is recorded (this can also be done if, as sometimes happens, the pointer 'sticks' on one or other of the stops).

The levelling system should then be checked. Because astatic systems are asymmetric, the effect of a levelling error depends on the direction of tilt. A slight error on the cross-level, which is at right angles to the balance arm, produces a reading of gravity field multiplied by the cosine of the tilt angle (an error of about 0.015 mGal for a tilt of 0.01°). Off-levelling a correctly adjusted cross-level will therefore reduce the reading, regardless of the direction of offset. To check that this actually happens, the meter should be set up and read normally and the cross-level should then be offset by equal amounts in both directions. The pointer should move roughly the same distance in the same direction in each case. Meters are usually considered usable if the movements are at least in the same direction, but otherwise the level must be adjusted.

If the long-level, in the plane of the balance arm, is in error, reading *sensitivity* (the amount the pointer moves for a given rotation of the dial) is affected. The recommended sensitivity and instructions for resetting will be found in the handbook. The actual sensitivity can be estimated by moving the dial by a set amount and noting the pointer movement. After adjustment,

levels often take a few days to settle into their new positions, during which time they must be rechecked with special care.

2.2.6 Meter calibration

The readings from LaCoste manual meters are combinations of values read from a dial and numbers displayed on a mechanical counter. The sensitivity of G-meters is such that the final figure read from the dial corresponds to approximately 0.01 mGal.

Readings are converted to gravity units using calibration factors specific to the individual instrument. The factors quoted by the manufacturers require the insertion of a decimal point somewhere in the reading, its position depending on whether results are desired in mGal or g.u. They are not affected by changes in reading sensitivity but may alter slowly with time and should be checked regularly. This can be done by the manufacturers or by using calibration ranges of known gravity interval. Calibration ranges usually involve gravity changes of about 50 mGal, which is within the range of even the most limited-range meters, and almost always make use of the rapid change of gravity field with elevation. An elevation difference of about 250 m is generally necessary, although in some cases local gravity gradients can also play a part. Travel times between top and bottom stations should normally be less than 15 minutes, and the two stations should be well marked and described. A *run* should consist of at least an ABAB tie (Section 1.6.2), giving two estimates of gravity difference. If these differ by more than 0.02 mGal, more links should be added.

D-meters, which have separate fine and coarse adjustment dials, can be checked over different sections of their fine ranges by slightly altering the coarse setting. Most meters need a little time to stabilise after coarse adjustment, but if this is allowed it may be possible to identify minor irregularities in the calibration curve. With G-meters, only one part of the curve can be monitored on any one calibration range and, because of slight irregularities in the pitch of the adjustment screws, different meters used on the same range may give results that differ consistently by a few hundredths of a milliGal.

Commercial manufacture of manual gravity meters has now ceased. They are still in use in large numbers and may be available second-hand, but are steadily being replaced by automatic models.

2.2.7 Automatic gravity meters – the CG-5

In the automatic CG-5 (Figure 2.4), the reading system, batteries and vacuum-insulated quartz sensor are all contained within a single housing and carrying case, which considerably simplifies field operations. The elimination of the cable, which in LaCoste meters links the instrument to the

Figure 2.4 *The Scintrex CG-5 automatic gravity meter.*

battery, has removed one cause of potentially expensive accidents. Operator reading errors are also eliminated, because readings are stored in flash memory instead of (or as well as, because of the need to record ancillary data) in field notebooks. The instrument also incorporates a rather basic GPS that can be used to record meter position, but not elevation. Levelling is achieved by first rough-levelling the detachable tripod and then refining the levelling with the instrument itself. The sequencing in the use of the levelling screws is similar to that described in Section 2.2.4 for manual gravity meters. Readings are taken simply by pressing a key and waiting. The length of the wait is determined by the accuracy required, but only the most detailed surveys, demanding microGal accuracy, require reading times of more than one minute.

The CG-5 can be used for a full working day at normal temperate-zone temperatures, provided that the two internal lithium-ion batteries are fully charged at the start. These batteries have low leakage rates when disconnected and high power-to-weight ratios, but also have their problems. They

deteriorate with age more quickly than other rechargeable batteries, and overheating or overcharging can destroy them, or even make them explode. They can also be damaged if they are allowed to discharge completely. The CG-5 gives an audible warning as full discharge is approached, and the charger supplied is designed to prevent overcharging.

Surveying with the CG-5 is not always just a matter of switching on and starting work. The manufacturer recommends that the instrument should be stored 'on-heat', but if it has been powered down, it requires 4 hours to reach operating temperature and a further 48 hours to fully stabilise. Clearly, it is better to bring it fully operational to the survey area, but this may not be possible if the travel time exceeds half a day. Transport by air presents additional problems. The regulations covering lithium-ion batteries make it illegal for the meter to be in the hold baggage, and this can be useful when dealing with ground staff who want to see the meter handed over to the tender mercies of their baggage handlers. However, although the capacity of the batteries is in principle low enough for them to be accepted as cabin baggage, their mere presence means that the meter can be rejected simply 'at the discretion of the airline'.

All the techniques used, in surveys with manual meters, to control drift and convert meter readings to absolute values by the use of base stations must be used also in surveys with automatic meters.

2.3 Gravity Reductions

In gravity work, more than in any other branch of geophysics, large and (in principle) calculable effects are produced by sources that are not of direct geological interest. These effects are removed by *reductions* that involve the sequential calculation of a number of recognised quantities. In each case a positive effect is one that increases the magnitude of the measured field and the sign of the reduction is opposite to that of the effect it is designed to remove.

2.3.1 Latitude correction

Latitude corrections are usually made by subtracting the *normal* gravity, calculated from the International Gravity Formula, from the *observed* or absolute gravity. For surveys not tied to the absolute reference system, local latitude corrections may be made by selecting an arbitrary base and using the theoretical north-south gradient of about $0.812 \sin 2\lambda$ mGal km^{-1}.

2.3.2 Free-air correction

The remainder left after subtracting the normal from the observed gravity will be due in part to the height of the gravity station above the sea-level reference surface. An increase in height implies an increase in distance from

the Earth's centre of mass and the effect is therefore negative (i.e. the *free-air correction* is positive) for stations above sea level. It varies slightly with latitude and elevation, according to the equation:

$$\delta g = (0.308\,769\,1 - 0.000\,439\,8\,\sin^2\lambda)\,h + 0.000\,000\,072\,125\,h^2\,\text{mGal}$$

An average value of $0.3086\,\text{mGal}\,\text{m}^{-1}$ is often used, regardless of latitude. The quantity obtained after applying both the latitude and free-air corrections is termed the *free-air anomaly* or *free-air gravity*.

2.3.3 Bouguer correction

Since topographic masses are irregularly distributed, their effects are difficult to calculate precisely and approximation is necessary. The simplest approach assumes that topography can be represented by a flat plate, with constant density and thickness equal to the height of the gravity station above the reference surface, extending to infinity in all directions. This *Bouguer plate* produces a gravity field equal to $2\pi\rho Gh$, where h is the plate thickness and ρ the density. The correction is equal to $0.1119\,\text{mGal}\,\text{m}^{-1}$ at the standard $2.67\,\text{Mg}\,\text{m}^{-3}$ density.

The Bouguer effect is positive and the correction is therefore negative. Since it is only about one-third of the free-air correction, the net effect of an increase in height is a reduction in field. The combined correction is positive and equal to about $0.2\,\text{mGal}\,\text{m}^{-1}$, so elevations must be known to $0.5\,\text{cm}$ to make full use of the microGal sensitivity of the CG-5 gravity meter.

Because Bouguer corrections depend on assumed densities as well as measured heights, they are fundamentally different from free-air corrections, and combining the two into unified elevation corrections can be misleading. It is also sometimes stated that the combined corrections reduce gravity values to those that would have been obtained had the readings been made on the reference surface, with all the topography removed. This is not true. In Figure 2.5, the effect of the mass 'M' recorded at the observation point P is not altered by these corrections. It remains the effect of a body a distance 1.5 h below P, not a distance 0.5 h below P′. Still more obviously, the corrections do not eliminate the effect of the mass 'm', which is above the reference surface, since the Bouguer correction assumes constant density. Bouguer gravity is determined at the points where measurements are made, and this fact must be taken into account during interpretation.

2.3.4 Terrain corrections

In areas of high relief, detailed topographic corrections must be made. Although it would be possible to correct directly for the entire topography above the reference surface in one pass without first making the Bouguer
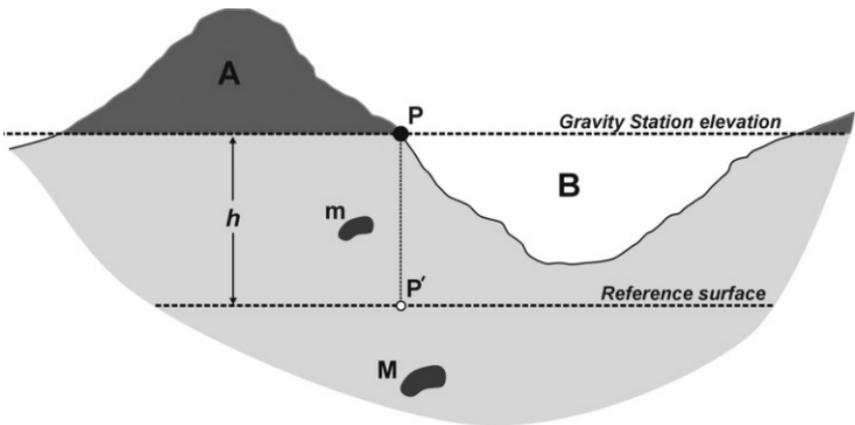
Figure 2.5 *Bouguer and terrain corrections. Even after the application of the Bouguer and free-air corrections, the gravity effects of the masses M and m will appear on the maps as they are measured at the station point P, and not as they would be measured at P′ on the reference surface. Terrain corrections, because they are made for the deviations of the topography from a surface through the gravity station and parallel to sea level, and not from sea level itself, are always positive (see discussion in text).*

correction, it is simpler to calculate the *Bouguer gravity* and then correct for deviations from the Bouguer plate.

A peculiarity of the two-stage approach is that the second-stage corrections are always positive. In Figure 2.5, the topographic mass (A) above the gravity station exerts an upward pull on the gravity meter, the effect is negative and the correction is positive. The valley (B), on the other hand, occupies a region that the Bouguer correction assumed to be filled with rock that would exert a downward gravitational pull. This rock does not exist. The terrain correction must compensate for an overcorrection by the Bouguer plate and is again positive.

Terrain corrections can be extremely tedious. To make them manually, a transparent *Hammer chart* is centred on the gravity station on the topographic map (Figure 2.6) and the difference between the average height of the terrain and the station height is estimated for each compartment. The corresponding corrections are then obtained from tables (see Appendix). Computers can simplify this process but require terrain data in digital form and may be equally time-consuming unless a digital terrain model (*DTM*) already exists. The SRTM and ASTER topographic grids discussed in Section 1.3.2 are extremely valuable in this respect, but are not reliable where slopes are very steep. Corrections for the very near topography have to be
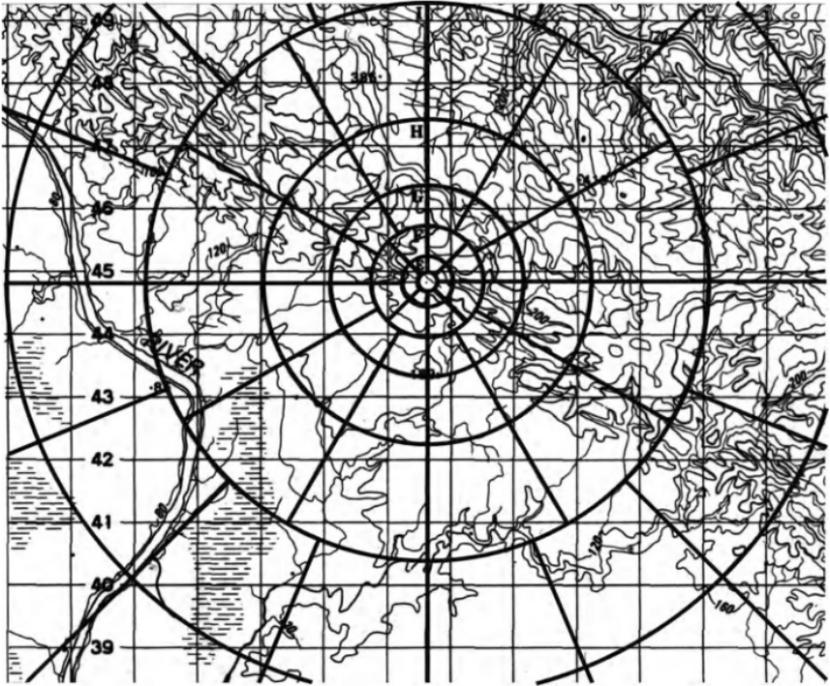
Figure 2.6 *Hammer chart (Zones E to I) overlaid on a topographic map. The difficulties in estimating average heights in the larger compartments are easily appreciated. The letters identifying the zones are difficult to see in this example but are clear when the overlay is removed from the map and viewed on its own.*

estimated in the field (see Section 2.4.3) but can often be minimised by careful selection of station location.

Adding terrain corrections to the simple Bouguer gravity produces a quantity often known as the *extended* or *complete Bouguer gravity*. Topographic densities are sometimes varied with geology in attempts to reduce terrain dependence still further.

2.4 Gravity Surveys

A gravity survey is a basically simple operation but few are completed wholly without problems, and in some cases the outcomes can only be described as disastrous. Most of the difficulties arise because gravity meters measure only differences in gravity field, so that readings have to be interrelated by links to a common reference system.

2.4.1 Survey principles

A gravity survey consists of a number of *loops*, each of which begins and ends with readings at the same point, the *drift base* (see Section 1.6). The size of the loop is usually dictated by the need to monitor drift and will vary with the transport being used and the accuracy desired; 2-hour loops are common in very detailed work where the drift base is never more than a short walk away. At least one station of the reference network must be occupied in the course of each loop, and operations are simplified if this is also the drift base for that loop. In principle, the network can be allowed to emerge gradually as the work proceeds, but if it is completed and adjusted early, absolute values can be calculated as soon as each field station has been occupied, allowing possible errors to be identified while there is still time for field checks to be made. There is also much to be gained from the early overview of the whole survey area that can be obtained while the network is being set up, and practical advantages in establishing bases while not under the pressure to maximise the daily total of new stations that characterises the routine production phase of most surveys.

A small survey may use an arbitrary base without any tie to an absolute system. Problems will arise only if such a survey has later to be linked to others or added to a national database. This often happens eventually, and use of a purely local reference may be a false economy.

2.4.2 Base stations

The criteria used in selecting reference bases differ from those for normal stations. Provided that exact reoccupation is possible, large terrain effects can be tolerated. These may make it inadvisable to use the value interpretatively, in which case the elevation is not needed either. On the other hand, since the overall survey accuracy depends on repeated base readings, easy access and quiet environments are important. Traffic noise and other strong vibrations can invalidate base (or any other) readings. Also, the general principles outlined in Section 1.6 apply to gravity bases, and descriptions should be provided in the form of sketch plans permitting reoccupation exactly in elevation and to within a few centimetres in position (Figure 2.7).

2.4.3 Station positioning

The sites of field stations must also be chosen with care. Except in detailed surveys where stations are at fixed intervals along traverses, observers in the field have some, and often considerable, freedom of choice. They also have the responsibility for estimating terrain corrections within about 50 metres from the reading point, where features too small to be shown on any
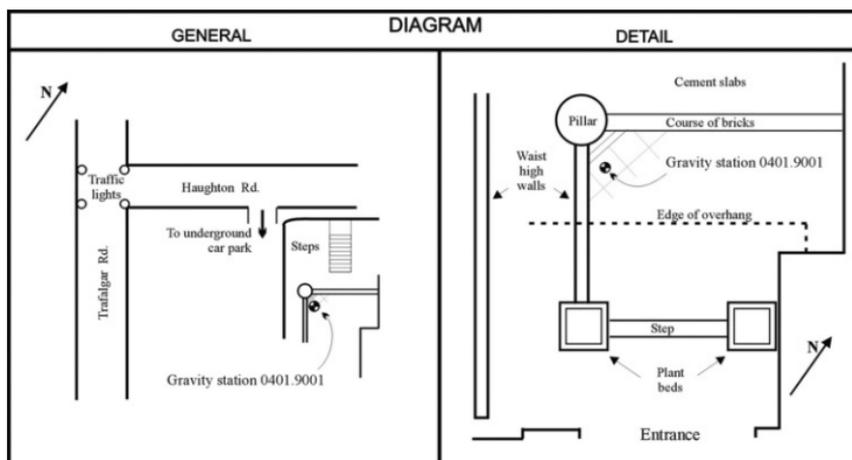
Figure 2.7 *Gravity base station sketches. Two sketches, at different scales, together with a short written description, are usually needed to ensure the station can be reoccupied quickly and accurately.*

topographic map can be gravitationally significant. Corrections can be estimated in the field using a truncated graticule such as that in Figure 2.8, which covers the Hammer Zones B and C only. Height differences of less than 30 cm in Zone B and 130 cm in Zone C can be ignored since they produce effects of less than a microGal per compartment. The charts can also be used qualitatively, to select reading points where overall terrain corrections will be small. The CG-5 gravity meter can calculate corrections for the B, C and D zones, provided that the necessary topographic information is entered before the reading sequence is begun.

The effect of a normal survey vehicle is detectable only if the observer actually crawls underneath it, and most modern buildings produce similarly small effects. Old, thick-walled structures may need to be treated with more respect (Figure 2.9). Subsurface cavities, whether cellars, mine-workings or natural caverns, can produce anomalies of well over 0.1 mGal. The gravity method is sometimes used in cavity detection but where this is not the object of the survey it is obviously important that stations are not sited where such effects may occur.

2.4.4 Tidal effects

Before meter drift can be estimated (but after conversion of the reading from meter scale divisions to mGal) allowance must be made for *Earth tides.* These are background variations due to changes in the relative positions
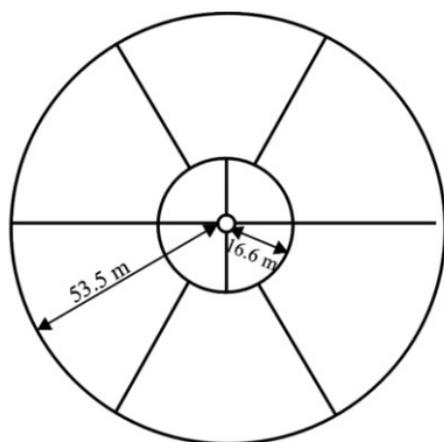
Area -------------------

Date -------------------

Station ----------------

Observer ---------------

Note: Terrain should be flat in central zone A (radius 2m)

ZONE B (2.0–16.6 m)	
Terrain correction (mGal)	Height difference (metres)
0.001	0.3–0.6
0.002	0.6–0.8
0.003	0.8–0.9
0.004	0.9–1.0
0.005	1.0–1.1
0.01	1.1–2.1
0.02	2.1–2.7
0.03	2.7–3.6
0.04	3.6–4.3
0.05	4.3–4.9

ZONE C (16.6–53.5 m)	
Terrain correction (mGal)	Height difference (metres)
0.001	1.3–2.3
0.002	2.3–3.0
0.003	3.0–3.5
0.004	3.5–4.0
0.005	4.0–4.4
0.01	4.4–7.3
0.02	7.3–9.7
0.03	9.7–11.9
0.04	11.9–13.7
0.05	13.7–15.5

Figure 2.8 *Field observer's Hammer Chart, for Zones B and C.*

of the Earth, Moon and Sun, and follow linked 12- and 24-hour cycles superimposed on a cycle related to the lunar month (Figure 2.10). Swings are largest at new and full moons, when the Earth, Moon and Sun are in line, when changes of more than $50\,\mu$Gal may occur within an hour and total changes may exceed $250\,\mu$Gal. The assumption of linearity made in correcting for drift may fail if tidal effects are not first removed.

Earth tides are predictable, at least at the 0.01 mGal level required for most gravity surveys, and corrections can be calculated using widely available
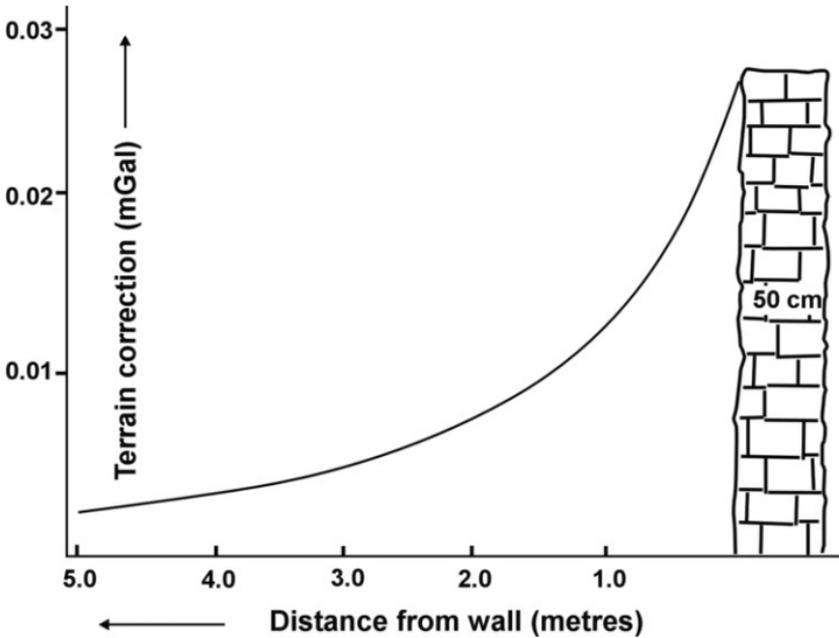
Figure 2.9 *Effect of a half-metre thick stone wall on the gravity field.*

computer programs (or the CG-5 inbuilt program). However, the Earth itself deforms in response to tidal forces, with a secondary effect on surface gravity field, and tidal calculations use a correction for Earth elasticity that is a global average. In extremely accurate work, automatic base stations may be needed to record the actual variations.

Figure 2.10 *Typical (actually 5 to 15 January 1986) tidal variations, in* µ*Gal.*

2.4.5   Drift corrections

Absolute observed gravities are obtained by adding the absolute value at the drift base to the drift-corrected gravity differences.

To correct manually for instrument drift, readings are first tidally corrected and the corrected initial reading at the drift base is then subtracted from every other corrected reading in turn. Th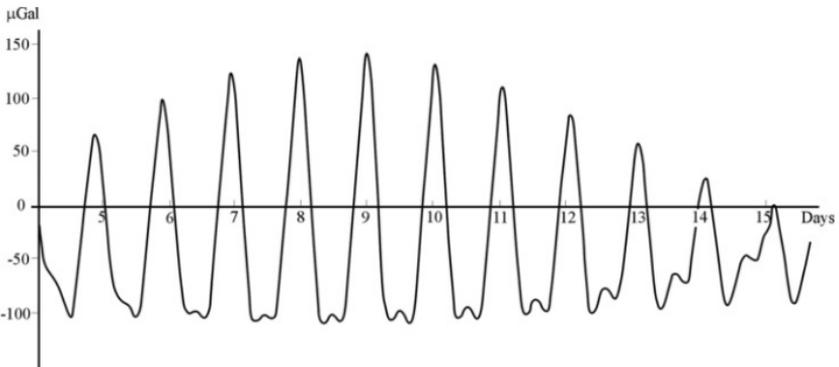e result of doing this to the final reading at the drift base gives the total drift. The *pro rata* corrections to the other stations can then be calculated or estimated graphically to the necessary accuracy. The drift is assumed to be linear and the sign of the correction is dictated by the requirement that, after correction, the relative values for all occupations of the drift-base should be zero.

Short-term drift is dependent mainly on temperature, and the assumption that it has been linear in the time between two base readings is unlikely to be true if large changes in temperature have occurred and then been wholly or partly reversed during that time. In the CG-5, the internal temperature is recorded and compensated automatically.

2.4.6   Barometric pressure and water-level corrections

Gravity meters will 'drift' in response to changes in atmospheric pressure, but magnitudes have been steadily reduced by improved instrument design. For the CG-5, the quoted value of pressure sensitivity is now only 0.015 μGal per millibar ($0.15\,\mu Gal\,kPa^{-1}$), so that typical diurnal changes of several millibars have negligible effects. On the other hand, an elevation change of only 10 metres produces a change in pressure of about a millibar, so that successive readings at points differing in elevation by hundreds or thousands of metres can be measurably affected by pressure-induced apparent drifts. It is hard to imagine such surveys requiring microGal accuracies.

Large pressure changes imply changes in atmospheric loading, and a millibar increase in pressure will reduce the real gravity field by 0.3–0.4 μGal. The lack of a universally applicable formula for calculating this effect is another reason why, in very-high-accuracy work, frequent visits to base stations are advisable. Altitude-dependent pressure changes are taken into account in the 1980 version of the International Gravity Formula (see Section 2.1.1).

Readings taken on openwork piers or jetties are affected by tidal changes in sea level. If the sea actually comes in under a reading point, the effect could be as much as 0.04 mGal for a metre change in sea level. Occasionally base stations have been established on such structures for use by research vessels, but there can be few other reasons for reading a gravity meter on them. A reading taken on a mole (rock-fill jetty) or on a low cliff dropping steeply to the sea could also be affected, perhaps by as much as 0.02 mGal per metre of sea-level change.

In areas subject to heavy seasonal rainfall or snow melt, the level of groundwater in the soil or in surface water may change significantly. The effect is small and dependent on porosity, but is unlikely to exceed 0.01 mGal for a metre change in water table (which is not likely to occur during a single day's surveying).

2.4.7 Elevation control

The ways in which the elevations of gravity survey points are determined depend on the purpose of the survey. An elevation error of 5 cm produces a 0.01 mGal error in Bouguer gravity, so 0.01 mGal contours, which are at the limit of what is achievable with modern meters, require at least 1 cm relative accuracies. This is still best achieved using optical levelling. Contours at the 0.1 mGal level require elevation control to ± 10 cm, which can be obtained with real-time kinetic or differential GPS (see Chapter 15). Barometric levelling or direct reference to sea-level and tide tables may be adequate for the 5 or 10 mGal contours common in regional surveys.

Measuring elevation is an important and often time-consuming part of a gravity survey, and advantage should be taken of any 'free' levelling that has been done for other purposes, e.g. surveyed seismic or resistivity lines.

2.4.8 Error summary

An understanding of the environmental and other influences on reading accuracy is essential, especially in microgravity surveys. The listing in Table 2.1 of typical effect magnitudes also summarises the actions that can be taken to minimise those effects. The impressive reading accuracies quoted by instrument manufacturers ($\sim 2\,\mu$Gal for the CG-5 and CG-3) are almost irrelevant unless all other effects are taken into account. In microgravity surveys the time between base station readings should not exceed 1–2 hours.

2.4.9 Field notebooks

At each station, the number, time and reading must be recorded, and automatic meters incorporate data loggers that do this at a keystroke. Other information, for example positional information (unless there is a sufficiently accurate built-in GPS receiver) and elevation data from barometers, must be recorded in field notebooks. Any factors that might affect readings, such as heavy vibrations from machinery, traffic, livestock or people, unstable ground or the possible presence of underground cavities, should be noted in a *Remarks* column. Comments on weather conditions may also be useful, even if only as indicating the observer's likely state of mind. Where local terrain corrections are only occasionally significant, estimates may also be entered as 'Remarks', but individual terrain-correction sheets may

Table 2.1 *Typical magnitudes of sources of error, and corresponding mitigating steps*

| Source of errors	Typical magnitude	Mitigating steps
Seismic (earthquakes)	Possibly hundreds of mGal at stations near earthquake epicentres, but typically 0.1 mGal for remote earthquakes	Survey work may have to be suspended for periods of minutes to several hours
Microseismicity related to wave action on shorelines, tree root movements, vehicle movement, on-site drilling or construction activities	0.01 to 0.1 mGal	Digitally stack readings and use a statistical measure (e.g. convergence of means) to decide when a reading is representative. It may be necessary to occupy a station for 5 or more minutes
Wind induced vibration	0.01 to 0.1 mGal	Use a wind break
Station elevation (the use of inappropriate technology to measure elevations can lead to errors that are significant for any given survey type)	Total effect about $0.2 \, \text{mGal} \, \text{m}^{-1}$. Errors proportional to errors in elevation	Use optical levelling for <1-cm accuracies, RTK-GPS for 10-cm accuracies and GPS/barometers for typical regional surveys
Instrument height (poor field technique can easily result in 2–3 cm errors)	Up to 10 µGal (implied by free-air effect of $0.3 \, \text{mGal} \, \text{m}^{-1}$)	Take great care to accurately measure to a standard point on the meter
Atmospheric pressure changes	$\sim -0.35 \, \mu\text{Gal} \, \text{mbar}^{-1}$ (hPa) of atmospheric loading	Monitor pressure changes. Increase frequency of repeats at base stations
Soft footing	Can cause 5 µGal or more reading error due to time-variant off-levelling	Use an automatic meter and stand well away when it is accumulating readings on soft ground
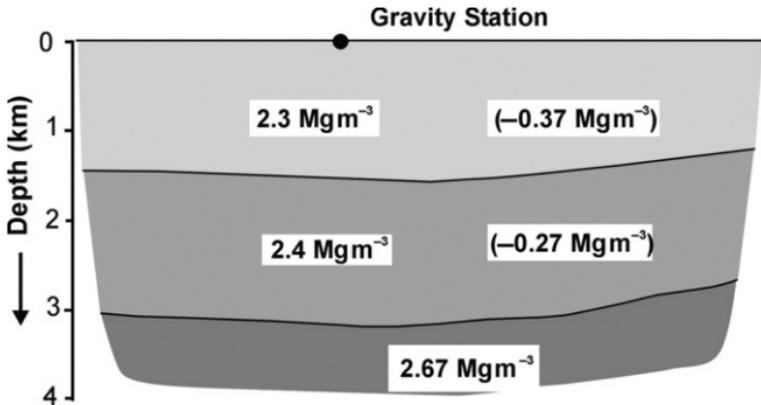
Figure 2.11 *Sedimentary basin model suitable for Bouguer plate methods of approximate interpretation. Modelling is usually done in terms of density contrasts with basement, here assigned the standard 2.67 Mg m⁻³ crustal density. (See Example 2.1).*

be needed for each station in rugged areas. Additional notebook columns may be reserved for tidal and drift corrections, since drift should be calculated each day, but these calculations are now usually made within the meters themselves, or on laptop PCs or programmable calculators, and not manually in field notebooks.

Each loop should be annotated with the observer's name or initials, the meter serial number and calibration factor, and the base station number and gravity value. It is also useful to record the difference between local and 'Universal' time (GMT) on each sheet, as a reminder for when tidal corrections are being calculated.

Gravity data are expensive to acquire and deserve to be treated with respect. The general rules of Section 1.5.2 should be scrupulously observed.

Example 2.1

In Figure 2.11, if the standard crustal density is taken to be $2.67\,\text{Mg m}^{-3}$, the effect of the upper sediment layer, 1.5 km thick, would be approximately $1.5 \times 0.37 \times 40 = 22$ mGal at the centre of the basin.

The effect of the deeper sediments, 1.6 km thick, would be approximately $1.6 \times 0.27 \times 40 = 17$ mGal.

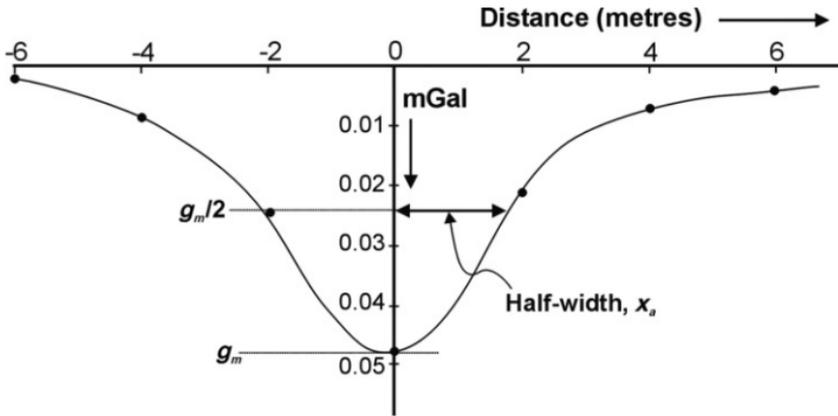The total (negative) anomaly would thus be about 39 mGal.

Figure 2.12 *Detailed Bouguer anomaly profile over a subsurface cavity, illustrating the definition of the 'half-width'. (See Examples 2.2 and 2.3).*

2.5 Field Interpretation

Gravity results are usually interpreted by calculating the fields produced by geological models and comparing these with the actual data. This requires a computer, and until recently was rarely done in the field. Even now, an appreciation of the effects associated with a few simple bodies can help an observer when actually on traverse or otherwise temporarily severed from his or her laptop to assess the validity and significance of the data being collected. This can sometimes lead to a vital decision to infill with additional stations being taken at a time when this can be done quickly and economically.

Example 2.2

Interpreting the anomaly of Figure 2.12 as due to a roughly *spherical* air-filled cavity in rock of density $2.5\,\mathrm{Mg\,m^{-3}}$ and working in metres:

$\Delta\rho = -2.5$ and the anomaly amplitude $= \Delta g = 0.05$ mGal.
Anomaly half-width $= 2$ m. Therefore, the depth to the sphere centre $=$
$\quad h = 2 \times 4/3 = 2.7$ m.

$$r^3 = (\Delta g \times h^2)/(0.028 \times \Delta\rho) = (0.05 \times 2.7^2)/(0.028 \times 2.5) = 5.2.$$

i.e. $r = 1.7$ m.

Example 2.3

Interpreting the anomaly in Figure 2.12 as due to a roughly *cylindrical* air-filled cavity in rock of density $2.5\,\text{Mg m}^{-3}$ and working in metres:

$\Delta\rho = -2.5$ and the anomaly amplitude $= \Delta g = 0.05$ mGal.
Anomaly half-width $= 2\,\text{m}$. Therefore, the depth to the cylinder centre $= h = 2\,\text{m}$.

$$r^2 = (\Delta g \times h)/(0.04 \times \Delta\rho) = (0.05 \times 2)/(0.04 \times 2.5) = 1.$$

i.e. $r = 1\,\text{m}$.

2.5.1 The Bouguer plate

The Bouguer plate provides the simplest possible interpretational model. An easily memorised rule-of-thumb is that the gravity effect of a slab of material 1 km thick and $1.0\,\text{Mg m}^{-3}$ denser than its surroundings is about 40 mGal. This is true even when the upper surface of the slab is some distance below the reading point (as in the case of the second layer in Figure 2.11), provided that the distance of the station from the nearest edge of the slab is large compared with the distance to its lower surface. The effect varies in direct proportion to both thickness and density contrast.

2.5.2 Spheres and cylinders

Less extensive bodies can be modelled by homogeneous spheres or by homogeneous cylinders with circular cross-sections and horizontal axes. If, in Figure 2.12, the anomaly is due to a sphere, radius r measured at a point immediately above its centre, the maximum anomalous field is:

$$\Delta g = 4\pi.\Delta\rho.Gr^3/3h^2$$

The factor $4\pi G/3$ is about 28, if lengths are measured in kilometres, or 0.028 if lengths are measured in metres. The depth, h, of the centre of the sphere is roughly equal to four-thirds of the half-width of the anomaly.

If, however, the source in Figure 2.12 can be modelled as an infinite horizontal cylinder of circular cross-section (an example of a *two-dimensional* source), the maximum field is:

$$\Delta g = 2\pi.\Delta\rho.Gr^2/h$$

The factor $2\pi G$ is about 40, if lengths are measured in kilometres, or 0.04 if lengths are measured in metres. The depth, h, of the axis of the cylinder is equal to the half-width of the anomaly.

2.5.3 Nettleton's method for direct determination of density

Density information is crucial to understanding gravity anomalies, but is not easily obtained. Samples collected in the field may be more weathered, and therefore less dense, than the bulk of the rock mass they are assumed to represent, and the reduction in density may be accentuated by the loss of some of the pore water. Estimates may also be obtained, rarely and expensively, from borehole gravity or from radiometric well logging, but such data will normally only be available where the work is being done in support of exploration for hydrocarbons.

In some cases the bulk density of the topography can be estimated directly from gravity data. The method assumes that the correct value is the one that removes the effect of topography from the gravity map when corrections are made. The method can work only if there is no real gravity anomaly associated with the topography; it will fail if, for example, a hill is the surface expression of a dense igneous plug or a dipping limestone bed (Figure 2.13).
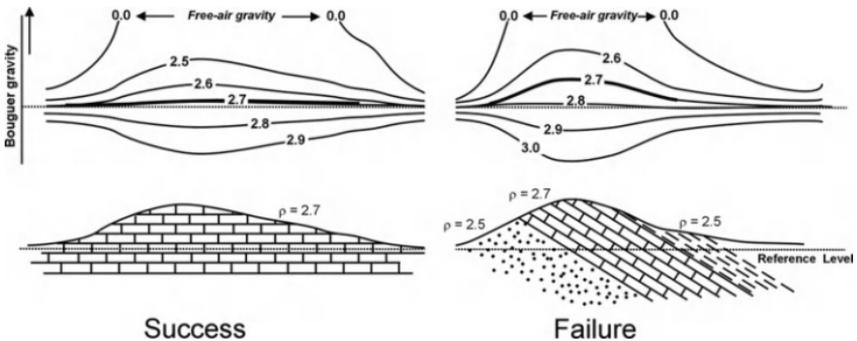

Figure 2.13 *Success and failure with Nettleton's method of determining density. The numbers show the densities (in $Mg\,m^{-3}$) used to calculate the Bouguer corrections for the respective Bouguer gravity curves. The thickened curves show the shapes that would have been recorded had it been possible to remove the topography and make the measurements at the level of the reference surface. In the 'failure' case there is a genuine anomaly due to the limestone layer sandwiched between the sandstones and shales, and using the 'level' profile would lead to an erroneously high estimate of the average topographic density. Note that the free-air gravity curve, equivalent to a Bouguer curve with zero topographic density, would show a strong positive correlation with topography for almost any underlying geology.*

Nettleton's method (described in Nettleton, L.L. 1976: *Gravity and Magnetics in Oil Prospecting*. McGraw Hill, New York) can be applied to a profile or to all the gravity stations in an area. In the latter case, a computer may be used to determine the density that produces the least correlation between topography and the corrected anomaly map. Even though these calculations are normally done by the interpreters, field observers should understand the technique, since they may have opportunities to take additional readings for density control.

3

MAGNETIC METHOD

Compasses and dip-needles were used in the Middle Ages to find magnetite ores in Sweden, making the magnetic method the oldest of all applied geophysical techniques. It is still one of the most widely used, even though significant magnetic effects are produced by only a very small number of minerals.

Magnetic field strengths are now usually measured in *nanotesla* (nT). The term '*gamma*', originally defining a unit equal to 10^{-5} gauss and numerically equal to the nT, is still occasionally used

3.1 Magnetic Properties
Although governed by the same fundamental equations, magnetic and gravity surveys are very different. The magnetic properties of adjacent rock masses may differ by several orders of magnitude rather than a small percentage (see Table 1.2).

3.1.1 Poles, dipoles and magnetisation
An isolated magnetic pole would, if it existed, produce a field obeying the inverse-square law. In reality, the fundamental magnetic source is the dipole (see Section 1.2.5) but, since a line of dipoles end-to-end produces the same effect as positive and negative poles isolated at opposite ends of the line (Figure 3.1), the pole concept is often useful.

A dipole placed in a magnetic field tends to rotate, and so is said to have a *magnetic moment.* The moment of the simple magnet of Figure 3.1, which is effectively a positive pole, *m*, at a distance 2*L* from a negative pole with strength −*m*, is equal to 2*Lm*. The magnetisation of a solid body is defined by its magnetic moment per unit volume and is a vector, having direction as well as magnitude.

3.1.2 Susceptibility
A body placed in a magnetic field acquires a magnetisation, *M*, that, if small, is proportional to the field:
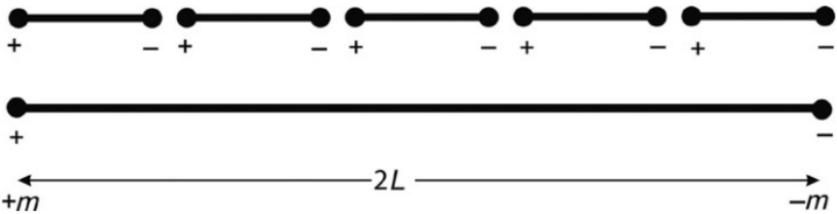
$$M = kH$$

Figure 3.1 *Combination of magnetic dipoles to form an extended magnet. The positive and negative poles at the ends of adjacent dipoles cancel each other out. The pole strength of the magnet is the same as that of the constituent dipoles, but its magnetic moment is equal to its length multiplied by that pole strength.*

The *susceptibility*, *k*, is small for most natural materials, and may be either negative (diamagnetism) or positive (paramagnetism). The fields produced by such materials are usually considered too small to affect survey magnetometers. Modern high-sensitivity magnetometers are creating exceptions to this rule, but most useful magnetic anomalies are still due to the small number of *ferro-* or *ferri-magnetic* substances in which the molecular magnets are held parallel by intermolecular *exchange forces.* Below the *Curie temperature*, these forces are strong enough to overcome the effects of thermal agitation. Magnetite, pyrrhotite and maghemite, all of which have Curie temperatures of about 600°C, are the only important naturally occurring strongly magnetic minerals and, of the three, magnetite is by far the most common. Hematite, the most abundant iron mineral, normally has a very small susceptibility, and many iron ore deposits do not produce significant magnetic anomalies.

3.1.3 Susceptibilities of rocks and minerals
The susceptibility of a rock usually depends on its magnetite content. Sediments and acid igneous rocks have small susceptibilities whereas basalts, dolerites, gabbros and serpentinites are usually strongly magnetic. Weathering generally reduces susceptibility because magnetite is oxidised to hematite, but some laterites are magnetic because of the presence of maghemite and remanently magnetised hematite. The susceptibilities, in rationalised SI units, of some common rocks and minerals are given in Table 1.2. Negative values, indicating diamagnetism, are observed only with very pure materials, since diamagnetic moments are nearly always swamped by any paramagnetism present.

The magnetic properties of highly magnetic rocks tend to be extremely variable and their magnetisation is not strictly proportional to the applied field. Quoted susceptibilities are for Earth-average field strengths.

3.1.4 Remanence

Ferro- and ferri-magnetic materials may have permanent as well as induced magnetic moments, so that their magnetisations are not necessarily in the direction of the Earth's field. The *Konigsberger ratio*, of the permanent magnetisation to the magnetisation that would be induced in an Earth-standard field of 50 000 nT, is generally large in highly magnetic rocks and small in weakly magnetic ones, but is occasionally extraordinarily high ($>10\,000$) in hematite, despite the low susceptibility. Magnetic anomalies due entirely to remanent magnetisation are therefore very occasionally produced by hematite ores.

3.2 The Magnetic Field of the Earth

The magnetic fields of geological bodies are superimposed on a background of the Earth's main field. Variations in magnitude and direction of this field influence both the magnitudes and the shapes of local anomalies.

The terms *North* and *South* that are commonly used to describe magnetic polarity are replaced in geophysics by *positive* and *negative*. The direction of a magnetic field is conventionally defined as the direction in which a unit positive pole would move, and geophysicists give little thought to whether it is the North or South magnetic pole that is positive.

3.2.1 The main field of the Earth

The Earth's main magnetic field originates in electric currents circulating in the liquid outer core, but can be largely modelled by a dipole source at the Earth's centre. Distortions in the dipole field extending over regions thousands of kilometres across can be thought of as caused by a relatively small number of subsidiary dipoles at the core-mantle boundary.

The variations with latitude of the magnitude and direction of an ideal dipole field aligned along the Earth's spin axis are shown in Figure 3.2. Near the Equator the dip angles change almost twice as fast as the latitude angles. Neither the magnetic equator, which links points of zero magnetic dip on the Earth's surface, nor the magnetic poles coincide exactly with their geographic equivalents (Figure 3.3), and the main dipole would have to be inclined at about 11° to the spin axis to explain the Earth's actual field. The North Magnetic Pole has, in the last ten years, moved from northern Canada into the Arctic Ocean, and the South Magnetic Pole is currently in the Southern Ocean at about 65°S 138°E. Differences between the directions of true and magnetic north are known as declinations, presumably because a compass needle *ought* to point north but *declines* to do so.

Dip angles estimated from the global map (Figure 3.3) can be used to obtain rough estimates of magnetic latitudes and hence (using Figure 3.2) of regional gradients. This approach is useful in determining whether a regional

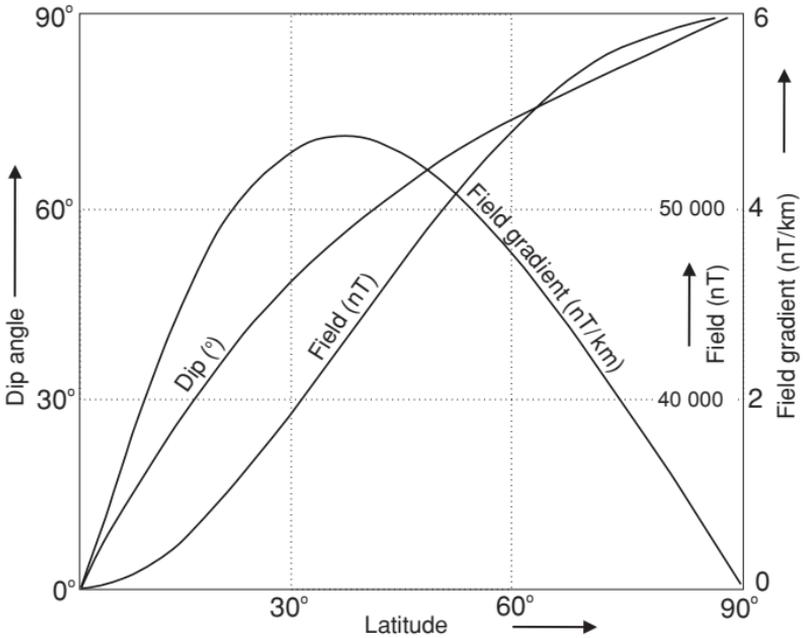*Figure 3.2* *Variation in intensity, dip and regional gradient for an ideal dipole aligned along the Earth's spin axis and producing a polar field of 60 000 nT (and an equatorial field of 30 000 nT).*
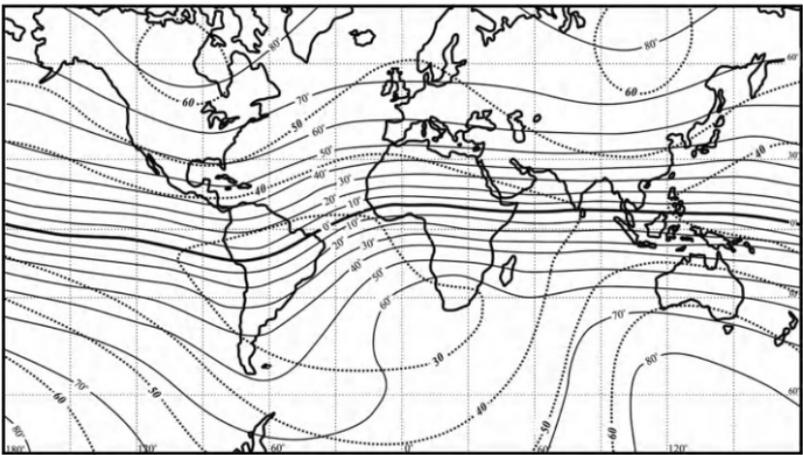

Figure 3.3 *Dip (continuous lines, values in degrees) and intensity (dotted lines, values in thousands of nT) of the Earth's magnetic field. The thickened continuous line is the magnetic equator.*

68

gradient is likely to be significant but gives only approximate correction factors, because of the existence of very considerable local variations. Gradients are roughly parallel to the local *magnetic* north arrow, so that corrections have E–W as well as N–S components. In ground surveys, where anomalies of many tens of nanotesla are commonplace, regional corrections, which amount to only a few nanotesla per kilometre, are often neglected.

3.2.2   The International Geomagnetic Reference Field (IGRF)

The Earth's main field varies not only with latitude and longitude but also with time. Between 1980 and 2002 the field strength in the central Atlantic decreased by about 6%. These *secular* changes are described by the empirical International Geomagnetic Reference Field (IGRF) equations, which are defined to order $N = 10$ by 120 spherical harmonic coefficients, supplemented to order $N = 8$ by a predictive secular variation model. The shortest wavelength present is about 3000 km. IGRFs provide reasonable representations of the actual fields in well-surveyed areas, where they can be used to calculate regional corrections, but discrepancies of as much as 250 nT can occur in areas from which little information was available at the time of formulation. Since 2000, the accuracies of the IGRFs have been greatly improved by incorporating data from the Danish Oersted and German CHAMP satellites.

Secular changes are adequately predicted by extrapolation from past observations for only a few years into the future, and IGRFs are updated every five years and are also revised retrospectively to give definitive models (DGRFs). These time-dependent corrections are vital for comparing airborne or marine surveys carried out years or even months apart but are less important in ground surveys, where base stations can be reoccupied.

3.2.3   Diurnal variations

The Earth's field also varies because of changes in the strengths and directions of currents circulating in the ionosphere. The variations, with peak-to-trough amplitudes of the order of a few tens of nanotesla in mid-latitudes, tend to be directly related to local solar time, because the upper atmosphere is ionised by solar radiation. In the normal *solar-quiet* (Sq) pattern, the background field is almost constant during the night but decreases between dawn and about 11 a.m., increases again until about 4 p.m. and then slowly falls back to the overnight value (Figure 3.4). However, at points up to a few hundred km apart, amplitude differences of more than 20% due to differences in crustal conductivity may be more important than time dependency. Shorter period, horizontally polarised and roughly sinusoidal *micropulsations* are significant only for surveys that are to be contoured at less than 5 nT.
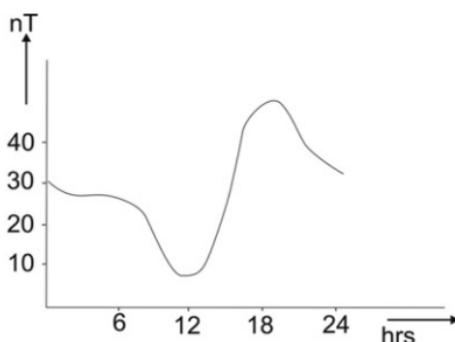
Figure 3.4 *Typical 'quiet day' magnetic field variation at mid-latitudes.*

The magnetic field within about $5°$ of the magnetic equator is strongly influenced by the *equatorial electrojet*, a band more than 1000 km wide of high conductivity in the ionosphere. Diurnal variations in the affected regions may be well in excess of 100 nT and may differ by 10 to 20 nT at points only a few tens of kilometres apart. In polar regions, similarly severe short-period fluctuations can be explained by the existence of *auroral electrojets*. It is therefore particularly important in these regions that background variations are monitored continuously. Returning to a base station at intervals of one or two hours may be quite insufficient.

3.2.4 Magnetic storms

Short-term auroral effects are special cases of the irregular disturbances (Ds and Dst) known as *magnetic storms*. These are produced by sunspot and solar flare activity and, despite the name, are not meteorological, often occurring on clear, cloudless days. There is usually a sudden onset, during which the field may change by hundreds of nanotesla in a few minutes, followed by a slower, erratic return to normality. Time scales vary widely but the effects can persist for hours and sometimes days. Micropulsations are generally at their strongest in the days immediately following a storm, when components with periods of a few tens of seconds can have amplitudes of as much as 5 nT.

Ionospheric prediction services in many countries give advance warning of the general probability of storms but not of their detailed patterns, and the field changes in both time and space are too rapid for corrections to be applied, unless high-frequency (10 or 20 Hz) base stations are used. Otherwise, the only solution is to wait until the storm is over. Aeromagnetic data are severely affected by quite small irregularities, and for contract purposes *technical magnetic storms* may be defined, sometimes as departures from linearity in the diurnal curve of as little as 5 nT in an hour. Similar criteria

may have to be applied in those archaeological surveys in which only a single sensor is used.

3.2.5 Geological effects

The Curie temperatures (in the range 500–600°C) of the geologically important magnetic materials are reached in the lower parts of normal continental crust but below the Moho beneath the oceans. The upper mantle is only weakly magnetic, so that the effective lower limit of rock magnetic sources is the Curie isotherm beneath the continents and the Moho beneath the oceans.

Massive magnetite deposits can produce magnetic fields of as much as 200 000 nT, which is several times the magnitude of the Earth's normal field. Because of the dipolar nature of magnetic sources, these and all other magnetic anomalies have positive and negative parts, and directional magnetometers may actually record negative fields in extreme cases. Anomalies of this size are unusual, but basalt dykes and flows and some larger basic intrusions can produce fields of thousands and occasionally tens of thousands of nT. Anomalous fields of more than 1000 nT are otherwise rare, even in areas of outcropping crystalline basement. Sedimentary rocks generally produce changes of less than 10 nT.

In some tropical areas, occasionally large magnetic fields are produced by maghemite formed as nodular growths in laterites. The nodules may later weather out to form ironstone gravels, which give rise to high noise levels in ground surveys. The factors that control the formation of maghemite rather than the commoner, non-magnetic form of hematite are not yet fully understood. Human activity can also change iron oxidation states and produce small (<10 nT) anomalies in natural materials, for example by deliberate heating of clays in brick and pottery manufacture or simply by disturbing soil horizons.

3.2.6 Man-made sources

Iron and steel are ferromagnetic and are usually strongly magnetised. Even quite small steel objects can produce fields of hundreds of nT, although for small objects these fields decrease rapidly with distance. In investigations of 'brownfield' sites and in the search for unexploded ordnance (UXO), man-made objects may be the targets and their associated magnetic fields may be very useful, but in geological work they are not, and high-cut filters may have to be applied to survey results. In extreme cases, meaningful results may be impossible to obtain.

Most power lines transmit electricity as alternating current, producing alternating magnetic fields that may interfere with instrument operation but are unlikely to give rise to readings that are both repeatable and spurious. However, where the distances involved are very large, electric power is often

transmitted using direct current, producing fields of several tens of thousands of nT that may extend for more than a kilometre on either side of the power lines. Useful magnetic surveys are simply impossible in such areas.

3.3 Magnetic Instruments

Early magnetometers were compass needles mounted on horizontal axes, and measured vertical fields. These *torsion magnetometers* were in common use until about 1960, when they began to be replaced by fluxgate, proton precession and alkali vapour magnetometers. Instruments of all three types are still used, are available with built-in data loggers and can be set to record automatically at fixed time intervals. All three can be used singly or, in tandem, as *gradiometers*, although care must then be taken with precession instruments to ensure that the polarising field from one sensor does not affect the other.

3.3.1 Proton precession magnetometer

Proton precession magnetometers make use of the small magnetic moment of the hydrogen nucleus (proton). The sensing element consists of a 'bottle' containing a low-freezing-point hydrocarbon fluid about which is wound a coil of copper wire. Although many fluids *can* be used, the manufacturer's recommendation, usually for high-purity decane, should always be followed if the bottle has to be topped-up. A *polarising* current of the order of 1 amp or more is passed through the coil, creating a strong magnetic field, along which the moments of the protons in the hydrogen atoms tend to align.

When the current is switched off, the protons realign to the direction of the Earth's field. Quantum theory describes this reorientation as occurring as an abrupt 'flip', with the emission of a quantum of electromagnetic energy. In classical mechanics, the protons *precess* about the field direction, as a gyroscope precesses about the Earth's gravity field, at a frequency proportional to the field strength, emitting electromagnetic waves as they do so. Both theories link the wave frequency to the external field via two of the most accurately known of all physical quantities, Planck's constant and the proton magnetic moment. In principle the proton magnetometer is capable of almost any desired accuracy, but in practice the need for short reading times and reasonable polarisation currents sets the limit at about 0.1 nT. In the Earth's field of about 50 000 nT, the precession frequency is about 2000 Hz, and sophisticated phase-sensitive circuitry is needed to measure to this accuracy in the half or one second that is all that modern geophysicists will tolerate.

Proton magnetometers may give erratic readings in strong field gradients or because of interference from power lines and radio transmitters, or even from eddy currents induced in nearby conductors by the termination of the polarising current. Also, they can only measure total fields, which may

cause problems in interpreting large anomalies where the direction as well as magnitude of the field changes rapidly from place to place. However, these are only minor drawbacks, and 1 and 0.1 nT proton magnetometers are probably still the commonest instruments in ground surveys. They have the advantage of producing drift-free absolute measurements, but corrections must still be made for diurnal variations. Since the protons align themselves, the sensor need not be precisely orientated and can be mounted on a pole and carried well away from both the observer and from small magnetic sources at ground level (see Figure 1.8). To obtain adequate signals it is necessary only to ensure that the polarising field is at a high angle to the measured field; i.e. it should be horizontal in high latitudes and vertical in low latitudes. The manufacturer's instructions need to be read carefully before mounting the sensor, since the polarising field is not necessarily directed along the axis of the sensor housing.

3.3.2   High-sensitivity (alkali vapour) magnetometers

Proton magnetometers can be made more sensitive using the Overhauser effect, in which a VHF radio wave acts on paramagnetic material added to the bottle fluid. This increases the precession signal by several orders of magnitude, considerably improving the signal/noise ratio. However, high sensitivity is now more commonly achieved using electron magnetic moments, which are about 2000 times larger than proton moments. Effectively 'free' electrons are required, and these occur in the outer electron 'shells' of alkali metal (usually caesium) vapours. The principle is similar to that of the proton magnetometer, in that transitions between energy states are observed, but the much higher energy differences imply much higher frequencies, which can be measured with much smaller percentage errors. The actual measurement processes are quite complicated, involving the raising of electrons to high energy states by laser beams ('optical pumping') and then determining the frequency of the HF radio signal that triggers a transition to a lower state. This is all, however, invisible to the user. The effects of electrical noise and high field gradients are less serious than with proton precession instruments, and measuring times are very short. Readings can be routinely obtained every tenth of a second (approximately every 5–10 cm at walking speeds), which can be important in archaeology, where very high rates of coverage are required, and can be achieved using a non-magnetic trolley with a trigger actuated by the rotations of the wheels.

Alkali vapour magnetometers are slightly direction-sensitive. Readings cannot be obtained if the sensor is orientated within a few degrees of either the direction of the magnetic field or at right angles to it. This is not a significant limitation in most circumstances, and the rather slow acceptance of these instruments for ground surveys has had more to do with their

relatively high cost and the fact that in most geological applications the high sensitivity is of little use. However, absolute field changes of less than 1 nT may be significant in archaeology, and sensitivity is always important in gradiometry. The greater tolerance of high field-gradients can also be an advantage in engineering/environmental surveys.

3.3.3 Fluxgate magnetometers

The sensing elements of fluxgate magnetometers consist of one or more cores of magnetic alloy, which are magnetised to saturation by alternating current passed through coils wound around them. The changes in the electrical properties of the circuits as the core magnetisations pass from the unsaturated to the saturated state can be converted into voltages proportional to the external magnetic field along the core axes. Measurements are thus of the magnetic field component in whichever direction the sensor is pointed. This is the vertical in most ground surveys, since this is the most easily determined direction.

Fluxgates do not measure absolute fields and therefore require calibration. They are also subject to thermal drift, because the magnetic properties of the cores and, to a lesser extent, the electrical properties of the circuits, vary with temperature. Early ground instruments sacrificed thermal insulation for portability and were often accurate to only 10 or 20 nT. Readings were displayed, rather crudely, by the position of a needle on a graduated dial. Despite some claims to the contrary, such sensitivity is inadequate for almost all ground survey work.

One problem with portable fluxgates is that, because they require orientation at each station, the observer must be holding the sensor when a reading is taken. Ensuring that people are completely non-magnetic is never easy, and battery packs, in particular, can be sources of noise. Any batteries supplied by the manufacturer should be completely non-magnetic but they, and any subsequent replacements, should be carefully checked.

Fluxgates are now used mainly in archaeological surveys, where price is important and the necessary large numbers of readings must be obtained more quickly than is possible with proton instruments. Also, measurements often have to be made close to the ground, which may be difficult with a proton magnetometer because of its sensitivity to field gradients and electrical interference. Fluxgate sensors are usually paired, with fixed vertical separations of between 50 and 100 cm, and often only the differences in the readings are recorded. Such instruments are usually known as *gradiometers*, but the operation of the inverse-square or inverse-cube laws ensures that sources that produce measurable anomalies are at depths comparable with sensor separations. It is easier to understand the resulting maps if the instruments are thought of as *differential magnetometers* (Figure 3.5).
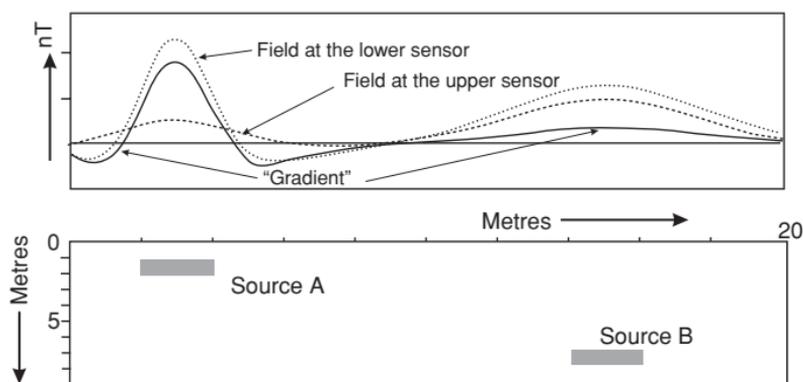
Figure 3.5 *Inverse-cube law effects in magnetic gradiometry. The dotted curve shows the magnetic effects of the two bodies as measured at the ground surface, and the dashed curve shows the effects 1 metre above the surface. The solid curve shows the differential effect. In the case of Source A, the difference ('gradient') anomaly has 80% of the amplitude of the anomaly measured at ground level. In the case of the deeper (but also stronger) Source B, the total field anomaly amplitudes at the two sensors are much more similar and the difference anomaly is therefore small.*

Use of two sensors minimises thermal drift effects, reduces the effect of errors in orientation, emphasises local sources and virtually eliminates the effects of diurnal variations, including micropulsations. It is, however, necessary to ensure that the two are very precisely aligned and are in thermal equilibrium with each other and the environment. Three-component fluxgates can eliminate the need for precise orientation or, alternatively, can provide information on field direction as well as field strength.

3.4 Magnetic Surveys
Although absolute numerical readings are obtained (and can be repeated) at a keystroke with proton and caesium magnetometers, faulty magnetic maps can still be produced if simple precautions are ignored. For example, all base locations, whether used for repeat readings or for continuous diurnal monitoring, should be checked for field gradients. A point should not be used as a base if moving the sensor by a metre produces a significant change.

3.4.1 Starting a survey
The first stage in any magnetic survey is to check the magnetometers (and the operators). Operators can be potent sources of magnetic noise, although the problems are much less acute when sensors are on long poles than when they are carried in backpacks or when, as with fluxgates, they must be

held close to the body. Compasses, pocket knives and geological hammers are all detectable at distances of less than about a metre, and users of high-sensitivity magnetometers may need to visit tailors (and cobblers) for non-magnetic clothing. Survey vehicles can be detectable at distances of up to 20 m. The safe distance should be determined before starting survey work.

Absolute magnetometers should all give the same reading when read at the same time and in the same place. Differences between instruments manufactured prior to 1980 were often greater than 10 nT, but now are seldom more than 1 or 2 nT. Sensors can be placed very close together and may even touch when checks are being made, but proton magnetometers cannot be read exactly simultaneously in this way because the two polarising fields would interfere.

Large discrepancies and very variable proton magnetometer readings usually indicate poor tuning. The correct tuning range can be roughly identified using global maps (see Figure 3.3) but final checks should be made in the field. Near-identical readings should be obtained if the setting is varied over a range of about 10 000 nT about its optimum position (e.g. 47 000 in Example 3.1). Manual versions are generally rather coarsely tuneable in steps of a few thousand nT, but greater accuracy is possible with microprocessor control. It is partly this finer tuning that is now allowing proton magnetometers to be routinely read to 0.1 nT. They may be programmed to respond to faulty tuning or high gradients by refusing to display the digit beyond the decimal point.

Example 3.1: Proton magnetometer tuning (manual model)

Tuning setting		Readings	
30 000	31 077	31 013	31 118
32 000	32 770	32 788	32 775
34 000	35 055	34 762	34 844
36 000	37 481	37 786	37 305
38 000	42 952	40 973	41 810
41 000	47 151	47 158	47 159
44 000	47 160	47 158	47 159
47 000	47 171	47 169	47 169
50 000	47 168	47 175	47 173
53 000	47 169	47 169	47 169
56 000	53 552	54 602	54 432
60 000	59 036	59 292	58 886
64 000	65 517	65 517	65 517

The readings at the 64,000 nT setting show that repeatability alone is no guarantee of correct tuning. It is the range of settings over which the circuits can lock to the precession signal that provides the crucial evidence.

3.4.2   Monitoring diurnal variation

Diurnal corrections are essential except in differential or gradient surveys. If only a single instrument is available, corrections have to rely on repeated visits to a base or sub-base, ideally at intervals of less than an hour. A more complete diurnal curve can be constructed if a second, fixed, magnetometer is used to obtain readings at 3 to 5 minute intervals. This need not be of the same type as the field instrument. A cheaper proton magnetometer can provide adequate diurnal control for surveys carried out using a (more expensive) caesium vapour instrument.

In principle, frequent base reoccupations are unnecessary when an automatic base station is operating. It is, however, poor practice to rely entirely on the automatic record, since the field data will then be difficult, if not impossible, to correct if the base instrument fails. Problems can occur even when a base instrument is manually operated, but they are especially likely if it is unattended. The battery drain is rather high in automatic instruments and the transition from operational to unworkable can occur suddenly and without warning. Readings already stored are preserved, but diurnal control is lost for the rest of the day.

Obviously, bases should be remote from possible sources of magnetic interference (especially temporary sources such as traffic), and should be describable for future use. Special vigilance is needed if field and diurnal instruments are later linked by a data-exchange line and corrections are made automatically. Unless the diurnal curve is actually plotted and examined, absurdities in the diurnal data (as might be caused by an inquisitive passer-by driving up to the base) may appear, in reverse, as anomalies in the field data.

3.4.3   Field procedures – total field 'point' surveys

Except in gradiometer surveys, a diurnal magnetometer should be set up at the start of each day. The first reading of the field magnetometer should be at a base or sub-base, and should be made at the same time as a reading, either automatic or manual, of the diurnal magnetometer. This does not necessarily require the two instruments to be adjacent.

High-sensitivity magnetometers, with their 0.1 second reading times, can be used to obtain essentially continuous profiles, but proton magnetometers require the operator to stop for each reading. All such readings should be repeated, and the two should differ by not more than 1 nT. Greater differences may indicate high field gradients, which may need further investigation. Large differences between readings at adjacent stations call for *infill* at intermediate points. It is obviously desirable that the operator notices this, and fills in immediately.

At each station the location, time and reading must be recorded, and also any relevant topographic or geological information and details of any visible

or suspected magnetic sources. Unless the grid is already well mapped, the notebook should also contain enough information for the line locations to be verified using maps or air-photos.

At the end of the day, a final reading should be made at the base first occupied. This should again be timed to coincide with a reading of the diurnal magnetometer. If the field readings are being recorded manually, it is good practice to transcribe the diurnal values for the times of the field readings into the field notebook, which then contains a complete record of the day's work.

3.4.4 Standard values

A diurnal curve records the way in which field strength has varied at the fixed base, and data processing is simplified if this base is at the same point throughout a survey. A *standard value* (SV) must be allocated to this point, preferably by the end of the first day of survey work. The choice is to some extent arbitrary. If the measured value varies between 32 380 and 32 410 nT, it could be convenient to adopt 32 400 nT as the SV, even though this was neither the mean nor the most common reading.

In large survey areas it may be necessary to establish sub-bases (see Section 1.6) and determine their SVs. The underlying principle is that if, at some given time, the base magnetometer reading is actually equal to the base SV, then identical instruments at all other bases and sub-bases would record the SVs at those points. The field readings are then processed so that this is also true of the values assigned to all survey points.

3.4.5 Processing magnetic data

During a survey, bases or sub-bases should be occupied at intervals of not more than 2 hours, so that data can be processed even if the diurnal record is lost or proves faulty. The ways in which such readings might be used to provide diurnal control, with or without an automatically recorded diurnal curve, are shown in Figure 3.6. The diurnal correction at any time is simply the difference between the SV at the diurnal station and the actual diurnal reading, but magnetic data can be corrected in two different ways using this fact. The more straightforward is to determine, by interpolation when necessary, the diurnal value at the time of a given field reading and to subtract this from the reading. The diurnal station SV can then be added to give the SV at the field station. If a computer is available, the whole operation can be automated.

This method is simple in principle and provides individual values at all field points but is tedious and error-prone if hundreds of stations have to
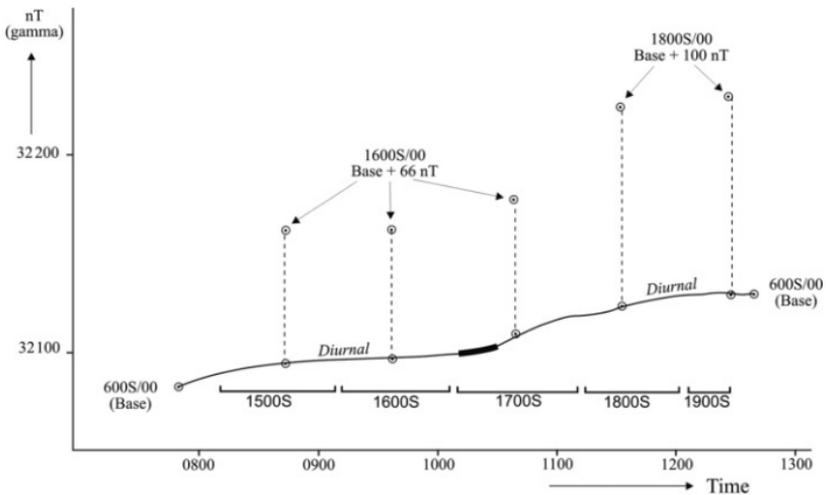
Figure 3.6 *Diurnal control, with variations monitored both by a memory instrument at the diurnal base and by repeat readings at various sub-bases. The time periods during which individual survey lines were being read are also shown. The shifts needed to make the sets of sub-base values fall on the diurnal curve provide estimates of the differences between the standard values (SVs) at the diurnal base and at the sub-bases. The greatest error introduced by using straight-line interpolation between the diurnal values derived from measurements at the sub-bases would have been about 5 nT and would have affected Line 1700S over the thickened part of the curve. Interpolation using a smooth curve instead of straight lines would have significantly reduced this error.*

be processed by hand each evening. If only a contour map is required, this can be based on profiles of uncorrected readings, as shown in Figure 3.7. Fewer calculations are needed, and errors and peculiarities in the data are immediately obvious.

Even if no computer is available to do the hard work, plotting magnetic profiles should be a field priority since these provide the best way of assessing the significance, or otherwise, of diurnal effects and noise. For example, the profile in Figure 3.7 shows very clearly that, with 100 nT contours, the 5 nT discrepancy shown in Figure 3.6 between the diurnal curves based on direct observation and base reoccupations is unimportant. It also shows the extent to which the contours leave significant magnetic features undefined.

If a computer is used to calculate corrected values at each field point, profiles should still be produced, for quality control, but using the corrected
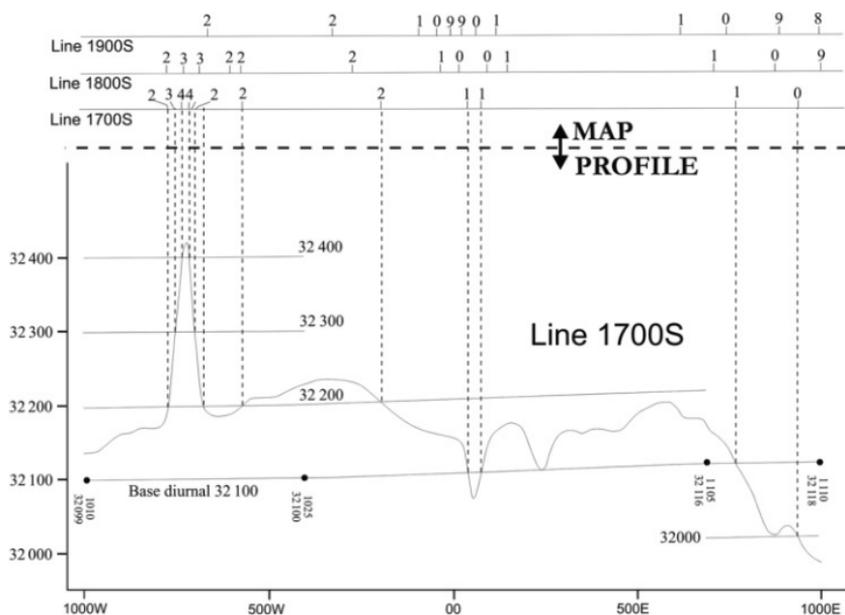
Figure 3.7 *PROFILE: Contour cuts at 100-nT intervals on uncorrected profile 1700S, by diurnal curve and parallel curves. The reference base standard value (SV) is 32 100 nT, and the points at which the diurnal curve intersects the profile therefore correspond to points on the ground where the corrected value of the magnetic field is also 32 100 nT. Curves parallel to the diurnal identify points at which the SVs differ from those at the diurnal base by integer multiples of 100 nT, which is an appropriate contour interval for mapping the data shown. MAP: Contour cut plot of Line 1700S and two adjacent lines. Only the 'cuts' need be plotted on the map and some may be omitted if they are very close together.*

data. The profiles that are so easily produced from the semi-continuous data provided by alkali-vapour instruments should be carefully examined for artefacts of the sorts described in Section 1.7.4.

3.4.6 Noise in ground magnetic surveys

Magnetic readings in populated areas are often affected by *cultural noise*, i.e. stray fields from pieces of iron and steel that are irrelevant to the survey objectives. Since these are often quite small and likely to be buried within a metre of the ground surface, the effects are very variable. Even if no obvious sources are visible, profiles obtained along roads are usually very distorted

compared to those along parallel traverses through open fields only 10 or 20 m away.

One approach to the noise problem is to try to take all readings well away from likely sources, noting in the field books where this has not been possible. Alternatively, the almost universal presence of ferrous noise can be accepted and the data can be filtered. For this method to be successful, many more readings must be taken than would be needed to define purely geological anomalies. The technique is becoming more popular with increasing use of data loggers, which discourage note-taking but allow vast numbers of readings to be taken and processed with little extra effort, and is most easily used with alkali vapour and fluxgate instruments, which read virtually continuously. It is, however, only safe to dispense with notebooks for comments on individual stations if the survey grid as a whole is well surveyed, well described and accurately located.

High-cut filters cannot be used in archaeological surveys, since they are likely to remove the signals along with the noise. Data in such surveys are usually displayed as images in which each pixel corresponds to an individual reading point and is coloured or grey-shaded according to the value (see Figure 1.12). Low-cut filters may be used prior to display to emphasise the short-wavelength features.

3.5   Simple Magnetic Interpretation

Field interpretation of magnetic data allows areas needing infill or checking to be identified and then revisited immediately and at little cost. Good interpretation requires profiles, which preserve all the detail of the original readings, and contour maps, which allow trends and patterns to be identified. Fortunately, the now almost ubiquitous laptop computer has reduced the work involved in contouring (providing the necessary programs have been loaded).

3.5.1   Forms of magnetic anomaly

The shapes of magnetic anomalies vary dramatically with the dip of the Earth's field, as well as with variations in the shapes of the source bodies and their directions of magnetisation. Simple sketches can be used to obtain rough visual estimates of the anomaly produced by any magnetised body.

Figure 3.8a shows an irregular mass magnetised by induction in a field dipping at about $45°$. Since the field direction defines the direction in which a positive pole would move, the effect of the external field is to produce the distribution of poles shown. The secondary field due to these poles is indicated by the dashed lines of force. Field direction is determined by the simple rule that like poles repel.

If the secondary field is small, the directions of the total and background fields will be similar and the anomalous field will not be detected near C
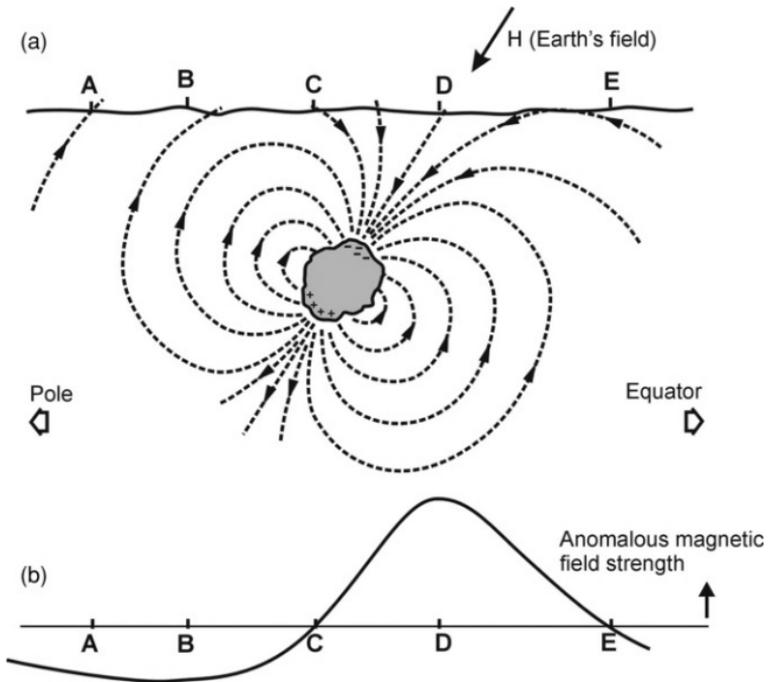
Figure 3.8 *Mid-latitude total field anomaly due to induced magnetisation. (a) The induced field. (b) The anomaly profile, derived as described in the text.*

and E, where it is at right angles to the Earth's field. The anomaly will be positive between these points and negative for considerable distances beyond them. The anomaly maximum will be near D and there will be a quite strong, but broader and negative beyond C, with a minimum near B but detectable well beyond A. The peak on the profile is thus offset towards the magnetic equator (Figure 3.8b). Applying this sketching technique at the magnetic equator, where the inducing field is horizontal, shows that the total-field anomaly there would be negative and centred over the body and would have positive side-lobes to north and south.

Because each positive magnetic pole is somewhere balanced by a negative pole, the net flux involved in any anomaly is zero. The fields from the positive and negative poles cancel out above the central parts of a uniform and horizontal magnetised sheet, and only the edges are magnetically detectable. Strongly magnetised but flat-lying bodies may consequently produce little or no anomaly.
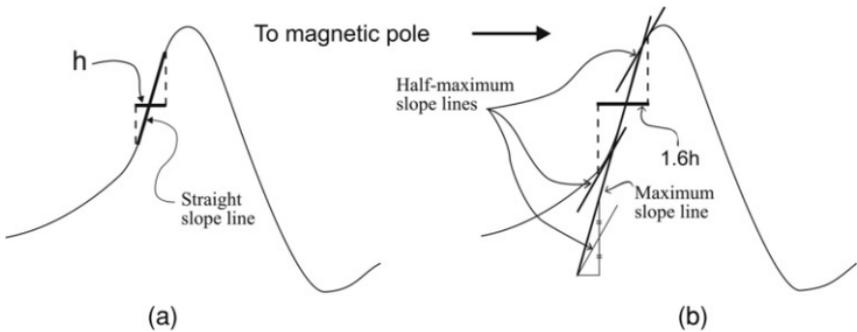
Figure 3.9 *Simple depth estimation. (a) Straight slope method. The distance over which the variation appears linear is (very) roughly equal to the depth to the top of the magnetised body. (b) Peters' method. The distance between the contact points of the half-slope tangents is (very) roughly equal to 1.6 times the depth to the top of the magnetised body. Note the construction triangle used to graphically evaluate the half-slope gradient.*

3.5.2 'Rule of thumb' depth estimation

Depth estimation is one of the main objectives of magnetic interpretation. Simple rules give depths to the tops of source bodies that are usually correct to within about 30%, which is adequate for preliminary assessment of field results.

In Figure 3.9a the part of the anomaly profile over which the variation is almost linear on the side nearest the magnetic equator is emphasised by a thickened line. The depths to the top surfaces of bodies of many different shapes would be approximately equal to the horizontal extent of this straight-line section. This method is quick and effective but is open to the objection that it relies on an optical illusion, since there is actually no straight-line segment of the curve.

In the slightly more complicated *Peters' method,* a tangent is drawn to the profile at the point of steepest slope, again on the side nearest the equator, and lines with half this slope are drawn using the geometrical construction of Figure 3.9b. The two points at which the half-slope lines are tangents to the anomaly curve are found by eye or with a parallel ruler, and the horizontal distance between them is measured. This distance is divided by 1.6 to give a rough depth to the top of the source body.

Peters' method relies on model studies that show that the true factor generally lies between about 1.2 and 2.0, with values close to 1.6 being common for thin, steeply dipping bodies of considerable strike extent. Results are usually very similar to those obtained using the straight slope. In all cases
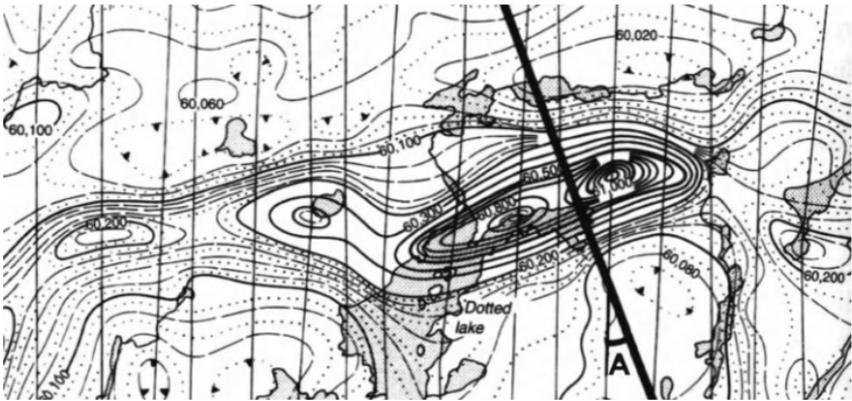
Figure 3.10 Effect of strike. A depth estimate on a profile recorded along a traverse line (i.e. one of the set of continuous, approximately straight lines) must be multiplied by the cosine of the angle A between the traverse lines and a line drawn at right angles to the magnetic contours. The example is from an aeromagnetic map (from northern Canada) but the same principle applies in ground surveys.

the profile must either be measured along a line at right angles to the strike of the anomaly or else the depth estimate must be multiplied by the cosine of the intersection angle (A in Figure 3.10).

4

RADIOMETRIC SURVEYS

The radioactivity of rocks can be measured using gamma-ray scintillation counters (scintillometers) and spectrometers. Although most radiometric instruments were developed with uranium search in mind, other uses were soon found. Amongst these were regional geological mapping and correlation, exploration for some industrial minerals and *in situ* determinations of phosphates. The same instruments can also be used to track the movement of artificial radioactive *tracers* deliberately introduced into ground water, and to assess health risks from natural and artificial radiation sources. Radon gas 'alpha' detectors are important in public health applications and have some exploration uses.

4.1 Natural Radiation

Spontaneous radioactive decay occurs when an unstable atomic nucleus loses energy by emitting alpha, beta and gamma radiation. Alpha and beta 'rays' are actually particles. Quantum theory tells us that gamma rays, which are high-energy electromagnetic waves, can also be treated for many purposes as if composed of particles.

4.1.1 Alpha particles

An alpha particle consists of two protons held together by two neutrons to form a stable helium nucleus. Emission of alpha particles is the main process in radioactive decay, resulting in decreases of four in atomic mass and two in atomic number. The particles have large kinetic energies but are rapidly slowed by collisions with other atomic nuclei. At *thermal* energies they soon gain two orbital electrons and become indistinguishable from other helium atoms. The average distance travelled in solid rock before this occurs is measured in fractions of a millimetre, and even in air is only a few centimetres.

4.1.2 Beta particles

Beta particles are electrons ejected from atomic nuclei. They differ from other electrons only in having higher kinetic energies, and so cease to be identifiable after being slowed by multiple collisions. Energy is lost most

rapidly in collisions with other electrons. In solids or liquids the average range of a beta particle is measured in centimetres.

4.1.3 Gamma radiation

At the high frequencies involved, the electromagnetic 'gamma rays' are best treated as consisting of particles, known as *photons*, with energies, measured in electron-volts (eV), proportional to frequencies. The boundary between gamma photons and the less energetic X-rays, which are emitted by electrons orbiting the atomic nucleus, occurs at about 0.1 MeV (a frequency of about 0.25×10^{20} Hz).

Because they are electrically neutral, photons penetrate much greater thicknesses of rock than do either alpha or beta particles and are consequently the most geophysically useful form of radiation. Even so, approximately 90% of the gamma radiation detected above bare rock will come from within 20 to 30 cm of the surface, and even above soil only 10% will come from below about 50 cm. A metre of water will absorb about 97% of the radiation travelling through it. Attenuation in air is frequency dependent and (for once) it is the higher-frequency, higher-energy radiation that has the greater penetrating power. Half of a 1-MeV flux is absorbed by about 90 m of air but it takes about 160 m of air to absorb half of a 3-MeV flux. Either figure implies that atmospheric absorption can be generally ignored in ground surveys.

4.1.4 Radioactivity of rocks

Gamma rays provide information on the presence of unstable atomic nuclei. The average number of decays in a given time will be directly proportional to the number of atoms of the unstable element present. The rate of decrease in mass of a radioactive material therefore obeys an exponential law governed by a *half-life* (see Section 1.2.6).

Elements with short half-lives occur in nature because they are formed in decay series that begin with very long-lived isotopes, sometimes termed *primeval*. The principal primeval isotopes are ^{40}K, ^{232}Th, ^{235}U and ^{238}U, which are mainly concentrated in acid igneous rocks and in sediments deposited as evaporites or in reducing environments. Other primevals, such as ^{48}Ca, ^{50}V and ^{58}Ni, are either rare or only very weakly radioactive.

4.1.5 Radioactive decay series

The main radioactive decay schemes are shown in Table 4.1. The isotope ^{40}K, which forms about 0.0118% of naturally occurring potassium, decays in a single stage, either by beta emission to form ^{40}Ca, or by electron capture (K-capture) to form ^{40}Ar. The argon nucleus is left in an excited state but

Table 4.1 *Natural radioactive decay of ^{238}U, ^{232}Th and ^{40}K*

Parent	Mode	Daughter	Half-life	γ energy (MeV) and % yield[a]
^{238}U	α	^{234}Th	4.5×10^9 years	0.09(15) 0.6(7) 0.3(7)
^{234}Th	β	^{234}Pa	24.1 days	1.01(2) 0.77(1) 0.04(3)
^{234}Pa	β	^{234}U	1.18 min	0.05(28)
^{234}U	α	^{230}Th	2.6×10^5 years	
^{230}Th	α	^{226}Ra	8×10^4 years	
^{226}Ra	α	^{222}Rn	1600 years	0.19(4)
^{222}Rn	α	^{218}Po	3.82 days	
^{218}Po	α	^{214}Pb	3.05 min	
^{214}Pb	β	^{214}Bi	26.8 min	0.35(44) 0.24(11) 0.29(24) 0.05(2)
^{214}Bi	β	^{214}Po	17.9 min	2.43(2) 2.20(6) 1.76(19) 1.38(7) 1.24(7)[b]
^{214}Po	α	^{210}Pb	1.6×10^{-4} s	
^{210}Pb	β	^{210}Bi	19.4 years	
^{210}Bi	β	^{210}Po	5.0 days	0.04(4)
^{210}Po	α	^{206}Pb	138.4 days	
^{232}Th	α	^{228}Ra	1.4×10^{10} years	0.06(24)
^{228}Ra	β	^{228}Ac	6.7 years	
^{228}Ac	β	^{228}Th	6.1 h	1.64(13) 1.59(12) 0.99(25) 0.97(18) 0.34(11)[b]
^{228}Th	α	^{224}Ra	1.9 years	
^{224}Ra	α	^{220}Rn	3.64 days	
^{220}Rn	α	^{216}Po	54.5 s	
^{216}Po	α	^{212}Pb	0.16 s	
^{212}Pb	β	^{212}Bi	10.6 h	0.30(5) 0.24(82) 0.18(1) 0.12(2)[b]
^{212}Bi	β	^{212}Po (66%)	40 min	1.18(1) 0.83(8) 0.73(10)
	α	^{208}Tl (34%)	97.3 min	
^{212}Po	α	^{208}Pb	0.3×10^{-6} s	
^{208}Tl	β	^{208}Pb	3.1 min	2.62(100) 0.86(14) 0.58(83) 0.51(25)[b]
^{40}K	β	^{40}Ca (89%)	1.47×10^9 years	
	K-capture	^{40}Ar (11%)	1.17×10^{10} years	1.46(11)

[a] Percent yields (shown in parentheses) indicate number of decays, out of each 100, that produce photons of the energy specified. The total may exceed 100, since some single decay events produce more than one photon.

[b] Photons of numerous additional energies are emitted in these events.

Note also that decay chain branches that involve less than 10% of the parent element are not shown.

settles down with the emission of a 1.46-MeV photon. The half-life of ^{40}K is 1.47×10^9 years for beta decay and 11.7×10^9 years for K-capture.

The other important primeval radioisotopes decay into nuclei that are themselves unstable. As with ^{40}K, there may be more than one possible decay mode, and the decay chains are quite complicated. All, however, end in stable isotopes of lead. The decay chains for ^{238}U and ^{232}Th are shown in Table 4.1. ^{235}U makes up only about 0.7% of naturally occurring uranium but is less stable than ^{238}U and responsible for almost 5% of uranium radiation. Even so, it may be ignored for most geophysical purposes.

Not all decay events produce significant radiation. The first stage in the decay of ^{232}Th involves only weak gamma activity, and the strongest radiation in the chain (a 2.615-MeV photon, the most energetic radiation to come from any terrestrial source) comes from the decay of ^{208}Tl, near the end.

In the ^{238}U chain, ^{214}Bi is notable for the numbers and energies of the gamma photons produced. Those at 1.76 MeV are taken as diagnostic of the presence of uranium but the gaseous radon isotope, ^{222}Rn, which precedes ^{214}Bi in the chain, has a half-life of nearly 4 days and so can disperse quite widely from a primary uranium source. Gaseous dispersion has much less effect in thorium decay because the half-life of ^{220}Rn is less than a minute.

4.1.6 Radioactive equilibria

If a large amount of a primeval isotope is present, and if all the daughter products remain where they are formed, an equilibrium will eventually be established in which, for each element, the same number of atoms are created in a given time as decay. Only the concentrations of the two end members of the series change.

In equilibrium decay, each member of the chain loses mass at a rate equal to the mass of the element present, multiplied by the appropriate decay constant. Equilibrium masses are therefore inversely proportional to decay constants. If more (less) of an element is present than is required for equilibrium, decay will be faster (slower) than the equilibrium rate until equilibrium is re-established.

Equilibrium can be disrupted if gaseous or soluble intermediate products have half-lives long enough to allow them to disperse before they decay. Dispersion of radon from uranium ores notably disrupts equilibrium, and the primary source of a 'uranium' (actually ^{214}Bi) anomaly may be hard to find. *Roll-front* uranium ores are notorious for the separation between the uranium concentrations and the zones of peak radioactivity.

4.1.7 Natural gamma-ray spectra

Natural gamma-ray energies range from above 3 MeV for *cosmic* (mainly solar) radiation down to the 0.12 MeV of the most energetic X-rays. A typical
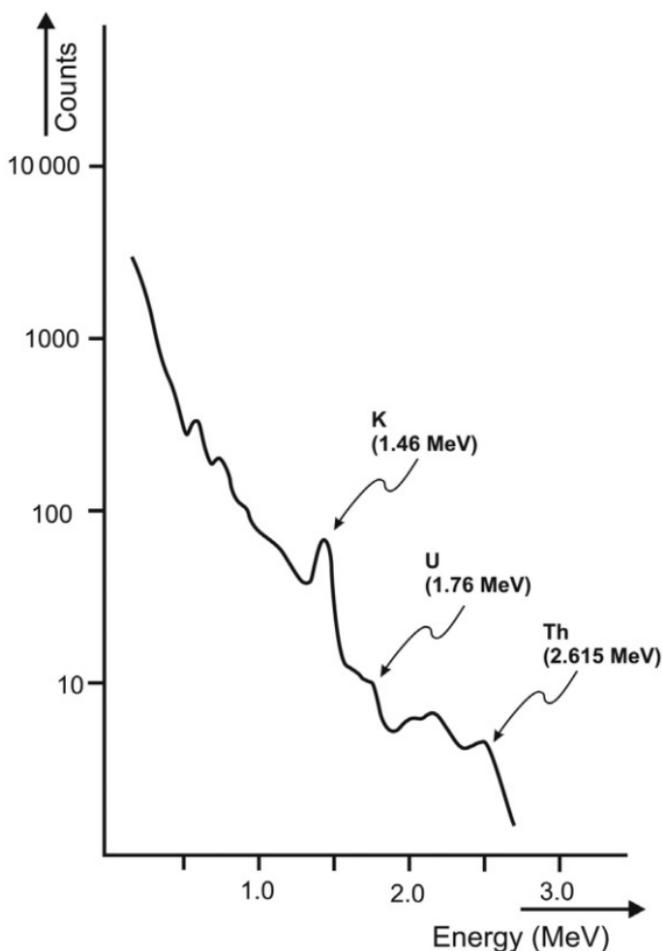
Figure 4.1 *A natural gamma-ray spectrum. Note that the vertical scale (numbers of counts) is logarithmic.*

measured spectrum is shown in Figure 4.1. The individual peaks correspond to specific decay events. Their widths are determined by the kinetic energies of the decaying nuclei and by errors in measurement.

The background curve upon which the peaks are superimposed represents terrestrial and cosmic radiation that has been *scattered*, i.e. has lost some of its initial energy. Photons with energies above 1 MeV may produce electron-positron pairs when passing close to atomic nuclei, and the positrons will interact with other electrons to produce less energetic photons. At lower

energies, photons may lose energy in removing orbital electrons from atoms (*Compton scattering*). Eventually, at X-ray energies, photons may be totally absorbed when ejecting electrons (*photoelectric effect*).

4.2   Radiation Detectors

The earliest radiation detectors relied on the ability of radiation to ionise low-pressure gas and initiate electrical discharges between electrodes maintained at high potential differences. These *Geiger–Müller* counters are now obsolete. They responded mainly to alpha particles and suffered long 'dead' periods after each count, during which no new events could be detected.

4.2.1   Scintillometers

Some materials absorb gamma rays and convert their energy into flashes of light. The most widely used detector is a sodium iodide crystal *activated* by the addition of a small amount of thallium. The light can be detected by photomultiplier tubes (PMTs) that convert the energy into electric current. The whole sequence occupies only a few microseconds and corrections for 'dead time', which in some instruments are made automatically, are required only at very high count rates.

A scintillometer consists of a crystal, one or more PMTs, a power supply (which must provide several hundred volts for the PMTs), and some counting circuitry. The results may be displayed digitally, but quite commonly are shown on the dials of analogue *rate meters*. Some instruments produce audible clicks each time a gamma photon is detected, or have alarms that are triggered when the count rate exceeds a predetermined threshold, so that the dial need not be continually observed.

Radioactive decay is a statistical process. It is impossible to predict when an individual atom will decay but the *average* number of events observed at a given point in a given time will be roughly constant, with some variations about the mean. The continuous averaging of a rate-meter is controlled by a time constant, and if this is too small, the needle will be in continual motion and readings will be difficult to take. If it is too large, the response will be slow and narrow anomalies may be overlooked. Where a digital display is used, a fixed count-time is selected that must be long enough to produce a statistically valid result (see Section 4.3.1).

The sensitivity of a scintillometer depends almost entirely on crystal size; larger crystals record more events. Count rates are thus relative, not absolute, and many instruments are designed to be compatible with several different crystals, chosen on the basis of cost, time available for survey work and accuracy required.

Similar instruments with similar crystals should read roughly the same in the same places, but even this must be checked, since radioactive contaminants near and within the crystals can cause readings to differ. Different scintillometers may record different count rates because crystals are usually shielded so that they detect radiation from one direction only, and even supposedly identical instruments may have different apertures. If it is essential that comparable data be obtained, portable radioactive sources can be used for calibration and also to check the extent of shielding. Such comparisons are, strictly speaking, valid only at the specific (usually fairly low) gamma-ray energy of the test source.

4.2.2   Gamma-ray spectrometers

The energy of the gamma photon that produces a scintillation event can be estimated if a *pulse-height analyser* is incorporated in the PMT circuitry. Events with energies within certain predetermined energy *windows* or above preselected energy *thresholds* can then be counted separately, or the entire gamma-ray flux can be observed at a series of narrow adjoining windows (channels) to obtain a curve such as that shown in Figure 4.1. Strictly speaking, the term *spectrometer* should be reserved for instruments with many channels (typically 256 or more) but in practice it is used for any instrument that has some degree of energy discrimination. Often there are only four channels, one for total count and one each for the ^{208}Tl peak at 2.62 MeV (for thorium), ^{214}Bi at 1.76 MeV (for uranium) and ^{40}K at 1.46 MeV (for potassium). Typical windows for these peaks might extend from 2.42 to 2.82 MeV, from 1.66 to 1.86 MeV and from 1.36 to 1.56 MeV respectively.

Because the characteristics of the measuring circuits vary slowly over time (and also, more rapidly, with temperature), older instruments require regular calibration to check the positions of spectrometer windows or thresholds, using portable sources that produce gamma rays of a single energy. An additional form of calibration is needed if spectrometer results are to be converted directly into radioelement concentrations. In many countries, calibration sites have been established where instrument performance can be checked over concrete pads containing known concentrations of various radioelements. 'Null' pads allow background to be estimated. If several instruments are to be used in a single survey, it is wise, even if they have all been calibrated, to compare them in the actual survey area before attempting to reduce all results to common equivalent readings. Bases at which this has been done should be described for the benefit of later workers. Portable calibration pads are available but are not easily transported. The most modern

instruments are self-stabilising, i.e. they can identify and 'lock-on' to the main naturally occurring peaks.

4.2.3 Stripping ratios

Provided that they are in equilibrium with their daughter products, concentrations of the three parent radioelements can be estimated from spectrometer readings, but corrections must be made for gamma rays scattered from other parts of the spectrum. The thorium peak must be corrected for cosmic radiation, and also for the 2.43-MeV radiation from ^{214}Bi in the uranium decay chain, which overlaps into the commonly used 'thorium' window. The uranium count must in turn be corrected for thorium, and the potassium count for thorium and uranium. The correction process is known as *stripping*.

Stripping factors vary from detector to detector, primarily with variations in crystal size. They are listed in instrument handbooks and in some cases can be applied by built-in circuitry so that the corrected results can be displayed directly, Since abundance estimates will be correct only if equilibrium, which is assumed when automatic corrections are made, actually exists, it is generally preferable to record actual count rates in the field and make corrections later.

4.3   Radiometric Surveys

Ground radiometric surveys can be frustrating. Because of the shielding effect of even thin layers of rock or soil, radioactive minerals may be hard to detect in rocks that are only patchily exposed at the surface. Reliance on stations placed at uniform distances along a traverse may be unwise, and the field observer must be more than usually aware of his or her environment.

4.3.1   Reading times

Accurate radiometric data can be obtained only by occupying each station long enough for the statistical variations to average out. What this implies will depend on the count levels themselves and must be determined by actual experiment. The percent statistical error is equal to about $100/\sqrt{n}$, where n is the number of counts, and so is about 30% for ten counts and only 1% for 10 000. A period of time that is fully adequate for total count readings may not be sufficient for readings on the K, U and Th channels.

If areas where count rates are low are in any case of no interest, there is little point in wasting time obtaining accurate data and it may be sufficient to cover the ground at a slow walk listening to an audio signal or waiting for an alarm to sound. The rate of progress should be such that the narrowest source of interest would not be crossed completely in a time equal to the time constant selected.

Even when a spectrometer is used, it is usual to record only total count in the first instance, reserving the more time-consuming spectral readings for areas of total-count anomaly. There are, of course, risks with this approach, as the concentration of one radioelement might decrease where another increased, but this would be unusual.

4.3.2 Radiometric assays

If a bare rock surface is available, a gamma-ray spectrometer can be used for quantitative thorium, uranium and potassium assays. The rock should be dry, so that absorption by moisture, either on or below the surface, is not a factor. Observations must be taken over sufficiently long periods for statistical fluctuations to be smoothed out, which in practice means accumulating at least 1000 counts. Each count takes a few microseconds, and at 10 000 cps the instrument would be 'dead' for several tens of milliseconds in each second. Corrections are therefore needed for 'dead' time when working with very radioactive material.

Radioelement concentrations are estimated either by inserting the observed count rates into equations, provided by the manufacturers, that are specific to the instrument and crystal being used, or by comparison with 'pad' calibrations.

4.3.3 Corrections for background variations

Atmospheric radon, cosmic radiation and radioactive particles attached to the instrument itself produce background radiation unrelated to survey objectives. The background contributions are usually less than 10% of the total count and are often ignored in ground surveys. If corrections are necessary, either because very subtle variations are being observed or precise assay work is being done, their magnitude can be estimated by taking readings either in the middle of a body of water at least 1 m deep and at least 10 m across or with the detector shielded from the ground by a lead sheet. Neither of these methods is likely to be very convenient, and sometimes 'background' is defined simply (and possibly unreliably) as the lowest reading obtained anywhere in the survey area. Variations in background, due mainly to changes in atmospheric humidity (wet air absorbs radiation far more efficiently than dry), can be monitored using a fixed detector in this location.

The level of background radiation due to radioactive material in the detector itself (including the crystal) should be constant over long periods and can be measured by placing it in a totally shielded environment, but in practice this is likely to be difficult to arrange. The correction is usually trivial and it is far more important to ensure that dirt, which might be contaminated, is not allowed to remain smeared on the detector housing.

An observer is an important possible source of spurious radiation, especially if the sensor is carried in a backpack. In these circumstances the absorption of radiation by the observer's body must also be taken into account, usually by direct experiment. Watches with radioactive luminous dials are now rare, but compasses need to be carefully checked. Obviously, a calibration source should not be carried.

A small amount of radioactive material is included in the vacuum chambers of some gravity meters to prevent build-ups of static electricity on the springs. Radiometric and gravity surveys are occasionally done together, and absurd conclusions have been reached.

4.3.4 Recording radiometric data

Because gamma radiation is strongly absorbed by both rock and soil, comprehensive notes should be taken during radiometric surveys. Radiation comes from a very thin surface layer, and a source with a lateral extent that is small compared with the distance to the detector will produce only a small anomaly. If, on the other hand, the source is extensive and at the surface, the height of the detector should not greatly affect the count rate. Generally, this condition (2π geometry) is achieved if the lateral extent of the source is ten or more times its distance below the detector. Some other possible source geometries and factors for correction to standard 2π values are shown in Figure 4.2.

Source *geometry* is always important in radiometric surveys and especially so in assay work. Any departures from 2π geometry must be noted, together with details of soil cover. If bare rock cannot be seen, some attempt should be made to decide whether the overburden developed *in situ* (and is therefore likely to be radiometrically similar to the bedrock) or was transported into place, and to estimate its thickness. Weather conditions can also be important. In particular, since absorption is much greater in wet than in
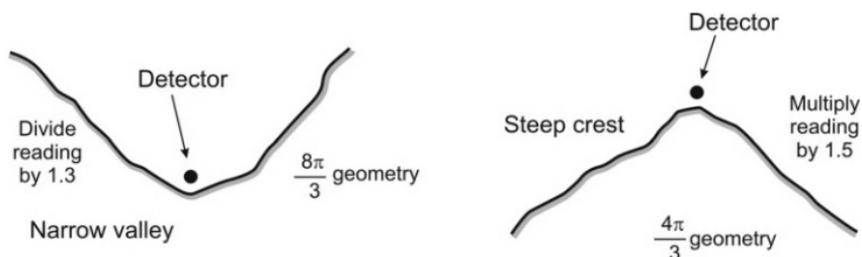

Figure 4.2 *Geometries and correction factors for radiometric surveys over ground with uniform concentrations of radioactive elements.*

dry soil, recent rain and the presence of puddles of standing water should always be noted.

The ways in which readings were taken, including time constants or count periods, must be recorded. Field parameters should not be varied in the course of a survey and the location of the sensor (e.g. whether hand-held or in a backpack) should be specified.

4.3.5 Alpha-particle monitors

Because radiation is absorbed by almost trivial thicknesses of overburden, exploration is sometimes based on soil-gas sampling. Radon gas, which is a copious source of alpha particles, diffuses readily through rocks and soil to the surface and also dissolves in groundwater, and can therefore act as a pathfinder to 'blind' uranium mineralization. The detection methods were first developed for public health applications, since the gas can be a health hazard if allowed to accumulate in enclosed spaces.

In uranium exploration a number of special factors have to be taken into consideration. Dispersal routes can be complex, and locating a primary source is not necessarily easy. Variations in temperature, atmospheric pressure and humidity can cause the flow of gas to the surface to vary by factors of three or more in the course of a day and by an order of magnitude over longer periods, so that, ideally, monitoring should extend over days or even weeks in the hope that these effects will average out. Exploration practicalities usually set limits of one or two weeks, and it is obviously important that all the detectors are in place for the same lengths of time. Methods that rely on measuring absorbed or adsorbed radon are unsuitable because the 3.8-day radon half-life implies that, regardless of the length of the sampling period, the values recorded will relate only to the previous few days. The traces left by alpha particles when passing through some materials provide the basis of true cumulative methods. The most common detector is a strip of plastic film made of CR39 polycarbonate, enclosed in a small chamber to which only filtered gas has access. The tracks are subsequently made visible by etching with a caustic solution

In a radon survey, hundreds or even thousands of detectors may have to be buried across the area of interest, under conditions that are, as nearly as possible, identical. A hand or, preferably, motorised auger is needed to prepare the large number of holes required within an acceptable period of time, and these must be closed in ways that allow the detectors to be quickly retrieved. Backfilled soil can be contained by plastic sheeting, and then simply lifted out, or inverted plastic cups can be attached to the bases of wooden or plastic plugs of the same diameter as the hole (Figure 4.3). Concentrations of soil gas increase with depth, reaching a plateau at about one metre below surface in typical loose soils. Ideally, therefore, the holes
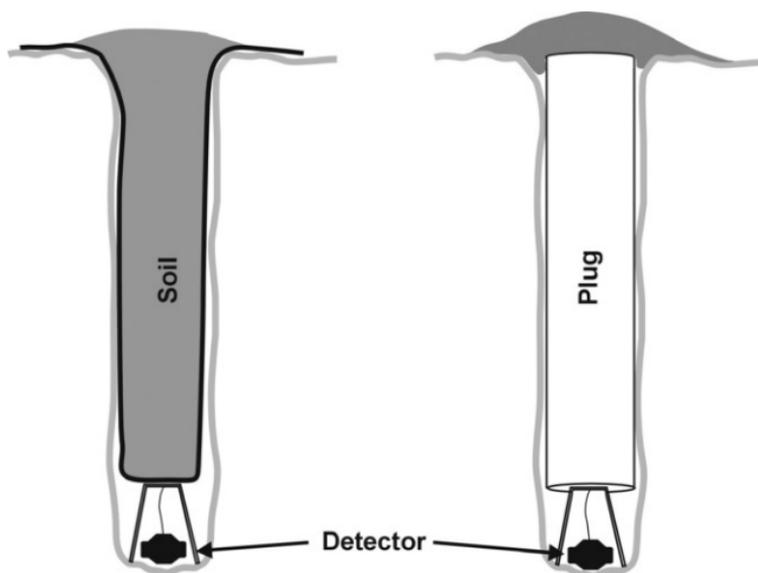
Figure 4.3 Radiation detectors in their holes. The detector is shielded by a simple inverted plastic drinking cup.

should be a metre deep, but it is more important that all the placements are very similar than that they conform to some supposed sampling optimum.

Once retrieved, the detector films are placed in containers that shield them from further exposure to ionising radiation and are transported as quickly as possible to a specialist laboratory. Fortunately, the tedious business of track counting has now been almost completely automated, but the results still require careful interpretation in the light of underlying soil and rock structures.

5

ELECTRIC CURRENT METHODS: GENERAL CONSIDERATIONS

Many geophysical methods rely on measurements of the voltages or magnetic fields associated with electric currents flowing in the ground. Some of these currents are natural, being sustained by oxidation-reduction reactions in the subsurface (self-potential methods) or by variations in ionospheric and atmospheric magnetic fields (magnetotellurics). The currents used in resistivity, induced polarisation (IP) and most electromagnetic (EM) methods are generated artificially.

Currents can be made to flow in the ground by direct injection, by capacitive coupling or by electromagnetic induction. The term 'galvanic' is used where currents are either capacitively coupled or are injected via electrodes. Methods using this second approach are usually referred to as direct current, or DC, even though in practice the current is reversed at intervals of between 1 and 0.05 seconds (1 to 20 Hz) to cancel some forms of natural background noise. In capacitively coupled resistivity (CCR), the currents, typically with frequencies of between 10 and 30 kHz, are sourced from coaxial cables that rest on, but are insulated from, the ground. In EM surveys the currents are driven inductively by time-varying magnetic fields generated by coils or long wires that are not in contact with the ground.

Because all electrical methods respond to the same intrinsic properties of materials (resistivity, conductivity and chargeability), and because of overlaps between them (and because what is noise in one type of survey may be signal in another), the basic electrical concepts are all introduced in this chapter. DC and CCR methods are then discussed in Chapter 6. Natural potential (self-potential or *SP*) and *induced polarisation* (*IP*) methods are covered in Chapter 7, and Chapters 8 and 9 deal with EM surveys using, respectively, local and remote (often natural) sources. Many modern transmitters and receivers can be used interchangeably for DC, IP and EM surveys.

5.1 Resistivity and Conductivity

Metals and most metallic sulphides conduct electricity efficiently by flow of electrons. Electrical methods are therefore important in the search for

sulphide ores, as well as in environmental investigations where metallic objects are often the targets. Graphite is a good 'electronic' conductor and, since it is not itself a useful mineral, is a source of noise in mineral exploration. Conduction in most other types of rock is by ions in the pore waters, and electrical methods are important in water resource investigations.

5.1.1 Ohm's Law and resistivity

The current that flows in a conductor is in most cases proportional to the voltage across it, i.e.:

$$V = IR$$

This is *Ohm's Law*. The constant of proportionality, *R*, is known as the resistance and is measured in ohms when current (*I*) is in amps and voltage (*V*) is in volts. The reciprocal, conductance, is measured in Siemens, also known as mhos.

The resistance of a metre cube of a material to current flowing between opposite faces defines the resistivity (ρ) of that material and is measured in ohm-metres (Ω m). The reciprocal, conductivity (σ), is expressed in Siemens per metre (S m^{-1}) or mhos per metre. The resistance of a rectangular block to current flow between opposite faces is proportional to the resistivity and to the distance *x* between the faces, and inversely proportional to their cross-sectional area, *A*, i.e.:

$$R = \rho(x/A)$$

Isotropic materials have the same resistivity in all directions. Most rocks are reasonably isotropic but strongly laminated slates and shales are more resistive across the laminations than parallel to them.

Geophysicists working with galvanic methods generally talk about *resistivity*, while those working with induction methods talk about *conductivity*. Both quantities are, in fact, complex (using the word in its strict mathematical sense), involving both amplitude and phase (see Section 5.2.3). Amplitudes reflect the bulk resistance of the ground, while phase is determined by the ground's ability to store electric charge, i.e. the *chargeability*.

5.1.2 Electrical resistivities of rocks and minerals

The resistivity ranges for some common rocks and minerals are listed in the fourth column of Table 1.2. Conductivity ranges can, of course, be calculated from these values but, for convenience, are listed in the fifth column. Table 1.2 emphasises the possible ambiguities in interpretation, implying, for example, that the resistivity contrast between a wet sand and underlying competent limestone bedrock may be the reverse of that between
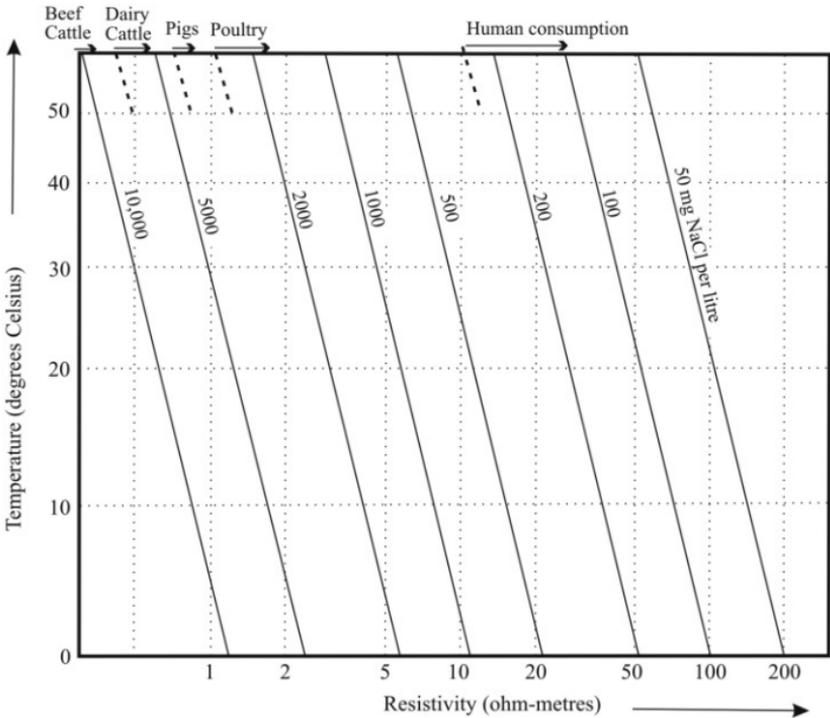
Figure 5.1 *Variation of water resistivity with concentration of dissolved NaCl. The uses that can be made of waters of various salinities are also indicated.*

a dry sand and underlying weathered limestone. Bulk resistivities of more than $10\,000\,\Omega\,m$ or less than $1\,\Omega\,m$ are rarely encountered in field surveys.

Most rock-forming minerals are very poor conductors, and subsurface currents are carried mainly by ions in the pore waters. Pure water ionises only very slightly, so water in its pure state is almost non-conductive; the electrical conductivity of pore waters depends on the presence of dissolved salts, mainly sodium chloride (Figure 5.1). However, clay minerals are ionically active and clays conduct well if even slightly moist.

In many rocks, resistivity is roughly equal to the resistivity of the pore fluids divided by the fractional porosity. A closer approximation is provided by *Archie's Law*, which states that:

$$\rho = a.\rho_w/P^m$$

where ρ is the bulk resistivity of a saturated porous medium, P is the fractional porosity, ρ_w is the pore fluid resistivity, and m and a are empirical

quantities determined by the geometry of the pores. The *m* parameter varies between about 1.2 and 1.8 according to the shape of the matrix grains. Archie's Law departures from linearity are small for common values of porosity (Figure 5.2).
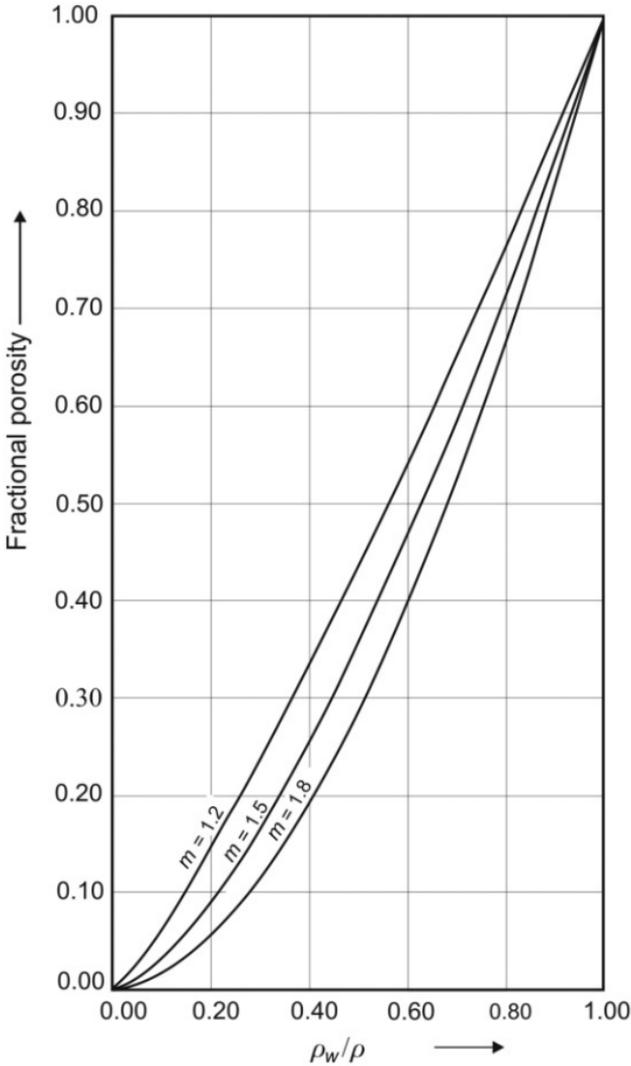
Figure 5.2 *Variation, according to Archie's Law, of bulk resistivity, ρ, for a rock with an insulating matrix and pore-water resistivity ρ_w. The index, m, is about 1.2 for spherical grains and about 1.8 for platey or tabular materials.*

5.1.3 Apparent resistivity

A single electrical measurement tells us very little. The most that can be extracted from it is the resistivity of a completely homogeneous ground (also described as a *homogeneous half space*) that would produce the same result under the same survey conditions. This quantity is known as the *apparent resistivity* (ρ_α). Variations in apparent resistivity or its reciprocal, *apparent conductivity,* provide the raw material for interpretation in most electrical surveys.

Where electromagnetic methods are being used to detect very good conductors such as sulphide ores or steel drums, locating the target is more important than determining its precise electrical properties. Since it is difficult to separate the effects of target size from target conductivity for small targets, results are sometimes presented in terms of a *conductivity-thickness product*.

5.1.4 Overburden effects

Build-ups of salts in the soil produce high conductivities in near-surface layers in many arid tropical areas. Conductive overburdens will effectively short-circuit any currents produced by sources situated at or above the ground surface and therefore pose problems for all electrical methods. Continuous-wave electromagnetic methods are the most severely affected.

Highly resistive surface layers are obstacles in surveys using electrodes but may actually be advantageous in EM surveys, because attenuation is reduced and the depth of investigation is increased. Capacitive coupling can also be used, provided that the very resistive layer is less than about a metre thick.

5.1.5 Anisotropy

Most elementary analysis of electrical data (and certainly most analyses performed in the field) assume that resistivity is the same in all directions. This is usually true where current is carried by ions in pore waters, but not necessarily in other cases. For example, a graphitic shale usually conducts electricity much more readily along the bedding planes than across them. Currents and the electric fields that drive them are vectors but in anisotropic media are not necessarily in the same direction. The resistivity or conductivity that describes their relationship is therefore a tensor.

Currents may also be driven by alternating magnetic fields, as described in Section 5.2, and there is therefore a magnetic resistivity/conductivity tensor as well as an electric one. The use of full tensor information is still uncommon, but where it is required by the interpreters, field crews will find their lives made considerably more complicated, because more sensors will

have to be deployed, and more measurements will have to be made, at each field point.

5.2 Varying Currents

Varying electrical currents circulating in *transmitter* circuits can generate subsurface currents without actual physical contact, using either inductive or capacitive coupling. Such methods, essential in airborne work, can also be useful on the ground, since making direct electrical contact is always tedious and may be impossible on concrete, asphalt, ice or permafrost.

5.2.1 Induction

In the late 1800s, Ampere discovered that a circular magnetic field existed around a current passing through a wire. A little later, Faraday realised that, conversely, a changing magnetic field could induce current in a wire loop. Maxwell combined and expanded these principles to provide a complete description of the interactions between EM fields and conductors. Geophysical interpretation of EM data is about visualising Ampere's and Faraday's laws at work in the subsurface.

In a changing magnetic field, a voltage (electromotive force, or *emf*) is induced at right angles to the direction of the change, and currents will flow in any nearby conductors that form parts of closed circuits. The equations governing this phenomenon are relatively simple, but geological conductors are very complicated and for theoretical analyses the induced currents, known as *eddy currents*, are approximated by greatly simplified models.

Eddy current magnitudes are determined by the rates of change of current flow in the inducing circuits and by a geometrical parameter known as the *mutual inductance*. Mutual inductances are large, and conductors are said to be well coupled, if there are long adjacent conduction paths, if the changes in the inducing magnetic field take place at right angles to directions of easy current flow, and if magnetic materials are present to enhance field strengths.

When current changes in a circuit, an opposing emf is induced in that circuit. As a result, a tightly wound coil strongly resists current change and is said to have a high *impedance* and a large *self-inductance*.

5.2.2 Permeability and permittivity

Electrical conductivity and resistivity are adequate for describing the ways in which charges move in conductors in response to steady fields. With varying fields, other factors become important, and Maxwell's *constitutive equations* describe how materials respond to the fields that constitute an electromagnetic wave. Permittivity, ε, describes how constrained charges

(**d**) move in response to an electric field (**E**):

$$\mathbf{d} = \varepsilon.\mathbf{E}$$

Magnetic permeability, μ, describes how atomic and molecular magnetic moments (**b**) respond to a magnetic field (**H**):

$$\mathbf{b} = \mu.\mathbf{H}$$

The quantities in bold are vectors. The equations require even empty space to have both permittivity, ε_0, and permeability, μ_0, with values, in SI units, of 8.85×10^{-12} farads per metre and $4\pi \times 10^{-7}$ henries per metre respectively. Complications then arise in treating other media because it is possible to use either absolute values, ε and μ, or relative values ε_r and μ_r, such that:

$$\varepsilon = \varepsilon_0 \, \varepsilon_r \quad \text{and} \quad \mu = \mu_0 \, \mu_r$$

This would not be a problem, except that often the subscripts (and the word 'relative') are omitted and it is not always easy to know whether absolute or relative quantities are being used. A good general rule is that, unless otherwise stated or indicated, symbols without subscripts in equations are likely to refer to absolute values, whereas numerical values are nearly always relative. Permittivity is most important in radar (GPR) work, where the relative values are known also as the dielectric constants (Chapter 10).

5.2.3 Phase

The behaviour of alternating fields in the subsurface is fundamental to much of EM surveying. These fields can be described using sine and cosine functions.

If, in Figure 5.3, the line OP, length **a**, rotates anticlockwise from an initial horizontal position at a constant angular velocity ω, then, at any time t, OR will equal **a**.$\sin \omega t$ and OS will equal **a**.$\cos \omega t$. These two functions can be represented graphically as sine and cosine curves, or *waves*, where the horizontal axis represents time and the vertical axis represents displacement. Such waves are termed *sinusoidal*. The *period* of the sinusoidal wave is the time that it takes OP to describe the complete circle, i.e. rotate through an angle of 360° or 2π radians, and is equal to $2\pi/\omega$ if ω is measured in radians per second. The two waves have the same maximum amplitude (usually referred to as simply 'the amplitude'), which is equal to **a**. They differ only by an angular displacement, referred to as a difference in *phase*, of 90° or $\pi/2$ radians. The cosine wave in Figure 5.3 is said to *lead* the sine wave
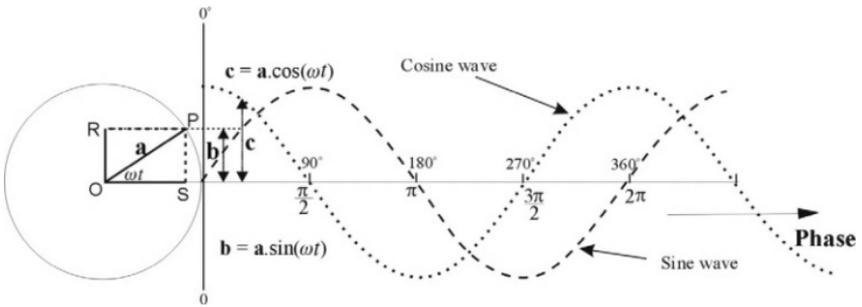
Figure 5.3 *Generation of sinusoidal waves by an anticlockwise-rotating arm. See text for explanation of symbols.*

by $\pi/2$ radians, and the sine wave is said to *lag* the cosine wave by $\pi/2$ radians (it looks, at first sight, as if it should be the other way round, but what the graphs actually show is that the cosine wave reaches its maximum amplitude a quarter of a cycle before the sine wave).

A completely general *sinusoidal* wave is described by the equation:

$$z = a \cdot \sin(\omega t + \varphi)$$

where φ is the *phase angle*. Any curve of this type can be resolved into separate sine and cosine waves with amplitudes s and c related to the original amplitude and phase by the equations:

$$a = \sqrt{(s^2 + c^2)} \quad \text{and} \quad \tan \varphi = s/c$$

These equations can be derived from the picture in Figure 5.3 of a rotating arm, but in this case OP would begin its rotation (at time $t = 0$) with an initial phase angle φ, which translates into a difference in time equal to φ/ω. A minor complication is that in some geophysical applications (notably in seismic processing), phase angles are referenced to the symmetrical cosine wave rather than the anti-symmetric sine wave.

In EM surveys, the induced currents and their associated secondary magnetic fields differ in phase from the primary field and can therefore be resolved into components that are in-phase and 90° (or $\pi/2$ radians) out of phase with the primary. The *out-of-phase* component can also (more accurately and less confusingly) be described as being in *phase-quadrature* with the primary signal. The fact that the two components can be represented by vectors drawn at right angles (orthogonal) to each other also allows them to be described in terms of the mathematics of complex numbers, which uses

the 'imaginary' square root of -1 (conventionally indicated by either 'i' or 'j', according to taste). The in-phase and quadrature components are then described as being respectively *real* and *imaginary*. A knob labelled 'I' on an EM instrument may therefore control either the 'In-phase' or the 'Imaginary' signal, according to the manufacturer's whim. Yet another reason for reading the handbook.

Since electromagnetic waves travel through air at the speed of light and not instantaneously, phase alters with distance from the transmitter. The small distances between transmitters and receivers in most geophysical surveys ensure that these shifts are negligible and can be ignored.

5.2.4 Transients

As an alternative to sinusoidal signals, currents circulating in transmitter coils or wires can be terminated abruptly. *Transient electromagnetic* (*TEM*) methods are effectively multi-frequency, because a square wave contains elements of all the odd harmonics of the *fundamental* up to theoretically infinite frequency (Figure 5.4). Most of the many advantages of TEM over continuous wave (CW) methods derive from the fact that TEM measurements are made after the termination of the primary current. There is thus
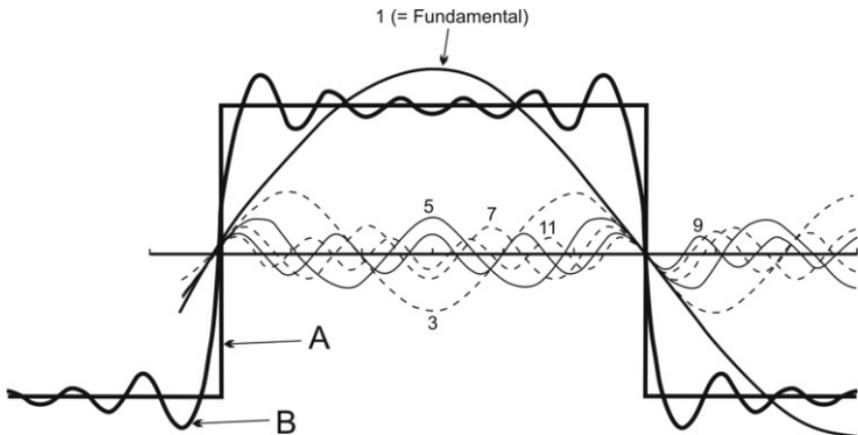
Figure 5.4 *The square wave (A) as a multi-frequency sinusoid. A good approximation (B) to the wave can be obtained by summing the first five odd harmonics (integer multiples 3, 5, 7, 9 and 11) and the fundamental frequency (1). The amplitudes required for each of these component waves can be determined using the techniques of Fourier analysis. The addition of higher odd harmonics with appropriate amplitudes would further improve the approximation.*

no possibility of part of the primary field 'leaking' into secondary field measurements, either electronically or because of errors in coil positioning.

5.2.5 Depth penetration

The currents that electromagnetic fields cause to flow in nearby conductors extract energy from the fields and therefore reduce penetration. The attenuation of a planar alternating wave in a continuous conductor obeys an exponential law (see Section 1.2.6) governed by an attenuation constant (α) given by:

$$\alpha = \omega \left[\mu\, \varepsilon \left\{ \left(\sqrt{(1 + \sigma^2/\omega^2 \varepsilon_a^2)} \right) - 1 \right\} /2 \right]^{1/2}$$

Here μ and ε are the absolute values of, respectively, magnetic permeability and electrical permittivity, and ω ($= 2\pi f$) is the *angular frequency*. The reciprocal of the attenuation constant is known as the *skin depth*, and is equal to the distance over which the signal falls to $1/e$ of its original value. Since e, the base of natural logarithms, is approximately equal to 2.718, signal strength decreases by almost two-thirds over a single skin depth.

The rather daunting attenuation equation simplifies considerably under common limiting conditions. At the frequencies used in EM surveys, the factor $\sigma^2/\omega^2\varepsilon^2$ is much larger than 1 and the quantity inside the square brackets reduces to $\sigma/\omega\varepsilon$, which, with a little further manipulation, implies that α is equal to $\sqrt{(\mu\sigma\omega/2)}$. If, as is usually the case, the value of the magnetic permeability is close to the value in free space, then:

$$\alpha = \frac{\sqrt{\sigma f}}{500}$$

The wavelengths of electromagnetic waves in Earth materials are approximately equal to the skin depths (Figure 5.5) multiplied by 2π. The depth of investigation in situations where skin depth is the limiting factor, i.e. where planar or quasi-planar waves are involved, is commonly quoted as being equal to the skin depth divided by $\sqrt{2}$, i.e. to about $350/\sqrt{(\sigma f)}$. Ideally, surveys should be designed so that the skin depth is twice the depth of the deepest target object. This is not the whole story, because inverse-cube law attenuation has to be taken into account for small dipole sources, and inverse-square law attenuation for long wires. An important concept in systems where both transmitters and receivers are small coils is that of *induction number*, which is equal to the coil separation divided by the skin depth. At low induction numbers, i.e. in situations where the coil separation
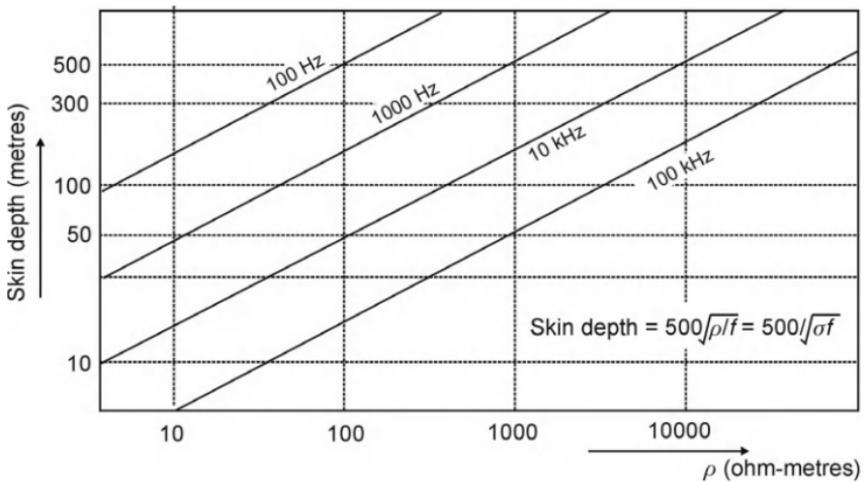
Figure 5.5 *Variation in skin depth,* d*, with frequency,* f*, resistivity,* ρ*, and conductivity,* σ*.*

is much smaller than the skin depth, the separation is the dominant factor controlling depth penetration.

The higher the frequency, the more precisely small targets will be located and resolved, but the skin-depth equation shows that penetration is reduced at higher frequencies. In CW surveys, additional information can be obtained by using two or more frequencies. TEM surveys are inherently multi-frequency.

6
RESISTIVITY METHODS

Nomenclature in electrical geophysics can be confusing. The flow of current in the so-called *direct current* (DC) surveys is usually reversed at intervals of one or two seconds, and CCR surveys, in which alternating currents are introduced into the ground by capacitive coupling, have more in common with DC than EM methods, despite the use of kHz frequencies. These two methods are discussed in this chapter.

6.1 DC Survey Fundamentals

Surface resistivity methods are based on the principle that the electric potentials (voltages) measured around a current-carrying electrode are affected by the electrical resistivities of the underlying materials.

6.1.1 Apparent resistivity

The 'obvious' method of measuring ground resistivity, by simultaneously passing current through and measuring voltage between a single pair of grounded electrodes, does not work, because of contact resistances, which depend on such things as ground moisture and contact area and may amount to thousands of ohms. The problem can be avoided if, as illustrated in Figure 6.1, voltages are measured between a second pair of electrodes using a high-impedance voltmeter that draws virtually no current. The voltage drop through the voltage electrodes is then negligible, and although the resistances at the current electrodes limit current flow, they do not affect resistivity calculations. Usually the current electrodes are in line with the voltage electrodes but they can be located anywhere. A geometric factor is needed to convert the readings obtained with these four-electrode *arrays* to resistivity.

The result of any single measurement with any array could be interpreted as due to a homogeneous ground with constant resistivity. The geometric factors used to calculate this *apparent resistivity*, ρ_α, can be derived by applying, to each current-electrode/voltage-electrode pair, the formula
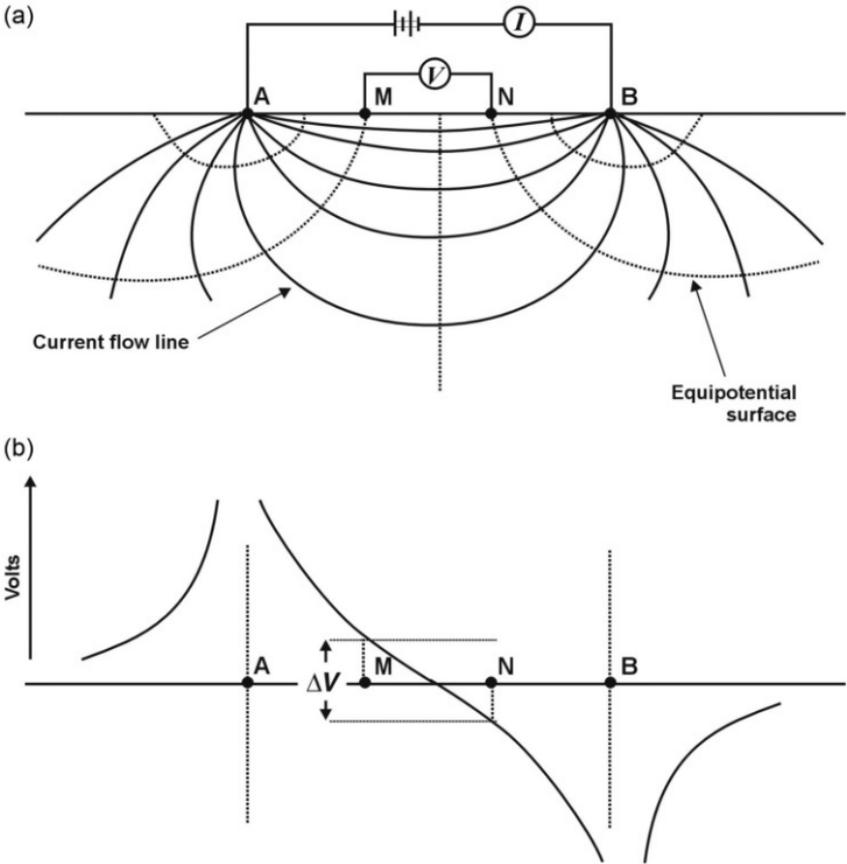
$$V = \rho I/2\pi a$$

Figure 6.1 (a) Current flow (continuous lines) between electrodes A and B and equipotential surfaces (dotted lines, orthogonal to the current lines) in a homogeneous half-space. (b) Voltage drops. The potential ΔV is measured between the electrodes M and N, positioned (in this case) as for a Wenner array.

which defines the electric potential at a distance *a* from a point electrode at the surface of a *uniform half-space* (homogeneous ground) of resistivity ρ. The current *I* may be positive (if into the ground) or negative, and the potential at any point is equal to the sum of the contributions from the individual current electrodes.

The difference between the summed potentials at the two voltage electrodes M and N is equal to:

$$\Delta V = I\rho/2\pi(1/[MA] - 1/[NA] - 1/[MB] + 1/[NB])$$

110

where a current I flows between current electrodes at A and B and the voltage electrodes are at M and N (Figure 6.1a), and the quantities in square brackets represent inter-electrode distances. These distances are always the actual distances between electrodes, whether or not they lie in a straight line. The quantity in brackets is often denoted by $1/K$, and the apparent resistivity can then be calculated from the *basic array equation*:

$$\rho_a = 2\pi K (V/I)$$

where V and I are measured and K is the *geometric factor,* expressed in metres, which is determined by the electrode arrangement.

Geometric factors are not affected by interchanging current and voltage electrodes, but voltage electrode separations are normally kept small to minimise the effects of natural potentials.

6.1.2 Electrode arrays

Figure 6.2 shows some common electrode arrays and their geometric factors. The names are those in general use and may upset pedants. A dipole, for example, *should* consist of two electrodes separated by a distance that is negligible compared to the distance to any other electrode. Application of the term to the dipole-dipole and pole-dipole arrays, where the distance to the next electrode is usually only between one and six times the 'dipole' spacing, is thus formally incorrect. Not many people worry about this.

The distance to a fixed electrode 'at infinity' should be at least 10, and ideally 30, times the distance between any two mobile electrodes. The long cables required can impede field work, and can also act as aerials, picking up stray electromagnetic signals that can affect the readings (*inductive noise*).

6.1.3 Array descriptions (see Figure 6.2)

Wenner array: Very widely used, and supported by a vast amount of inter-
   pretational literature and computer packages. The 'standard' array against
   which others are often assessed.
Two-electrode (pole-pole) array: Theoretically interesting since it is possible
   to calculate from readings taken with it along a traverse the results that
   would be obtained from any other type of array, providing coverage
   is adequate. However, the noise that accumulates when large numbers
   of results obtained with closely spaced electrodes are added together
   prevents any practical use being made of this fact. The array is very
   popular in archaeological work because it lends itself to rapid one-man
   operation. As the *normal* array, it is one of the standards in electrical
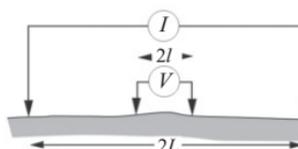   well-logging.

(a) Wenner

$$\rho_a = 2\pi a \frac{V}{I}$$
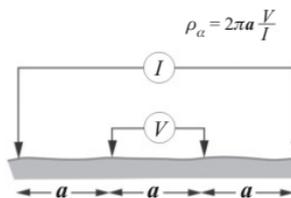
(b) Two-electrode (pole-pole)
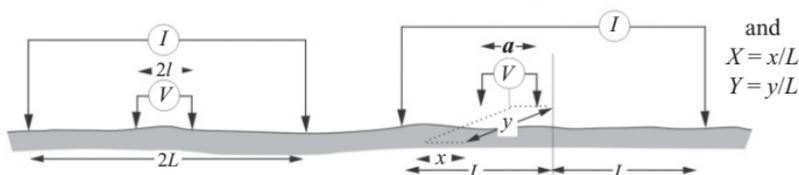
$$\rho_a = 2\pi a \frac{V}{I}$$
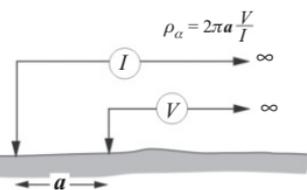
(c) Schlumberger

Exact $\quad \rho_a = \pi \frac{L^2 - l^2}{2l} \frac{V}{I}$

Ideal dipole '$a = 2l$' $\rho_a = \pi \frac{L^2}{a} \frac{V}{I}$

(d) Gradient

Ideal dipole 'a' $\quad \rho_a = 2\pi \frac{L^2}{Ca} \frac{V}{I}$

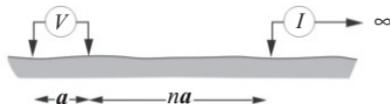where $C = \left[\dfrac{1-X}{[Y^2 + (1-X)^2]^{3/2}} + \dfrac{1+X}{[Y^2 + (1+X)^2]^{3/2}} \right]$

and

$X = x/L$
$Y = y/L$

(e) Dipole-dipole

$$\rho_a = \pi n(n+1)(n+2) a \frac{V}{I}$$

(f) Pole-dipole

$$\rho_a = 2\pi n(n+1) a \frac{V}{I}$$

(g) Square array

No factor

$$\rho_a = \frac{2\pi a}{2 - \sqrt{2}} \frac{V}{I}$$
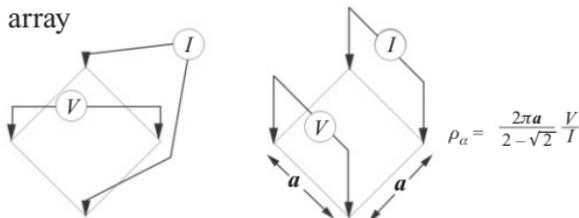
Figure 6.2 Common electrode arrays and their geometric factors. There is no factor for the diagonal square array, because no voltage difference is observed over homogeneous ground.

112
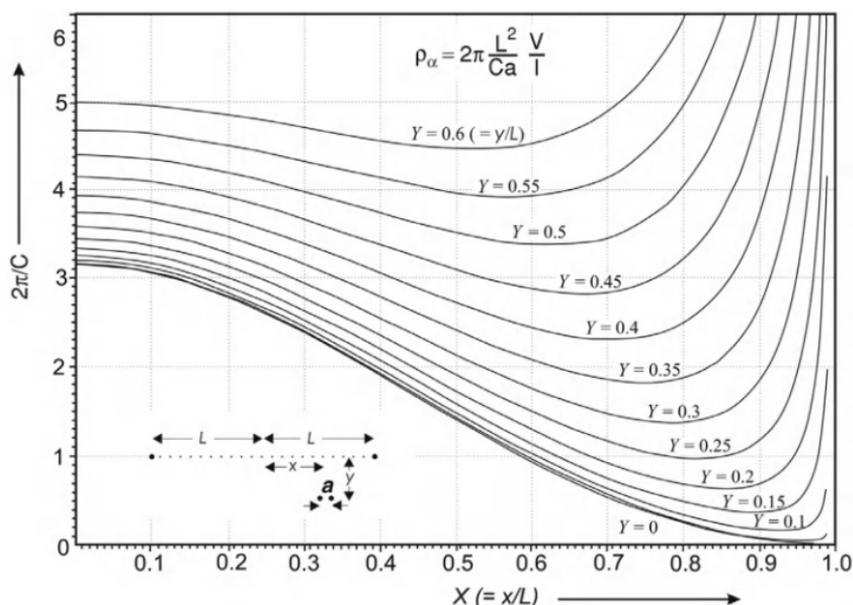
Figure 6.3 *Approximate variation in gradient-array response with voltage dipole position. Array total length* 2L, *voltage dipole length* **a**. *The factor* C *is defined in Figure 6.2d. The distances* x *and* y *define the position of the centre of the voltage dipole, as measured along the centre-line and at right angles to the centre-line respectively. For* x = 0 *and* y = 0, C = 2 *and the gradient array and Schlumberger array 'ideal dipole' equations are identical. The approximate equation should be used only where the curves are roughly horizontal, and it is generally preferable to use the basic array equation and the exact factors. This is easily done using a computer spreadsheet program.*

Schlumberger array: The only array to rival the Wenner in availability of interpretational material, all of which relates to the 'ideal' array with negligible distance between the inner electrodes. Favoured, along with the Wenner, for electrical depth-sounding work.

Gradient array: Widely used for reconnaissance, especially in induced polarisation (IP) surveys. Provided that sufficiently powerful generators are available, large numbers of readings can be taken on parallel traverses without moving the current electrodes. Figure 6.3 shows how the geometric factor given in Figure 6.2d varies with the position of the voltage dipole.

Dipole-dipole (Eltran) array: Popular in IP work because the complete separation of current and voltage circuits reduces the vulnerability to inductive

noise. A considerable body of interpretational material is available. Information from different depths is obtained by changing n. In principle, the larger the value of n, the deeper the penetration of the current path sampled. Results are usually plotted as pseudo-sections (see Section 7.5.2).

Pole-dipole array: Produces asymmetric anomalies that are more difficult to interpret than those produced by symmetric arrays. Peaks are displaced from the centres of conductive or chargeable bodies and electrode positions have to be recorded with special care. Values are usually plotted at the point midway between the moving voltage electrodes but this is not a universally agreed standard. Results can be displayed as pseudo-sections, with depth penetration varied by varying n.

Square array: Four electrodes positioned at the corners of a square are variously combined into voltage and current pairs. Depth soundings are made by expanding the square. The entire array is moved laterally when traversing. Inconvenient, but can provide an experienced interpreter with vital information about ground anisotropy and inhomogeneity. Few published case histories or type curves.

Multi-electrode arrays (not shown)

Lee array: Resembles the Wenner array but has an additional central electrode. The voltage differences from the centre to the two 'normal' voltage electrodes give a measure of ground inhomogeneity. The two values can be summed for application of the Wenner formula.

Offset Wenner: Similar to the Lee array but with all five electrodes the same distance apart (see Figure 6.9). Measurements made using the four right-hand and the four left-hand electrodes separately as standard Wenner arrays are averaged to give apparent resistivity and differenced to provide a measure of ground inhomogeneity.

Focused arrays: Multi-electrode arrays have been designed to focus current into the ground to give deep penetration without large expansion. Arguably, this is an attempt to do the impossible, and such arrays should be used only under the guidance of an experienced interpreter.

6.1.4 Signal-contribution sections

Current-flow patterns for one and two-layers are shown in Figure 6.4. Near-surface inhomogeneities strongly influence the choice of array. Their effects are graphically illustrated by contours of the *signal contributions* made by each unit volume of ground to the measured voltage, and hence to the apparent resistivity (Figure 6.5). For linear arrays, these contours have the same appearance in any plane, whether vertical, horizontal or dipping,
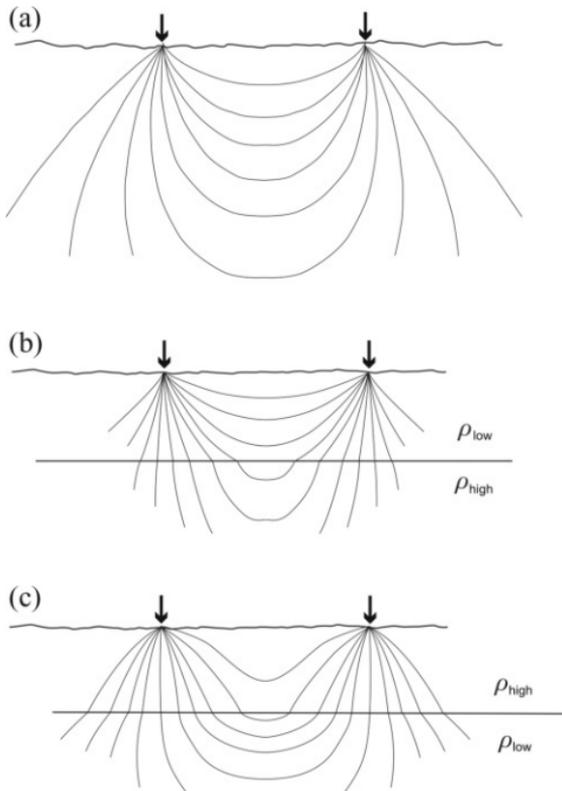
Figure 6.4 *Current flow patterns for (a) uniform half-space; (b) two-layer ground with lower resistivity (ρ_{low}) in upper layer; and (c) two-layer ground with higher resistivity (ρ_{high}) in upper layer.*

through the line of electrodes (i.e. the contours are semicircles when the array is viewed end on).

A reasonable first reaction to Figure 6.5 is that useful resistivity surveys are impossible, as the contributions from regions close to the electrodes are very large. Some disillusioned clients would endorse this view. However, the variations in sign imply that a conductive near-surface layer will in some places increase and in other places decrease the apparent resistivity. In homogeneous ground these effects can cancel quite precisely.

When a Wenner or dipole-dipole array is expanded, all the electrodes are moved and the contributions from near-surface bodies vary from reading to reading. With a Schlumberger array, near-surface effects vary much less, provided that only the outer electrodes are moved, and for this reason the
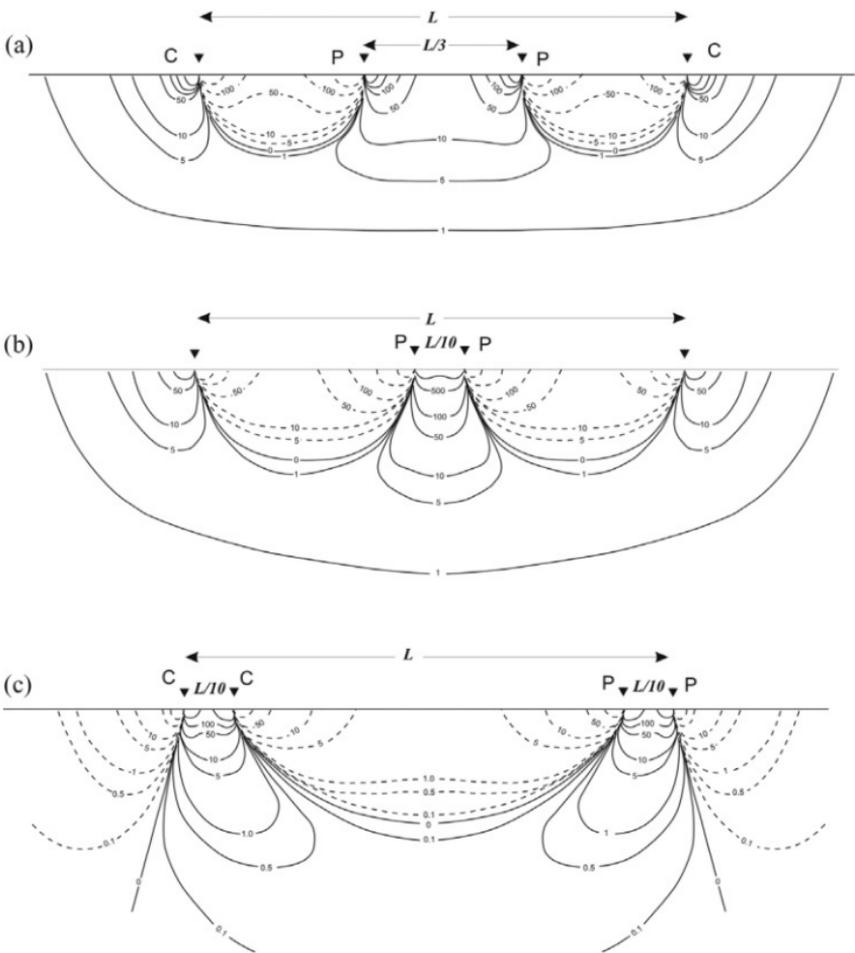
Figure 6.5 *Signal contribution sections for (a) Wenner, (b) Schlumberger and (c) dipole-dipole arrays. The contours show the relative contributions to the signal from unit volumes of homogeneous ground. Dashed lines indicate negative values. Reproduced by permission of Dr R. Barker.*

array is often preferred for depth sounding. However, offset techniques (see Section 6.4.3) allow excellent results to be obtained with the Wenner.

Near-surface effects may be large when a gradient or two-electrode array is used for profiling, but they are also very local. Large numbers of readings are collected, and smoothing filters can be applied.

6.1.5 Depth penetration

Arrays are usually chosen at least partly for their depth penetration, which is almost impossible to define because the depth to which a given fraction of current penetrates depends on the layering as well as on the separation between the current electrodes. Voltage electrode positions determine which part of the current field is sampled, and the penetrations of the Wenner and Schlumberger arrays are thus likely to be very similar for similar total array lengths. For either array, the expansion at which the effect of a deep interface beneath a homogeneous upper layer becomes evident depends on the resistivity contrast, but is of the order of a quarter of the spacing between the outer electrodes.

For any array, there is a depth at which the effect of a thin layer of different resistivity in otherwise homogeneous ground is a maximum. This would be expected to be less than the depth at which the interface in a two-layer Earth first becomes apparent, and plots of the effects of such a layer, shown in Figure 6.6 for the Wenner, Schlumberger and dipole-dipole arrays, confirm this. By this criterion, the Wenner is the least, and the dipole-dipole the most, penetrative array. However, the Wenner curve is the most sharply peaked, suggesting better resolving power. This is confirmed by the signal-contribution contours (Figure 6.5), which are slightly flatter at depth for the Wenner than for the Schlumberger array, indicating that the Wenner locates flat-lying interfaces more accurately. The signal-contribution contours for the dipole-dipole array are almost vertical at considerable depths in some places, indicating that the array is best suited to mapping lateral rather than vertical changes.

6.2 DC Practicalities

The currents used in surveys described as 'direct current', or *DC*, are never actually unidirectional. Reversing the direction of flow allows the effects of unidirectional natural currents to be eliminated by simply summing and averaging the results obtained in the two directions.

DC surveys require current generators, voltmeters, ammeters and electrical contact with the ground. Cables and electrodes are cheap but vital parts of the systems, and it is with these that much of the noise is associated.

6.2.1 Metal electrodes

The electrodes used to inject current into the ground are nearly always metal stakes, which in dry ground may have to be hammered in to depths of more than 50 cm (but not to more than 10% of the inter-electrode spacing) and be watered to improve contact. Where contact is very poor, salt water and multiple stakes may be used. If this fails to lower contact resistances
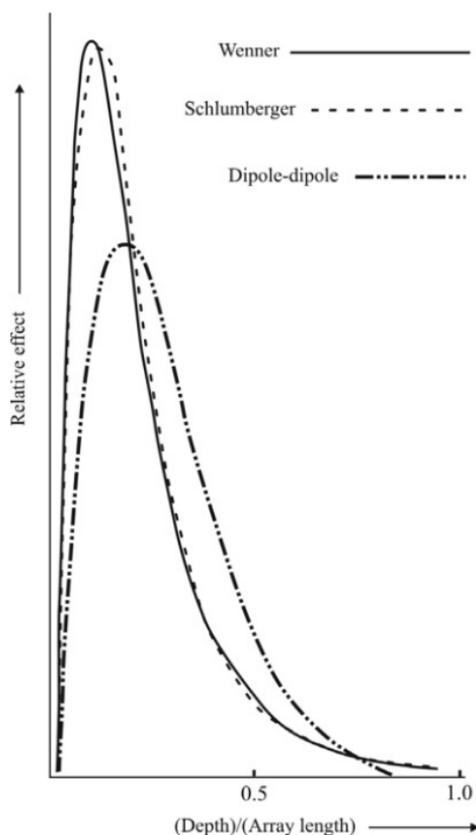
Figure 6.6 *Relative effect of a thin, horizontal high-resistance bed in other-wise homogeneous ground. The areas under the curves have been made equal, concealing the fact that the voltage observed using the Schlumberger array will be somewhat less, and using the dipole-dipole array very much less, than with the Wenner array.*

sufficiently, a pit can be dug, lined with aluminium foil, wetted with salt water and backfilled with soil. In arid conditions, detergent can be added to the water to break down surface tension and improve contact with the ground. A useful method of improving contact in coarse-grained materials such as gravel is to fill holes with a clay slurry before inserting the metal stake. In extreme cases, holes may have to be blasted through highly resistive caliche or laterite surface layers.

Metal stake electrodes come in many forms. Lengths of drill steel are excellent if the ground is stony and heavy hammering necessary. Pointed lengths of angle-iron are only slightly less robust and have larger contact

areas. If the ground is soft and the main consideration is speed, large numbers of metal tent pegs can be pushed in along a traverse line by an advance party.

Polarisation voltages are generated wherever metals are in contact with the ground water. Their magnitudes depend on the metals concerned and they are relatively small when electrodes are made of materials such as stainless steel. Polarisation voltages are unidirectional, and the routine reversal of current flow generally achieves acceptable levels of cancellation of these effects in conventional DC surveys.

6.2.2 Non-polarising electrodes

Polarisation voltages would be serious sources of noise in *SP* surveys, which involve the measurement of small natural (and unidirectional) potentials, and in induced polarisation (*IP*) surveys (see Chapter 7), and non-polarising electrodes must therefore be used. Their design relies on the fact that no contact potential exists at an interface between a metal and a saturated solution of one of its own salts. The commonest type consists of a copper rod in contact with a saturated solution of copper sulphate. The rod is attached to the lid of a container (*pot*), which has a porous base of wood, or, more commonly, unglazed earthenware. Contact with the ground is made via the solution that leaks through the base. Some solid copper sulphate should be kept in the pot to ensure saturation and the temptation to 'top up' with fresh water must be resisted, because voltages will be generated if any part of the solution is less than saturated. The high resistance of these electrodes is not generally important because currents should not flow in voltage-measuring circuits.

Despite some theoretical advantages, non-polarising electrodes are seldom used in routine DC work. In induced polarisation surveys it may very occasionally be desirable to use non-polarising electrodes for current injection as well as voltage measurement, but not only does resistance then become a problem but the electrodes deteriorate rapidly due to electrolytic dissolution and deposition of copper.

Copper sulphate solution gets everywhere and rots everything and is therefore not liked. Non-polarising electrodes that use lead/lead chloride or tin/tin chloride combinations are more pleasant to handle, but considerably more expensive. Stainless steel polarises to only a small extent, and also makes efficient current electrodes, and is therefore favoured where the same electrode is sometimes to be part of a current pair and sometimes of a voltage pair.

6.2.3 Cables

The cables used in DC and IP surveys have traditionally been single-core, multi-strand copper wires insulated by plastic or rubber sleeving. Thickness is usually dictated by the need for mechanical strength rather than low

resistance, since contact resistances are nearly always very much higher than cable resistance. Steel reinforcement may be needed for long cables.

In virtually all DC surveys, two of the four cables will be long, and the good practice in cable handling described in Section 1.4.2 is essential if delays are to be avoided. Multicore cables that can be linked to multiple electrodes are becoming increasingly popular because, once the cable has been laid out and connected up, a series of readings with different combinations of current and voltage electrodes can be made using a selector switch.

Power lines can be sources of noise, and it may be necessary to keep the survey cables well away from their obvious or suspected locations. The 50 or 60 Hz power-line frequencies are very different from the 5 to 0.5 Hz frequencies at which current is reversed in most DC and IP surveys, but can affect very sensitive modern instruments, particularly in time-domain IP (see Section 7.3). Happily, the results produced are usually either absurd or non-existent, rather than misleading.

Cables are usually connected to electrodes by crocodile clips, since screw connections can be difficult to use and are easily damaged by careless hammer blows. Clips are, however, easily lost and every member of a field crew should carry at least one spare, a screwdriver and a small pair of pliers.

6.2.4 Generators and transmitters

The instruments that control and measure current in DC and IP surveys are known as *transmitters*. Most deliver square-wave currents, reversing the direction of flow with cycle times of between 0.2 and 2 seconds. The lower limit is set by the need to minimise inductive (electromagnetic) and capacitive effects, the upper by the need for an acceptable rate of coverage. Current levels must be either preset or monitored, since low currents may affect the validity of the results.

Power sources may be dry or rechargeable batteries, or motor generators. The hand-cranked generators (*Meggers*) once used for DC surveys are now very rare. Outputs of several kVA may be needed if current electrodes are more than one or two hundred metres apart, and the generators used are then not only not very portable but also supply power at levels that can be lethal. Stringent precautions must then be observed, not only in handling the electrodes but in ensuring the safety of passers-by and livestock along the whole lengths of the current cables. In at least one case (in Australia), a serious grass fire was caused by faulty insulation on a time-domain IP transmitter cable.

6.2.5   Receivers and detectors

The instruments that measure voltage in DC and IP surveys are known as *receivers*. The primary design requirement is for negligible current to be drawn from the ground. High-sensitivity moving-coil instruments and potentiometric (voltage balancing) circuits were originally used but have now been replaced by units based on field-effect transistors (FETs). Instruments designed exclusively for low-power DC surveys may have transmitter and receiver integrated into a single housing, and readings are then displayed directly in ohms. Separate units are almost inevitable where high power, from large battery packs or motor generators, is required, and in IP surveys, where direct receiver-transmitter interaction must be avoided at all costs.

To allow noise levels to be assessed and SP surveys to be carried out, most receivers allow voltages to be measured even when no current is being supplied. Voltage ranges, numbers of cycles (representing a compromise between speed of coverage and good signal-to-noise ratio) and read-out formats may have to be specified via front-panel key-pads or switches. Usually a displayed reading is updated as each cycle is completed, and the measurement should be accepted only after this has stabilised.

Error conditions, such as low current, low voltage and incorrect or missing connections, may be indicated on displays by numerical codes that are meaningless without the handbook. If all else fails, read it.

6.2.6   Noise in electrical surveys

Since electrodes can, in principle, be positioned on the ground surface to any desired degree of accuracy (although errors are always possible and become more likely as separations increase), and since most modern instruments provide current at one of a number of preset levels so that fluctuations in supply are generally small and unimportant, noise enters the apparent resistivity calculations almost entirely via the voltage measurements. One important factor is voltmeter sensitivity, but there may also be noise from natural voltages, which can vary with time and so be incompletely cancelled by reversing the current flow and averaging, and from induction in the cables. Large separations and long cables should be avoided if possible, but the most effective method of improving a signal/noise ratio is to increase the signal. Modern instruments often provide observers with direct readings of V/I, measured in ohms, and so tend to conceal voltage magnitudes. Small ohm values indicate small voltages, but current levels also have to be taken into account. There are physical limits to the amount of current any given instrument can supply, and it may be necessary to choose arrays with geometric factors that imply large voltages for a given current flow.

The Wenner and two-electrode arrays score more highly in this respect than most others.

The voltages measured using a Schlumberger array are always less than those for a Wenner array of the same overall length, because the separation between the voltage electrodes is always smaller. For the dipole-dipole array, the comparison depends upon the 'n' parameter, but even for $n = 1$ (i.e. for an array very similar to the Wenner in appearance), the signal strength is smaller by a factor of three.

Differences are even greater when the gradient and two-electrode reconnaissance arrays are compared. If the distances to the fixed electrodes are 30 times the dipole separation, the two-electrode voltage signal is more than 150 times the gradient array signal for the same current. However, the shorter gradient-array voltage cable is easier to handle and less vulnerable to inductive noise, and much larger currents can be used with safety because the current electrodes are not moved.

6.3 Resistivity Profiling

Resistivity traverses are used to detect lateral changes. Array parameters are kept constant and the penetration therefore varies only with changes in subsurface layering. Depth information can be obtained from a profile if only two layers, of known and constant resistivity, are involved since each value of apparent resistivity can then be converted into a depth using a two-layer type-curve (Figure 6.7). Such estimates should, however, be checked at regular intervals against the results from expanding-array soundings of the type discussed in Section 6.4.

6.3.1 Targets

The ideal traverse target is a steeply dipping contact between two rock types of very different resistivity, concealed under thin and relatively uniform overburden. Such targets do exist, especially in man-modified environments, but changes in apparent resistivity due to geological changes of interest are often small and must be distinguished from a background due to other geological sources. Gravel lenses in clays, ice lenses in Arctic tundra, and caves in limestone are all much more resistive than their surroundings but tend to be small and rather difficult to detect. Very good conductors, whether sulphide ore bodies or metal pipes, are usually more easily found using the electromagnetic methods described in Chapter 8.

6.3.2 Choice of array

The preferred arrays for resistivity traversing are those that can be most easily moved. The gradient array, which has only two mobile electrodes separated by a small distance and linked by the only moving cable, has much to
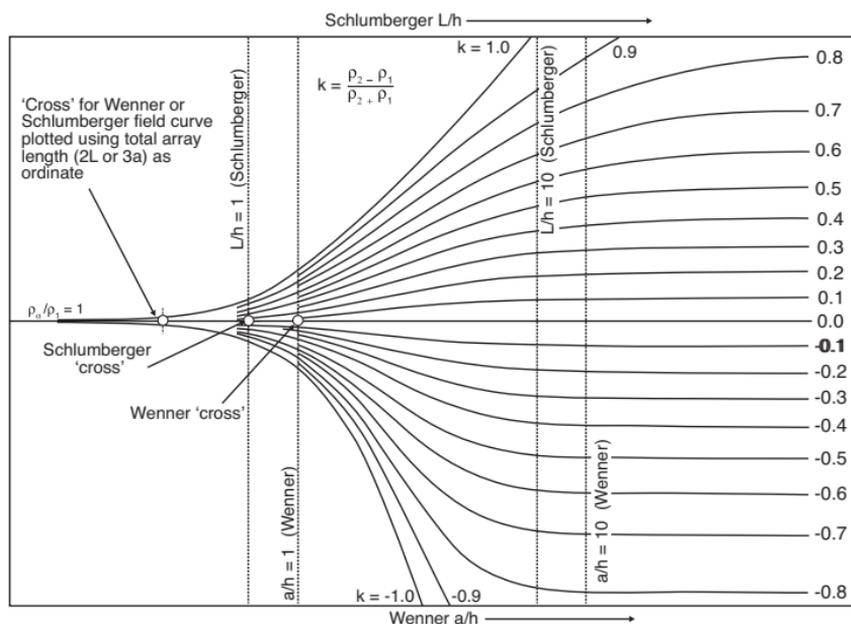
Figure 6.7 *Two-layer apparent resistivity type curves for the Wenner array, plotted on log-log paper. When matched to a field curve obtained over a two-layer Earth, the line* $a/h = 1$ *indicates the depth of the interface and the line* $\rho_\alpha/\rho_1 = 1$ *indicates the resistivity of upper layer. The value of* k *giving the best fit to the field curve allows the value* ρ_2 *of the lower layer resistivity to be calculated. The same curves can be used, to a good approximation, for Schlumberger depth-sounding, with the depth to the interface given by the line* $L/h = 1$.

recommend it. However, the area that can be covered with this array is small unless current is supplied by heavy motor generators. The two-electrode array has therefore now become the array of choice in archaeological work, where target depths are generally small. Care must be taking in handling the long cables to the electrodes 'at infinity', but large numbers of readings can be made very rapidly using a rigid frame on which the two electrodes, and often also the instrument and data logger, are mounted. Many of these frames now incorporate multiple electrodes and provide results for a number of different electrode combinations.

With the Wenner array, all four electrodes are moved, but mistakes are rare since all inter-electrode distances are the same. Entire traverses of cheap metal electrodes can be laid out in advance. Provided that 'direct' or very-low-frequency alternating current is used, so that induction is not a problem,

the work can be speeded up by cutting the cables to the desired lengths and binding them together, or by using purpose-designed multicore cables.

The dipole-dipole array is mainly used in IP work (see Chapter 7), where induction effects must be avoided at all costs. Four electrodes have to be moved and the observed voltages are usually small.

Resistivity anisotropy can be roughly assessed using the square array (Figure 6.2g), but more detail can be obtained by rotating a linear four-electrode array about a central point. A sweep through $180°$ is normally carried out, typically in $10°$ steps. Once a sweep is completed, the array is moved to another pivot point and the process is repeated. Information on fracture trends or water seepage pathways can be obtained using this approach.

6.3.3   Traverse field-notes

Since the array parameters remain constant when traversing, the array type, spacing and orientation, and very often the current settings and voltage ranges, can be noted on page headers. In principle, only station numbers, remarks and V/I readings need be recorded at individual stations, but any changes in current and voltage settings should also be noted since they affect reading reliability.

Notes should be made on changes in soil type, vegetation or topography and on cultivated or populated areas where non-geological effects may be encountered. These observations will usually be the responsibility of the instrument operator but, because local conditions around any of the electrodes might be important, handlers of remote electrodes may have to be involved. Since any note about an individual field point will tend to describe it in relation to the general environment, a general description and sketch map should be included. When using frame-mounted electrodes to obtain rapid, closely spaced readings, the results are usually recorded directly in a data logger and the general description and sketch become all important.

6.3.4   Displaying traverse data

The results of resistivity traversing are most effectively displayed as profiles, which preserve all the features of the original data. Profiles of resistivity and topography can be presented together, along with abbreviated versions of the field notes. Data collected on a number of traverses can be shown by plotting *stacked* profiles on a base map (see Section 1.5.10), but there will usually not then be room for much annotation.

Strike directions of resistive or conductive features are more clearly shown by contours than by stacked profiles. The locations of traverse lines and

data-points should always be shown on contour maps. Where anisotropy is present, maps of the same area that have been produced using arrays in different orientations can be very different.

6.4 Resistivity Depth-Sounding

Resistivity depth-soundings investigate layering, using arrays in which the distances between some or all of the electrodes are increased systematically. Apparent resistivities are plotted against expansion, on log-log paper (Figure 6.7). Although techniques have been developed for interpreting dipping layers, conventional depth-sounding works well only over interfaces that are roughly horizontal.

6.4.1 Choice of array

Since depth-sounding involves expansion about a centre point, the instruments generally stay in one place. Instrument portability is therefore less important than in profiling. The Wenner array is very popular, but for speed and convenience the Schlumberger array, in which only two electrodes are moved, is often preferred. Interpretational literature, computer programs and type curves are widely available for both arrays. The differences between Wenner and Schlumberger curves are usually below the level of observational noise (Figure 6.7).

Array orientation is often constrained by topography; that is, there may be only one direction in which electrodes can be taken a sufficient distance in a straight line. If there is a choice, an array should be expanded parallel to the probable strike direction, to minimise the effect of non-horizontal bedding. It is generally desirable to carry out a second, orthogonal expansion to check for directional effects, even if only a very limited line length can be obtained.

The dipole-dipole and two-electrode arrays are not used for ordinary DC sounding work. Dipole-dipole *depth pseudo-sections,* much used in IP surveys, are discussed in Section 7.5.2.

6.4.2 Using the Schlumberger array

Site selection, extremely important in all sounding work, is particularly critical with the Schlumberger array, which is very sensitive to conditions around the closely spaced inner electrodes. A location where the upper layer is very inhomogeneous is unsuitable for an array centre, and the offset Wenner array (see Section 6.4.3) may therefore be preferred for landfill sites.

The outer electrodes of a Schlumberger array are usually moved in steps that are approximately or accurately logarithmic. The half-spacing sequence 1, 1.5, 2, 3, 5, 7, 10, 15 . . . is convenient, but some interpretation programs
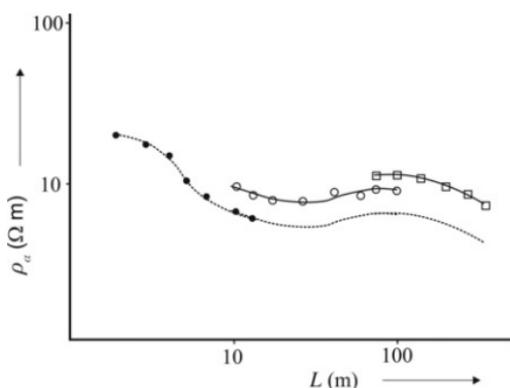
Figure 6.8 *Construction of a complete Schlumberger depth-sounding curve (dashed line) from overlapping segments obtained using different inner-electrode separations. Current electrode separation = 2L.*

require exact logarithmic spacing. The sequences for five and six readings to the decade are 1.58, 2.51, 3.98, 6.31, 10.0, 15.8 . . . and 1.47, 2.15, 3.16, 4.64, 6.81, 10.0, 14.7 . . . respectively. Curves drawn through readings at other separations can be resampled but there are obvious advantages in being able to use the field results directly.

Schlumberger apparent resistivities are usually calculated from the approximate equation of Figure 6.2c, which strictly applies only if the inner electrodes form an ideal dipole of negligible length. Although more accurate apparent resistivities can be obtained using the precise equation, interpretations are not necessarily more reliable, because all type curves are based on the ideal dipole.

As the Schlumberger array is expanded (by moving the outer electrodes), the voltage will eventually become too small to be accurately measured unless the inner electrodes are also moved farther apart. The sounding curve will thus consist of a number of separate segments (Figure 6.8). Even if the ground actually is divided into layers that are perfectly internally homogeneous, the segments will not join smoothly because the approximations made in using the dipole equation change each time the inner-electrode separation is changed. This effect is generally less important than the effect of ground inhomogeneities around the potential electrodes, and the segments may be linked for interpretation by moving them in their entirety parallel to the resistivity axis to form a continuous curve. To do this, overlap readings must be made. Ideally there should be at least three of these at each change, but two are more usual (Figure 6.8), and one is unfortunately the norm.
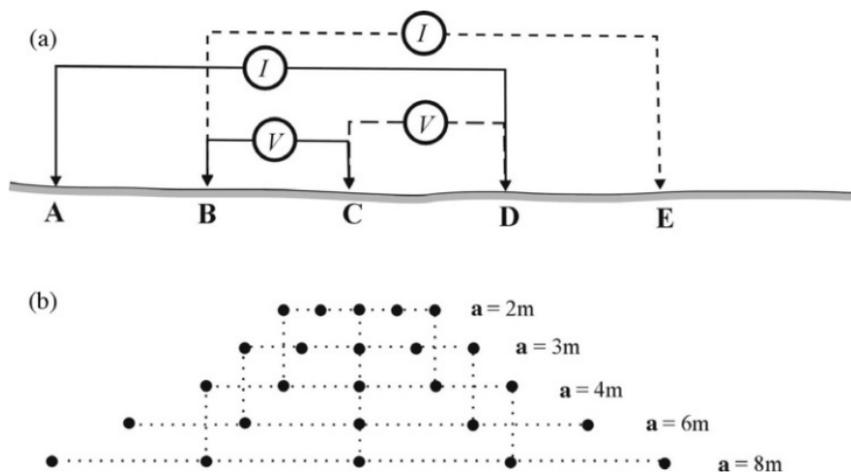
Figure 6.9 *Offset Wenner sounding, with variable inter-electrode spacing* **a**. *(a) Voltage readings are obtained between B and C when current is passed between A and D, and between C and D when current is passed between B and E. (b) Expansion system allowing reuse of electrode positions and efficient operation with multicore cables. The array in (a) is positioned for the 8-m readings in (b).*

6.4.3 Offset Wenner depth-sounding

Schlumberger interpretation is complicated by the segmentation of the sounding curve and by the use of an array that only approximates the conditions assumed in interpretation. With the Wenner array, on the other hand, near-surface conditions differ at all four electrodes for each reading, risking a high noise level. A much smoother sounding curve can be produced with an *offset Wenner* array of five equi-spaced electrodes, only four of which are used for any one reading (Figure 6.9a). Two readings are taken at each expansion and averaged to produce a curve in which local effects are suppressed. The differences between paired readings indicate the magnitudes of those effects.

The use of five electrodes complicates field work, but if expansion is based on doubling the previous spacing (Figure 6.9b), very quick and efficient operation is possible using multicore cables designed for this purpose.

6.4.4 Depth-sounding notebooks

In field notebooks, each sounding should be identified by location, orientation and array type. The general environment should be clearly described and any peculiarities, for example the reasons for the choice of a particular

orientation, should be given. Considerable variations in current strengths and voltage levels are likely, and range-switch settings should be recorded for each reading.

Generally, and particularly if a Schlumberger array is used, operators are able to see the inner electrodes. For comments on the outer electrode positions at large expansions, they must either rely on second-hand reports or personally inspect the whole length of the line.

6.4.5   Presentation of sounding data

The observer often has time while distant electrodes are being moved to calculate and plot apparent resistivities. Minor delays are in any case better than returning with uninterpretable results, and field plotting should be routine. All that is needed is a pocket calculator and a supply of log-log paper. A laptop in the field is often more trouble than it is worth, since all are expensive, most are fragile and few are waterproof.

Simple interpretation can be carried out using two-layer type curves (Figure 6.7) on transparent material. Books of three-layer curves are also available, but a full set of four-layer curves would fill a library. If an exact two-layer fit cannot be found using the type-curves available, a rough interpretation based on segment by segment matching will be the best that can be done. The process is controlled using auxiliary curves to define the allowable positions of the origin of the two-layer curve that is being fitted to the later segments of the field curve (Figure 6.10).

Step-by-step matching was the main interpretation method until about 1980. Computer-based interactive modelling is now possible, even in field camps, and gives more reliable results, but the step-by-step approach is still often used to suggest initial computer models.

6.5   Electrical Resistivity Imaging (ERI)

Electrical resistivity imaging uses arrays of electrodes at multiple separations to generate resistivity–depth cross-sections (pseudo-sections). Although in principle this could be done (laboriously) using basic simple equipment, in practice multiple electrodes are connected by multicore cable, and laptop or palm-top PCs are used to automatically control data collection. Logarithmic expansions are difficult to implement under these conditions, and linear expansion, to relatively small separations, is the norm. The systems are generally designed specifically for shallow surveys (<20 m depth).

6.5.1   Basic ERI

Terminology can be confusing. In Europe, Africa and Australia, 2D electrical resistivity imaging is sometimes referred to as electrical resistivity
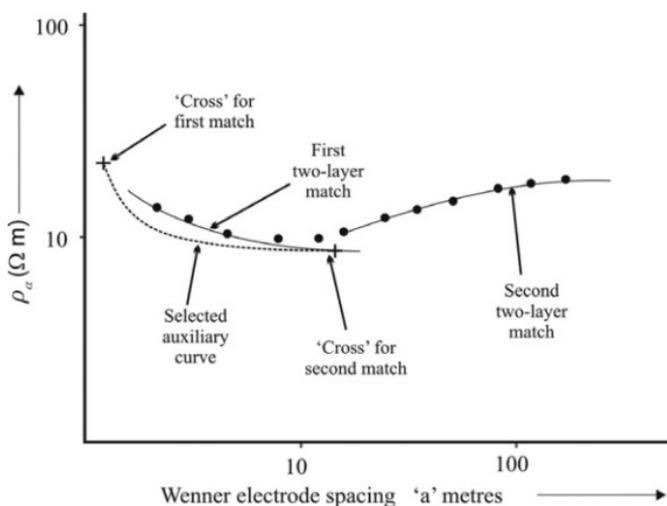
Figure 6.10 *Sequential curve matching. The curve produced by a low-resistivity layer between two layers of higher resistivity is interpreted by two applications of the two-layer curves of the type shown in Figure 6.7. In matching the deeper part of the curve, the intersection of the* $a/h = 1$ *and* $\rho_\alpha/\rho_1 = 1$ *lines (the 'cross') must lie on the line defined by the auxiliary curve.*

tomography (ERT), but ERT in North America is generally used only for measurements made using multiple parallel strings of electrodes, with electrodes selected in pairs from different strings. The data are then treated using tomographic reconstruction techniques. This is consistent with the meaning of tomography in both medical and seismological imaging, and for this reason is the definition favoured by the authors. The term electrical resistivity imaging (ERI) is here used specifically for resistivity data acquired using a single multi-electrode cable.

ERI data are collected along a traverse at a number of different separations that are multiples of a fundamental spacing. As shown in Figure 6.11, the results can be displayed as contoured *pseudo-sections* that give rough visual impressions of the way in which resistivity varies with depth. The data can also be *inverted* using finite element and least-squares inversion methods to produce so-called true resistivity sections with vertical scales in depth rather than electrode separation, giving more realistic images of actual resistivity variations. As a result of the increasing use of these techniques, the inadequacies of simple depth sounding have become much more widely recognised.
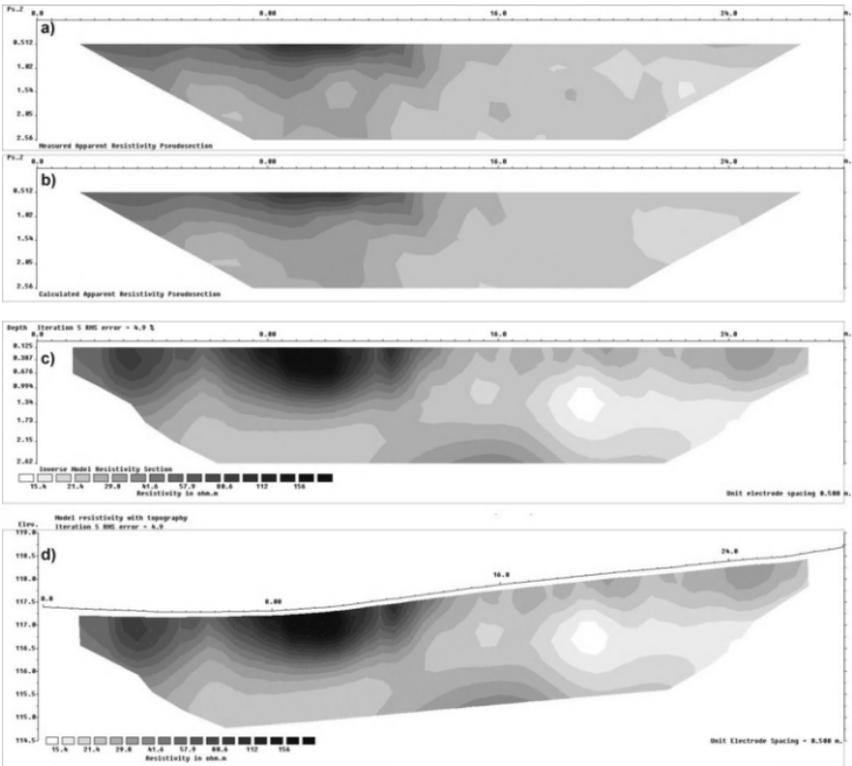
Figure 6.11 *(a) Wenner array apparent resistivity pseudo-section. (b) Calculated apparent resistivity pseudo-section derived from (c), the true resistivity model, also displayed (d) in relation to the actual topography, with a 2:1 vertical exaggeration. Dark colours indicate high resistivity. The high-resistivity region near the break of slope is a former clay pit, infilled with rubble and other industrial waste.*

The maximum depth of an ERI investigation is determined by the fundamental spacing between the electrodes and by the number of electrodes in the array. For a 64-electrode array with an electrode spacing of 2 m, this depth is approximately 20 m, but depends on ground resistivity. At any one site, fewer and fewer points are collected at each 'depth level' as the spacing between the active electrodes is increased, until at the final level only a single reading is obtained. To counteract this effect, the array must be 'rolled-along' the line of investigation. The additional line length needed to get the required coverage at the depth of interest can have a significant impact on productivity, and is an essential factor in survey design.
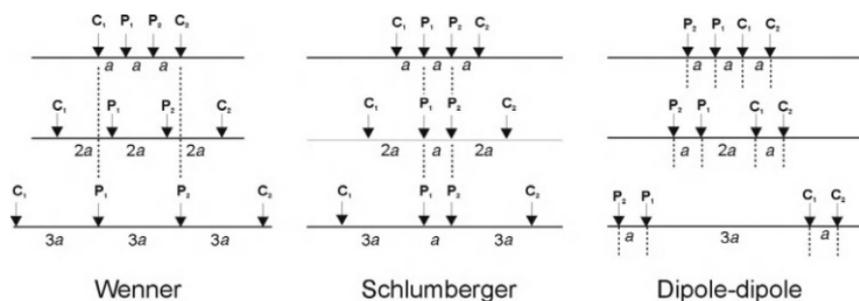
Figure 6.12 *Electrode layout patterns for three popular electrical resistivity imaging (ERI) arrays, illustrating the different switching considerations for multicore cables. With the Wenner spread, the original current electrodes can be used, after two expansions, as voltage electrodes. With the Schlumberger array, the voltage electrodes may not need to be switched at all, whereas switching for the dipole-dipole array must keep voltage and current electrodes in constant-separation pairs but can implement lateral shifting as well as expansion.*

6.5.2 Choice of array and depth of investigation

Identifying the array that best meets the survey objectives is a critical part of ERI survey design, but the basic principles are the same as those that govern the selection of arrays for profiling or depth-sounding – that is, the depth of investigation is largely controlled by the maximum length, the vertical resolution is controlled by the type of array, and lateral resolution is controlled by electrode separation. However, the choice also depends on the type of structure to be mapped, the sensitivity of the resistivity meter and the background noise level, and the ease with which automatic switching can be implemented via a multicore cable (Figure 6.12).

As a general rule, the Wenner array provides better vertical resolution of layers, and the dipole-dipole array better lateral resolution of steep boundaries. Because of the greater signal strength for a given current input, the Wenner array would be most appropriately deployed in noisy conditions and the dipole-dipole in quiet conditions where increased instrument sensitivity can be used to record relatively low signal strengths. The Schlumberger array is a good compromise for targets requiring both vertical and lateral resolution.

All inversion methods are attempts to find models for the subsurface with responses that agree with the measured data. Unsurprisingly, there are likely to be many models that produce calculated apparent resistivity values that agree equally well with a given set of measured values.

Geological features and man-made structures such as buried pits are three-dimensional in nature. Pole-pole, pole-dipole and dipole-dipole arrays are typically used for 3D surveys and continue to achieve good data coverage even near survey grid boundaries. Until recently, the collection and processing of 3D resistivity data was prohibitively costly, but with the development of multichannel resistivity meters to collect data faster and more powerful microcomputers to invert very large data sets, 3D surveys have become almost routine.

The advantages of ERI are only fully realised if purpose-designed equipment is used to reduce the time taken to collect the necessary large volumes of data, and with a modern 61-channel system the ground can be covered about 50 times faster than with a single-channel system.

6.5.3   Topographic effects

Topography can mask features in apparent resistivity data, and inversion can produce resistivity sections with spurious structure in areas of high relief. In these cases it is important to measure the relative heights of electrodes in the field and model the topographic effects.

Even where the topography does not significantly affect the resistivity values, it is important that it be recorded, because it provides additional information about the site being investigated. The pseudo-section in Figure 6.11d, which is referenced to the actual topography, is far more informative than the simpler version in Figure 6.11c.

6.5.4   Time-lapse measurements

Time-lapse ERI has many applications where situations change with time. These include, but are not limited to, seepage pathways in earth dams, leachate egress from landfills, remediation progress at clean-up sites, sink hole activity, seasonal variations in permafrost, movements of fresh water/ salt water interfaces in tidal areas, embankment stability and landslide risk. The standard approach is to separately model data obtained at different times and produce difference plots to highlight changes. The initial inverted model from the first data set obtained at the site may be used for reference or, if appropriate, comparisons may be made with models based on data from areas where the subsurface is, and remains, in a relatively unperturbed state.

Remote monitoring of permanent ERI installations using mobile phone or internet technologies can be a powerful and cost-effective alternative to regular site visits. It is likely that this technique will grow significantly in popularity over the next decade.

6.6 Capacitive Coupling

There have been many attempts to make galvanic resistivity profiling continuous or semi-continuous. Towed electrode arrays rely on devices such as spiked wheels to make contact with the ground, but achieving and maintaining sufficiently low contact resistances is a major challenge for these systems. They function poorly on surfaces that are frozen or very dry. An alternative approach is to use capacitive coupling, in which electrical fields due to alternating currents flowing in insulated conductors cause currents to flow in the ground without direct contact. The *aerials* of such systems can be freely dragged along the ground, either manually or mechanically, and resistivity can be measured continuously.

6.6.1 Capacitive principles

Capacitive coupling relies on the ability of alternating current to pass through a capacitor (Figure 6.13). In capacitively coupled resistivity (CCR), a cable or metal sheet forms one plate of the capacitor while the ground behaves as the other.

In a CCR survey, what would have been the electrodes in a conventional electrical survey are removed from the ground and insulated from it. If the power source is then connected, current will flow only until the electrical potentials produced by the charges on the current electrodes are equal and opposite to those produced by the source. The ability of the system to store charge in this way is termed its electrical *capacity* and is measured in farads.
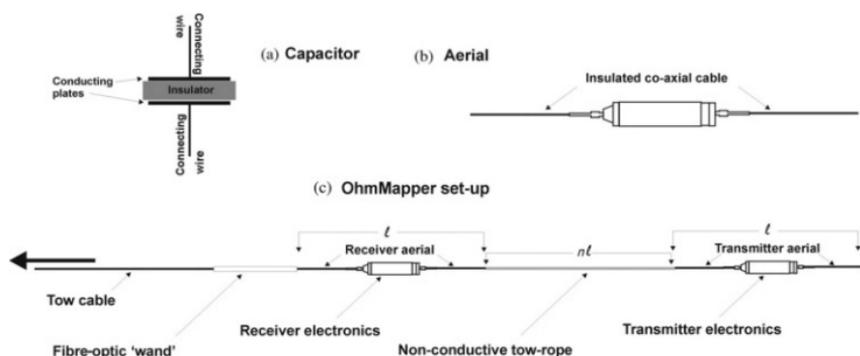

Figure 6.13 *Principle of capacitively coupled resistivity (CCR) operation. (a) Simple parallel plate capacitor capable of storing charge, and passing alternating current. (b) OhmMapper*[TM] *aerial (consisting of insulated co-axial cable) and electronics nacelle. (c) Complete OhmMapper*[TM] *assembly.*

The fact that the current electrodes, even when insulated from the ground, are charged, implies the existence of an electric field that can cause charged particles in the ground to move. This current flow would also be brief, persisting only until equal and opposite reverse potentials were established. If, however, the power-source polarity were then to be reversed, there would be a reverse flow of charge until a new equilibrium was established. An alternating voltage of sufficiently high frequency will thus cause alternating current to flow in the ground, despite the presence of the insulators. This is capacitive coupling. The current flow in the ground is also coupled to the receiver aerial, providing a signal.

6.6.2 Instrumentation

The Geometrics OhmMapper[TM] (Figure 6.14), which is probably the most widely used CCR instrument, can be used to illustrate the basic principles. A 16.5-kHz alternating current is supplied to a dipole aerial that, in standard configurations, is made up of a 2, 5, 10 or 20 m length of cable. The signal is received at a second, similar aerial towed ahead of the first and separated from it by a non-conductive linkage. The linkage is usually an integer

Figure 6.14 *Towing the OhmMapper*[TM]. *Photo reproduced courtesy of Geometrics Inc.*

multiple 'n' of the dipole lengths, so the system geometrically resembles a dipole-dipole galvanic array.

The dipole-dipole system is the industry standard for CCR surveys for practical reasons. Since every bit of unshielded wire contributes to the coupling, current flow must be confined to the aerials. The dipole-dipole configuration allows this, and minimises transmitter-receiver direct coupling, because the transmitter and receiver electronics and power sources are enclosed in nacelles situated at the midpoints of their respective aerials. The signal from the receiver electronics to the data logger is, in the target area close to the aerials, carried by optical pulses in a fibre-optic cable ('optical wand') and not by current flow.

Measurements are obtained at fixed time intervals and the data logger is, when the system is being dragged manually, strapped to the operator's belt, which also takes the strain on the cable. The logger display can show the resistivity profile as it develops, and also several of the preceding profiles. The precautions discussed in Section 1.3.3 regarding all 'continuous' measurements need to be observed to ensure data validity. Even when towed behind a vehicle, speeds should be kept below about 5 km/h.

6.6.3 CCR parameters

The factor used to convert resistance measurements to apparent resistivity for a DC galvanic dipole-dipole array was shown in Figure 6.2e. The equivalent for CCR line-source measurements is very different, being quoted by Geometrics as:

$$K = \frac{l\pi}{\ln\left[\left(\dfrac{b^2}{b^2-1}\right)^{2b}\left(\dfrac{b^2+2b}{(b+1)^2}\right)^{b+2}\left(\dfrac{b^2-2b}{(b-1)^2}\right)^{(b-2)}\right]}$$

where l is the dipole length and $b = 2(n+1)$ and ln indicates that the logarithm is being taken to the base 'e'.

The OhmMapper^{TM} only measures electric fields. The magnetic field is ignored, which is equivalent to assuming that the skin depth (Section 5.2.5) is greater than the transmitter–receiver separation. At the 16.5-kHz frequency, this implies that the separation should be, numerically, less than four times the square root of the resistivity. EM effects will then be negligible and apparent resistivities should be comparable to those obtained galvanically. However, dielectric effects can also influence results. Comparisons of CCR and galvanic apparent resistivities over permafrost have shown that the former may be only 25% of the latter, even though the two are comparable over unfrozen soils. This occurs because the permafrost, an insulator,

effectively increases the distance between the capacitor 'plates'. Similar effects would be observed over other highly resistive surface layers.

Only signal amplitudes are used by the OhmMapper[TM], but there will generally also be phase differences between the currents circulating in the receiving and transmitting aerials, and these could provide additional useful information.

The depth of investigation in a galvanic survey is determined mainly by the total length, L, of the array, and this is true of CCR surveys also. Rough rules of thumb are that if $n \geq 3$, then the investigation depth is equal to $L/5$, but if n is 2, the penetration drops to $L/5.7$, and for $n = 1$ the penetration is only equal to $L/7.2$. At the frequencies and separations characteristic of the OhmMapper[TM], there will often also be some element of skin-depth limitation.

Although the intervals between readings are very small in manually towed surveys, this does not provide equivalently high resolution, since this is determined by aerial length and separation.

6.6.4 Advantages and disadvantages of capacitive coupling

Capacitively coupled resistivity surveys allow resistivity data to be obtained very rapidly, even in areas where ground contact via electrodes would be difficult or impossible. Traverses can be repeated with different separations between the aerials, and commercially available inversion programs allow resistivity cross-sections to be constructed from multi-spaced data. However, as with all geophysical methods, there are problems, both practical and theoretical.

Results of CCR will be reliable only if the coupling between the ground and the aerials remains reasonably constant, so changes due to surface irregularities are sources of noise. These are minimised by weighting the aerials but with obvious disadvantages for one-man operations. Considerable force is needed to drag the system over anything but the smoothest terrain, and especially uphill.

External EM noise sources can affect data quality. However, signal frequencies lie within a narrow band, and noise sources such as power lines or telluric currents actually affect CCR less than galvanic systems. Both are adversely affected by long linear conductors that run parallel to survey lines.

7
SP AND IP

Natural, unidirectional currents flow in the ground and produce voltage (self-potential, or SP) anomalies that can amount to several hundreds of millivolts. They have applications in exploration for massive sulphides, and in some engineering and environmental work.

Artificial currents flowing in the ground can cause some parts of the rock mass to become electrically polarised. The process is analogous to charging a capacitor or a car battery, and both capacitive and electrochemical effects are involved. If the current suddenly ceases, the polarisation cells discharge over periods of several seconds, producing currents, voltages and magnetic fields that can be detected at the surface. Disseminated sulphide minerals can produce large effects of this type, and *induced polarisation* (*IP*) techniques are therefore widely used in exploring for base metals. Arrays are similar to those used in conventional resistivity work. The gradient and dipole-dipole arrays are especially popular (for reconnaissance and detailed work respectively) because current and voltage cables can be widely separated to minimise electromagnetic inductive noise (*cross-talk*). More exotic systems, such as the pole-dipole array, are also used.

7.1 SP Surveys
SP surveys were at one time popular in mineral exploration because of their low cost and simplicity. They are now little used, because some near-surface ore bodies that are readily detected by other electrical methods produce no SP anomaly. In contrast, the popularity of SP methods in detecting water seepage pathways through containment structures is growing.

7.1.1 Origins of natural potentials
Natural potentials of as much as 1.8 V have been observed where alunite weathers to sulphuric acid, but the negative anomalies produced by sulphide ore bodies and graphite are generally less than 500 mV. The conductor should extend from the zone of oxidation near the surface to the reducing environment below the water table, thus providing a low-resistance path for oxidation-reduction currents (Figure 7.1).
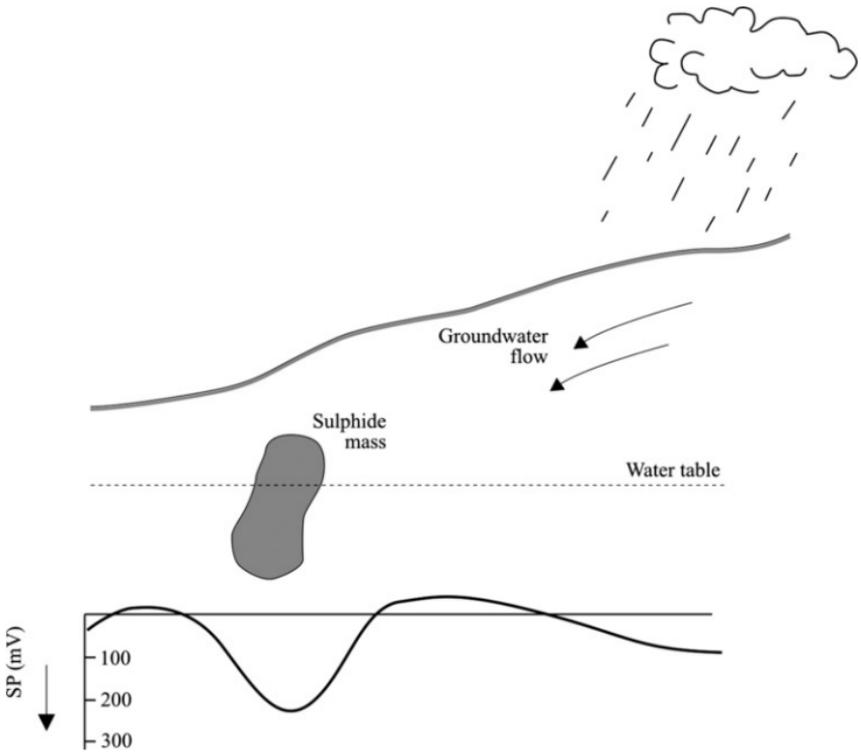
Figure 7.1 *Sources of self-potential (SP) effects. The sulphide mass straddling the water table concentrates the flow of oxidation-reduction currents, producing a negative anomaly at the surface. The downslope flow of groundwater after rain produces a temporary SP, in this case inversely correlated with topography.*

Small potentials, seldom exceeding 100 mV and usually very much less, may accompany ground-water flow and water seepage. Polarity depends on rock composition and on the mobilities and chemical properties of the ions in the pore waters, but usually the region towards which ground water is flowing becomes more electropositive than the source area. These *streaming potentials* are sometimes useful in hydrogeology and in leak detection but can make mineral exploration SP surveys inadvisable for up to a week after heavy rain.

Movements of steam or hot water can explain most of the SPs associated with geothermal systems, but small (<10 mV) voltages, which may be positive or negative, are produced directly by temperature differences. Geothermal SP anomalies tend to be broad (perhaps several kilometres across) and have amplitudes of less than 100 mV, so very high accuracies are needed.

The small alternating currents induced in the Earth by variations in the ionospheric component of the magnetic field and by thunderstorms are discussed in Chapter 9. Only the long-period components of the associated voltages, seldom amounting to more than 5 mV, can be detected by the DC voltmeters used in SP surveys. If, as is very occasionally the case, such voltages are significant, the survey should be repeated at different times of day so that results can be averaged.

7.1.2 SP surveys

Voltmeters used for SP work must have millivolt sensitivity and very high impedance so that the currents drawn from the ground are negligible. Copper/copper-sulphate 'pot' electrodes (see Section 6.2.2) are generally used, linked to the meters by lengths of insulated copper wire.

An SP survey can be carried out by using two electrodes separated by a small constant distance, commonly 5 or 10 m, to measure average field gradients. The method is useful if cable is limited, but errors tend to accumulate and coverage is slow because the voltmeter and both electrodes must be moved for each reading. More commonly, voltages are measured in relation to a fixed base. One electrode and the meter remain at this point and only the second electrode is moved. Sub-bases must be established if the cable is about to run out or if distances become too great for easy communication. Voltages measured from a base and a sub-base can be related provided that the potential difference between the two bases is accurately known.

Figure 7.2 shows how a secondary base can be established. The cable is almost fully extended at field point B, but it is still possible to obtain a reading at the next point, C, using the original base at A. After differences have been measured between A and both B and C, the field electrode is left at C and the base electrode is moved to B. The potential difference between A and B is thus estimated both by direct measurement and by subtracting the B to C voltage from the directly measured A to C voltage. The average difference can be added to values obtained with the base at B to obtain values relative to A.
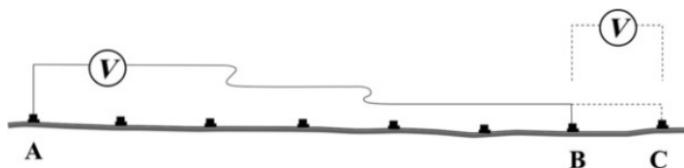

Figure 7.2 *Moving base in an SP survey. The value relative to A at the new base (B) is measured directly and also indirectly by measurements of the voltage at the field point C relative to both bases.*

7.1.3 Errors and precautions

If two estimates of a base/sub-base difference disagree by more than 3 mV, work should be stopped until the reason has been determined. Usually it will be found that copper sulphate solution has either leaked away or become undersaturated. Electrodes should be checked every two to three hours by placing them on the ground a few centimetres apart. The voltage difference should not exceed 1 or 2 mV.

Accumulation of errors in large surveys can be minimised by working in closed and interconnecting loops around each of which the voltages should sum to zero (see Section 1.6.3).

7.2 Polarisation Fundamentals

Induced polarisation (IP) methods are perhaps the most popular of all geophysical techniques in mineral exploration, being the only ones responsive to low-grade disseminated mineralisation. There are two main and incompletely understood mechanisms of rock polarisation and three main ways in which polarisation effects can be measured. In theory the results obtained by the different techniques are equivalent but there are practical differences.

7.2.1 Membrane polarisation

The surfaces of clays and some other platey or fibrous minerals are negatively charged and cause *membrane polarisation* in rocks with small pore spaces. Positive ions in the formation waters in such rocks congregate near the pore walls, forming *electrical double layers*. If an electric field is applied, the positive ion clouds are distorted and negative ions move into them and are trapped, producing concentration gradients that impede current flow. When the applied field is removed, a reverse current flows to restore the original equilibrium.

7.2.2 Electrode polarisation

The static *contact potentials* that exist between metallic conductors and electrolytes were discussed in Section 6.2.2. Additional *over-voltages* are produced whenever currents flow. This *electrode polarisation* occurs not merely at artificial electrodes but wherever grains of electronically conducting minerals are in contact with the ground water. The degree of polarisation is determined by the surface area, rather than the volume, of the conductor present, and polarisation methods are thus exceptionally well suited to exploration for sulphides in disseminated *porphyry* ores. Strong anomalies are also usually associated with massive sulphide mineralisation, because of surrounding disseminated haloes.

Although, for equivalent areas of active surface, electrode polarisation is the stronger mechanism, clays are much more abundant than sulphides and

most observed IP effects are due to membrane polarisation. Changes in ion concentration (e.g. salinity levels) affect both types of polarisation. Non-ionic fluids (e.g. oil contamination) can also change polarisation behaviour.

7.2.3 The square wave in chargeable ground

When a steady current flowing in the ground is suddenly terminated, the voltage V_o between any two grounded electrodes drops abruptly to a small *polarisation voltage V_p* and then declines asymptotically to zero. Similarly, when current is applied to the ground, the measured voltage first rises rapidly and then approaches V_o asymptotically (Figure 7.3). Although in theory V_o is never reached, in practice the difference is not detectable after about a second.

Chargeability is formally defined as the polarisation voltage developed across a unit cube energised by a unit current, and is thus in some ways analogous to magnetic susceptibility. The *apparent chargeability* of an entire rock mass is defined, in terms of the square wave shown in Figure 7.3, as the ratio of V_p to V_o. This is a pure number but, in order to avoid very small values, it is generally multiplied by a thousand and quoted in millivolts per volt.

The V_p/V_o ratio cannot be measured directly because electromagnetic transients are dominant in the first tenth of a second after the original current has ceased to flow, and the practical definition of time-domain chargeability, in terms of the decay voltage at some specified delay time, is only tenuously linked to the theoretical definition. Not only do different instruments use different delays, but it was originally essential and is still quite common to measure an area under the decay curve using integrating circuitry, rather than an instantaneous voltage. The results then depend on the length of the integration period as well as on the delay and are quoted in milliseconds.

7.2.4 Frequency effects

Figure 7.3 also shows that if a current were to be terminated almost immediately after being introduced, a lower apparent resistivity, ρ_{hf}, equal to $2\pi(V_o - V_p)/I$ multiplied by the array geometrical factor, would be calculated. The IP frequency effect is defined as the difference between the 'high frequency' and 'DC' resistivities, divided by the high-frequency value. This is multiplied by 100 to give an easily handled whole number, the *Percent Frequency Effect* (*PFE*). The theoretical relationship between the PFE and the chargeability is given by:

$$M = [PFE]/(100 + [PFE])$$

Figure 7.3 illustrates the derivation of this relationship.

Because of electromagnetic transients, the theoretical PFE cannot be measured and the practical value depends on the frequencies used. To cancel
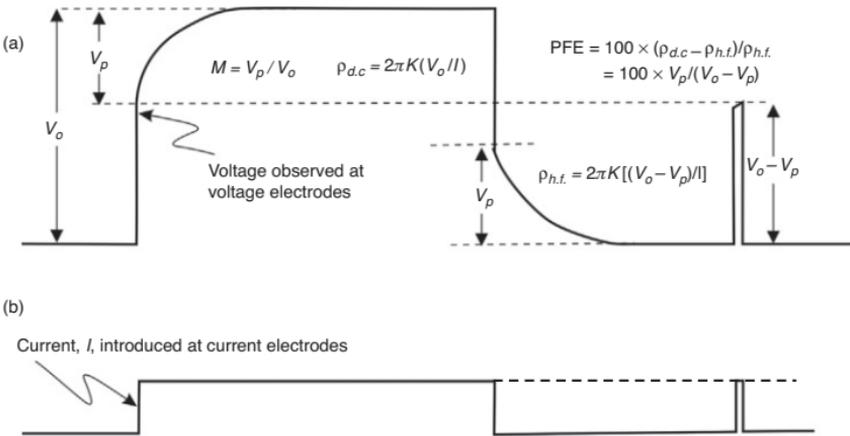
Figure 7.3 *(a) Ground responses to equal-amplitude square-wave and spike impulse currents, for an array with a geometric factor K (see Section 6.1.1). The voltage* V_p *(discussed in the text) is seldom more than a small percentage of the steady-state measured voltage* V_o. *If the current is terminated shortly after being introduced, as in any half-cycle of a high-frequency alternating square wave, the measured voltage barely rises above* $V_o - V_p$. *The 'steady state' and impulse currents introduced at the current electrodes are shown in (b).*

telluric and SP noise, 'DC' measurements are taken with current reversed at intervals of the order of a few seconds, while the 'high' frequencies are usually kept below 10 Hz to minimise electromagnetic induction.

7.2.5 Metal factors
A PFE can be divided by the DC resistivity to give a quantity which, multiplied by 1000, 2000 or 2000π, produces a number of a convenient size known as the *metal factor*. Metal factors emphasise rock volumes that are both polarisable and conductive and may therefore be assumed to have a significant sulphide (or graphite) content. Although this may be useful when searching for massive sulphides, low resistivity is irrelevant and can be actually misleading in exploration for disseminated deposits. As usual when factors that should be considered separately are combined, the result is confusion, not clarification.

7.2.6 Phase
The square-wave current of Figure 7.3 can be resolved by Fourier analysis into sinusoidal components of different amplitudes and frequencies

(see Section 5.2.4). The asymmetry of the voltage curve implies frequency-dependent phase shifts between the applied current and the measured voltage. In *spectral* IP surveys, these shifts are measured, in milliradians, over a range of frequencies.

7.3 Time-Domain IP Surveys

Large primary voltages are needed to produce measurable IP effects. Current electrodes can be plain metal stakes but non-polarising electrodes must be used to detect the few millivolts of transient signal.

7.3.1 Time-domain transmitters

A time-domain transmitter requires a power source, which may be a large motor generator or a rechargeable battery. Voltage levels are usually selectable within a range from 100 to 500 V. Current levels, which may be controlled through a current limiter, must be recorded if apparent resistivities are to be calculated as well as chargeabilities

Current direction is alternated to minimise the effects of natural voltages, and cycle times can generally be varied from 2 to 16 seconds. One second each for polarisation and reading is not usually sufficient for reliable results, while cycle times of more than 8 seconds unreasonably prolong the survey.

7.3.2 Time-domain receivers

A time-domain receiver measures primary voltage and one or more decay voltages or integrations. It may also be possible to record the SP, so that chargeability, resistivity and SP data can be gathered together.

Early *Newmont* receivers integrated from 0.45 to 1.1 seconds after current termination. The SP was first balanced out manually and the primary voltage was then *normalised* by adjusting an amplifier control until a galvanometer needle swung between defined limits. This automatically ratioed V_p to V_o for the M (chargeability) values recorded by a second needle. Experienced operators acquired a 'feel' for the shape of the decay curve from the rates of needle movement and were often able to recognise electromagnetic transients where these persisted into the period used for voltage sampling.

Instruments with dials have now been replaced by instruments with keyboards and display screens, but with purely digital instruments the diagnostic information provided by a moving needle is lost and enough cycles must be observed for statistical reduction of noise effects. Digital systems allow more parameters to be recorded and very short integration periods, equivalent to instantaneous readings, to be used. Natural SPs are now compensated (*backed-off* or *bucked-out*) automatically rather than manually. Memory circuits store data and minimise note-taking.

In a time-domain IP survey, the cycle period of the transmitter must be manually entered into the receiver so that it can lock on to the transmissions without use of a reference cable (which could carry inductive noise). Synchronisation can also be achieved using GPS timing signals (see Section 15.2.5). Cycle times of 4, 8 or 16 seconds are now generally favoured. Changing the cycle time can produce quite large differences in apparent chargeability, even for similar delay times, and chargeabilities recorded by different instruments are only vaguely related.

7.3.3 Decay-curve analysis

With readings taken at several different delay times, curve analysis can be attempted. A method first suggested for use with Huntec receivers assumed that each decay curve was a combination of two exponential decays, corresponding to electrode and membrane polarisations, which could be isolated mathematically. This is far too drastic a simplification of the actual physical processes, and the separation, using a limited number of readings, of two exponential functions that have been added together is in any case virtually impossible in the presence of even small amounts of noise. Nonetheless, research continues into the controls on decay-curve shapes, and chargeabilities should be recorded at as many decay times as are conveniently possible in areas of interesting anomaly. In non-anomalous areas a single value generally suffices.

7.4 Frequency-Domain Surveys

Quite small currents and voltages can be used for resistivity measurements, and frequency-domain transmitters can therefore be lighter and more portable than their time-domain equivalents. Special care has to be taken when positioning cables, to minimise electromagnetic coupling. Coupling is increased by increasing the spacing within or between dipoles, by increasing frequency and by conductive overburden. Unfortunately, field crews may have very limited control over this final factor (although swampy ground can sometimes be avoided), and are also forced to use large electrode separations if deep targets are being sought.

7.4.1 Frequency-domain transmitters

Square waves are commonly used for work in the frequency domain as well as in the time domain, and most modern IP transmitters can be used for both. Measuring resistivity at two frequencies in separate operations is time-consuming and does not allow precise cancellation of random noise, and complex waveforms are therefore used to take effectively simultaneous readings at two different frequencies. Simple square waves may be used if

the receiver can analyse the voltage waveform to extract the high-frequency effects.

7.4.2 Frequency/phase receivers

Sophisticated receivers are needed to analyse waveforms and extract frequency effects from either single- or dual-frequency transmissions, but this sophistication is normally not apparent to an operator recording PFEs from a front panel display.

To measure phase differences for multi-frequency (spectral) IP surveys, a common time reference for transmitter and receiver is essential. Reference cables are operationally inconvenient and could increase inductive coupling, and crystal clocks that could be synchronised at the start of a day's work were used to avoid this problem. These were designed to drift no more than a fraction of a millisecond in 24 hours, but the continual updating possible using GPS signals is now generally preferred.

7.4.3 Phase measurements

A typical spectral IP plot is shown in Figure 7.4. The frequency at which the maximum phase shift occurs is dependent on grain size, being higher for fine-grained conductors. The sharper the peak, the more uniform the grain size. Most attempts to distinguish between different types of IP source are now based on analysis of these spectral curves, since grain size may be correlated with mineral type. However, exploration programmes soon reach the point at which further theoretical analysis of IP curves is less effective than drilling a few holes.

At high frequencies, increases in phase shifts are caused by electromagnetic coupling. Simple *decoupling* calculations involve readings at three different frequencies and assume a quadratic relationship (i.e. $\varphi = A + Bf + Cf^2$) between phase-shift and frequency. The three readings allow this equation to be solved for A, the *zero-frequency* phase shift value. At most survey points only the value of A will be worth recording, but at points that are clearly anomalous an entire phase spectrum, using many more than the three basic frequencies, may be stored for further processing.

7.4.4 Comparison of time- and frequency-domain methods

The relationship between polarisation and current is not precisely linear. This not only limits the extent to which time, frequency and phase measurements can be interrelated, but can also affect comparisons between different surveys of the same type. The effects generally do not exceed a small percentage, but provide yet another reason for the very qualitative nature of most IP interpretation.
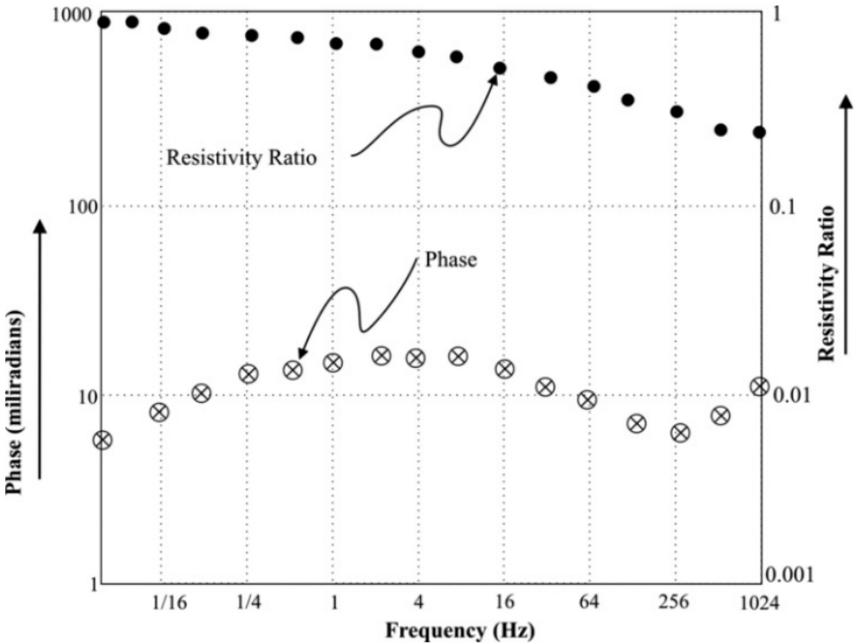
145

Figure 7.4 *Typical plot of induced polarisation (IP) phase and amplitude against frequency.*

The relative merits of time- and frequency-domain IP have long been argued, especially by rival instrument manufacturers. Time-domain surveys are essentially multi-frequency and the shapes of decay curves provide information equivalent to that obtained by measurements at several different frequencies in frequency-domain or phase work. It is, moreover, generally conceded that PFEs and phase shifts are more vulnerable to electromagnetic interference than are time-domain chargeabilities, and that the additional readings needed if correction factors are to be calculated take additional time and demand more sophisticated instruments. However, frequency-domain surveys require smaller currents and voltages and may be preferred as safer and using more portable instruments. The final choice between the two usually depends on personal preference and instrument availability.

7.5 IP Data

The methods used to display IP data vary with the array. Profiles or contour maps are used for gradient arrays, while dipole-dipole and pole-dipole data are almost always presented as pseudo-sections. In all surveys, the spacing

between the voltage electrodes should not be very much greater than the width of the smallest target that would be of interest.

7.5.1 Gradient-array data

Current paths are roughly horizontal in the central areas investigated using gradient-arrays, and chargeable bodies will be horizontally polarised. Profiles can be interpreted by methods analogous to those used for magnetic data, with approximate depths estimated using the techniques of Section 3.5.2.

7.5.2 Dipole-dipole data

Dipole-dipole traverses at a constant n value can be used to construct profiles, but multi-spaced results are almost always displayed as pseudo-sections (Figure 7.5). The relationships between the positions of highs on pseudo-sections and source body locations are even less simple with dipole-dipole than with Wenner arrays (see Section 6.5.1). In particular, the very common *pant's leg* anomaly (Figure 7.5) is usually produced by a near-surface body with little extent in depth, since every measurement made with either the current or the voltage dipole near the body will record high chargeability. Anomaly shapes are thus very dependent on electrode positions, and the directions of apparent dip are not necessarily the directions of dip of the chargeable bodies. Even qualitative interpretation requires considerable experience as well as familiarity with model studies.
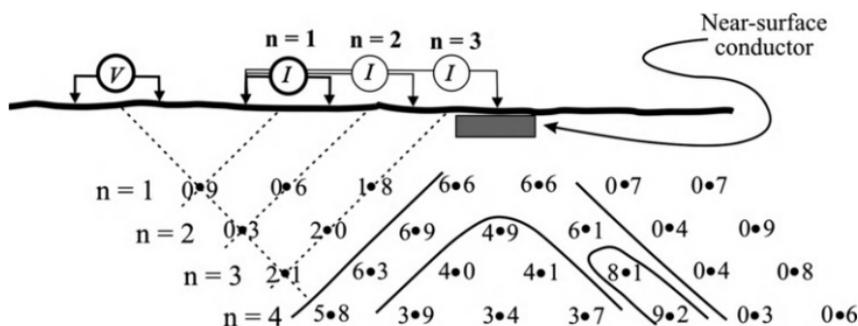
Figure 7.5 *Pseudo-section construction. The three different positions of the current dipole correspond to three different multiples of the basic spacing. Measured values [of induced polarisation (IP) or resistivity] are plotted at the intersections of lines sloping at 45° from the dipole centres. The plotting 'point' often doubles as a decimal point for IP values. The 'pant's leg' anomaly shown is typical of those produced by small, shallow bodies.*

Pseudo-sections are nearly always plotted in relation to horizontal base-lines, even in rugged terrain. Referencing them to topographic profiles (using construction lines similar to those of Figure 7.5 but at 45° to the actual ground surface) has its dangers, since it might be taken as implying much closer correlations with true subsurface distributions of resistivity and chargeability than actually exist. However, steep and varied slopes do influence dipole-dipole results and it is arguably better that they be displayed than ignored.

7.5.3   Negative IPs and masking

Negative IP effects can be caused by power or telephone cables or, as shown by signal contribution sections (Figure 6.5), by lateral inhomogeneities. Layering can also produce negative values, and can conceal deeper sources, most readily if both the surface and target layers are more conductive than the rocks in between. The penetration achieved in such cases may be very small, and total array lengths may need to be ten or more times the desired exploration depth.

Interactions between conduction and charge in the Earth are very complex, and interpreters may need resistivity data that are more reliable than those provided by the dipole-dipole array, which performs poorly in defining layering. A small number of Wenner or Schlumberger expansions, carried out specifically to map resistivity, can be invaluable. Also, any changes in surface conditions that might correlate with changes in surface conductivity should be noted. The response from an ore-body will be quite different beneath a bare rock ridge and an adjacent swamp.

8

ELECTROMAGNETIC METHODS

Electromagnetic (EM) induction, which is a source of noise in resistivity and IP surveys (Chapters 6 and 7), is the basis of a number of geophysical methods. These were originally used mainly in the search for conductive sulphide ores but are now being increasingly used for depth-sounding and area geological mapping. Because a small conductive mass within a poorly conductive environment has a greater effect on induction than on 'DC' resistivity, and also because the responses to conductors of similar size and shape are proportional to their conductivities, discussions of EM methods tend to focus on conductivity (σ), the reciprocal of resistivity, rather than on resistivity itself.

There are two limiting situations. In the one, eddy currents are induced in small conductive bodies embedded in insulators, producing discrete anomalies that provide information on body location and conductivity. In the other, the effects produced at the surface by horizontal currents induced in horizontally layered media can be interpreted in terms of apparent conductivity. Most real situations involve combinations of layered and discrete conductors, making greater demands on interpreters, and sometimes also on field crews.

Wave effects are important only at frequencies above about 10 kHz, and the methods can otherwise be most easily understood in terms of varying current flow in conductors and varying magnetic fields in space. Where the change in the inducing primary magnetic field is produced by the flow of sinusoidal alternating current in a wire or coil, the method is described as continuous wave (CWEM). In transient electromagnetic (TEM) methods, induction arises from the abrupt termination of current flow.

8.1 Two-Coil CW Systems

A current-carrying wire is surrounded by circular, concentric lines of magnetic field. Bent into a small loop, the wire produces a magnetic dipole field (Figure 1.5), which can be varied by alternating the current. This varying magnetic field causes currents to flow in nearby conductors (see Section 5.2.1).
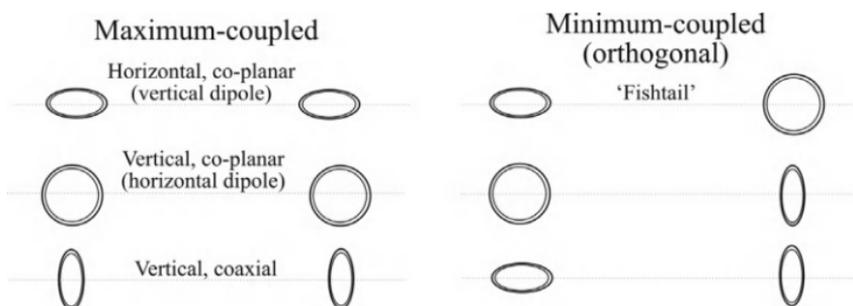
Figure 8.1 *Coil systems for electromagnetic surveys.*

8.1.1 System descriptions

In both CW and TEM surveys, sources are (usually) and receivers are (almost always) wire loops or coils. Small coil sources produce dipole magnetic fields that vary in strength and direction as described in Section 1.2.5. Anomaly amplitudes depend on the coil magnetic moments, which are proportional to the number of turns in the coils, the coil areas and the currents circulating. Anomaly shapes depend on system geometry as well as on the nature of the conductor.

Coils are described as horizontal or vertical according to the plane in which the windings lie. 'Horizontal' coils have vertical axes and are alternatively described as *vertical dipoles*. Systems are also characterised by whether the receiver and transmitter coils are *orthogonal* (at right angles to each other), *co-planar* or *co-axial*, and by whether the coupling between them is a maximum, a minimum or variable (Figure 8.1).

Co-planar and co-axial coils are maximum-coupled, because in both cases the primary flux from the transmitter acts along the axis of the receiver coil. Maximum-coupled systems are affected only slightly by small misalignments but, because a strong in-phase field is detected even in the absence of a conductor, they are very sensitive to changes in coil separation. Orthogonal coils are minimum-coupled. The primary field is not detected and small changes in separation have little effect, but large errors are produced by even slight misalignments. In the field it is easier to maintain a constant coil separation than a constant relative orientation, and this is one reason for favouring maximum coupling.

Dip-angle systems, in which the receiver coil is rotated to determine the dip of the resultant field, were once very popular but are now generally limited to the *shoot-back* instruments used in rugged terrain. Shoot-back receiver and transmitter coils are identical and are linked to electronic units

that can both transmit and receive. Topographic effects are cancelled by measuring and averaging the receiver coil dip angles with first one and then the other coil held horizontal and used as the transmitter.

8.1.2 Slingram

Most ground EM systems use horizontal co-planar coils ('horizontal loops'), usually with a shielded cable carrying a phase-reference signal from transmitter to receiver. The sight of two operators, loaded with bulky apparatus and linked by an umbilical cord, struggling across rough ground and through thick scrub, has provided light entertainment on many surveys. Very sensibly, some instruments allow the reference cable to be also used for voice communication. Fortunately, memory units have not (yet) been added to record the conversations.

The Swedish term *Slingram* is often applied to horizontal-loop systems but without any general agreement as to whether it is the fact that there are two mobile coils, or that they are horizontal and co-planar, or that they are linked by a reference cable, which makes the term applicable.

8.1.3 Response functions

In a Slingram survey, the electromagnetic *response* of a body, *s*, is proportional to its mutual inductances (coupling) with the transmitter and receiver coils (indicated respectively by M_{ts} and M_{sr}) and inversely proportional to its self-inductance (L), which limits eddy current flow. Anomalies are generally expressed as percentages of the primary field and are therefore also inversely proportional to the mutual inductance (M_{tr}) between transmitter and receiver, which determines the primary field strength. These four parameters can be combined in a single *coupling factor*, $M_{ts}M_{sr}/M_{tr}L$.

Anomalies also depend on a *response parameter* that involves frequency, self-inductance (always closely related to the linear dimensions of the body) and resistance. Response curves (Figure 8.2) illustrate both how the in-phase and quadrature responses would vary over targets of different resistivity using a fixed-frequency system, and over a single target as frequency is varied. The quadrature field dominates at low frequencies but is very small at high frequencies, where the distinction between good and merely moderate conductors tends to disappear.

Most single-frequency systems (except, as discussed in Section 8.2, those used for conductivity mapping) operate below 1000 Hz, and even the multi-frequency systems that are now the norm generally work entirely below 5000 Hz. Narrow, poor-quality conductors may produce measurable anomalies only at the highest frequency or not at all.
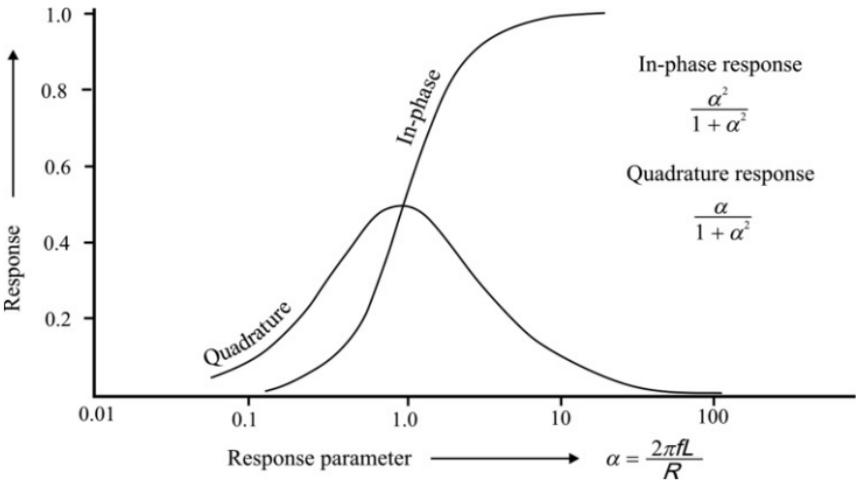
Figure 8.2 *Response of a horizontal-loop electromagnetic (EM) system to a vertical loop target, as a function of the response parameter (α). L is the loop self-inductance,* R *its resistance and* f *is the frequency. Note that the frequency scale is logarithmic. Curves for more complex targets have the same general form, and in-phase/quadrature ratios can therefore be used qualitatively as guides to conductivity.*

8.1.4 Slingram practicalities

The coil separation in a Slingram survey should be adjusted to the desired depth of penetration. The greater the separation, the greater the effective penetration, because the primary field coupling (M_{tr}) is reduced more by the increase than is either M_{ts} or M_{sr} (Figure 8.3). The maximum depth of
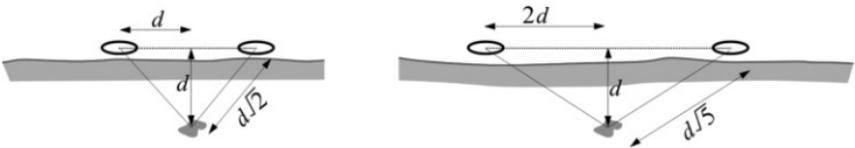
Figure 8.3 *Spacing and penetration. When the two coils are moved apart, the fractional change in distance between them is greater than between either and the conductor at depth. The increased separation thus increases the anomalous field as a percentage of the primary. In the example, in which the coil separation is initially equal to twice the depth of the conducting body, doubling the coil separation increases the coil to target distances by only about 60%.*

investigation of a Slingram system is often quoted as being roughly equal to twice the coil separation, provided that this is less than the skin depth (see Figure 5.5), but this ignores the effects of target size and conductivity and may be unduly optimistic.

Because signals in Slingram surveys are referenced to primary field strengths, the *100% level* must be verified at the start of each day by reading at the standard survey spacing on ground that is level and believed to be non-anomalous. This check has to be carried out even with instruments that have fixed settings for allowable separations, because drift is a continual problem. Finding such ground can be a major problem.

A check must also be made for any leakage of the primary signal into the quadrature channel (*phase-mixing*). Instrument manuals describe how to test for this condition and how to make any necessary adjustments. Receivers and transmitters must, of course, be tuned to the same frequency for sensible readings to be obtained. Care is needed, because a receiver can be seriously damaged if a transmitter tuned to its frequency is operated close by.

Figure 8.4 shows the horizontal-loop anomaly over a thin, steeply dipping conductor. No anomaly is detected by a horizontal receiving coil immediately above the body, because the secondary field there is horizontal. Similarly, there will be no anomaly when the transmitter coil is vertically above the body because no significant eddy currents will be induced. The strongest (negative) secondary fields will be observed when the conductor lies mid-way between the two coils. Coupling depends on target orientation, and lines should be laid out across the expected strike. Oblique intersections produce poorly defined anomalies that may be difficult to interpret.

Readings obtained with mobile transmitter and receiver coils are plotted at the mid-points. This is reasonable because in most cases where relative coil-orientations are fixed, anomaly profiles are symmetrical over symmetrical bodies and are not affected by interchanging receiver and transmitter. Even where this is not completely true, recording mid-points is less likely to lead to confusion than recording either transmitter or receiver coil positions.

In all EM work, care must be taken to record any environmental variations that might affect the results. These include obvious actual conductors and also features such as roads, alongside which artificial conductors are often buried. Power and telephone lines cause special problems because they broadcast noise that, although different in fundamental frequency, may have harmonics that pass through the rejection (*notch*) filters. It is important to check that the rejection frequencies are appropriate (i.e. 60 Hz in most of the Americas and 50 Hz nearly everywhere else).

Ground conditions should also be noted, since variations in overburden conductivity can drastically affect anomaly shapes as well as signal
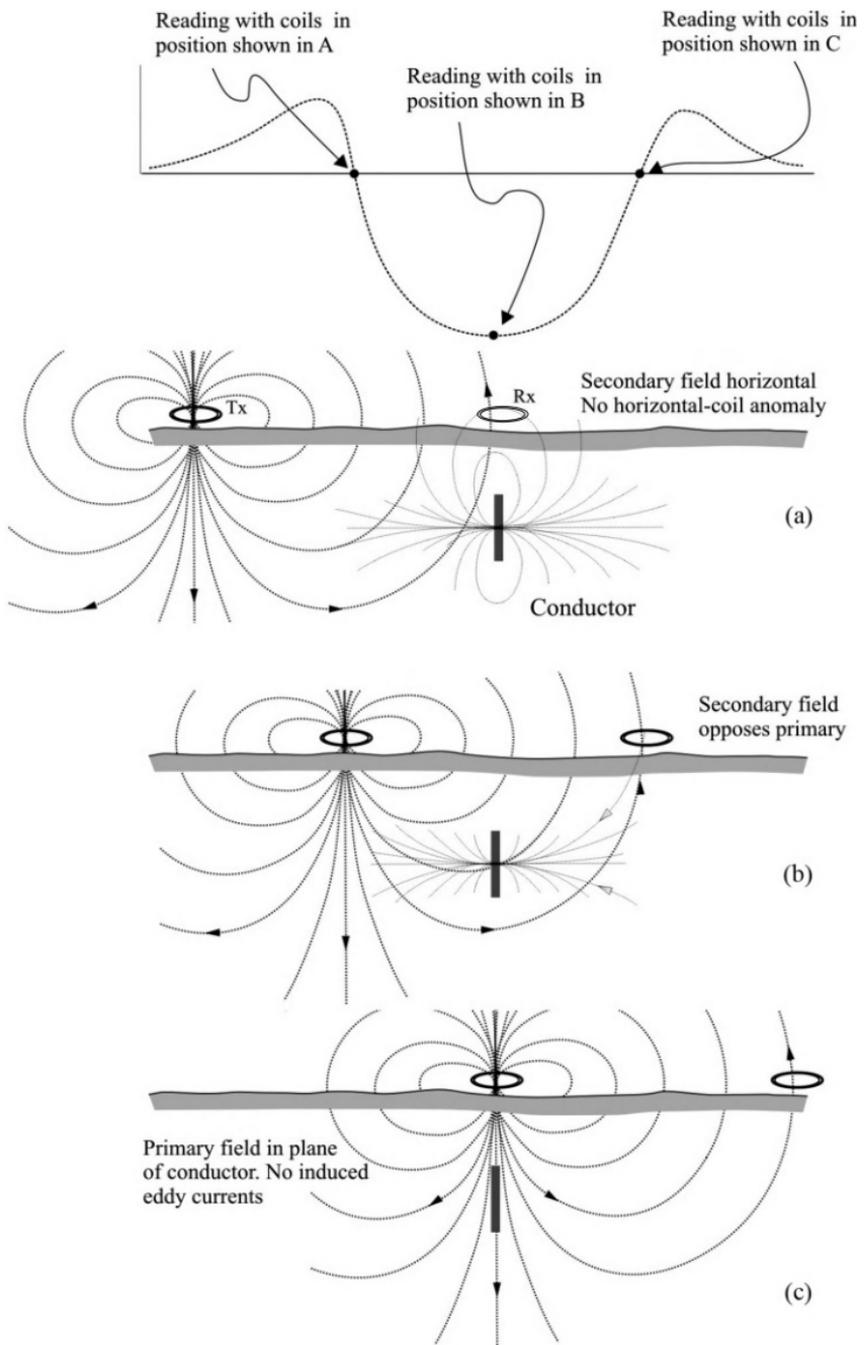
Reading with coils in position shown in A

Reading with coils in position shown in B

Reading with coils in position shown in C

Rx

Tx

Secondary field horizontal
No horizontal-coil anomaly

Conductor

(a)

Secondary field opposes primary

(b)

Primary field in plane of conductor. No induced eddy currents

(c)

Figure 8.4 Horizontal loop anomaly across a steeply dipping conductive sheet. The anomaly width is largely determined by coil separation, not by target width. Over a dipping sheet, the area between a side-lobe and the distance axis would be greater on the down-dip side.

$F(p) = A.(1 - 3\sin^2\phi)/r^3$

$F(r) = 2A.\sin\phi/r^3$

Rx

$F(t) = A.\cos\phi/r^3$

r

Tx

ϕ
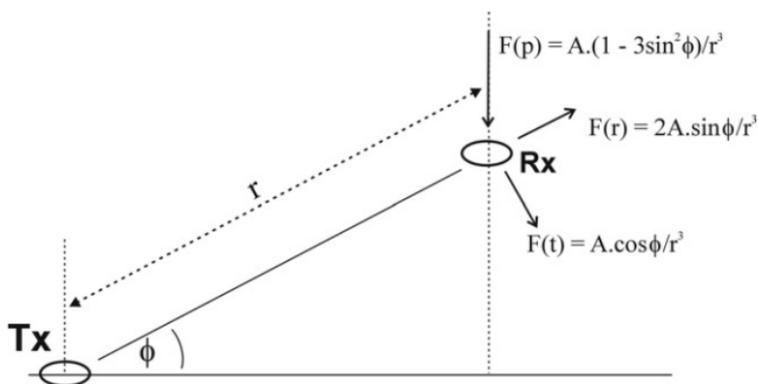
Figure 8.5 *Field components due to a current-carrying loop (Tx) acting as a magnetic dipole source. A is a constant dependent on coil moment and current strength, and* F(r) *and* F(t) *are, respectively, the radial and tangential components. The 'primary' field,* F(p), *measured by a horizontal receiver coil (Rx) is obtained by adding the vertical components of each.*

penetration. In hot, dry countries, salts in the overburden can produce surface conductivities so high that CWEM is ineffective and has been superseded by TEM.

8.1.5 Effects of coil separation

Changes in coupling between transmitter and receiver can produce spurious in-phase anomalies. The field at a distance r from a coil can be described in terms of radial and tangential components $F(r)$ and $F(t)$, as shown in Figure 8.5. The amplitude factor A depends on coil dimensions and current strength.

For co-planar coils, $F(r)$ is zero because φ is zero and the measured field, F, is equal to $F(t)$. The inverse-cube law for dipole sources then implies that, if the separation is actually $r(1 + \varepsilon)$, then:

$$F = F_O/(1 + \varepsilon)^3$$

where F_o is the field strength at the intended spacing. If ε is small, this can be written as:

$$F = F_O(1 - 3\varepsilon)$$

155
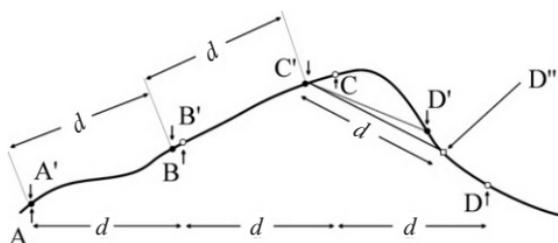
Figure 8.6 *Secant chaining and slope chaining. Down arrows and closed circles show locations of stations separated by intervals of* d *metres measured along slope. Up arrows and open circles show locations of secant-chained stations, separated by* d *metres horizontally. Between C and D the topographic 'wavelength' is less than the station spacing and the straight line separation of C′ from D′ is less than the 'along-slope'* d*. The 'correct' slope position, measured from C′, is at D″.*

Thus, for small errors, the percentage error in the in-phase component is three times the percentage error in distance. Since real anomalies of only a few percent can be important, separations must be kept very constant.

8.1.6 Surveys on slopes

On sloping ground, the distances between survey pegs may be measured either horizontally (*secant chaining*) or along slope (Figure 8.6). If along-slope distances are used in reasonably gentle terrain, coil separations should be constant, but it is difficult to keep coils co-planar without a clear line of sight and simpler to hold them horizontal. The field $F(p)$ along the receiver axis is then equal to the co-planar field for that separation (r in Figure 8.5) multiplied by $(1 - 3\sin^2\varphi)$, where φ is the slope angle (Figure 8.5). The factor $(1 - 3\sin^2\varphi)$ is always less than 1 (coils really are maximum-coupled when co-planar) and the correction multiplier $1/(1 - 3\sin^2\varphi)$ becomes infinite when the slope is $35°$ because the primary field is then horizontal (cf. Figure 1.5).

If secant-chaining is used, the distances along slope between coils are proportional to the secant ($=1/\text{cosine}$) of the slope angle. For truly co-planar coils, the ratio of the 'normal' to the 'slope' field is therefore $\cos^3\varphi$ and the correction factor is $\sec^3\varphi$. However, as noted, ensuring that the coils are, in fact, co-planar is not easy, and often they are held horizontal. The combined correction factor in this case is $\sec^3\varphi/(1 - 3\sin^2\varphi)$ (Figure 8.7).

Separations in rugged terrain can differ from their nominal values if the coil separation is greater than the topographic 'wavelength' (Figure 8.6). Accurate surveying is essential in such areas and field crews may
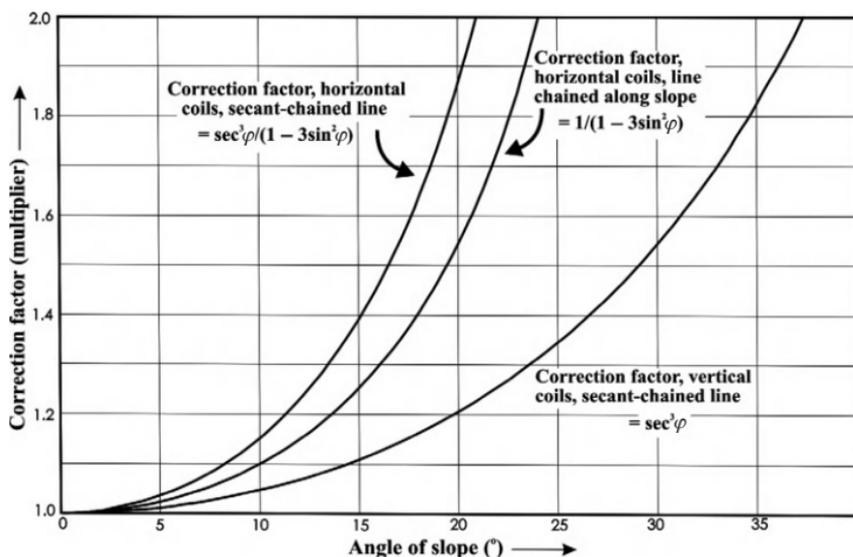
Figure 8.7 *Slope corrections for a two-coil system calibrated for use in co-planar modes. Readings should be multiplied by the appropriate factors.*

need to carry lists of the coil tilts required at each station. Instruments that incorporate tilt meters and communication circuits are virtually essential, but even so errors are depressingly common and noise levels tend to be high. A simpler option is to use the vertical co-planar configuration, which does not rely on a horizontal reference, but penetration is reduced and less interpretational material is available. Moreover, not all systems can be easily used in this way.

8.1.7 Applying the corrections

For any coupling error, whether caused by distance or tilt, the in-phase field that would be observed with no conductors present can be expressed as a percentage of the maximum-coupled field, F_o.

A field calculated to be 92% of F_o because of non-maximum coupling can be converted to 100% *either* by adding 8% *or* by multiplying the actual reading by 100/92. If the reading obtained actually were 92%, these two operations would produce identical results of 100%. If, however, there were a superimposed secondary field (e.g. if the actual reading were 80%), addition would correct only the primary field (converting 80 to 88% and indicating the presence of a 12% anomaly). Multiplication would apply a correction to the secondary field also and would indicate a 13% anomaly. Neither procedure

is actually 'right', but the principles illustrated in Figure 8.3 apply, i.e. the deeper the conductor, the less the effect of a distance error on the secondary field. Since any conductor that can be detected is likely to be quite near the surface, correction by multiplication is generally more satisfactory, but in most circumstances the differences will be trivial.

Coupling errors cause fewer problems if only quadrature fields are observed, since these are anomalous by definition (although, as Figure 8.2 shows, they may be small for very good as well as poor conductors). For small targets, rough corrections can be made using the in-phase multipliers but there is little point in doing this in the field. The detailed problems caused by changes in coupling between a transmitter, a receiver and a third conductor can, thankfully, be left to the interpreter, provided the field notes describe the system configurations and topography precisely.

8.2 CWEM Conductivity Mapping

Provided that the *induction number*, which is equal to the transmitter–receiver spacing divided by the skin depth, is significantly less than unity, Slingram-type systems can be used for rapid conductivity mapping, operating in the region (towards the left of the graphs of Figure 8.2) where the in-phase response is negligible but there is a significant quadrature signal.

8.2.1 Ground conductivity measurement

Low induction numbers imply either low conductivity or low frequency, or both. Under these conditions and regardless of coil orientation, small induced currents, phase shifted by approximately $90°$ with respect to the primary signal, flow horizontally in homogeneous or horizontally stratified ground. Their magnitudes at any depth are determined only by the conductivity at that depth and by the natural attenuation, and not by the current flow at any other depth. The ratio between the in-phase (primary) and quadrature (secondary) magnetic fields is approximately proportional to the conventionally defined apparent resistivity of the ground.

Figure 8.8 shows how current flow varies with depth for plane-horizontal and plane-vertical coil systems at low induction numbers (slightly confusingly, Geonics, the main manufacturer of instruments designed specifically for conductivity mapping, focuses on dipole orientation rather than coil orientation, describing horizontal coils as vertical dipoles). The responses observed with vertical co-planar coils, and hence the apparent conductivity estimates, are dominated by the surface layer, and this is a good reason for preferring the alternative, horizontal coil, configuration. However, the orientations of the secondary-field vectors are such that vertical coils are less sensitive to coil misalignments than are horizontal coils, and it is also
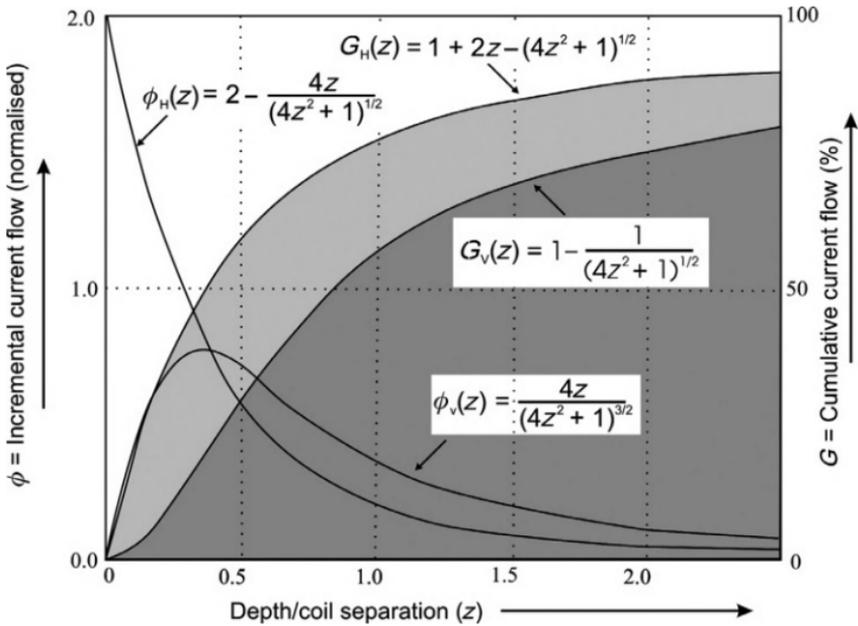
Figure 8.8 *Variation of induced current with depth in homogeneous ground, for co-planar coil systems operating at low induction numbers. 'Filled' curves show total current flowing in the region between the surface and the plane at depth, as a fraction of total current flow. Incremental curves are normalised. Subscripts $_V$ and $_H$ refer to horizontal and vertical* dipoles, *following the Geonics terminology used with the EM31 and EM34.*

less tiring to keep the large coils typical of this work vertical and resting on the ground. Moreover, the low-induction-number approximation holds over a wider range of values when vertical coils are used (see Figure 8.11).

The independence of current flows at different levels implies that the curves of Figure 8.8, which strictly speaking are for homogeneous media, can be used to calculate the theoretical apparent resistivity of a layered medium. Using this principle, layering can to some extent be investigated by raising or lowering the coils within the zero-conductivity air 'layer'. In principle it could also be investigated by using a range of frequencies, but the range would have to be very wide and the low-induction-number condition could not always be satisfied. Methods such as TEM (see Section 8.4) or CSAMT/MT (see Section 9.2) that are inherently *broadband* are preferable for depth-sounding.

8.2.2 Instrumentation

The Geonics EM31 (Figure 8.9) is an example of a co-planar coil instrument that can be used, at some risk to life and limb on difficult sites, by one operator to obtain rapid estimates of apparent resistivity. Man-made conductors such as buried drums and cables may also be detected. The instrument is designed to be used with the coils horizontal, giving, at low induction numbers, a penetration of about 6 m and a radius of investigation of about 3 m with the fixed 3.7 m coil spacing. This compares very favourably with the 20 to 30 m total length of a Wenner array with similar penetration (see Section 6.1.3). Measurements can also be made (although not easily), with the coils vertical, halving the penetration. A shorter, and therefore more manoeuvrable version, the EM31-SH, is only 2 m long and provides better resolution, but only about 4 m of penetration.

Both versions of the EM31 operate at 9.8 kHz. The more powerful, two-person, Geonics EM34-3 uses frequencies of 0.4, 1.6 and 6.4 kHz with spacings of 40, 20 and 10 m respectively. The frequency is quadrupled each time the coil separation is halved, so the induction number remains the same. Coil separation is monitored using the in-phase signal, and phase is monitored using a signal sent along a reference cable. Penetrations are 15, 30

Figure 8.9 *EM31 in use in open country.*

and 60 m for horizontal coils, and 7.5, 15 and 30 m for the more convenient but more surface-influenced vertical coils.

Both the EM31 and the EM34-3 are calibrated to read apparent conductivity directly in mS m^{-1}, but the low-induction-number conditions on which these values are based are not always achieved in practice, since relatively high frequencies are used to ensure that there will be a measurable signal in most ground conditions. Figure 8.10, which is based on Technical Notes available on the Geonics website, shows the ways in which the indicated conductivity would deviate from the true conductivity of a homogeneous half-space as this decreases, for both the EM31 and the EM34-3, and defines the conditions under which the deviations become appreciable. The indicated conductivity is always lower than the true conductivity; i.e. the indicated resistivity is always higher than the true resistivity. For the EM34-3 in the vertical-dipole (horizontal-coil) configuration, the deviation is large (i.e. a factor of 50%) at the high but perfectly possible conductivity of 100 mS m^{-1} (resistivity 10 Ω m) and has become catastrophic by about 500 mS m^{-1} (2 Ω m). The situation is much better for the EM34-3 in the horizontal-dipole (vertical-coil) configuration, and for the EM31 in both configurations. These curves can be used to correct readings obtained over homogeneous ground and, in principle, also those obtained over layered ground. However, the concepts of apparent resistivity and apparent conductivity are themselves somewhat nebulous and in most cases survey results are only used qualitatively (as, e.g., in Figure 8.11).

Selected induction number values are also indicated on Figure 8.10, for both the EM31 and the EM34-3. These demonstrate the difficulty of formulating a general rule governing the induction number value at which the approximation breaks down.

8.3 Fixed-Source Methods

CWEM surveys can be carried out using long wires, set up in fixed locations, instead of coils as sources. There are many possible system geometries but the general principles remain the same.

8.3.1 The Biot–Savart Law

The fields produced by straight, current-carrying wires can be calculated by repeated applications of the *Biot-Savart Law* (Figure 8.12). The application of this relationship to four wires forming a rectangular loop is illustrated in Figure 8.13. If the measurement point is outside the loop, vectors that do not cut any side of the loop have negative signs.

The Slingram anomaly of Figure 8.4 was symmetrical because the receiver and transmitter coil were moved over the body in turn. If the source, whether a coil or a straight wire, were to be fixed and therefore not pass over the
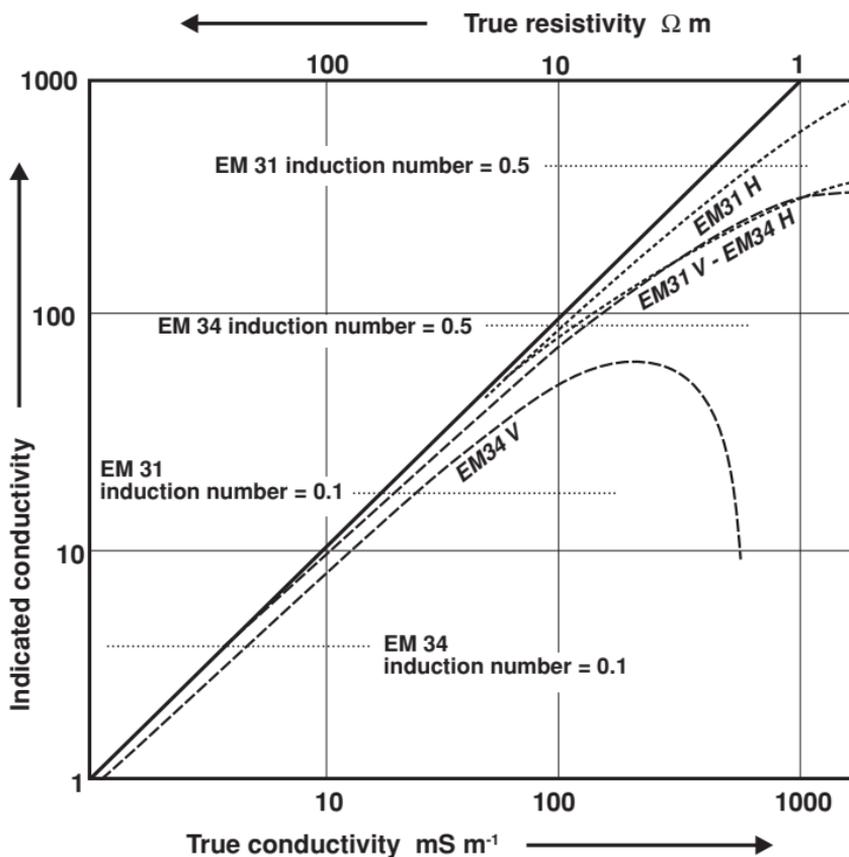
Figure 8.10 *Plots of indicated against true ground conductivity for the EM34-3 (all separations, dashed curves) and the 'long' EM31 (dotted curves), based on Geonics Technical Notes 6 and 8. For both instruments, the deviation from true conductivity is smaller when they are used with the dipoles horizontal (H-curves, i.e. loops vertical) than when they are vertical (V-curves, i.e. loops horizontal). The vertical dipole curve for the EM31 and horizontal dipole curve for the EM34-3 almost coincide. The EM31-SH curve is not shown but deviates least, since it operates, under all conditions, at little more than half the induction number of the EM31. The logarithmic scales tend to disguise the magnitudes of the deviations.*
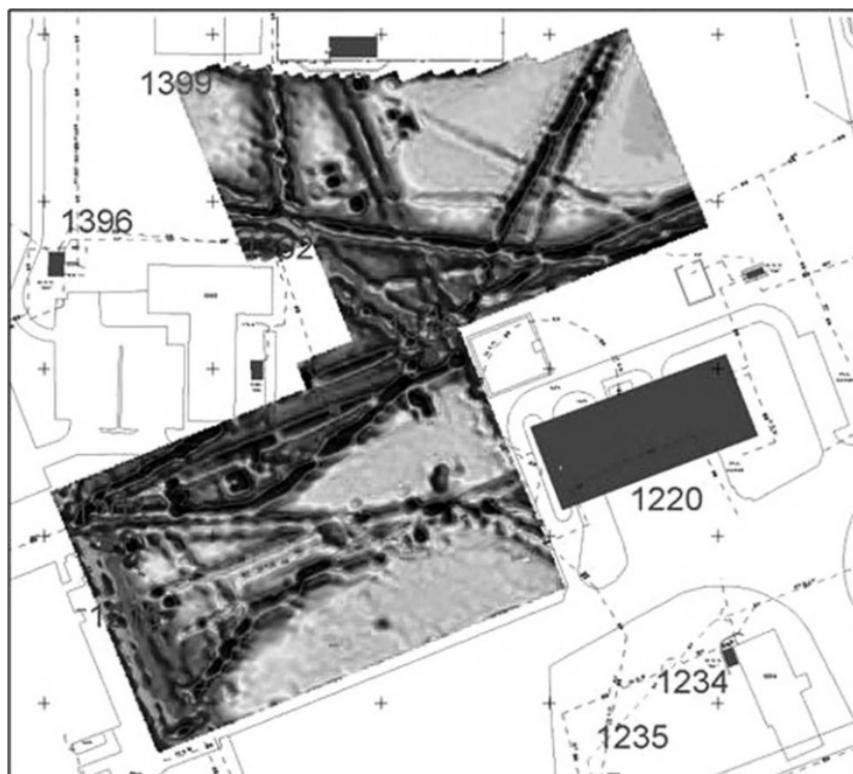
Figure 8.11 *Results of a detailed EM31 resistivity survey, plotted as a shaded-relief image. The aim of the survey was to locate buried pipes and cables in a redevelopment site. Fewer than half the objects actually located had been previously known to the developers.*

conductor, then effectively only a half of the Slingram anomaly would be produced; that is, a steeply dipping body would have a zero response when a horizontal receiver coil was immediately above it and the anomaly would be anti-symmetric. Fixed source systems often measure dip-angles or their tangents, which are the ratios of the vertical to the horizontal fields.

8.3.2 Turam

In the now almost obsolete *Turam* (Swedish, 'two coil') method, the field produced by a fixed extended source was recorded at two receiving coils separated by a distance of the order of 10 m. Anomalies were assessed by calculating *reduced ratios* equal to the actual ratios of the signal amplitudes through the two coils divided by the *normal* ratios that would have been observed over non-conductive terrain. The phase differences between the
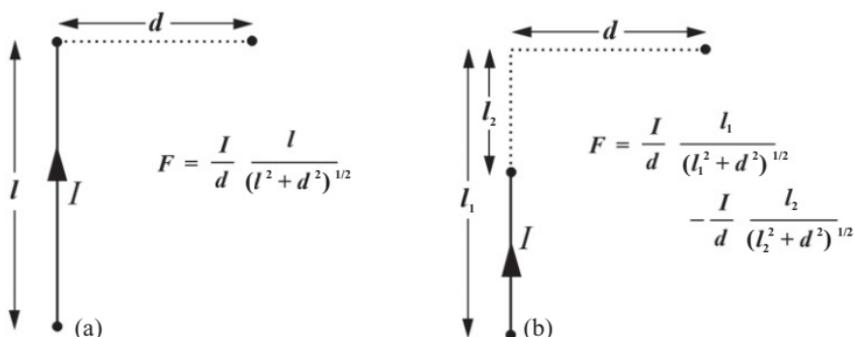
Figure 8.12 *The Biot–Savart Law. (a) Basic geometry for the Biot–Savart equation. Current* I *flows in the wire, length* l, *and the field* (F) *is measured at a point offset a perpendicular distance d from one end of the wire. (b) Application of the Biot–Savart equation when the line from the observation point to the end of the wire is not perpendicular to the wire. The effect is identical to that of two wires positioned so that the equation can be applied, but in one of which (in this case) there is a negative current* (–I).

currents in the two coils were also measured, any non-zero value being anomalous. There was no receiver-transmitter reference cable, but absolute phases and ratios relative to a single base could be calculated if each successive reading was taken with the trailing coil placed in the position just vacated by the leading coil. The method is now seldom used for CWEM surveys, but large fixed sources are common in TEM work (Section 8.4).
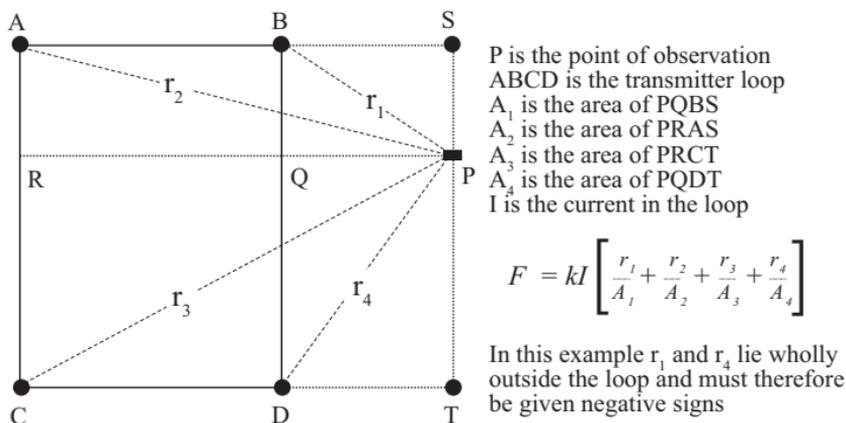
Figure 8.13 *Application of the Biot–Savart Law to a rectangular loop.*

The fields produced by fixed sources can be observed at distances so large that the effects of the ionosphere on signal propagation become significant, and the Biot–Savart Law no longer applies. These *Controlled-Source Audio-Magnetotelluric* (CSAMT) methods are discussed in Chapter 9.

8.4   Transient Electromagnetics

Transient electromagnetic (TEM) systems provide multi-frequency data by repeatedly sampling the transient magnetic fields that persist after a transmitter current has been terminated. A modified square wave of the type shown in Figure 8.14 flows in the transmitter circuits, and transients are induced in the ground on both the upgoing and downgoing *ramps*. Only currents induced during the downgoing ramps are used, since only they can be observed in the absence of the primary field. Ideally, the up-ramp transients should be small and decay quickly, and the up-ramp is often *tapered* to reduce induction. In contrast, on the down-ramp the current flow is terminated as quickly as
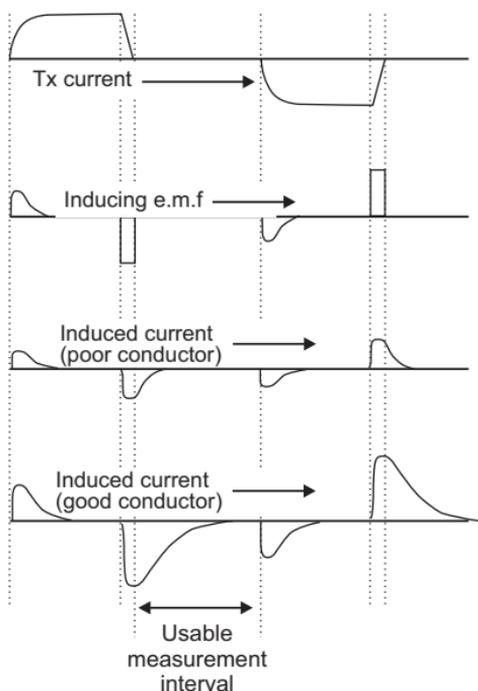
Figure 8.14 Transient electromagnetic (TEM) transmitted and received signals.

possible, to maximise induction. Transmitter self-induction must be minimised, and single-turn loops are therefore preferred to multi-turn coils.

8.4.1 TEM survey parameters

A system in which the primary field is not present when secondary fields are being measured can use very high powers, and TEM systems are popular in areas where overburden conductivities are high and penetration is skin-depth limited. TEM surveys can use the 'Slingram' coil configuration but, since measurements are made when no primary field is present, it is also possible to use the transmitter loop, which may have sides of 100 m or more, to receive the secondary field. Alternatively, a smaller receiver coil can be positioned within the transmitter loop, a technique that can be used in CWEM surveys only with very large loops because of the strong coupling to the primary field.

Most commercial systems can use several different loop configurations that differ in portability. Most can also accommodate a variety of sampling patterns. The decay curves for geological conductors often conform to a power-law equation, $s = t^{-k}$, where s is the signal strength, t the time and k is a constant in the range, for simple layered Earths, from 1.5 to 3.5. The very good conductors typical of many unexploded ordnance (UXO) and engineering surveys are better described by exponential decay curves, $e^{-t/\tau}$, where τ is a constant, while the curves associated with natural electronic conductors such as sulphide orebodies tend not to conform to any simple law. The signals from electronic conductors persist to long delay times, and sampling strategies have to take account of this fact. Typically, samples will be collected at between 20 and 40 discrete times up to maximum delays of, sometimes, more than 100 ms. The sampling points would normally be spaced logarithmically along the decay curve, so that in a mineral exploration application, a quarter of the total would be concentrated in the first half-millisecond. In engineering, site investigation and UXO surveys, sampling would not normally be extended beyond 10 ms after current termination. Signals are routinely averaged over several thousand cycles at each field point.

An alternative approach is provided by the UTEM™ system, in which current with a precisely triangular waveform and a fundamental frequency of between 25 and 100 Hz is circulated in a large rectangular loop. In the absence of ground conductivity, the received signal, proportional to the time-derivative of the magnetic field, is a square wave. Deviations from this in the vertical magnetic and horizontal electric fields are observed by sampling at eight time delays.

In mineral exploration, TEM data are usually presented as profiles for individual delay times (Figure 8.15). The results at short delays are
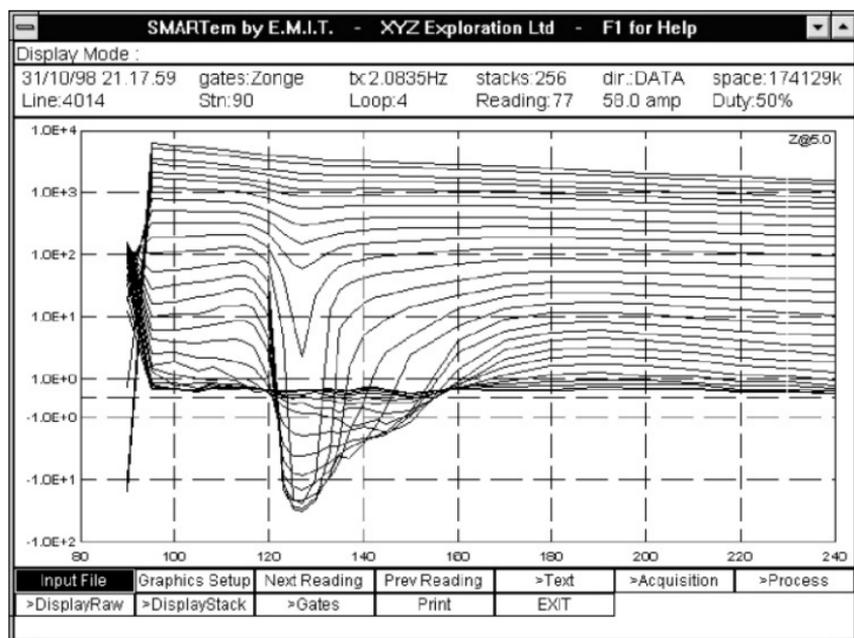
Figure 8.15 *Transient electromagnetic (TEM) system profile display. The image shows a screen dump from a SMARTem survey of a ~200-m profile with a single dominant anomaly at intermediate delay times. Each curve corresponds to a single delay time (gate). Horizontal axis is time and vertical axis is signal strength.*

dominated by eddy-currents in large-volume, relatively poor conductors. These die away quite rapidly, and the later parts of the decay curves are controlled by currents circulating in any good conductors present.

8.4.2 TEM depth-sounding

Transient electromagnetic methods were originally developed to overcome some of the disadvantages of CWEM methods in mineral exploration but are now also being widely used for depth-sounding. In homogeneous or horizontally layered ground, the termination of current flow in the transmitter loop induces a similar current loop or ring in the adjacent ground. This current then decays, inducing a further current ring with a slightly greater radius at a slightly greater depth. The induced current thus progresses through the subsurface as an expanding 'smoke ring' (Figure 8.16), and the associated magnetic fields at progressively later times are determined by current flow (and hence by resistivity) at progressively greater
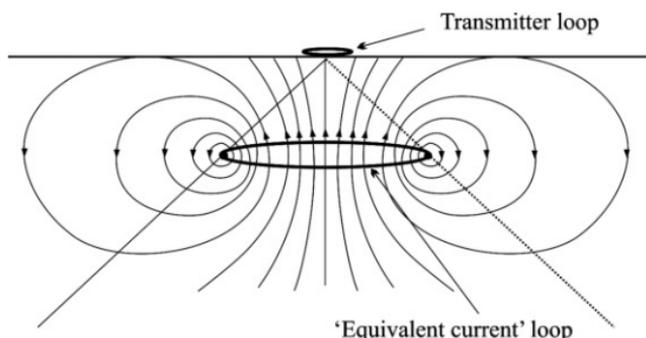
Figure 8.16 *The transient electromagnetic (TEM) 'expanding smoke ring' in a layered medium. The equivalent current loop defines the location of the maximum circulating current at some time after the termination of current flow in the transmitter loop. The slant lines define the cone within which the loop expands. The arrows are on the lines of magnetic field associated with the induced loop.*

depths. The *diffusion depth*, in metres, after a time t is equal to $40\sqrt{(t/\sigma)}$ metres, where σ is in $S\,m^{-1}$. The equivalent depth of investigation, conventionally obtained by dividing by $\sqrt{2}$, is equal to $28\sqrt{(t/\sigma)}$ metres. The apparent resistivity, in ohm-metres, is given by:

$$\rho_\alpha = 6.322 \times 10^{-3}[(I.A_{Tx}.A_{Rx})/V]^{2/3}.t^{-5/3}$$

where A_{Tx} and A_{Rx} are, respectively, the moments (areas multiplied by coil turns) of the transmitter and receiver loops, I is the transmitter current in amps and V is the receiver voltage, in volts.

Transient electromagnetic surveys with 100 metre transmitter loops have been used to obtain estimates of resistivity down to depths of several hundred metres, something requiring arrays several kilometres in length if conventional DC methods are used. However, the large loops necessary for deep penetration are inevitably difficult to handle and can be moved only slowly. Deep penetration may also require generators capable of producing, and loops capable of handling, currents of as much as 50 amperes.

If localised good conductors, whether buried oil drums or sulphide orebodies, are present, the effects of the eddy currents induced in them will dominate the late parts of decay curves and may prevent valid depth-sounding data from being obtained. A relatively minor shift in position of the transmitter and receiver loops may be all that is needed to solve the problem.

8.4.3 TEM and CWEM

Transient electromagnetic and CWEM methods are theoretically equivalent but differ in their advantages and disadvantages because the principal sources of noise are quite different.

Noise in CWEM surveys arises mainly from variations in the coupling between transmitter and receiver coils; hence, the separations and relative orientations of the coils must either be kept constant or, if this is not possible, must be very accurately measured. The receiver circuitry must be very stable, but even so it is difficult to ensure that the initial 100% (for the in-phase channel) and 0% (for the quadrature channel) levels do not drift significantly during the course of a day. The primary field is the ultimate source of all these forms of noise, and the situation can therefore not be improved merely by increasing transmitter power. In TEM surveys, on the other hand, the secondary fields due to ground conductors are measured at times when no primary field exists, and coupling noise is therefore negligible. Moreover, the very sharp termination of transmitter current provides a timing reference that is inherently easier to use than the rather poorly defined maxima or zero-crossings of a sinusoidal wave.

The most important sources of noise in TEM surveys are external natural and artificial field variations. The effect of these can be reduced by increasing the strength of the primary field and by N-fold repetition to achieve a $\sqrt{N}$ improvement in signal to noise ratio (see Section 1.5.7). There are, however, practical limits. Transmitter-loop magnetic moments depend on current strengths and loop areas, neither of which can be increased indefinitely. Safety and generator power, in particular, set fairly strict limits to usable current magnitudes. Multiple repetitions are not a problem in shallow work, where virtually all the useful information is contained in the first few milliseconds of the decay curve, but can be time-consuming in deep work, where measurements may have to be extended to time delays of as much as half a second. Moreover, repetition rates must be adjusted so that power-line noise (which is systematic) is cancelled and not enhanced, and the number of repetitions must be adequate for this purpose. It may take more than ten minutes to obtain satisfactory data at a single point when sounding to depths of more than a hundred metres. This does, of course, compare very favourably with the time needed to obtain soundings to similar depths with Wenner or Schlumberger arrays, but the equipment required is considerably more expensive.

In Slingram surveys, resolution is determined by the spacing between the transmitter and receiver coils. Both CWEM and TEM surveys may use the Slingram configuration but with TEM it is also possible to co-locate the receiver and transmitter coils, giving very high resolving power. TEM is thus much more suitable than CWEM for precisely locating very small

targets. Most modern metal detectors, including 'super metal detectors' such as the Geonics EM63, which was designed specifically for UXO detection at depths of a few metres, use TEM principles.

8.4.4 TEM and IP

TEM superficially resembles the time-domain IP methods discussed in Chapter 7. The most obvious difference is that currents in most IP surveys are injected directly into the ground and not induced by magnetic fields. However, at least one IP method does use induction and a more fundamental difference lies in the time scales.

Time-domain IP systems usually sample only at delays of between 100 ms and 2 s, and so avoid most of the EM effects, which occur mainly at delays of less than 200 ms. There is thus a small region of overlap between the two methods, and some frequency-domain or phase IP units are designed to work over the whole range of frequencies from DC to tens of kilohertz to obtain conductivity spectra (see Section 7.4.2). However, it is usually possible to regard the EM and IP phenomena as completely separate and to avoid working in regions, either of frequency or time delay, where both are significant.

9

REMOTE-SOURCE ELECTROMAGNETICS

Some electromagnetic surveys use sources that are at very large distances from the detectors. The actual separation then becomes irrelevant, and the penetration achieved is limited by frequency and conductivity/resistivity, via the skin-depth (see Section 5.2.5).

Natural electromagnetic radiation covers a broad range of frequencies. Longer wavelengths (lower frequencies) are associated with ionospheric micropulsations, but most of the radiation used in audio-frequency magnetotelluric (AMT or AFMAG surveys, at above 1 Hz) comes from distant thunderstorms. The strength of these waves varies considerably with both time and frequency, and natural radiation surveys may be supplemented or replaced by methods using signals produced by distant artificial sources.

Not all the artificial signals used are transmitted with geophysics in mind. The construction, from about 1960 onwards, of large numbers of high-power military transmitters working in the 15-25-kHz *very low frequency* (VLF) band gave rise to an entire subset of exploration geophysics, and also led to interest in the geophysical use of more conventional (i.e. higher frequency) radio signals. Many of the military transmitters have now been decommissioned and only the radio-frequency systems continue in geophysical use, for applications that require high precision and little penetration. The VLF band is still very important, but is observed within the wider context of the entire natural frequency spectrum.

9.1 Natural Electromagnetic Radiation

Currents produced in the ionosphere by the impact of the protons of the 'solar wind' dominate the terrestrial electromagnetic spectrum at frequencies below 1 Hz. There is a 'dead band' between about 0.5 and 5 Hz (Figure 9.1), beyond which the signals (*sferics*), at frequencies up to about 20 kHz, are mainly generated by thunderstorms in the tropical rainbelts of Africa, South America and Indonesia. A single sferic typically consists of an initial higher frequency (VLF) oscillatory portion with maximum spectral power in the 4–10 kHz range, followed by a long ELF tail with power concentrated below
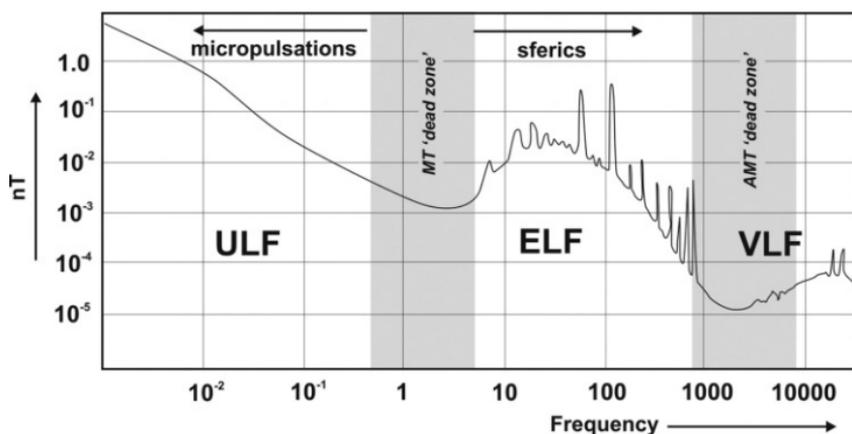
Figure 9.1 *The low-frequency natural radiation spectrum. The peaks in the spectrum above 1 Hz are the Schumann resonances in the Earth-ionosphere waveguide. Their exact location varies in a daily cycle controlled by the height of the base of the ionosphere. The abbreviations VLF, ELF and ULF stand for, respectively, very, extra and ultra low frequency.*

about 1 kHz (Figure 9.2). There is thus a second minimum in the natural radiation spectrum centred at around 2 kHz (Figure 9.1), due in part to the strong attenuation of these frequencies in the Earth-ionosphere waveguide. The time between the peaks increases with distance from the source event and, inevitably, the amplitude decreases. In some cases the VLF component is lacking and the signals are then virtually invisible on long-period plots of the sort shown in Figure 9.2a.
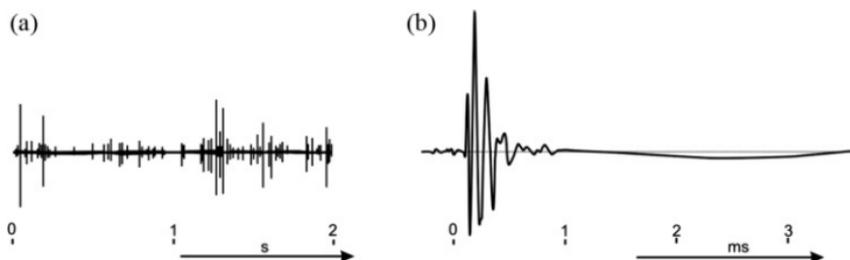
Figure 9.2 *Sferics. (a) Typical natural radiation recorded over a period of two seconds. The individual sferics appear as simple spikes at this scale. (b) A single sferic signal, consisting of an initial very low frequency (VLF) oscillation and a longer extra low frequency (ELF) 'tail'.*

9.1.1 Audio-frequency natural signals

Of all natural electrical phenomena, sferics are the most useful in geophysical exploration, because their wide frequency range allows conductivity to be investigated over depths ranging from a few metres to several kilometres. Modern AMT (audio-magnetotelluric) methods differ from the earlier AFMAG (audio-frequency magnetics, now generally confined to airborne surveys) in that measurements are made of electrical as well as magnetic field components.

Unfortunately, like so many things that come for free, natural signals are not always reliable. Thunderstorm activity varies both seasonally and on a daily basis, with a minimum in most temperate zones during the late morning. Signals are stronger in the evening and at night, times that are not very convenient for survey work. They are also strongest during the summer, which unfortunately is the blackfly season in high northern latitudes. Low signal strengths and high noise levels were major obstacles to acquiring good data in the early days of AMT, and are still factors today, despite the enormous improvements in equipment sensitivity.

9.1.2 AF wave propagation

Regardless of whether they are produced naturally or artificially, electromagnetic waves from distant sources propagate in the spherical wave-guide formed by the surface of the Earth and the base of the ionosphere. Attenuation is determined by the physical properties of the two conducting media and their separation. The properties of the ionosphere, and the height of its base, vary in a 24-hour cycle, since the effect of solar radiation is to lower the base to about 60 km and render it more diffuse. At night, the base rises to about 90 km and is more sharply defined, and attenuation is reduced. The best signals are obtained when the entire path, from thunderstorm to receiver, is in darkness.

If the Earth were a perfect conductor, the electric field vector in the sferic wave would be at right angles to the ground surface, and the magnetic vector would be horizontal and at right angles to the propagation direction. Such a wave would be magnetically *horizontally polarised* and electrically *vertically polarised*. However, the ground has a finite resistance, and sferic waves can penetrate some distance into it. In homogeneous conducting ground, this downgoing wavefront is planar and horizontal, i.e. there is a horizontal electrical field, E_x, as well as a horizontal magnetic field, H_y. The ratio between these two fields is a measure of the subsurface conductivity, governed by the Cagniard equation:

$$\rho = (E_x/H_y)^2/5f$$

Ground resistivity in many sediment-covered areas is a function of depth; i.e., the subsurface is horizontally layered or one-dimensional (1D). The Cagniard equation then gives an *apparent resistivity* and the variation of resistivity with depth can be investigated by using a range of frequencies. The earliest applications of this principle (in *magnetotelluric*, or *MT*, surveys) used the long-wavelength ionospheric radiation to probe the deep crust, but interpretations were often made in terms of oversimplified 1D models, and large errors were common. The more sophisticated 2D and even 3D models made possible by modern computers and justified by the much improved instrument sensitivities produce much better results. Although complicated by the existence of the MT 'dead-band' (Figure 9.3), this form of resistivity depth-sounding is now finding applications in oil industry studies of sedimentary basins.
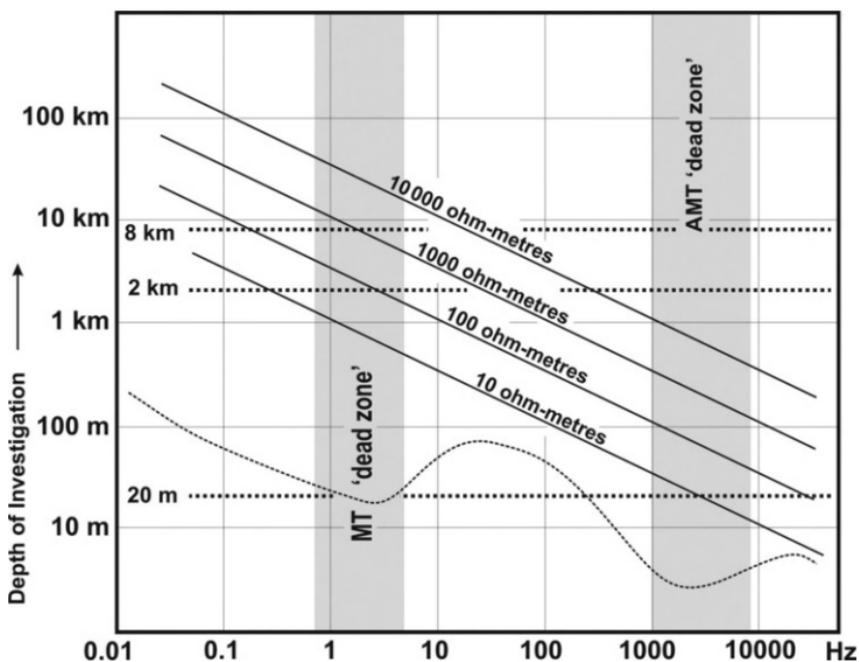
Figure 9.3 *Variation of depth of investigation with frequency. The dotted curve is a simplified version of the curve in Figure 9.1, with the resonance peaks removed. Depth of investigation is defined as equal to the skin-depth divided by $\sqrt{2}$, or about $350\sqrt{(\rho/f)}$. Note the implications of the existence of the audio-magnetotelluric (AMT) dead zone for mineral exploration at depths of 20 m to 2 km, and of the MT dead zone for oil exploration at depths of 2–8 km.*

9.1.3 Magnetic field effects

Advances in instrumentation and in the power of computers to solve complex equations have also led to increased use of AMT in mineral exploration. Almost by definition, mineral deposits are associated with lateral inhomogeneities, and it is no longer sufficient simply to measure the horizontal field components. The eddy currents induced by alternating magnetic fields in subsurface conductors produce secondary magnetic fields at the same frequency as the primary but with different phases, and generally are directed so as to oppose the changes in the primary field. The basis of the technique is most easily demonstrated in terms of the (now virtually obsolete) methods that used single-frequency VLF signals from military transmitters.

Figure 9.4 shows the effect of a thin, sheet-like vertical conductor on a VLF wave. Immediately over the conductor the secondary magnetic field is horizontal and reinforces the primary field, but on either side it has oppositely
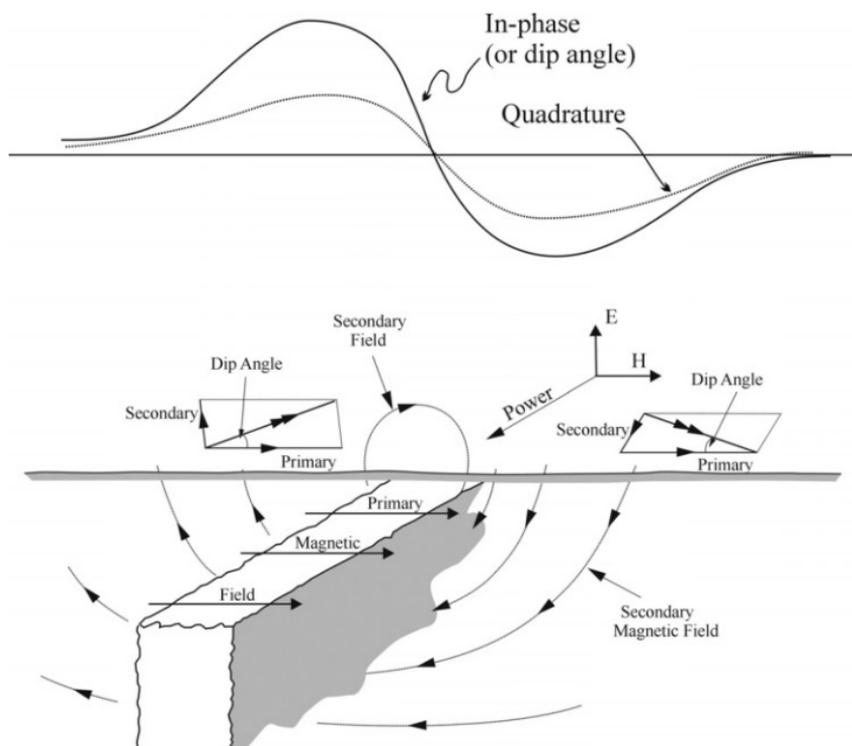
Figure 9.4 *Very low frequency (VLF) magnetic component anomaly over a vertical conducting sheet striking towards the transmitter. An arbitrary sign convention is used to distinguish the maximum from the minimum.*

175
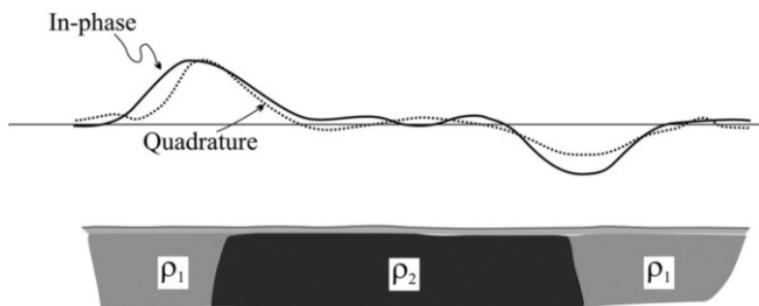
Figure 9.5 *Very low frequency (VLF) magnetic field anomalies at the margins of an extended conductor. Sign convention as for Figure 9.4.*

directed vertical components. Because the field strengths at the military transmitters were, inevitably, unknown, the results were usually presented in the form of profiles of either the dip-angle or the dip-angle tangent (ratio of the vertical to the horizontal field) of the magnetic component. For a vertical sheet, the anomaly would be anti-symmetric, but a dipping sheet would produce an anomaly in which maximum and minimum had different amplitudes. A wider conductor would be characterised by a greater separation between the maximum and minimum and, in the limiting case, a steeply dipping interface would be marked by either a simple maximum or a simple minimum (Figure 9.5).

9.1.4 Coupling

The magnetic-component response shown in Figure 9.4 depends critically on conductor orientation. This is also true in conventional EM surveys, but EM traverses can be laid out at right angles to the probable geological strike, automatically optimising responses. In VLF work the traverse direction was almost irrelevant, the critical parameter being the relationship between the strike of the conductor and the bearing of the transmitting station. A body striking towards the transmitter was said to be *well-coupled*, since it was at right angles to the magnetic vector and induction was maximised. In other cases less current would flow, reducing the strength of the secondary field. If the probable strikes of the conductors in a given area were either variable or unknown, two transmitters, on bearings roughly at right angles and distinguished by slightly different operating frequencies, were used to produce separate VLF maps.

Coupling is more problematic in AMT surveys, because the source location is not only unknown but also variable. The only (partial and not always satisfactory) solution is to record at each station for an extended period and

hope that the primary signals will be reasonably well coupled to the target conductors for at least some of the time.

9.1.5 Elliptical polarisation

The vertical and horizontal fields in Figure 9.4 differ not only in amplitude but also in phase. If there is a 90° difference in phase between fields of the same frequency that are at right angles to each other, the result will be *elliptical polarisation* (Figure 9.6b). However, the secondary field generally has a horizontal component that will combine with the primary field to produce a field that is still horizontal but differs from its components in both phase and magnitude. Combining this with the vertical component of the secondary field produces a resultant that is tilted as well as elliptically polarised. Because the secondary field has a horizontal component, the tangent of the tilt angle is not identical to the ratio of the vertical secondary field to the primary and, because of the tilt, the quadrature component of the vertical secondary field does not define the length of the minor axis of the ellipse. This seems complicated, but dip-angle data are usually only interpreted qualitatively. Where quantitative interpretation is attempted, it is usually based on physical or computer model studies, the results of which can be expressed in terms of any of the quantities measured in the field.
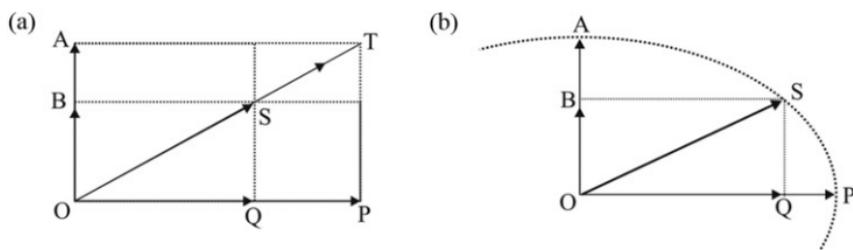
Figure 9.6 *Combination of vertical and horizontal magnetic field vectors. (a) Horizontal and vertical fields in-phase; the vertical vector has its maximum value OA when the horizontal vector has its maximum value OP and the resultant has its maximum value OT. At any other time, as when the vertical vector has value OB and the horizontal vector has value OQ, the resultant OS is directed along OT but with lesser amplitude. All three are zero at the same time. (b) Phase-quadrature: the vertical vector is zero when the horizontal vector has its maximum value OP, and has its maximum value OA when the horizontal value is zero. At other times, represented by OB, OQ and OS, the tip of the resultant describes an ellipse.*

9.1.6 The tipper

Magnetic-component results in most AMT surveys are presented in terms of a measure of the tilt of the polarisation ellipse known as the *tipper*, formally defined by the equation:

$$H_z = T_x \cdot H_x + T_y \cdot H_y$$

where H_z is the vertical magnetic field component at the selected frequency, H_x and H_y are the field components in any two horizontal directions at right angles to each other, and T_x and T_y are the corresponding tipper components. All the quantities in this equation are *complex*, involving the square root of -1 (see Section 5.2.3), and the expression above therefore represents two equations rather than just one. They can be solved by squaring and adding. Conventionally, the positive root is always taken so that, unlike the dip-angle, the tipper is always positive.

Two steeply dipping conductors close to each other produce a tipper anomaly similar to the sum of the anomalies that would have been produced by each body individually. Where, however, one of the bodies is steeply dipping and the other flat-lying, the results are more difficult to anticipate. One important case of this type is provided by the presence of conductive overburden, which influences, and can actually reverse, the phase of the secondary field.

9.1.7 MT practicalities

Magnetotelluric and AMT surveys require only receivers, not transmitters, and in many cases these are multi-purpose instruments that can equally well be used for resistivity, IP, CWEM and TEM surveys. In principle, three electrical and three magnetic field components could be measured at each site, but the E2-H3 configuration (Figure 9.7) is more common. The vertical electric field is the most difficult to measure and is frequently ignored, but the Barringer E-Phase airborne VLF system, which was operated successfully during the 1970s, used this component as a phase reference, and its use may again become fashionable in the future. The comparatively long time needed to collect stable AMT data has led to the use of multiple receivers, and the cost of doing so can be kept down by 'slaving' several relatively simple units with reduced numbers of recording channels to a single full-channel set.

It is often assumed that the vertical magnetic field is more stable than the other components, and can therefore be measured at more widely spaced points. Wherever electrical fields are to be measured, the sensors are pairs of grounded porous-pot (non-polarising) electrodes (see Section 6.2.2) separated by distances of 50–100 m. Magnetic sensors are SQUID
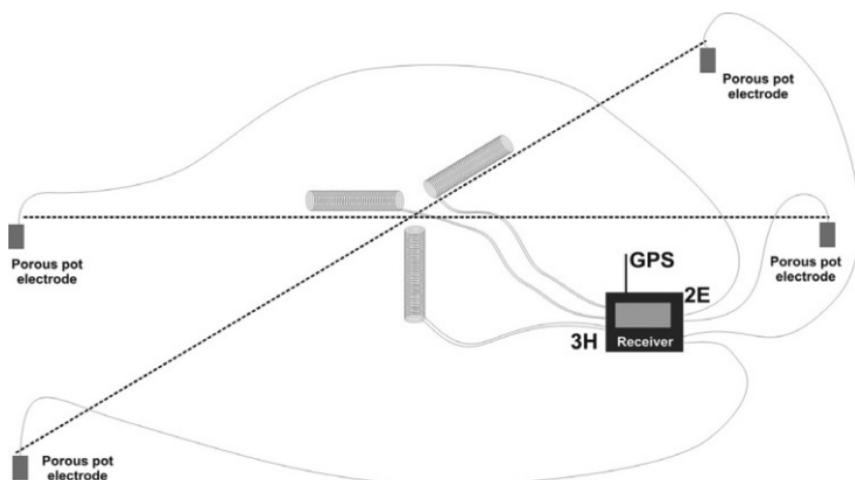
Figure 9.7 *Typical field set-up for an E2-H3 magnetotelluric survey.*

(superconducting) magnetometers or, more commonly, induction coils wound on ferrite cores 1–2 m long and 10–20 cm in diameter. Because propagation directions are both variable and unknown, all horizontal field measurements require two sensors deployed at right angles to each other. The multi-directional characteristics of the natural field are sometimes cited as an MT 'advantage', removing strike-dependency, but in practice all the events recorded during a typical measurement interval are likely to come from roughly the same region of thunderstorm activity. A second survey, at a time when the main activity is in a different province (e.g. in Africa, as opposed to Southeast Asia) might produce a somewhat different map.

One of the most difficult decisions required in MT and AMT work concerns the length of time needed for an individual measurement. A typical extended record of magnetic field variations shows a pattern of long-wavelength oscillations on which are superimposed numerous 'spike' events (Figure 9.2). It is these spikes that contain most of the AMT signal, and in a 20-minute period their number may range from only a few dozen to more than 1000. The recording time, which must be long enough to record a sufficient number, will vary not only with locality but also, as noted in Section 9.1.1, with season and time of day. Half an hour would be considered a minimum in many circumstances, since the high noise levels demand high stacking ratios. It is the relatively long time required for each individual reading that has encouraged the use of multiple recorders, and an ideal way of using these would be to position them in daylight and record overnight.

However, such slow progress would be acceptable only in exceptional circumstances.

9.2 Controlled-Source Audio-Magnetotellurics (CSAMT)

The problems associated with the use of natural fields have led to the development of methods whereby similar but more stable signals are generated artificially. One application of controlled sources has been to 'fill-in' the dead-band gaps in the natural spectrum (Figure 9.1), which can be critical in some applications.

9.2.1 CSAMT principles

The source for a CSAMT survey is usually a long (2-km or more) grounded wire in which current is 'swept' through a range of frequencies, which may extend from as little as 0.1 Hz to as much as 100 kHz. Such signals are now routinely available from multi-purpose transmitters that can equally well be used for CWEM and TEM surveys, and often also for frequency-domain and time-domain IP. High powers are needed, because CSAMT measurements are made at greater distances from the sources than in other types of electrical surveys. Close to a source, the Biot–Savart equations (see Figure 8.12) apply, but transmission effects dominate at the *far-field* distances of, generally, a few kilometres and the wavefront in the subsurface can then be treated as planar and horizontal. Because the source direction is known, only the horizontal electrical field parallel to the source wire (E_x) and the horizontal magnetic field at right angles to it (H_y) need be measured to determine the Cagniard resistivity (see Section 9.1.2). However, the traditional AMT orthogonal detector arrays are needed if, as recommended with the Geometrics EH4 Stratagem, a dual-loop antenna is used to obtain full tensor resistivity information.

The sensors used in CSAMT surveys are similar to those used for MT and AMT. Since the magnetic field from a long transmitter wire laid out parallel to the regional strike usually varies only slowly over any area, reconnaissance surveys are often made using measurements at up to ten electric dipoles for every magnetic measurement.

The far field for CSAMT measurements is commonly considered to begin at a distance of three skin depths from a long-wire source, and is therefore frequency dependent. On a single sounding plot, the onset of intermediate-field conditions can usually be recognised by an implausibly steep gradient in the sounding curve (Figure 9.8). The effective depth of investigation is then usually equal to somewhere between a quarter and a fifth of the distance from the source. This simple relationship can be used in planning surveys, but plans may have to be modified in the light of actual field conditions. In principle, the Cagniard equation should not be used in intermediate and
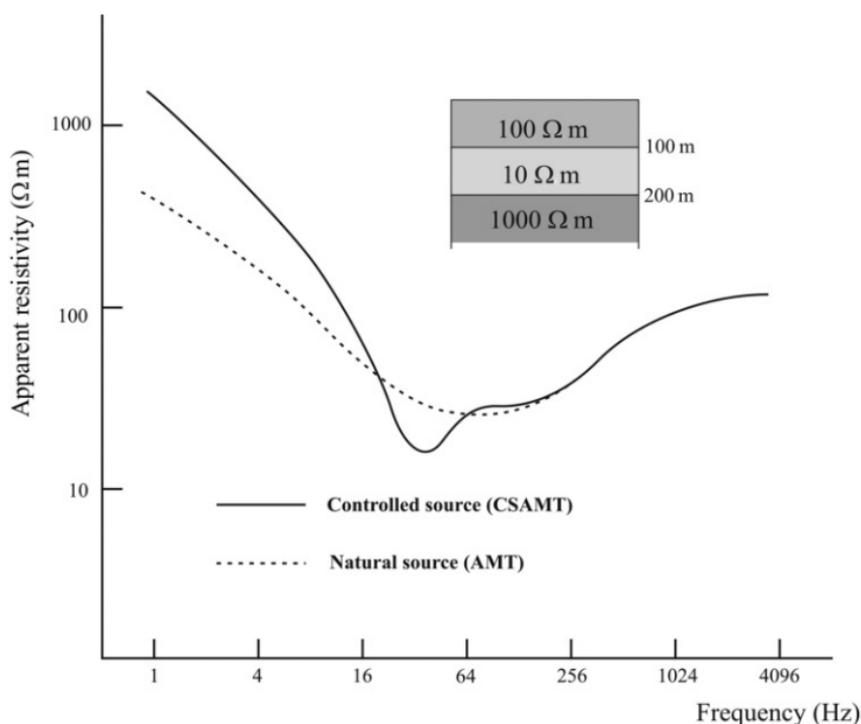
Figure 9.8 *Comparison of the results of audio-magnetotelluric (AMT) and controlled-source audio-magnetotelluric (CSAMT) soundings over a simple layered Earth (overburden 100 m thick with a resistivity of 100 Ω m overlying a 100-m thick low-resistivity layer resting on bedrock with 1000-Ω m resistivity). The AMT signal generates a plane wavefront at all frequencies but the CSAMT wavefront, from a grounded wire 2 km long and 8 km away, ceases to be effectively planar at a frequency of between 100 and 200 Hz. A similar transition from far-field to near-field conditions would be observed if the distance between source and receiver were gradually decreased over homogeneous ground and measurements were made at a single frequency. Illustration adapted from an original drawing by Kenneth Zonge.*

near-field interpretation, but quality control in the field is usually carried out using only the far-field approximation.

9.2.2 CSAMT data
In addition to the ratios of E_x to H_y, their phase difference (the *impedance phase*) is measured in most CSAMT surveys. Even in the field it is a simple
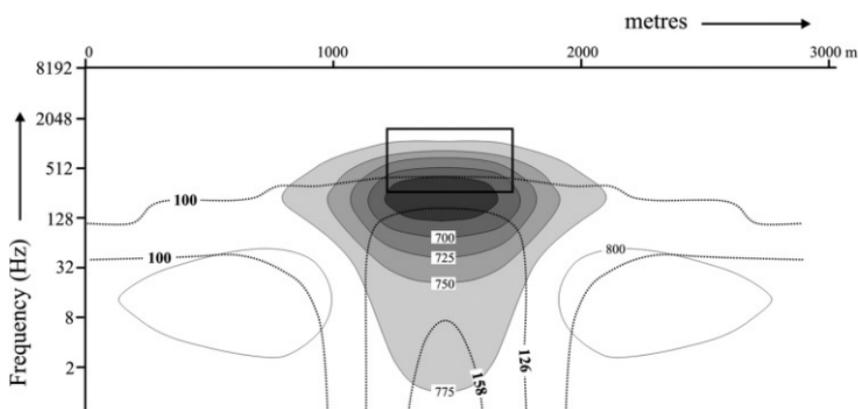
Figure 9.9 *Audio-magnetotelluric (AMT) response of a resistive (5000-Ω m) prism buried in a 100 Ω m medium; solid contours and shading map* E_x–H_y *phase differences ('impedance-phase') in milliradians (contour interval 25 milliradians). Dotted lines are apparent resistivity, in ohm-metres. The vertical scale is in frequency, not depth, and the prism (black solid outline) could be made to coincide with the phase anomaly peak by an arbitrary scale adjustment. In contrast to the apparent resistivity contours, the phase anomaly indicates a body with limited extent in depth. Illustration adapted from an original drawing by Kenneth Zonge.*

matter to plot Cagniard resistivity and phase-difference against frequency. The variation at a single point may be illustrated by a curve (Figure 9.8) but entire traverses can be studied using pseudo-sections (Figure 9.9). Programs are available that can be run on a laptop PC to carry out 1D (simple horizontal layered) and 2D inversions of Cagniard resistivities to provide estimates of resistivity layering. However, the fact that the depth of investigation is itself dependent on resistivity implies some degree of circularity in the calculations, and the modelling process is inherently ambiguous. To estimate the layer resistivity at a given depth, the average resistivities must be obtained down to at least three times that depth, and deep penetration surveys are possible only in resistive terrains.

Phase differences come into their own in investigating small sources. It may, as in the model study of Figure 9.9, be possible to see both the top and the base of a buried source using phase, even though only the top is visible on the corresponding resistivity pseudo-sections.

9.2.3 CSAMT practicalities
The use of controlled sources eliminates some of the problems associated with natural fields but introduces others. Very high currents are required if

long-wire sources are to generate sufficiently strong signals at the distances required by the far-field approximation, and it is seldom easy to find sites where wires carrying many amperes of current can be laid out on the ground safely (or even at all) over distances of kilometres. Even where this can be done, topographic irregularities may create significant distortions in the signal. Closed loop sources can be considerably smaller but require currents that are even larger (by factors of as much as ten) than those needed for line sources.

In the far-field, magnetic and electric field strengths decrease roughly as the inverse cube of distance from the transmitter. The dimensions of the receiver 'dipoles' affect target resolution as well as signal strength but, because signal strength is normally low at the distances at which AMT equations can be applied to CSAMT data, it may be impractical to use dipoles less than 20 m long. Even so, noise may exceed signal by a factor of ten or more, and observations must be extended over periods long enough to permit very high folds of stacking.

With every CSAMT system there is a distance, which is to a first approximation independent of both frequency and resistivity, beyond which the signals generated for a given input power become too weak to be usable. For every system there is also a distance, dependent on both frequency and resistivity, below which the far-field approximation can no longer be used (Figure 9.10). Inevitably there will, for any value of ground resistivity, be a frequency at which the two distances are identical, and this is the CSAMT limit for that system in that area. Signals can, of course, be transmitted at
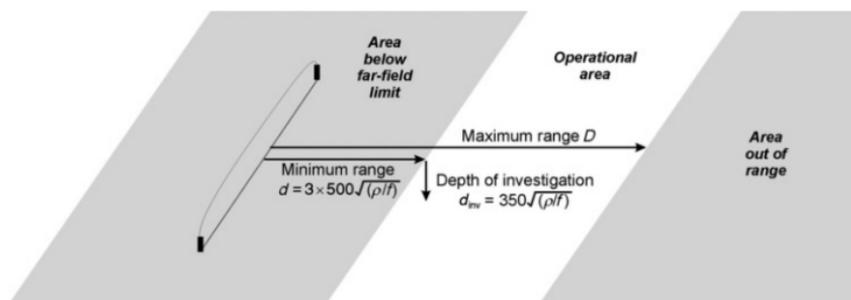
Figure 9.10 *Investigation constraints in controlled-source audio-magneto-telluric (CSAMT) surveys. The area within which survey work can be carried out lies between a maximum distance determined by signal strength and the minimum determined by the limit below which the Cagniard equation ceases to apply. The latter is frequency-dependent, and the operational zone therefore varies in width according to frequency* (f). ρ = *resistivity.*

lower frequencies, and will still contain information on ground resistivity, but the equations that have to be solved to extract the resistivity information become much more complex.

The CSAMT depth of investigation (d_{inv} in Figure 9.10) is closely related to the far-field limit (d). They are both linked to the skin-depth and proportional to the square root of resistivity divided by frequency. The greatest depth of investigation is achieved at the frequency at which the far-field limit and the range maximum (D) coincide. Its actual value does not depend explicitly on the frequency or resistivity but is, if the widely quoted far-field limit of three times the skin depth is correct, a little less than a quarter of the power-related system range (Figure 9.10). This provides a quick way to assess system capability. Thus, the range maxima quoted for the 'normal' and high-power versions of the Geometrics Stratagem are 400 m and 800 m respectively, implying maximum investigation depths of about 100 and 200 m. It would, of course, be logistically very difficult to carry out surveys that achieve this theoretical maximum penetration, because the source would have to be moved in concert with the receivers. As an alternative, CSAMT and MT/AMT surveys can be combined, with CSAMT used principally to fill in the frequency 'gaps' where natural signal strengths are inadequate.

CSAMT surveys require transmitters capable of providing signals over a wide range of frequencies, and time is needed to record data at each frequency in turn. An alternative approach is to use the inherently multifrequency characteristics of transient signals (see Section 5.2.4). This technique, known as LOTEM (*Long Offset Transient Electro-Magnetics*) is being used to a limited extent in a variety of applications.

10

GROUND PENETRATING RADAR

Radio-wave echo-sounding was first patented by Christian Hülsmeyer in 1904 as a device (the *telemobiloscope*) to prevent ships colliding. The term RAdio Detection And Ranging (RADAR) was not coined until 1934, by Robert Watson-Watt. The radar frequencies, from one to several thousand megahertz, were originally thought to be too high for useful ground penetration, but in the 1950s USAF pilots were crash landing on ice fields in Greenland because radar altimeter reflections from the base of the ice were misread as reflections from the landing surfaces. *Ground penetrating radar* (*GPR*) was born, and it was only a short step from its use in determining ice thickness to studies of permafrost. It was then realised that some penetration was being achieved into the deeper, unfrozen ground and that the depth of investigation, although unlikely to ever amount to more than a few tens of metres, could be increased by processing techniques almost identical to those applied to seismic reflection data.

Today GPR is one of the most extensively used of all geophysical methods. Numerous applications can be found in geological mapping, engineering, structural and archaeological investigations, as well as forensic and environmental surveys.

10.1 Radar Fundamentals

Ground penetrating radar uses electromagnetic waves with frequencies between 10 MHz and 4 GHz to detect changes in electrical properties. There are many resemblances to seismic (particularly seismic reflection) methods, although some terminologies seem to have been developed with the aim of disguising this fact (Figure 10.1). Transmission velocities are crucially important in interpretation, and are almost independent of frequency at radar frequencies for a wide range of materials. This contrasts with the situation below about 1 MHz, where conduction currents dominate and velocities and frequencies decrease together.

10.1.1 Radar parameters

Radar waves obey Maxwell's equations (see Section 5.2.2) but in GPR work it is usually assumed that the relative magnetic permeability, μ, of the
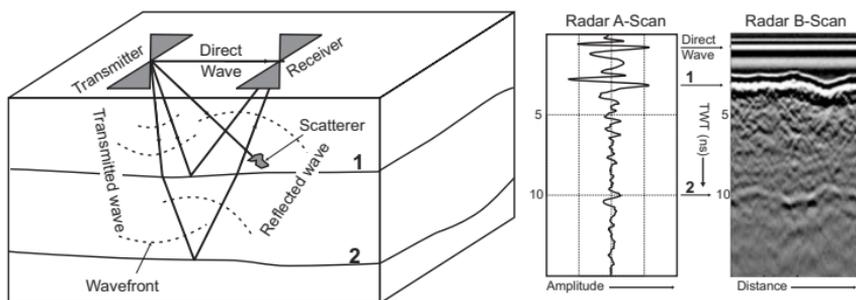
Figure 10.1 *Ground penetrating radar (GPR) measurement and display. Boundaries 1 and 2 separate layers with different electrical properties, and signals can also be 'scattered' from isolated objects. A-scans and B-scans are the radar equivalents of the 'traces' and 'sections' of seismic reflection work.*

ground is unity, and that the reflection, dispersion and attenuation of the radar signals are therefore caused only by changes in conductivity, σ, and relative electric permittivity, ε (which defines the ability of the medium to transmit or 'permit' an electric field). It is also usually assumed that these are scalar quantities, independent of the direction of the radiating field. These assumptions are not strictly true but are adequate for simple treatments.

The velocity of an electromagnetic wave in an insulator is given by:

$$V = c/\sqrt{(\mu \cdot \varepsilon)} \quad \text{or, taking } \mu = 1, \quad V = c/\sqrt{\varepsilon}$$

Here c ($=3 \times 10^8 \, \text{m s}^{-1}$, $300\,000 \, \text{km s}^{-1}$ or $0.30 \, \text{m ns}^{-1}$) is the velocity of light in *free* (empty) space. In conducting media there are further complications, which can be described in terms of a *complex permittivity* ($K = \varepsilon + j\sigma/\omega$), where $j = \sqrt{(-1)}$, and a *loss tangent* ($\tan\alpha = \sigma/\omega\varepsilon$), where ω ($=2\pi f$) is the *angular frequency*. Large loss tangents imply high signal attenuation.

Table 10.1 lists typical radar parameters for some common materials. Velocities are generally well below the free-space value. Electrical conductivities at radar frequencies differ, sometimes very considerably, from DC values, often increasing with frequency at roughly log-linear rates (Figure 10.2).

The radar wavelength in any material is equal to the wave velocity divided by the frequency, i.e.

$$\lambda = c/f\sqrt{\varepsilon}$$

Table 10.1 *Typical values of radar parameters for some common materials*

Material	Permittivity (ε)	Conductivity (σ) (mS m^{-1})	Velocity (V) (m ns^{-1})	Attenuation constant (α) (dB m^{-1})
Air	1	0	0.30	0
Ice	3–4	0.01	0.16	0.01
Fresh water	80	0.05	0.033	0.1
Salt water	80	3000	0.01	1000
Dry sand	3–5	0.01	0.15	0.01
Wet sand	20–30	0.01–1	0.06	0.03–0.3
Shales and clays	5–20	1–1000	0.08	1–100
Silts	5–30	1–100	0.07	1–100
Limestone	4–8	0.5–2.0	0.12	0.4–1
Granite	4–6	0.01–1	0.13	0.01–1
(Dry) salt	5–6	0.01–1	0.13	0.01–1
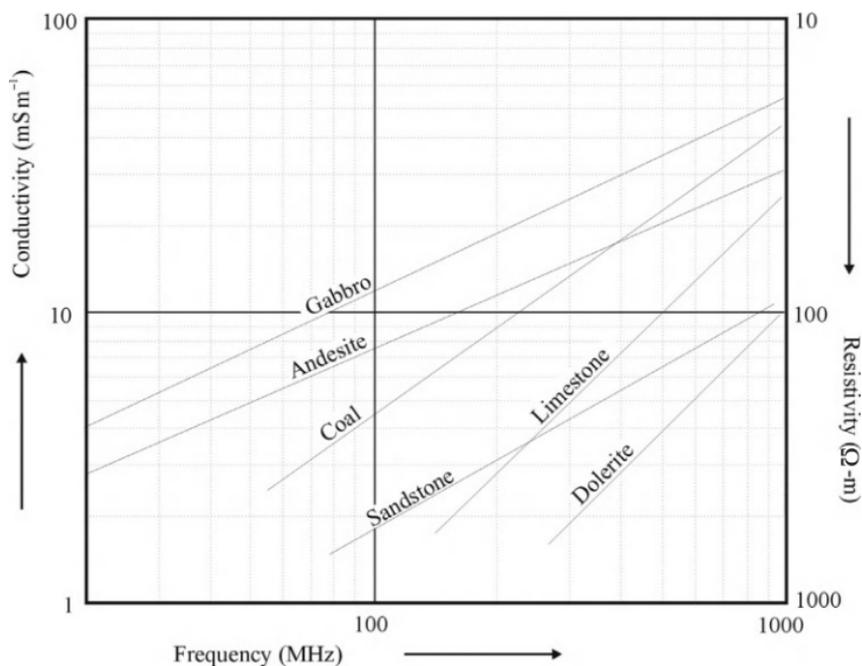
Figure 10.2 *Variation of conductivity at radar frequencies, after Turner (1992: GPR and the effects of conductivity: Exploration Geophysics 23, 381–386).*

The calculations should be straightforward but, because GPR velocities are usually quoted in m ns^{-1} and frequencies in MHz (Table 10.1), it is easy to lose a few powers of ten unless orders of magnitude are appreciated. The wavelength of a 100-MHz signal in air is 3 m, in rock with velocity 0.1 m ns^{-1} is 1 m, and in salt water, where $V = 0.01$ m ns^{-1}, only 10 cm. A useful way of remembering the relationship between the key radar parameters is provided by the $V/f/\lambda$ triangle:
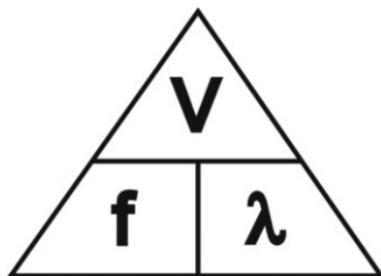
If the required wavelength and material dielectric properties are known, then the required signal frequency can be calculated by dividing the velocity by the wavelength. Similarly if the source frequency is known as well as the transmission velocity, then wavelength can be deduced by dividing the velocity by the frequency.

10.1.2 Reflection of radar pulses

If a GPR signal encounters a discontinuity in either permeability, conductivity or permittivity, some of the signal will be reflected. The amount is determined by the size of the target, the angle of incidence and the *amplitude reflection coefficient* for normal incidence on an infinite interface, which, for low-conductivity, non-magnetic materials, is given by:

$$\mathrm{RC} = (\sqrt{K_2} - \sqrt{K_1})/(\sqrt{K_1} + \sqrt{K_2}) = (V_1 - V_2)/(V_2 + V_1)$$

K_1 and K_2 are the complex permittivities of the host and target material respectively. The *power reflection coefficient*, equal to $(\mathrm{RC})^2$, is also sometimes used, but conceals the fact that there is a phase change of $180°$ on reflection from any boundary where permittivity decreases (and velocity therefore increases). Reflection coefficients in most geological materials are determined almost entirely by variations in water content, but conductivity dominates for metals.

Reflection power is also governed by the nature of the reflector surface. The strongest responses come from smooth surfaces, where specular reflection occurs (i.e. the angles of incidence and reflection are equal), but rough

surfaces scatter energy, reducing reflection amplitudes. Similarly, small targets generate only weak reflections. Success with GPR generally requires reflection of at least 1% of the incident wave (i.e. RC > 0.01) and the smallest lateral dimension of the target to be not less than one-tenth of its depth.

To resolve the thickness of any layer, the travel time through it must be greater than the pulse duration; that is, the thickness must be greater than the signal wavelength. If the layer is significantly less than one wavelength thick, interference effects (between reflections from its top and bottom) become important and thickness becomes difficult to measure. The layer can still be detected, as long as the net reflection amplitude exceeds the background noise level.

10.1.3 Attenuation

Radar waves are attenuated in the ground because they cause currents to flow, converting electrical energy to heat (*ohmic dissipation*), and signal amplitudes eventually fall below detectable levels. The exponential decay is governed by an attenuation constant that in most materials is approximately proportional to both frequency and conductivity and also, where this varies significantly from the free-space value, to magnetic permeability. The notorious failure of GPR to penetrate clay soils is due to the greater attenuation in more conductive media. The roughly linear variation of the attenuation constant with frequency over the frequency range of a typical pulse implies that pulse shapes will change with travel time, as the higher frequencies are preferentially attenuated.

GPR penetration is also limited by scattering, which can become more important than ohmic losses at frequencies above 1 GHz. Signals are scattered from localised objects if their dimensions are comparable with the GPR wavelength, and Rayleigh scattering, proportional to the fourth power of frequency, occurs for objects smaller than the wavelength. This is an important limitation on signal penetration in heterogeneous material.

In polarisable materials at radar frequencies, *displacement currents* that depend on permittivity and the rate of change of the EM field are dominant, because the displacements of charges within atoms and molecules, as opposed to their transport by movement of ions and electrons, become important. This is not the whole story, because even vacuum has a permittivity, but does at least explain the role of physical media.

The water molecule is highly polarisable, i.e. its electric dipole moment is easily aligned by an applied electric field, and water has a remarkably high relative permittivity ($\varepsilon = 80$; Table 10.1). It absorbs energy more strongly as frequency increases, at least up to the relaxation frequency of several GHz (depending on temperature and the degree of binding to soils). Even at 500 MHz, water losses can be seen in otherwise low loss materials.

10.1.4 Decibels

The performance of a radar system is often described in terms of processes involving amplifications and attenuations (*gains* and *losses*) measured in *decibels* (dB). If the power input to a system is **I** and the power output is **J**, then the gain, in dB, is equal to $10.\log_{10}(\textbf{J/I})$. A gain of 10 dB thus corresponds to a ten-fold, and a 20 dB gain to a hundred-fold, increase in signal power. Negative values indicate losses. The logarithmic unit allows the effect of passing a signal through a number of stages to be obtained by straightforward addition of the gains at each stage.

Log$_{10}$2 is equal to 0.301, and doubling the power is thus almost equivalent to a gain of 3 dB. This convenient approximation is so widely quoted that it sometimes seems to have become the (apparently totally arbitrary) definition of the decibel. Almost equally confusing is the popular use of decibels to measure absolute levels of sound. This conceals a seldom stated threshold (dB = 0) of 10^{-12} W m^{-2}, the commonly accepted minimum level of sound perceptible to the human ear. This *acoustic decibel* is, of course, irrelevant in radar work.

10.1.5 The radar range equation

Predicting, without an on-site test survey, whether useful results are likely to be obtained is probably no more difficult with GPR than with any other geophysical method (which means that it is very difficult) but, because the method is relatively new, the principles are less widely understood. The constraints can be divided into those related to instrument performance and those dependent on site conditions.

Instrument performance is dominated by the ratio between the power supplied by the transmitter and the minimum level of signal resolvable by the receiver. The signal loss during transmission through the ground, which is governed by the attenuation constant, is the most important factor but losses in the cables or optical fibres linking transmitters and receivers to their respective aerials, and due to the directional characteristics of the aerials, are significant. These are entered separately into the equations because they depend on the frequencies and aerials used, which can be changed in virtually all GPR units. If the sum, in dB, of all the instrumental factors is equal to *F*, then the *radar range equation* can be written as:

$$F = -10 \cdot \log_{10}[A\lambda^2 e^{-4ar}/16\pi^2 r^4]$$

where λ is the radar wavelength, *a* is the attenuation constant, *A* is a shape factor, with the dimensions of area, that characterises the target, and *r* is the *range*, i.e. the theoretical maximum depth at which the target can be detected. The logarithmic form, which makes the equation appear more
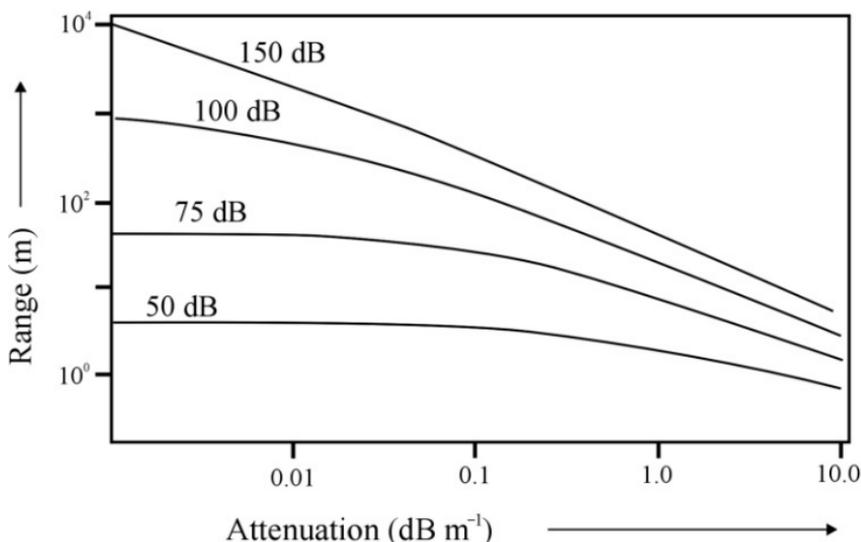
Figure 10.3 *'Typical' nomogram relating radar range to attenuation constant for various fixed values of system gain and spreading and attenuation losses.*

formidable than it really is, arises simply from the use of decibels. The factor $\log_{10}[\lambda^2/4\pi]$ is often included in the system parameters but the quantity remaining then has dimensions and care must be taken in choosing units.

For specular reflection from, respectively, a smooth plane and a rough surface, the shape factors are equal to $\pi r^2(RC)^2$ and $\pi \lambda r(RC)^2$, where RC is the reflection coefficient. The range equations become:

$$F = -10\log_{10}[(RC)^2\lambda^2 e^{-4ar}/16\pi r^2] \quad \text{and}$$
$$F = -10\log_{10}[(RC)^2\lambda^3 e^{-4ar}/32\pi r^3]$$

Neither can be solved directly for the range, which appears in both the exponent and the divisor. Computers can obtain numerical solutions but graphs provide another practical way of dealing with the problem (Figure 10.3). A rough rule that is sometimes applied where conductivity and/or attenuation are known is that the maximum depth of investigation will be somewhat less than 30 divided by the attenuation or 35 divided by the conductivity. The range equation usually provides a rather optimistic prediction of how a radar system will perform.

Noise in GPR work is caused by unwanted signal from various sources internal and external to the system. The ratio of the power received by an antenna to the power of the noise defines the signal to noise ratio (SNR).
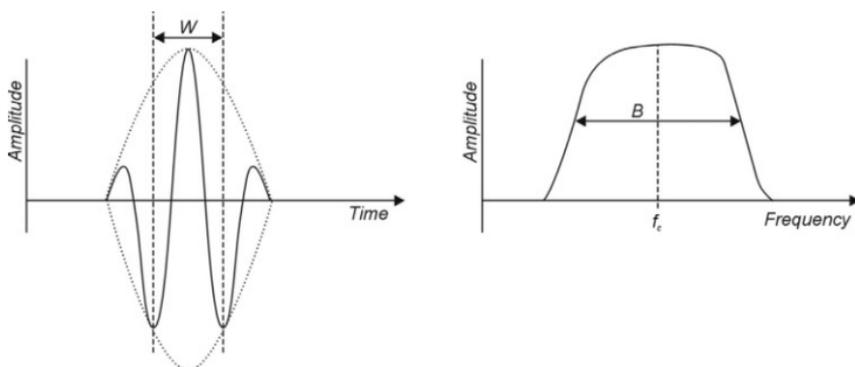
Figure 10.4 *Ultra wideband (UWB) signal in the form of a short-duration 'Ricker wavelet' with broad frequency content centred around the central frequency* f_c. W *is the pulse-width measured between the two side-lobe minima, and* B *is the bandwidth, measured between the 50% amplitude points. For an ideal Ricker wavelet,* $B = 0.78/W$.

10.1.6 Impulse radar

Impulse or ultra wideband (UWB) radar systems are the most common, employing short bursts of electromagnetic energy spanning a range of frequencies from about 50% below to 50% above some specified central frequency f_c. The bandwidth is thus numerically equal to the central frequency, and a typical 100-MHz signal has significant energy at frequencies as low as 50 MHz and as high as 150 MHz. The pulse-width, in seconds, is approximately equal to the reciprocal of the bandwidth, measured in Hz (Figure 10.4), and narrow pulses therefore imply large bandwidths.

The range resolution, ΔR, defines the minimum distance between interfaces that can be resolved by a radar signal and is related to the pulse width, W, and the transmission velocity, V, by the equation:

$$\Delta R \geq W \cdot V/4$$

Thus the range resolution of a survey can be improved by decreasing the pulse width, i.e. by increasing the signal bandwidth (Figure 10.5).

Lateral resolution is defined as the minimum horizontal distance that must exist between two objects if they are to be seen as separate. It is related to distance from the target, d, the pulse width, W, and the transmission velocity, V, by the equation:

$$\Delta L = \sqrt{(V \cdot d \cdot W/2)}$$

192
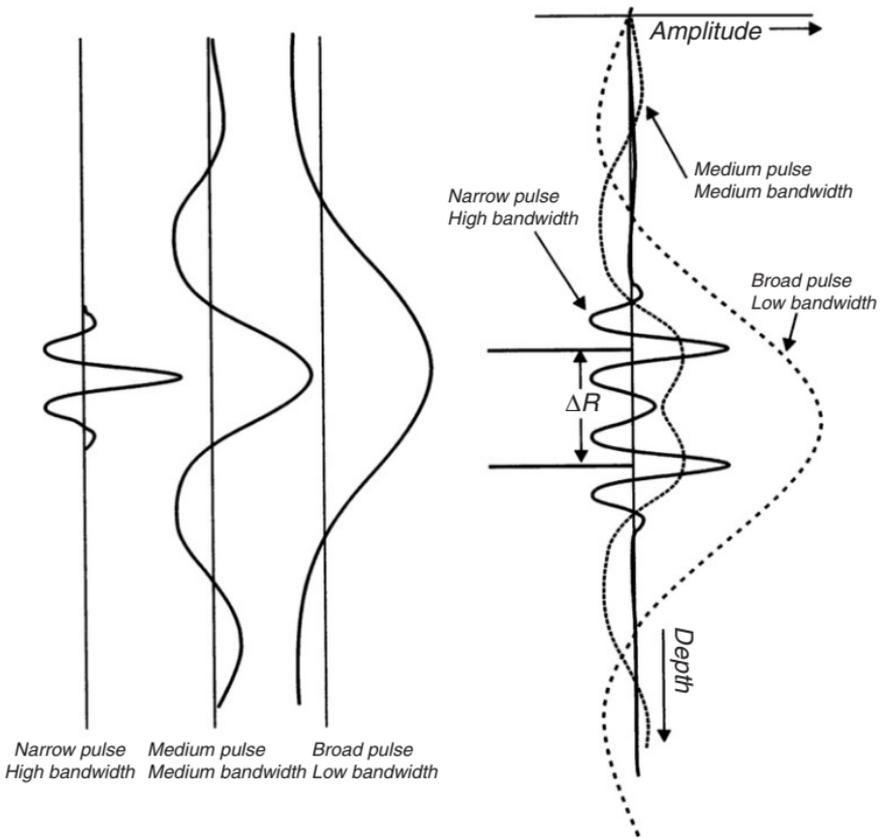
Figure 10.5 *Variation in range resolution (ΔR) with changes in bandwidth. The two interfaces, with reflection coefficients equal in both magnitude and sign, are easily resolved by the narrow pulse, poorly resolved but still separable by the pulse with a spatial width approximately equal to the separation and cannot be resolved by the broad pulse. The pulses, shown on the left, are modified Ricker wavelets (see Figure 10.4) with reduced sidelobes, but even these produce significant peaks and troughs in the vicinity of the interfaces at high bandwidths. The 'spatial width' is equal to pulse width (in time) multiplied by the transmission velocity.*

Depth enters into this equation because the zone of influence or 'footprint' of a radiated pulse (equivalent to the Fresnel zone in seismics) increases with depth, and also with velocity (Figure 10.6). The lower the velocity (higher the permittivity), the smaller the lateral dimension of a target that can be resolved, i.e. the radiated beam is narrower in higher permittivity materials.
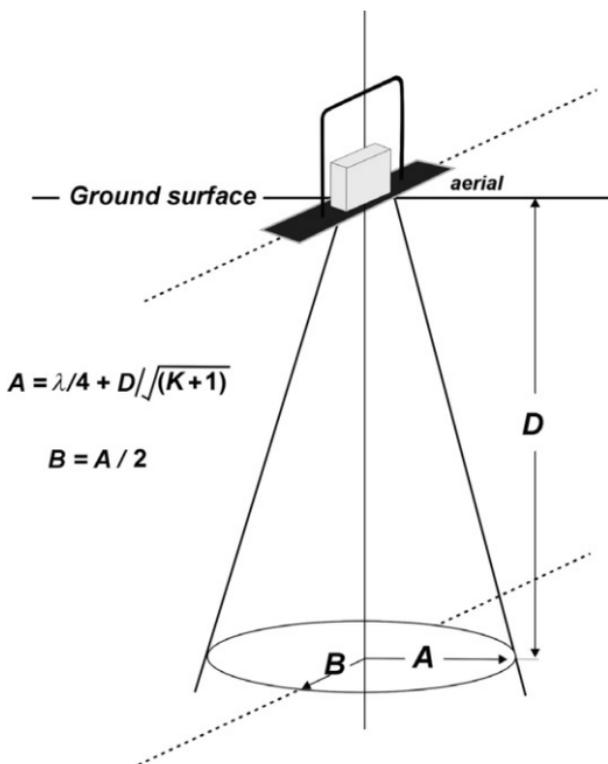
Figure 10.6 *Radiation footprint from a ground-coupled dipole antenna.* A *is the long axis of the ellipsoid-shaped footprint.* B *is the short axis.* D *is depth.*

10.1.7 Continuous-wave radar

Continuous-wave (stepped-frequency) radar systems are less common than impulse systems, and require more complicated antennas. They transmit series of sine waves with stepwise increasing frequency. Figure 10.7 shows an example of just one of these 'packets', or 'chirps'. Fourier transformation is then required to convert the data (collapse the phases) to produce the familiar narrow-pulse response.

Stepped-frequency data can be useful for characterising material properties, and a single system sweeping from say 100 MHz to 2 GHz in steps of 2 MHz (Figure 10.8) can provide detailed images of both shallow and deep features in a single pass. The trade-off comes in the relatively longer time required to collect the data. The *dwell-times* between each transmission can, and must, be defined by the user.
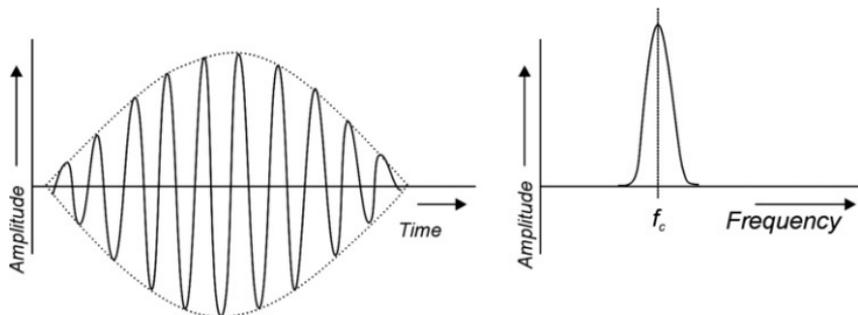
Figure 10.7 *Narrow bandwidth signal derived from a long oscillatory pulse (modulated sine wave) with a single dominant frequency,* f_c. *A series of these waves will be transmitted in a stepped-frequency radar system.*

10.1.8 Antennas

A multitude of antenna types are available, each designed to maximise the SNR with a specific function and scale of search in mind. Some of the commonest are shown schematically in Figure 10.9. All obey the general rule that the lower the frequency, the larger the antenna. A change in frequency thus normally implies a new antenna, but stepped-frequency systems avoid this by using more complicated designs such as the log-spiral.

Dipole and bow-tie antennas are operated close to the ground to maximise coupling, and in practice should not be more than 0.1–0.25 times the radar wavelength above it. Horn antennas, used for applications requiring higher frequencies, more directivity and limited depths of investigation, can be raised above ground by two or three times the wavelength or even more. The
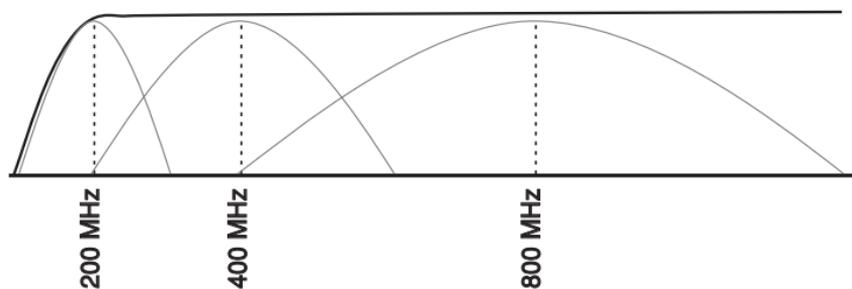
Figure 10.8 *Schematic view of 'continuous' spectral output (dark line) for a stepped-frequency radar system, compared with the spectral bandwidths for impulse ground penetrating radar (GPR). The latter requires multiple antennas to provide equivalent frequency coverage.*
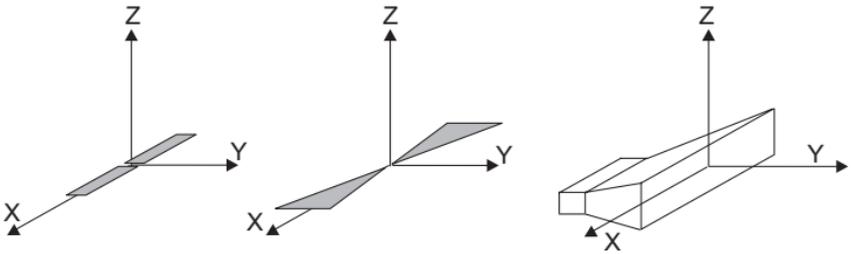
Figure 10.9 Schematic drawings showing, from left to right, dipole, bow-tie and horn antennas.

physical dimensions of a bow-tie or dipole antenna should be similar to the wavelength of the signal in the ground.

GPR principles are most simply discussed in terms of dipole antennas, which in free space radiate with cylindrical symmetry about, and zero intensity along, the dipole axis (Figure 10.10). This simple pattern is drastically modified by the ground. As shown in Figure 10.11, the angular position of a reflection point relative to the antenna strongly affects the strength of the received signal. The peaks in the H-plane pattern and the nulls in the E-plane pattern are governed by a critical angle that depends on the dielectric properties of the subsurface:

$$\theta_c = \sin^{-1}\left(\frac{1}{\sqrt{\varepsilon}}\right)$$
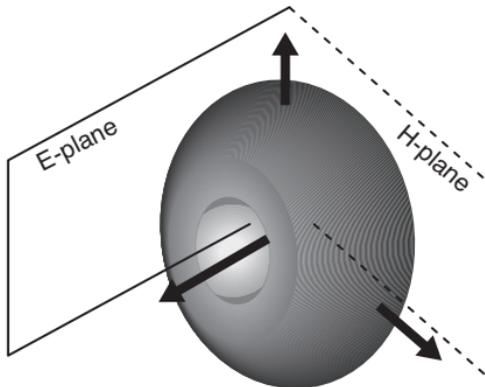
Figure 10.10 Idealised total free-space radiation pattern for the dipole antenna shown in Figure 10.9. The vertical axis of the dipole (x–z) is called the E-plane, or TM, and is orthogonal to the H-plane (TE). For an ideal dipole, no fields are produced off the ends of the antenna.

196
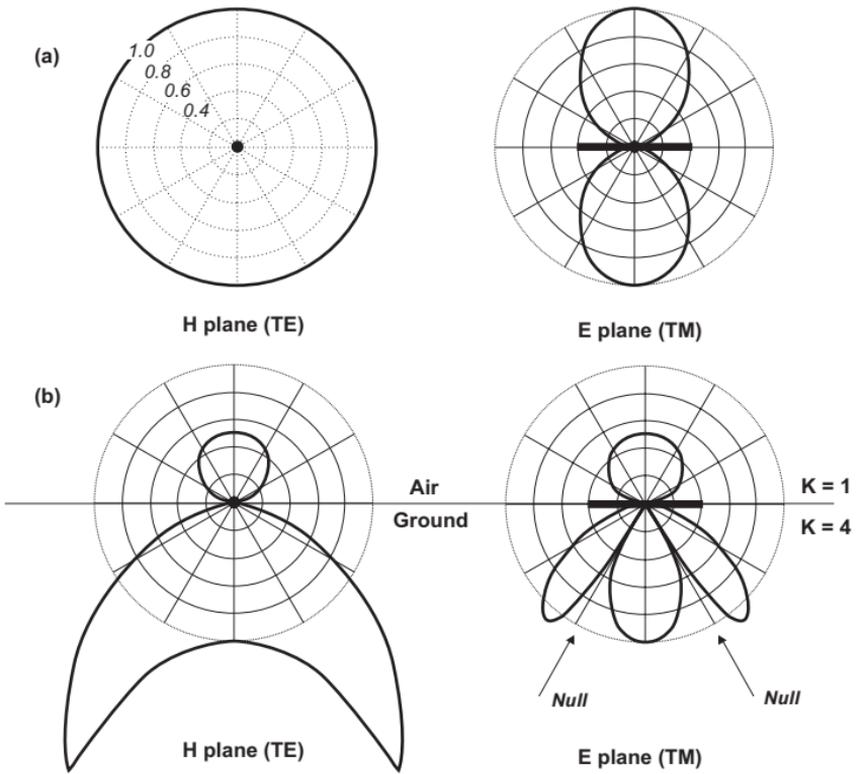
Figure 10.11 *Polar radiation plots showing radiation power in the H- and E-planes for a dipole antenna. (a) Radiation patterns in free space. (b) Radiation patterns with the antenna resting on ground with permittivity* $K = 4$*. In each case the plots are normalised to the power radiated vertically downwards from the antenna mid-point.*

If a reflector dips at the same angle as one of the nulls, little or no energy will be reflected. Awareness of this phenomenon is important, especially if the reflector orientation is poorly known (in which case it is advisable to collect bipolar data).

Antenna orientation is also important in relation to linear targets such as buried pipes. Dipole antennas are most sensitive to such objects when they lie parallel to the dipole long axis, as is the case with Target A in Figure 10.12. In contrast, Target B will be difficult to detect unless the antennas are turned through 90° (which could involve running a second survey along orthogonal survey lines). Some manufacturers produce GPR systems with in-built cross-polarised dipole antennas to achieve maximum detectability
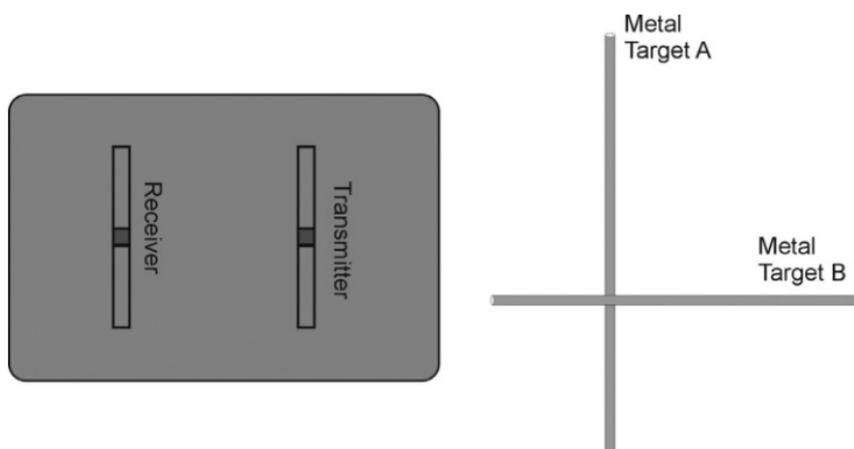
Figure 10.12 *Schematic showing the layout of parallel transmitter and receiver dipole antennas in a single housing in relation to metal targets A and B.*

(and productivity) in cases where the target orientation is unknown. However, if the object of the exercise is to image features below a metal pipeline or a concrete slab with reinforcement bars (*rebars*), then the antennas are best orientated perpendicular to the pipe or bar direction, to minimise their influence on the signal from the target.

As in Figure 10.12, most GPR systems utilise separate receiver and transmitter antennas, which may be mounted separately or integrated into a single module. Separability is desirable even if the spacing is to be kept constant in a given survey, because its optimum value depends on environment and target depth as well as on frequency and transmitter size. A single housing is, however, almost essential for continuous profiling, but it is then necessary to protect the receiver antenna from the high level of the transmitted signal. Reducing cross-coupling between transmitter and receiver antennas is an important aim of GPR system design.

Much effort is now going into designing antennas that are shielded to maximise the downward-looking signal and reduce the spurious responses from above-ground targets. No shielding is ever perfect and signal leakage can and will occur, despite claims to the contrary. These new antennas are mainly variants of resistively loaded dipoles and spiral and TEM horns.

10.1.9 Pulse stacking
An ideal GPR impulse system would transmit a pulse into the ground and display the convolved ground and receiver antenna response as a single

waveform over a selected time interval (measured in nanoseconds). Although feasible in acoustics and seismics, this has only recently become possible at radar frequencies, because Analogue to Digital Converters (ADCs) with enough dynamic range and sufficiently high sampling rates did not exist. To get around this, sequential sampling of repeated transmissions is widely used, with only one sample collected for each transmission. Thus, if 256 samples are being collected for every transmission, then 256 transmissions are required to collect a single sample at each sampling time (stack fold of 1). GPR systems generally have pulse-repetition frequencies (PRFs) of 100 kHz for central frequencies below 400 MHz, and the 1000-fold stack needed to increase the signal-to-noise ratio by 30 dB would require 256 000 transmissions and 2.6 s for a single scan. This is not a problem if antennas are being moved by hand, but it is not practical to collect more than 32–64 stacks per scan using continuously towed systems. This is a fundamental limit on the SNR. Practical and regulatory restrictions limit the scope for improving things by increasing transmitter power.

High-speed ADCs have recently become available that allow entire waveforms to be collected in a single pass. These offer the possibility of increasing the stack fold (up to a practical limit of about 10 000), and can provide an additional 30–40 dB for the same antenna. These systems are likely to become routine during the lifetime of this edition.

10.2 GPR Surveys

The key factors controlling the performance of a GPR system are the central frequency, f_c, and bandwidth, B (see Section 10.1.6). High frequencies are needed to resolve small objects but have limited penetration. Lower frequencies are required to detect deeper objects, which consequently must be larger. Table 10.2 provides typical centre frequency selections for a range

Table 10.2 *Indicative depths of investigation for different centre frequencies*

Centre frequency (MHz)	Approximate depth (m)
10	50
25	30
50	10
100	5
200	3
500	2
1000	1

Figure 10.13 *Examples of commonly available radar systems utilising ground-coupled dipole or bow-tie antennas.*

of target investigation depths. Relatively low permittivity materials are assumed for this table. For high permittivity and/or conductive materials, or materials containing a high number of scatterers (clutter), penetration will be significantly reduced.

10.2.1 Instrumentation

Ground penetrating radar systems (Figures 10.13 and 10.14) consist of control and recorder units (CRUs) linked to receiver and transmitter units, each of which is in turn linked to one or more antennas. Metal wires are less efficient conductors of alternating current at radar frequencies, and signals are usually transmitted to and from the CRUs via optical fibres. These have the advantage of immunity from electrical interference, but are more delicate than wires and less easily repaired when damaged. Some commercially available fibre cables are protected by black sleeving that makes them almost invisible when laid out on the ground, so plenty of spares must be taken into the field to replace breakages.

The settings on the CRU determine the radar frequency, the time period (*window*) during which data are to be recorded and the number of individual traces to be stacked. Central frequencies of 25, 50, 100 and 200 MHz would be typical for geological applications, with recordings made over time windows of between 32 and 2048 ns. The frequencies used for scanning engineering structures are usually in the range from 400 to 4000 MHz, over time periods of between 10 and 70 ns. Modern CRUs are equivalent to powerful personal computers, and much signal processing can

Figure 10.14 *Radar system with air-launched horn antennas being used to monitor railway trackbed ballast.*

be done in the field. In some early systems the CRU functions could be performed by any suitable laptop PC loaded with appropriate software, but this placed the user at the mercy of operating-system upgrades. Moreover, few laptops can survive being rained on.

10.2.2 Survey types

In most GPR surveys, the antenna separation is kept constant. This is *common-offset profiling* (Figure 10.15a). The alternative is to vary the distances of the two antennas from a fixed mid-point to obtain multi-fold *Common Mid-Point* (*CMP*) coverage (Figure 10.15b). CMP surveys allow velocities to be calculated from the variations in reflection time with offset but are slow, and therefore relatively rare, in GPR (as discussed in Section
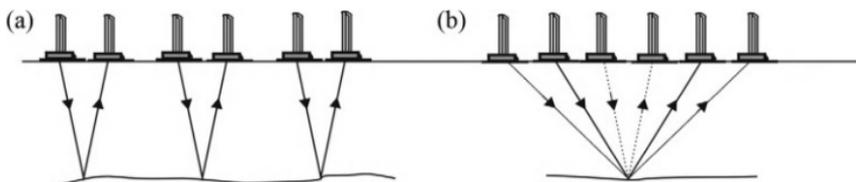
Figure 10.15 *(a) Fixed or constant offset profiling. (b) Common mid-point sounding.*

12.1.2, they are routine in seismic reflection, where cheap *geophone* detectors can be laid out in large numbers).

The advent during the last few years of multi-channel (multi-antenna or array) GPR systems promises to revolutionise many engineering, environmental and archaeological applications. Apart from the obvious productivity gains, the probability of detecting localised or linear small targets increases with the number of antennas deployed. The multi-channel systems now available, from several manufacturers, include single central-frequency UWB systems giving improved illumination and detectability of targets within a selected range, multiple central-frequency UWB systems providing improved illumination across a range of depths in a single pass, and stepped-frequency systems for providing improved illumination and detectability across a range of depths in a single pass.

10.2.3 Selecting survey parameters

Antenna separation, station separation, record length, transmission frequency and sampling frequency can all be varied to some extent in most GPR work. Transmission frequency is the most important single variable, since it constrains the values of many other parameters. Resolution criteria may, and usually do, conflict with penetration requirements, and field operators should at least know the extent to which, for the instruments they are using, the trade-off between penetration and resolution can be compensated for in processing.

Resolution is also affected by station spacing. Targets that would otherwise be resolved will not be properly defined if the distance between adjacent stations in a common offset profile is more than one quarter of the wavelength in the ground, i.e. approximately $75/f\sqrt{\varepsilon}$, where f is measured in MHz. Separations of about one-fifth of this value usually give good results (Figure 10.16) but smaller ones may be used for operational convenience. The various criteria are summarised in Table 10.3. Correct choices are particularly important if buried features must be mapped in real-time, since this implies interpretation in the field.

Radar signals are recorded digitally and must be sampled often enough to ensure that waveforms are fully defined. If there are fewer than two samples in each full period, *aliasing* (see Figure 1.7) will occur. Because the maximum frequency present in a GPR signal is approximately twice the nominal central frequency, the sampling rate should be at least four times the central frequency. A safety factor of two is usually added, giving a sampling frequency of 800 MHz for a 100-MHz signal (i.e. a sample interval of 1.25 ns). Most GPR systems are set up to collect many more samples than necessary. This can become important when survey speed/resolution tradeoffs are being considered.
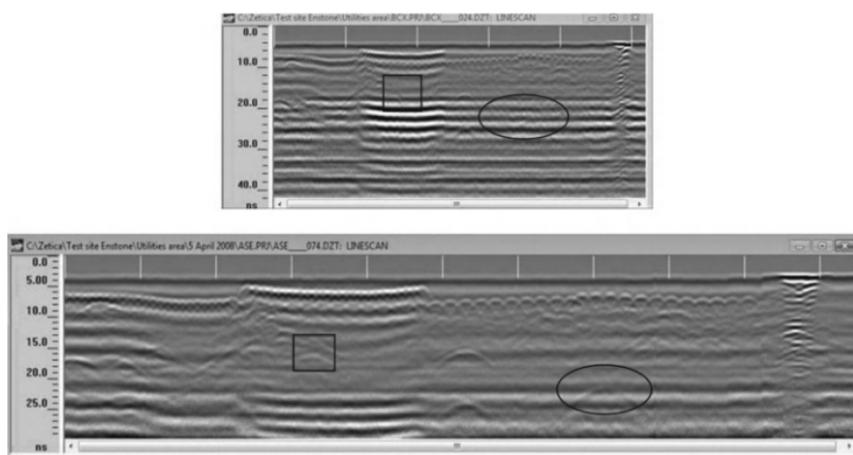
Figure 10.16 Radargrams from incorrectly and correctly set-up ground penetrating radar (GPR) systems. In the upper example the targeted features (outlined) are all but invisible. The improved set-up below, with 2.5 times the scans per metre and a 30% reduced range, reveals the subtle, but important, features sought.

The high repetition rates possible with radar systems allow large numbers of signals to be recorded at each transmitter/receiver set-up and to be stacked to reduce the effects of random noise. The decision as to how many repetitions should be used is one that, inevitably, must be made in the field. At the very least, the number of repetitions used should be the maximum that is possible before reading time begins significantly to affect productivity.

Dipolar transmitter and receiver antennas are most commonly set out side by side, but end-on and even broadside (orthogonal) configurations are also used. Antennas should be orientated parallel to the target strike-direction, if this is known (see Section 10.1.8). Freely available, one-dimensional (horizontally layered Earth) modelling software provides a useful way of testing assumptions, identifying multiple paths, reverberations and polarity changes, and confirming the ranges and optimum central frequencies required to resolve the expected targets (Figure 10.17).

10.2.4 Interference in GPR surveys

Even when depth penetration, reflectivity and resolution seem satisfactory, environmental problems can prevent success in a GPR survey. Radio transmitters are potential sources of interference, and powerful radio signals can overwhelm (*saturate*) receiver electronics. The presence of metal objects in the subsurface can also be disastrous if these are not the survey targets. Reflections can come from objects away to the side (*sideswipe*), and may be

Table 10.3 *Ground penetrating radar (GPR) survey design*

GPR survey design parameters	Rules of thumb	Remarks
GPR centre frequency	$\dfrac{150}{d\sqrt{\varepsilon}} = f_c$	Frequency in MHz, d and L in metres
Range	$1.3 * \dfrac{2^* D_{max}}{V_{min}}$	Time in nanoseconds (ns) D in metres, V in m ns^{-1}
Sampling intervals		Time in seconds, f_c in MHz. The number of samples required per scan is determined by dividing the range by the sample interval
Time	$\Delta t = \dfrac{1}{6 f_c}$	
Distance	$\Delta X_{min} < \dfrac{75}{f_c \varepsilon}$	Distance in metres, f_c in MHz. It is better to oversample. Data redundancy can be advantageous, e.g. when antenna-surface coupling is variable
Line spacing (LS)		Distances in metres
Linear trending targets	LS $< 0.5 *$ minimum distance of continuous linear segment	
Localised targets	$LS < A = \dfrac{D}{\sqrt{\varepsilon + 1}}$	
Antenna spacing (S)	$S = \dfrac{2^* D}{\sqrt{\varepsilon - 1}}$	Distance in metres $S = 20\%$ of target depth if conditions are unknown If $S \geq D$ then corrections must be applied to obtain $t = 0$
Antenna orientation	Normally E-field parallel to long axis of target	Orthogonal lines required in case of unknown trend direction of targets or unknown dips of layers

f_c, centre frequency.

t, time interval in nanoseconds.

d, spatial separation between reflectors.

ΔL, minimum length of heterogeneity (clutter); ignore if unknown.

ε, dielectric constant.

D, depth from surface to target, assuming antenna coupled to surface.

A, long axis radius of simplified elliptical GPR footprint (see Figure 10.7).

V, wave velocity in the medium.

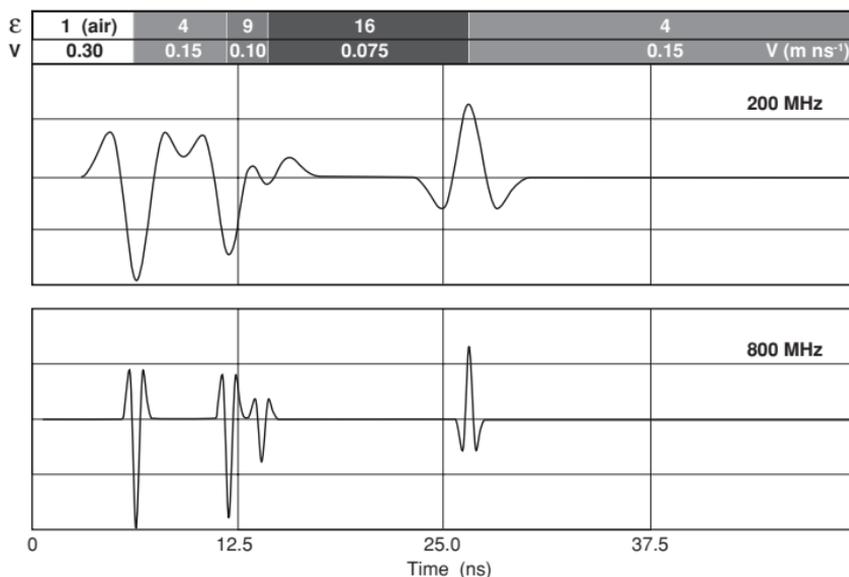| ε	1 (air)	4	9	16	4	
V	0.30	0.15	0.10	0.075	0.15	V (m ns⁻¹)

Figure 10.17 *'One-dimensional' A-scan responses at 200 and 900 MHz for a horizontally layered model in which a thin (0.1-m) layer is sandwiched between two layers with relative permittivities as shown (top). Note the greater resolution at the higher frequency, and the opposite polarity of the reflection from the top of the low-permittivity layer. Polarity can be a useful diagnostic for determining material type.*

very strong if the objects are metallic. Features actually at the surface can produce strong sideswipe because there is substantial radiation of energy along the ground/air interface if ground conductivity is high.

A common source of interference or low SNR is the presence of what is often termed *clutter*, i.e. inhomogeneities in the materials surrounding a target. The signal then loses significant energy by reflection from irrelevant features before reaching the target depth. This could, for example, occur in a survey over a pocket of coarse uncompacted gravels or industrial rubble. Lower frequencies than would otherwise be desirable may have to be used.

10.3   Data Processing

GPR data may need extensive processing. The reduction in size and increase in capacity of small computers has made it possible to process in the field, sometimes using the field instruments themselves. However, much time can be wasted for very little gain, and it is a useful principle to only process the raw data if there is a clearly identifiable benefit in doing so. For many types
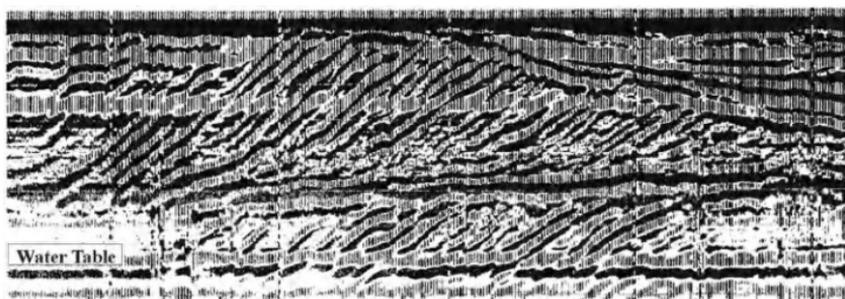
Figure 10.18 *SIR System-3 variable area ground penetrating radar (GPR) record (B-scan) showing sigmoidal oblique reflectors in the Lower Greensand. Reproduced by permission of Dr C. Bristow.*

of target, standard GPR systems produce field data that can be interpreted without further ado.

If processing is done, it is vital to preserve an audit trail so that any outcome can be traced back to source. With the world-wide proliferation in the use of GPR in civil engineering, it is only a matter of time before someone, somewhere is called on to defend an interpretation in a court of law.

10.3.1 Display of GPR data

A GPR 'trace' or A-scan (Figure 10.1) is recorded as a series of digital values equally spaced in time. An ideal A-scan would consist of a flat line punctuated by occasional 'events' produced by energy arriving at the surface after reflection from points vertically below the mid-point between the receiver and transmitter antenna (Figure 10.17). Display can be either as a grouping of A-scans (*wiggle trace*), or as *variable area* B-scans in which excursions on one side of the zero line are shaded. Colour is sometimes used to emphasise polarity.

On a normal B-scan, the horizontal axis represents distance and the vertical scale is in two-way reflection time (TWT). With most instruments, a B-scan formed from lightly processed traces is displayed on a screen in 'real time' as the antennas are moved along the traverse. This makes GPR one of the most interesting geophysical methods to actually use. The resemblance between B-scans and seismic reflection sections is extremely close (compare Figure 10.18 and Figure 12.8), and it is sometimes only the fact that the two-way time scale is in nanoseconds rather than milliseconds that shows that the section was produced by electromagnetic rather than seismic waves.

As in seismic reflection, radar work now includes 3D surveys, in which traverses are very closely spaced and the results are considered in terms of data volumes rather than individual sections. Presentations of the type
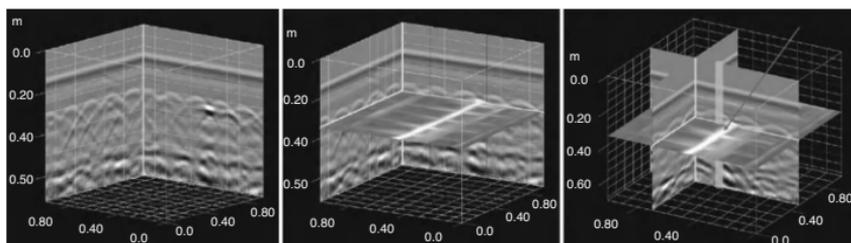
Figure 10.19 *Three-dimensional (3D) display of a high-resolution ground penetrating radar (GPR) data set (left) collected on a 0.1-m orthogonal grid to locate a safe area to drill a hole through a reinforced concrete slab. On the left, unprocessed data are presented as a 3D cube showing X and Y image sections. In the centre, a time slice (in the Z plane) has been selected to highlight a conduit cross-cutting the slab. The location of the recommended drill hole is shown (right) superimposed on the data cube. No data processing other than 3D presentation and selection of interpolation parameters was required, and nobody was hurt in the making of the hole.*

shown in Figure 10.19 are still relatively uncommon, but their use is likely to increase.

10.3.2 Migration

Water tables and sedimentary layering appear on GPR sections as continuous events (Figure 10.18), but pipes, cables, drums and UXO are usually
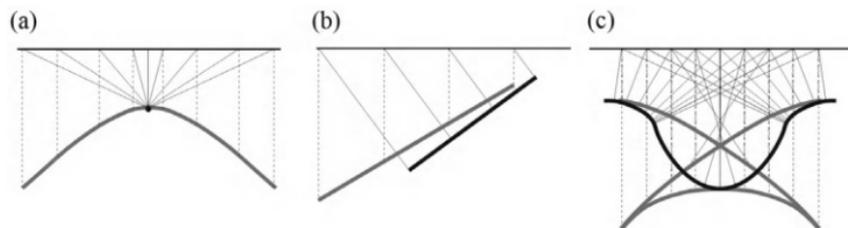
Figure 10.20 *Geometric distortion on radar sections. In each case the continuous lines show the actual reflection paths (for near-coincident receiver and transmitter antennas) and the dotted lines show the positions on the B-scans where the corresponding A-scans will be plotted. Thick grey lines show the plotted image, assuming no major velocity changes. (a) Diffraction pattern caused by a point reflector; (b) reduction in dip and lateral displacement of a dipping bed; (c) 'bow-tie' from a tight syncline. Examples of some of these features can be seen on the seismic section of Figure 12.8.*
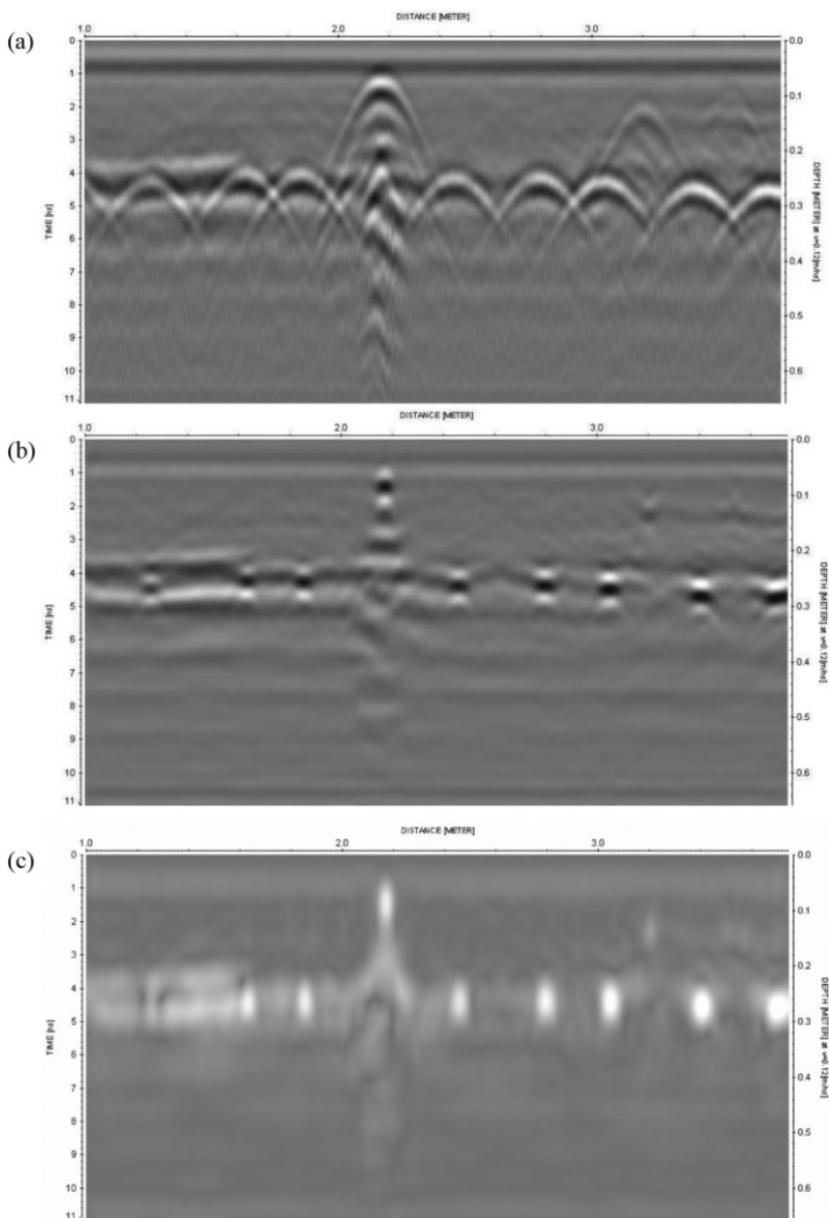
Figure 10.21 *Example of processing used to image steel reinforcing bars in a concrete slab. (a) The raw data, with hyperbolic responses from the bars. (b) Migration has been used to collapse the hyperbolic tails and highlight the reflection sources. (c) The energy envelope produced by a Hilbert transform.*

indicated by curved, upwardly convex *diffraction patterns* created by a form of sideswipe. The root cause of the problem is that A-scans are plotted out vertically to form B-scans, and therefore all events, even if from reflection along a slant path, appear vertically below the surface antenna positions. Small reflectors or sharp angles in continuous reflectors can reflect radar waves at many angles, producing diffraction patterns (Figure 10.20a). Provided that velocities vary only with depth, these patterns are hyperbolic. The source, provided that the traverse line actually passes over it, is located at the pattern apex.

Less dramatic distortions affect the positions of dipping reflectors, since reflections from these also travel along slant paths (Figure 10.20b). Tight, upwardly concave *synclines* can produce three reflections at a single surface point, giving rise to the peculiar features known as *bow-ties* (Figure 10.20c). All these distortions can be corrected by the *migration* programs first developed for use with seismic data, but these must work on numerous traces simultaneously and cannot easily be applied in the field.

It is almost true to say that radar applications are divided into two distinct groups. In the first, layered structures are imaged and mapped. In the second, diffraction patterns are sought that reveal the presence of limited highly reflective targets.

10.3.3 The processing sequence

As already noted, most of the processing techniques now being used for GPR data resemble those developed for seismic data, and seismic processing software has been used almost unmodified to enhance GPR results. There are differences in emphasis, largely because of the well controlled nature of the radar pulse and the general use of single-fold rather than CMP coverage, but these need not concern the operator in the field.

After stacking, GPR data are passed through low-cut filters, to remove noise due to inductive effects and the limitations in instrument frequency response, and high-cut filters, to eliminate noise spikes. The decrease in signal amplitude with time is then reversed by time-variant amplification. *Automatic gain control* (AGC) is used to do this in the field, producing records for quality control, but data are normally stored in unmodified form. Compensation, in the processing centre, for propagation effects, using SEC (*spherical and exponential compensation*) filters based on physical models of the subsurface, comes after any frequency-based filtering, because time-variant gain functions distort wavelets and must be applied with care if amplitude integrity is to be preserved. Migration algorithms operate on the entire data set rather than individual traces (Figure 10.21), and are usually the last to be applied.

11

SEISMIC METHODS: GENERAL CONSIDERATIONS

Seismic methods are the most effective, and the most expensive, of all the geophysical techniques used in investigating layered media. Features common to reflection and refraction surveys are discussed in this chapter. Chapter 12 is concerned with the special features of small-scale reflection work, and Chapter 13 with shallow refraction. Chapter 14 deals with the developing use of surface waves. Deep reflection surveys, which involve large field crews, bulky equipment and complex data processing, are beyond the scope of this book.

11.1 Seismic Waves

A seismic wave is acoustic energy (a sound wave) transmitted by vibration of rock particles. Low-energy waves are approximately elastic, leaving the rock mass unchanged by their passage, but close to a seismic source the rock may be shattered and permanently distorted.

11.1.1 Types of elastic wave

When a sound wave travels in air, the molecules oscillate backwards and forwards in the direction of energy transport. This *pressure* or 'push' wave thus travels as a series of compressions and rarefactions. In a solid medium it has the highest velocity of any of the possible wave motions and is therefore also known as the *primary* wave, or simply the *P-wave.*

Particles vibrating at right angles to the direction of energy flow (which can only happen in a solid) create an *S* (*shear*, 'shake' or, because of its relatively slow velocity, *secondary*) wave. In many consolidated rocks, the S-wave velocity is roughly half the P-wave velocity. It depends slightly on the plane in which the particles vibrate but these differences are not significant in small-scale surveys.

P- and S-waves are *body-waves* and expand within the main rock mass. Other waves, known as *Love waves,* are generated at interfaces, while particles at the Earth's surface can follow elliptical paths to create *Rayleigh waves*. Love and Rayleigh waves may carry a considerable proportion of the source energy but travel very slowly. Traditionally they were simply lumped

together as the *ground roll*, but are now finding applications in investigations of the shallow subsurface (see Chapter 14).

11.1.2 Seismic velocities

The 'seismic velocities' of rocks are the velocities at which sound waves travel through them. They are quite distinct from the continually varying velocities of the individual rock particles that are forced into oscillation by the transmitted energy.

Any elastic-wave velocity (V) can be expressed as the square-root of an elastic modulus divided by the square root of density (ρ). For P-waves the elongational elasticity, j, is appropriate, for S-waves it is the shear modulus, μ. The equations:

$$V_p = \sqrt{(j/\rho)} \qquad V_s = \sqrt{(\mu/\rho)}$$

suggest that high-density rocks should have low seismic velocities, but because elastic constants normally increase rapidly with density, the reverse is usually true. Salt is the only common rock with a high velocity but a low density.

If the density and P- and S-wave velocities of a rock mass are known, all its elastic constants can be calculated, since they are related by the equations:

$$(V_p/V_s)^2 = j/\mu = 2(1 - \sigma)/(1 - 2\sigma) \quad \text{i.e.}$$
$$\sigma = [2 - (V_p/V_s)^2]/2[1 - (V_p/V_s)^2]$$

and

$$j = q(1 - \sigma)/(1 + \sigma)(1 - 2\sigma) \qquad \mu = q/2(1 + \sigma)$$
$$K = q/3(1 - 2\sigma) \qquad j = K + 4\mu/3$$

where q is Young's modulus and K is the bulk modulus. The Poisson ratio, σ, is the ratio between the amount of shortening experienced by a cube of material in the direction of an applied compression and the expansion that then takes place at right angles to it. This ratio has a maximum value of 0.5 (for a completely incompressible material), at which point the V_p/V_s ratio becomes infinite. Solids can usually be compressed to some extent, but water and most other liquids are incompressible and S-waves do not propagate through them. The Poisson ratio in rocks seldom exceeds 0.35, and P-waves, always the fastest, travel rather less than twice as fast as S-waves in all but the most consolidated rocks (Figure 11.1).

Most seismic surveys use, and provide velocity estimates for, P-waves only, and these are rather rough guides to rock quality. Figure 11.2 shows the velocity ranges for common rocks and also *rippabilities*, defined by whether
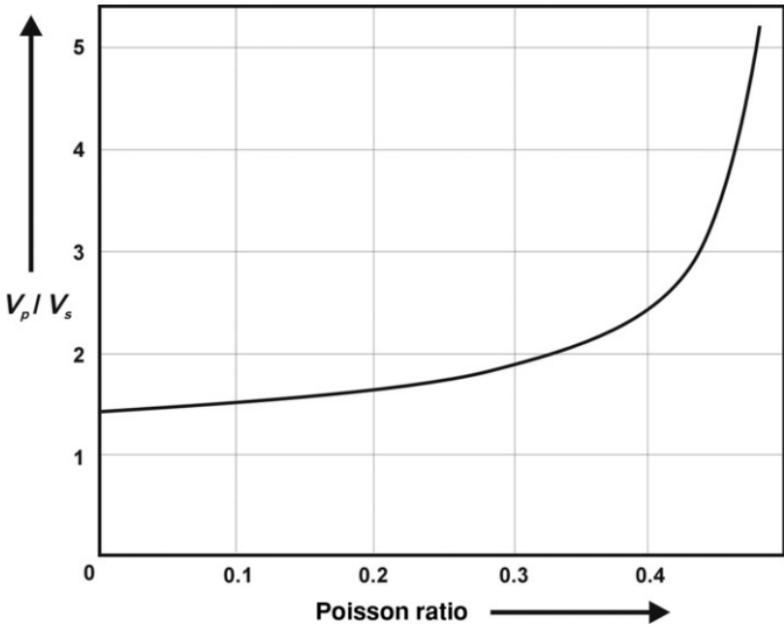
Figure 11.1 *Variation of V_p/V_s ratios with Poisson ratio.*

the rocks can be ripped apart by a spike mounted on the back of a bulldozer. S-waves are now being used in some shallow reflection surveys, where their slower velocities provide improved resolution (see Section 12.2.7).

11.1.3 Velocities and the time-average equation

Within quite broad limits, the seismic velocity in a mixture of different materials can be calculated by averaging the transit times (the reciprocals of velocities) through the pure constituents, weighted according to the relative amounts present. The principle can be used even when, as in Example 11.1, one of the constituents is a liquid.

Example 11.1 Seismic velocities in mixtures

The velocity of the P-wave in a sandstone that is 80% quartz and 20% water-filled porosity can be calculated from:

V_p (quartz) $= 5200 \, \mathrm{m \, s^{-1}}$ V_p (water) $= 1500 \, \mathrm{m \, s^{-1}}$
$1/V_p = 0.8/5200 + 0.2/1500 = 0.000\,287$
i.e. V_p(sandstone) is approximately $3500 \, \mathrm{m \, s^{-1}}$.
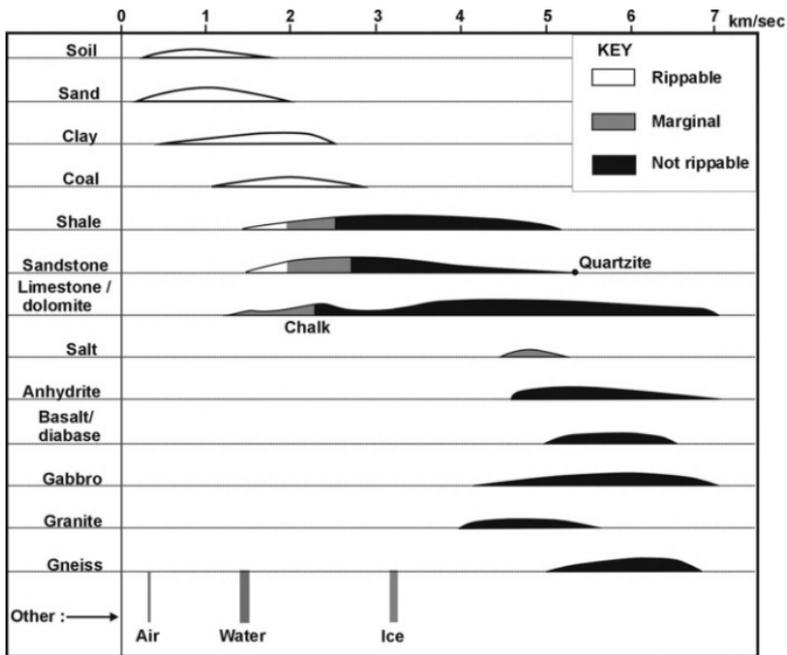
Figure 11.2 *Ranges of P-wave velocities and rippabilities in common rocks. The vertical axis for each rock type is intended to show approximately the relative numbers of samples that would show a given velocity.*

In dry rocks, the pore spaces are filled with air ($V = 330\,\text{m s}^{-1}$) rather than water. Time averaging cannot be applied quantitatively to gas-filled pores, but dry materials generally have very low P-wave velocities. If they are poorly consolidated and do not respond elastically, they may also strongly absorb S-waves. Poorly consolidated water-saturated materials generally have velocities only slightly greater than that of water, and the water table is often a prominent seismic interface in P-wave surveys. Because S-waves do not propagate in either gases or liquids, the water table is invisible in S-wave surveys (which can sometimes be an advantage).

Weathering normally increases porosity and therefore reduces rock velocities. This fact underlies the rippability ranges shown in Figure 11.2. Few fresh, consolidated rocks have velocities of less than about $2200\,\text{m s}^{-1}$, and rocks that are rippable are generally also at least partly weathered.

11.1.4 Ray-path diagrams

A seismic wave is completely described in terms of *wavefronts*, which define the points that the wave has reached at a given instant. However, only a small

part of a wavefront is of interest in any geophysical survey, since only a small part of the energy returns to the surface at points where detectors have been placed. It is convenient to identify the important travel paths by drawing seismic *rays*, to which the laws of geometrical optics can be applied, at right angles to the corresponding wavefronts. Ray-path theory works less well in seismology than in optics, because the most useful seismic wavelengths are between 25 and 200 m, and are thus comparable with survey dimensions and interface depths. Wave effects can be significant under these circumstances but field interpretation can nonetheless be based on ray-path approximations.

11.1.5 Reflection and refraction

When a seismic wave encounters an interface between two different rock types, some of the energy is reflected and the remainder continues on its way at a different angle, i.e. it is *refracted*. The law of reflection is very simple: the angle of reflection is equal to the angle of incidence (Figure 11.3a). Refraction, which is simply a consequence of the wavefronts expanding at different rates in media with different velocities, is governed by *Snell's Law*, which relates the angles of incidence and refraction to the seismic velocities in the two media:

$$\sin i / \sin r = V_1 / V_2$$

If V_2 is greater than V_1, refraction will be towards the interface. If $\sin i$ equals V_1/V_2, the refracted ray will be parallel to the interface and some of its energy will return to the surface as a planar wavefront known as the
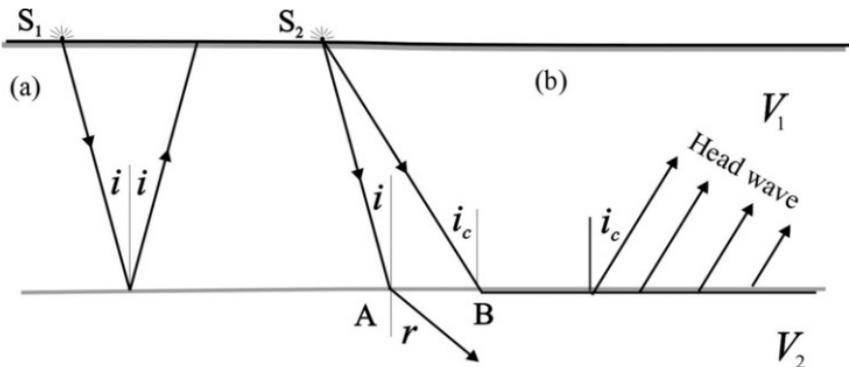

Figure 11.3 (a) Reflection and (b) refraction. Simple refraction occurs at A, critical refraction at B. The general angle of incidence is denoted by i and the angle of critical incidence by i_c. The angle of refraction, r, is 90° for critical incidence.

head-wave, which leaves the interface at the original angle of incidence (Figure 11.3b). This is the basis of the refraction methods discussed in Chapter 13. At greater angles of incidence there can be no refracted ray and all the energy is reflected.

When drawing ray paths for either reflected or critically refracted waves, allowance must be made for refraction at all shallower interfaces. Only the *normal-incidence* ray, which meets all interfaces at right angles, is not refracted.

11.2 Seismic Sources

The traditional seismic source is a small charge of dynamite. Impact and vibratory sources are now more popular but explosives are still quite commonly used.

11.2.1 Hammers

A 4- or 6-pound sledgehammer provides a versatile source for small-scale surveys. The useful energy produced depends on ground conditions as well as on strength and skill. Hammers can nearly always be used in refraction work on spreads 10 to 20 m long but very seldom if the energy has to travel more than 50 m.

The hammer is aimed at a flat plate, the purpose of which is not so much to improve the pulse (hitting the ground directly can sometimes provide more seismic energy) but to stop the hammer abruptly and so provide a definite and repeatable shot instant. Inch-thick aluminium or steel plates used to be favoured, but are now being replaced by thick rubber discs that last longer and are less painfully noisy. The first few hammer blows are often rather ineffective, as the plate needs to 'bed down' in the soil. Too much enthusiasm may later embed it so deeply that it has to be dug out.

11.2.2 Other impact sources

More powerful impact sources may be used in larger surveys. Weights of hundreds of kilograms can be raised by portable hoists or cranes and then dropped (Figure 11.4). The minimum release height is about 4 m, even if a shorter drop would provide ample energy, since the support rebounds when the weight is released and creates its own seismic wave-train. A long drop allows these vibrations to die away before the impact occurs. Tractor-mounted post-hole drivers, easily found in farming areas, are also convenient sources. The weight drops down a guide and is raised by a pulley system connected to the tractor power take-off. Another alternative is the Propelled Energy Generator (PEG), which uses an elastic band to propel a mass of a few tens of kg at high velocity from a height of just half a metre.
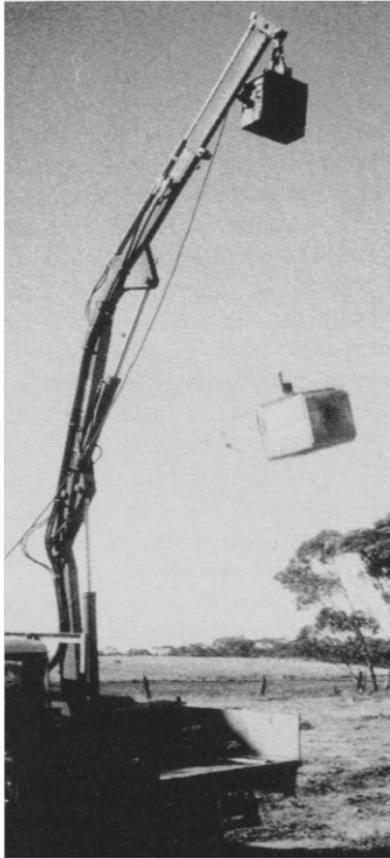
Figure 11.4 *Impact source. A half-ton weight being dropped from a portable crane during a survey of the low-velocity layer.*

Relatively small (70-kg) weights falling in evacuated tubes have sometimes been used. The upper surface of the weight is exposed to the air, and effectively several hundred extra kilograms of atmosphere are also dropped. The idea is elegant but the source is difficult to transport, because the tube must be strong and therefore heavy, and must be mounted on a trailer, together with a motor-driven compressor to pump out the air.

Vibration sources are widely used in large-scale reflection surveys but produce data that need extensive and complex processing. Advances in computing power have only recently made such processing feasible in small-scale reflection surveys. With the source shown in Figure 11.5, the vibrator is coupled to the ground and sweeps through a computer-selectable range

Figure 11.5 *Electromagnetic vibrator source.*

of frequencies. Typically the distance to the furthest geophone would be about 5 m, and the frequency range would be somewhere between 5 and 200 Hz, with actual values determined by site conditions. Vibration sources are becoming popular in surface-wave surveys (see Chapter 14).

11.2.3 Explosives

Almost any type of (safe) explosive can be used for seismic work, particularly if the shot holes are shallow and the charges will not be subject to unusual temperatures or pressures. Cord explosives, used in quarry blasting to introduce delays into firing sequences, are rather safer to handle than normal gelignite and can be fed into shot holes prepared by driving metal rods or crowbars into the ground. Detonators used on their own are excellent sources for shallow reflection surveys where high resolution is needed.

Much of the energy delivered by an explosion can be wasted in shattering rock near the shotpoint, and seismic waves are produced much more efficiently by shots fired in a metre or so of water. This effect is so marked that, if the shot position is not critical, it can be worth going tens or even hundreds of metres from the recording spread in order to put the charge in a river. In dry areas, significant improvements can be obtained by pouring water down shot holes.

Electrical firing is normal when using explosives but with ordinary detonators there is a short delay between the instant at which the filament burns

through, which usually provides the time reference, and the time at which the main charge explodes. *Zero-delay* detonators should be used for seismic work, and total delays through the entire system, including the recorders, should be routinely checked using a single detonator buried a few inches away from a geophone.

Explosives involve problems with safety, security and bureaucracy. They must be used in conformity with local regulations, which usually require separate secure and licensed stores for detonators and gelignite. In many countries the work must be supervised by a licensed shot-firer, and police permission is required almost everywhere. Despite these disadvantages, and despite the headaches that are instantly produced if gelignite comes into contact with bare skin, explosives are still used. They represent potential seismic energy in its most portable form and are virtually essential if signals are to be detected at distances of more than 50 m.

A variety of explosive-based methods are available that reduce the risks. Seismic waves can be generated by devices that fire lead slugs into the ground from shotgun-sized cartridges, but the energy supplied is relatively small and a firearms certificate may be needed, at least in the UK. Another approach is to use blank shotgun cartridges in a small auger that incorporates a firing chamber, combining the shot hole and the shot. However, this seldom provides more energy than a blow from a well-swung hammer, and is less easily repeated.

11.2.4 Safety

Large amounts of energy must be supplied to the ground if refractions are to be observed from depths of more than a few metres or reflections from depths of more than a few tens of metres, and such operations are inherently risky. The dangers are greatest with explosives – but nor is it safe to stand beneath a half-ton weight that is about to be released.

Explosives should only be used by experienced (and properly licensed) personnel. Even this does not necessarily eliminate danger, since experts in quarry blasting often lack experience in the special conditions of seismic surveys. If there is an accident, much of the blame will inevitably fall on the party chief who will, if he is wise, keep his own eye on safety even if specialists are employed for that purpose.

The basic security principle is that the shot-firer must be able to see the shot-point. Unfortunately, some seismographs have been designed so that the shot is triggered by the instrument operator, who can seldom see anything and who is in any case preoccupied with checking noise levels. If such an instrument is being used, it must at least be possible for firing to be prevented by someone who is far enough from the shot-point to be safe but close enough to see what is happening. This can be achieved if, after the

shot hole has been charged, the detonator is first connected to one end of an expendable cable 20 or 30 m long. Only when the shot-point is clear should the other end of this cable be connected to the cable from the firing unit. Firing can then be prevented at any time by pulling the two cables apart.

Unless 'sweaty' gelignite is being used (and the sight of oily nitroglycerine oozing out of the packets should be sufficient warning to even the least experienced), modern explosives are reasonably insensitive to both heat and shock. Detonators are the commonest causes of accidents. Although their explosive power is small, they have caused loss of fingers and even hands. If fired on their own as low-energy sources, they should always be placed in well-tamped holes, since damage or serious injury can be caused by fragments of the metal casing.

It is possible (although not common) for a detonator to be triggered by currents induced by powerlines or radio transmissions, but this is less likely if the leads are twisted together. Triggering by static electricity is prevented if the circuit is closed. The shorted, twisted ends of detonator leads should be parted only when the time comes to make the connection to the firing cable, which should itself be shorted at the far end. Explosives should not be handled at all when thunderstorms are about.

Explosive charges need to be matched to the holes available. Large charges may be used in deep holes with little obvious effect at the surface, but a hole less than 2 m deep will often blow out, scattering debris over a wide area. Only experience will allow safe distances to be estimated, and even experienced users can make mistakes; safety helmets and goggles should be worn and physical shelter such as a wall, a truck or a large tree should be available. Heavy blasting mats can reduce blow-outs, but their useful lives tend to be short and it is unwise to rely on them alone.

A point where a shot has been fired but no crater has formed should be regarded with suspicion. The concealed cavity may later collapse under the weight of a person, animal or vehicle, leading to interesting litigation.

11.2.5 Time-breaks

In any seismic survey, the time at which the seismic wave is initiated must be known. In some instruments this appears on the record as a break in one of the traces (the *shot-break* or *time-break*). On many instruments it now actually defines the start of the record.

Time-break pulses may be produced in many different ways. A geophone may be placed close to the source, although this is very hard on the geophone. Explosive sources are usually fired electrically, and the cessation of current flow in the detonator circuit can provide the required signal. Alternatively, a wire can be looped around the main explosive charge, to be broken at the
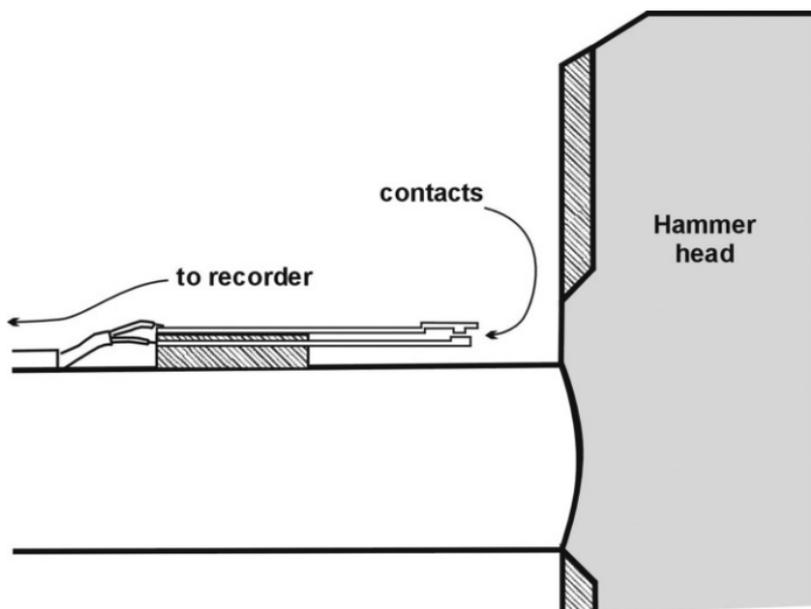
Figure 11.6 *'Post-office relay' impact switch on the back of a sledgehammer handle.*

shot instant. This technique can be used on the rare occasions when charges are fired using lit fuses.

Hammer surveys usually rely on making rather than breaking circuits. One method is to connect the hammer head to one side of the trigger circuit and the plate (assuming it is metal, not rubber) to the other. Although this sounds simple and foolproof, in practice the repeated shocks suffered by the various connections are too severe for long-term reliability. In any case, the plates themselves have rather short lives, after which new connections have to be made. It is more practical to mount a relay on the back of the hammer handle, just behind the head, that closes momentarily when the hammer hits the plate (Figure 11.6). It will close late, or not at all, if the hammer is used the wrong way round. Solid-state switches sold by some seismograph manufacturers give more repeatable results but are expensive and rather easily damaged.

The cable linking the trigger switch on a hammer to the recorder is always vulnerable, tending to snake across the plate just before impact. If it is cut, the culprit is traditionally required to both repair the damage and ease the thirst of all the witnesses.

Where the source is a heavy weight dropped from a considerable height, a relay switch can be attached to its top surface but may not trigger if the
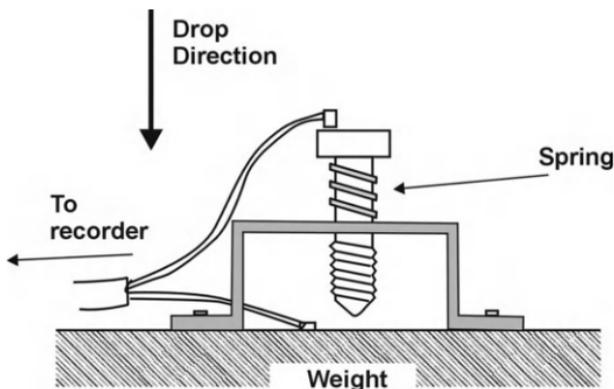
Figure 11.7 *Weight-drop contact switch. On impact the inertia of the bolt compresses the spring and contact is made with the upper surface of the weight.*

drop is not absolutely straight. A crude but more reliable home-made device that can be attached to any dropping weight is shown in Figure 11.7.

Time-break pulses may be strong enough to produce interference on other channels (*cross-talk*; see Section 11.3.5). Trigger cables and circuits should therefore be kept well away from data lines.

11.3 Detection of Seismic Waves

Land seismic detectors are known as *geophones*, marine detectors as *hydrophones*. Both convert mechanical energy into electrical signals. Geophones are usually positioned by pushing a spike screwed to the casing firmly into the ground, but it may be necessary to unscrew the spike and use some form of adhesive pad or putty when working on bare rock. Hydrophones are simply towed behind boats.

11.3.1 Geophones

A geophone consists of a coil wound on a high-permeability magnetic core and suspended by leaf springs in the field of a permanent magnet (Figure 11.8). If the coil moves relative to the magnet, voltages are induced and current will flow in any external circuit. The current is proportional to the velocity of the coil through the magnetic field, so that ground movements are recorded, not ground displacements. In most cases the coil is mounted so that it is free to vibrate vertically, since this gives the maximum sensitivity to P-waves rising steeply from subsurface interfaces, i.e. to reflected and critically refracted (but not direct) P-waves. These geophones, when normally
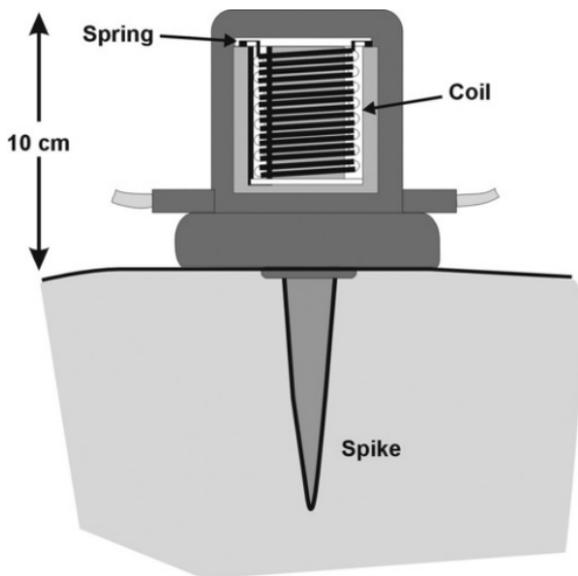
Figure 11.8 Moving coil geophone.

connected, give negative first-arrival pulses (*breaks*) for refractions and re-
flections, but may break either way for direct waves. In reflection work
using large offsets, or in refraction work where the velocity contrasts be-
tween overburden and deeper refractors are small, the rising wave-fronts
make relatively large angles with the vertical, and geophone discrimination
between S-waves and P-waves will be less good.

Geophone coils have resistances of the order of $400\,\Omega$, and damping
is largely determined by the impedance of the circuits to which they are
linked. The relative motion between coil and casing is also influenced by the
natural vibration frequency of the suspended system. At frequencies above
resonance, the response approximately replicates the ground motion, but
signals below the resonant frequency are heavily attenuated. Standard geo-
phones usually resonate at or below 10 Hz, which is well below the fre-
quencies useful in small-scale surveys. Response curves for a typical 10-Hz
phone are shown in Figure 11.9.

Geophones are remarkably rugged, which is just as well considering the
ways in which they are often treated. Even so, their useful lives will be
reduced if they are dumped unceremoniously from trucks into tangled heaps
on the ground. Frames can be bought or made to which they can be clipped
for carrying (Figure 11.10) and these can be good investments, but only if
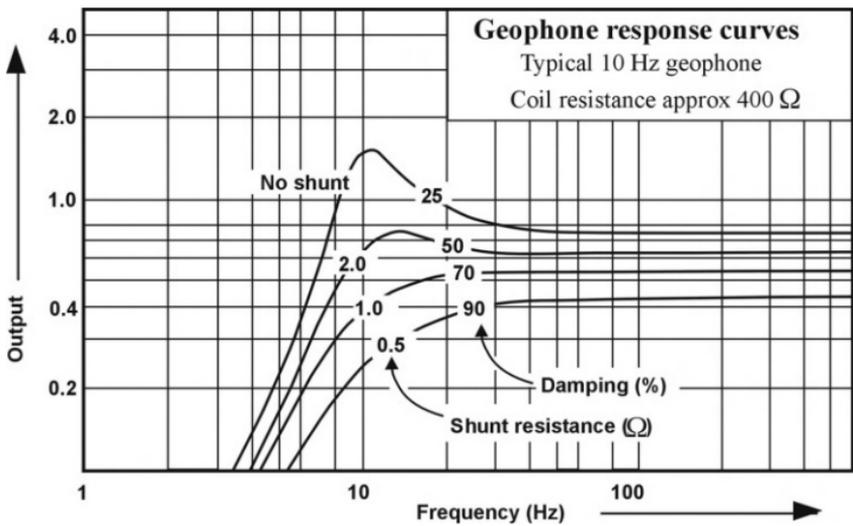actually used.

Figure 11.9 *Frequency response of a typical moving-coil geophone. The degree of damping depends on the value of the shunt resistance connected in parallel with the geophone, and also on the input resistance of the recorder. 'No shunt' corresponds to infinite shunt resistance.*

Figure 11.10 *Geophone carrying frame in use, Papua New Guinea.*

The energy that land field crews expend in setting out geophones individually by hand is regarded with amazement by marine crews, who simply drag their hydrophone streamers through the water. It is not easy to develop similar systems for land use, because of poorer coupling to the vibrating ground, higher background noise, electrical and mechanical cross-talk in the cables, and mechanical wear, but this is a rapidly developing field of design. At present, streamers are being used mainly in surface wave studies (see Chapter 14), where amplitudes may be an order of magnitude greater than for body waves and short spreads and closely spaced detectors are the norm.

11.3.2 Detection of S-waves

Although S-waves are regarded as noise in most seismic work, there are occasions when S-wave information is specifically sought. For example, both S- and P-wave velocities are required to determine elastic properties (see Section 11.1.2), and S-waves are sometimes preferred for high-resolution reflection surveys in noisy environments (see Section 12.2.7).

'S-wave' geophones have coils that move horizontally rather than vertically, the assumption being that any wavefronts of interest will be rising more or less vertically and that the S-wave vibrations will therefore be in the plane of the ground surface. Knowing this does not, of course, define the vibration direction, which may be determined by the way the source is operated but can be completely unknown. It is sometimes necessary to use two S-wave phones, at right angles to each other, in each location.

Because direct waves travel parallel to the ground surface, S-wave geophones are more sensitive to direct P-waves than to direct S-waves, just as P-wave geophones are sensitive to vertically polarised direct S-waves.

11.3.3 Detection in swamps and water

Normal geophones are rainproof rather than waterproof, and are connected to cables by open crocodile clips. Geophones are also available that are completely enclosed and sealed into waterproof cases, for use in swamps. These do not have external spikes but are shaped so that they can be easily pushed into mud.

Motion-sensitive instruments cannot be used in water. Piezo-electric *hydrophones* respond to variations in pressure rather than motion and are equally sensitive in all directions. Discrimination between P- and S-waves is not required since S-waves cannot travel through fluids.

11.3.4 Noise

Any vibration that is not part of the signal is *noise*. Noise is inevitable, and *coherent* noise is generated by the shot itself. S-waves, Love and Rayleigh waves and reflections from surface irregularities are all forms of coherent noise in P-wave surveys. In shallow refraction work these late-arriving, but

often energetic, waves usually prevent the use of any event other than the first arrival of energy.

Noise that is not generated by the shot is termed *random*. Movements of traffic, animals and people all generate random noise and can, to varying extents, be controlled. It should at least be possible to prevent the survey team contributing, by giving warning using a whistle or hooter.

Random noise is also produced by vegetation moving in the wind and disturbing the ground. The effects can be reduced by positioning geophones away from trees and bushes, and sometimes by clearing away smaller plants. Significant improvements can often be achieved by moving particularly noisy geophones a few centimetres. Placement is also important. It may not be easy to push a spike fully home in hard ground but a geophone that is just a couple of centimetres above the ground vibrates in the wind.

11.3.5   Seismic cables

Seismic signals are carried from geophones to recorders as varying electric currents, in cables that must contain twice as many individual wires as there are geophones. Wires are necessarily packed very closely and not only can external current carriers such as power and telephone cables induce currents, but also a very strong signal in one wire can be passed inductively to all the others. This '*cross-talk*' can be particularly severe from the strong signals produced by geophones close to the shot-point, and it may even be necessary to disconnect these to obtain good records on other channels.

The amount of cross-talk generally increases with the age of the cable, probably because of a gradual build-up of moisture inside the outer insulating cover. Eventually the cable has to be discarded.

Cables and plugs are the most vulnerable parts of a seismic system and are most at risk where they join. It is worth while being very careful. Re-soldering wires to a plug with 24 or more connections is neither easy nor interesting.

Most cables are double-ended, allowing either end to be connected to the receiver. If a wire is broken, only the connection to one end will be affected and the 'dead' channel may revive if the cable is reversed. All too often, however, other dead channels are discovered when this is done.

11.4   Recording Seismic Signals

Instruments that record seismic signals are known as *seismographs*. They range from timers that record only single events to complex units that digitise, filter and store signals from a number of detectors simultaneously.

11.4.1   Single-channel seismographs

Rudimentary seismic 'timers', which simply displayed the travel time of the first significant energy pulse, were once popular. More sophisticated

instruments had visual displays, on which the left-hand edge of the screen defined the shot or impact instant and the time range was switch or key-pad selected. Travel times were measured directly on screen, using a cursor that could be moved across the screen while the time corresponding to its position was displayed. Noise levels could be monitored by observing the trace in the absence of a source pulse. Hard-copy was rarely generated but a digital version of the signal could usually be stored in solid-state memory. A repeat signal could either replace this or be added to it. Any number n of signals could be summed (*stacked*) in this way, for a theoretical $\sqrt{n}$ improvement (*enhancement*) in signal/noise ratio.

Instruments that allowed signals to be displayed and summed were obviously superior to mere timers, and could be used to study events other than first arrivals. However, they were only useful in shallow refraction work, since it is almost impossible to distinguish between direct waves, refractions and reflections on a single trace. Hammer sources were universal, since it would have been expensive and inefficient to use an explosive charge to obtain such a small amount of data.

11.4.2 Multi-channel seismographs

Seismographs with 12 or 24 channels are generally used in shallow surveys, compared to the 96 or more channels typical of deep reflection work. With multiple channels, refractions and reflections can both be studied, and explosives can reasonably be used since the cost per shot is less important when each shot produces many traces. Powerful microcomputers are incorporated into most modern instruments, with high-capacity flash drives for data storage. Digital recording is virtually universal, display formats can be varied and individual traces can be selected for enhancement, replacement or preservation. Traces can be amplified after, as well as before, storage in memory, and time offsets can be used to display events that occur after long delay times. Filters can be used to reduce high-frequency random noise and also the long-period noise, of uncertain origin, that sometimes drives the traces from one or two geophones across the display, obscuring other traces. Bewildering numbers of acquisition and processing options are available via menu-driven software, and it is sometimes difficult, and time-consuming, to persuade these instruments to carry out routine, straightforward survey work. In practice, conditions in the field seldom allow full use to be made of all the options available, nor is it an efficient use of the machine to have it occupied with processing when it could be collecting fresh data. Digital storage allows the traces to be preserved for further processing on ordinary PCs, either at the field base or back in the office.

One disadvantage of analogue recording was that the range of amplitudes over which data could be recorded with acceptable accuracy (the *dynamic range*) was limited at low amplitudes by tape noise and at high amplitudes by magnetic saturation. The recorded signal amplitudes were therefore kept roughly constant by applying *automatic gain control* (AGC) before recording, but this always distorted the signal. Digital recording has eliminated the need for AGC, because of the very large dynamic range available when data are stored as fixed precision numbers plus exponents (Example 11.2).

Example 11.2 Dynamic range and digital recording

In digital systems, data are recorded as numerical values plus *exponents*, which are the powers of some other number by which the significant figures in the numerical value must be multiplied. Thus, the values

$$46\,789 \quad \text{and} \quad 0.000\,046\,789$$

which demand a very large dynamic range of an analogue storage system, can be written in engineering notation, which uses powers of 10, as:

$$4.6789\text{E} + 4 \quad \text{and} \quad 4.6789\text{E} - 5$$

The percentage accuracy is the same in both cases. Digital data are usually recorded in binary formats, the exponents being powers of 2 between -128 and $+127$, i.e. between roughly 10^{-38} and 10^{+38}.

11.4.3 Seismograph records

Most modern seismographs can provide optional hard-copy as well as screen displays. The refraction survey record in Figure 13.2 shows the signals recorded by 24 geophones at points successively further from the source, with the traces from distant geophones amplified more to compensate for attenuation. Inevitably, amplifying the signal also amplifies the noise.

In the field, arrival times can be estimated from the screen but this is never easy and seldom convenient. On the other hand, hard copies produced directly from the instrument are often of rather poor quality. This is especially true of dot-matrix outputs such as those in Figure 13.11, where the matrix size causes irregularities in what should be smooth curves. It is worthwhile having a reasonable printer at the field base, coupled to a laptop PC loaded with processing software. It would be unwise, however, not to produce, and preserve, the field hard-copy

12
SEISMIC REFLECTION

The seismic reflection method absorbs more than 90% of the money spent worldwide on applied geophysics. Most surveys are aimed at defining oil-bearing structures at depths of thousands of metres using hundreds or even thousands of detectors, and are beyond the scope of this book. However, some reflection work is done by small field crews probing to depths of, at most, a few hundred metres. The instruments used in these surveys were originally very simple but may now have as much inbuilt processing power as the massive processing laboratories of 30 years ago, and field operators need to have some understanding of the reasons why they are being presented with so many options.

12.1 Reflection Theory
Ray-path diagrams, introduced in Chapter 11, provide insights into the timing of reflection events but give no indication of amplitudes.

12.1.1 Reflection coefficients and acoustic impedances
The *acoustic impedance* of a rock, usually denoted by $\mathbf{I}$, is equal to its density multiplied by the seismic P-wave velocity. If a seismic wavefront is *normally incident* (incident at right angles) on a planar interface between two rock layers with impedances $\mathbf{I}_1$ and $\mathbf{I}_2$, then the *reflection coefficient* (RC), i.e. the ratio of the amplitude of the reflected wave to the amplitude of the incident wave, is given by:

$$\mathrm{RC} = (\mathbf{I}_2 - \mathbf{I}_1)/(\mathbf{I}_2 + \mathbf{I}_1)$$

If $\mathbf{I}_1$ is greater than $\mathbf{I}_2$, the coefficient is negative and the wave is reflected with phase reversed, i.e. a negative pulse will be returned if a positive pulse was transmitted, and vice versa.

The amount of energy reflected first decreases and then increases as the angle of incidence increases. If the velocity is greater in the second medium than in the first, there is ultimately total reflection and no transmitted wave (see Section 11.1.5). However, most small-scale surveys use waves reflected at nearly normal incidence.
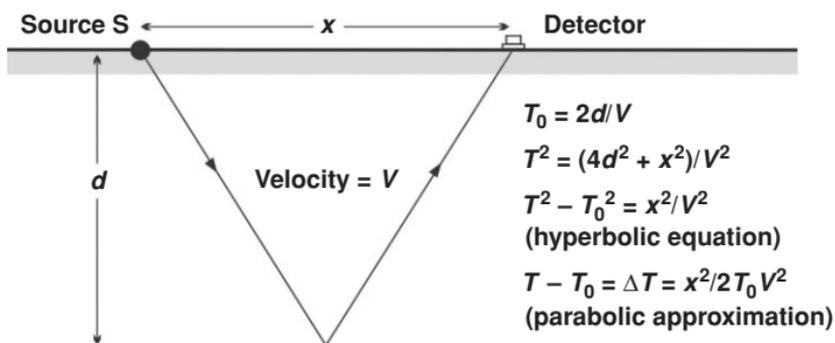
Figure 12.1 *Derivation of the normal moveout equation for a horizontal reflector.* T_0, *normal incidence time;* d, *depth;* x, *distance from source.*

12.1.2 Normal moveout

The true normal-incidence ray cannot be used in survey work, since a geophone at a source point would probably be damaged by the source and would certainly be set into such violent oscillation that the whole record would be unusable. Detectors are therefore offset from sources and geometric corrections must be made to travel times.

Figure 12.1 shows reflection from a horizontal interface, depth *d*, to a geophone at a distance *x* from the source. The exact *hyperbolic* equation linking the travel time *T* and the normal incidence time T_0 is established by application of the Pythagoras theorem. For small offsets, the exact equation can be replaced by the *parabolic* approximation, which gives the *normal moveout* (NMO), $T - T_0$, directly as a function of velocity, reflection time and offset:

$$\Delta T = T - T_0 = x^2/2V^2T_0$$

Since *V* usually increases with depth and T_0 always does, NMO decreases (i.e. NMO curves flatten) with depth.

Curved alignments of reflection events can be seen on many multi-channel records (Figure 12.2). Curvature is the most reliable way of distinguishing shallow reflections from refractions.

12.1.3 Dix velocity

If there are several different layers above a reflector, the NMO equation will give the *root-mean-square* (RMS) velocity defined as:

$$V_{\text{RMS}}^2 = \left(V_1^2 t_1 + V_2^2 t_2 \cdots + \cdots V_n^2 t_n \right) / \mathbf{T_n}$$
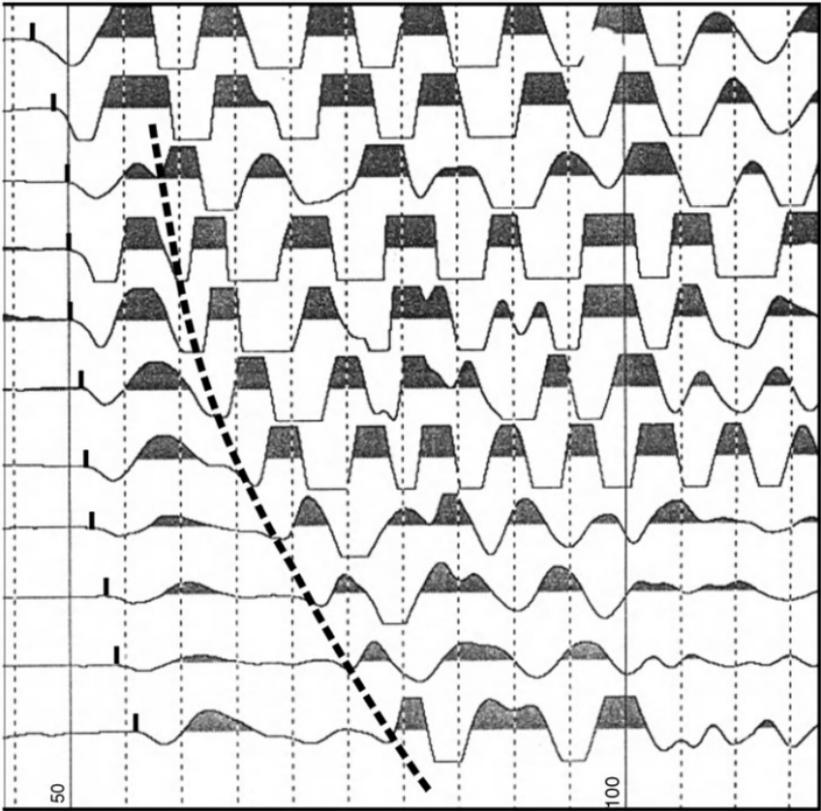
Figure 12.2 *Enhancement seismograph record showing curved alignments of reflections (black dashed line). The earlier, negative polarity, picked event was produced by the direct wave. The variable area presentation used is popular for reflection work since it emphasises trace-to-trace correlations, although some information is lost where traces overlap.*

where t_n is the transit time through the nth layer at velocity V_n, and $\mathbf{T_n}$ is the total transit time to the base of the nth layer. Interval velocities can be calculated from RMS velocities using the *Dix formula*:

$$V_{\text{DIX}}^2 = \left(V_{n-1}^2 \mathbf{T}_{n-1} - V_n^2 \mathbf{T}_n \right) / (\mathbf{T}_{n-1} - \mathbf{T}_n)$$

The subcripts $n-1$ and n denote, respectively, the top and bottom of the nth layer. *RMS* velocities are normally slightly higher than true average velocities, since squaring the high velocities increases their influence on the average. Significant errors can arise if RMS velocities are used directly to make depth estimates but if the interfaces concerned are not horizontal these

Figure 12.3 *Effect of dip on a single-fold record. Rays are reflected from the dipping interface as if derived from the image point S' at depth $2d.cos(\alpha)$ below the surface, where d is the perpendicular distance from the shot-point to the interface. The normal incidence travel time is $2d/V$ but the shortest travel time is for the ray that is vertical after reflection. An identical moveout hyperbola would be produced by a shot at point G and a horizontal interface at depth $d.cos(\alpha)$.*

may well be less than the errors caused by using the NMO equation in the first place. Dix conversion may not help very much in these cases.

12.1.4 Effect of dip

If the source is placed at the centre of the geophone spread, the curves obtained over horizontal interfaces will be symmetrical about the source point. If, however, the reflector has a uniform dip, α, the reduction in travel path on the updip side of the shot compensates to some extent for the offset, and some travel times will be less than the normal-incidence time (Figure 12.3). The minimum time $2d.cos(\alpha)/V$ is recorded at a distance $2d.sin(\alpha)$ from the shot, on the updip side. The reflected ray rises vertically to this point, about which the moveout curve is symmetrical. Dip effects in shallow reflection surveys are detectable only for very large dips or very long spreads.

12.1.5 Multiple reflections

A wave reflected upwards with high amplitude from a subsurface interface can be reflected down again from the ground surface and then back from the same interface. This is a simple *multiple*. Two strong reflectors can generate *peg-leg* and *intraformational* multiples (Figure 12.4).

Multiples are difficult to identify with certainty on single traces. They can sometimes be recognised on multi-channel records because they have

Figure 12.4 *Multiple reflections. (a) Simple multiple. (b) Peg-leg. (c) Intra-formational multiple.*

moveouts appropriate to shallow reflectors and simple time relationships with their primaries. For example, the arrival time of the simple multiple in Figure 12.4a is approximately double the arrival time of the primary.

12.2 Reflection Surveys

Reflected waves are never first arrivals, so clear-cut reflection events are seldom seen. Oil-industry techniques for improving signal-to-noise ratios can now be also applied in shallow work, and simplified versions of the programs used are incorporated in the software supplied with the latest generation of 12- and 24-channel seismographs.

12.2.1 Spread lengths

The distance from the source to the nearest geophone in a shallow reflection survey is usually dictated by the strength of the source (and the need to protect the geophone) and may be as little as 2 m when a hammer is being used. Even with explosives or heavy weight drops, minimum offsets of more than about 10 m are unusual when observing shallow reflections.

A reflection spread can be much shorter than a refraction spread used to probe to similar depths, but with powerful sources and multi-channel recording, the furthest geophone may be more than 100 m from the source. The optimum spread length can be determined only by experiment, since the most important factors are the arrival times of the noise trains associated with the direct wave and any strong refracted waves. Field work should begin with tests specifically designed to examine these arrivals, generally by using elongated spreads.

12.2.2 Arrays

Ideally, reflected energy should not arrive until after the direct wave, ground-roll and refractions have passed, but this may not be possible if the depth of investigation is very small. In such cases, several geophones may be connected to each recording channel in arrays. Reflected waves, which travel

almost vertically, will reach all the geophones in an array almost simultaneously but the direct wave will arrive at different times and the signals it produces may interfere destructively.

The efficiency with which a wave is attenuated by an array is defined by its *relative effect* (RE) compared to the effect of the same number of geophones placed together at the array centre. The variation of the RE with *apparent wavelength*, **L** (here measured in multiples of the geophone spacing and which for the direct wave is equal to the true wavelength), for a linear array of five geophones equally spaced on a line directed towards the shot-point, is shown in Figure 12.5. An array of geophones equispaced at 2 m can be used as an example. Strong attenuation occurs for **L** values between about 1.2 and 7, i.e. between actual apparent wavelengths of 2.4 and 14 m. A 500 ms^{-1}
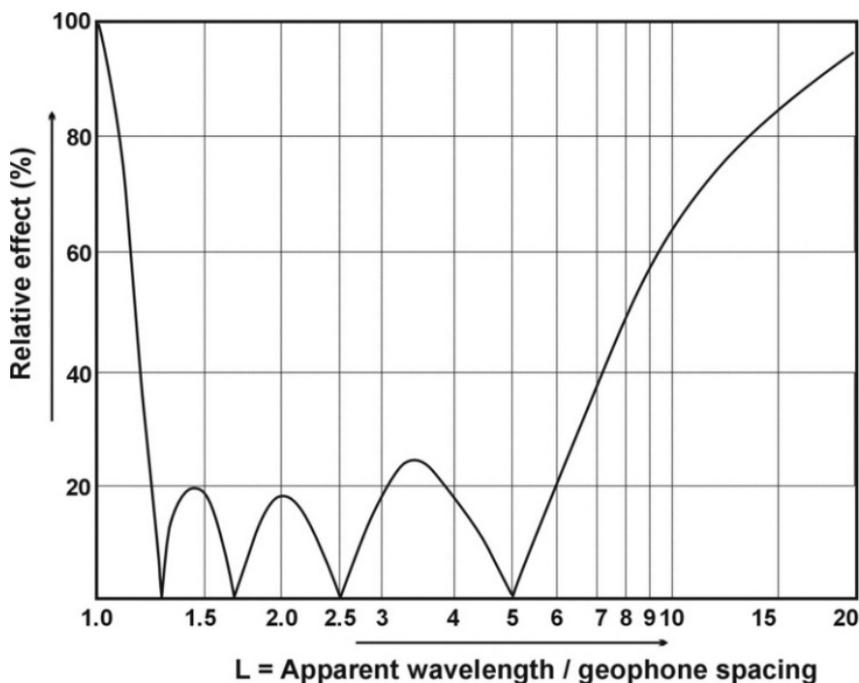
Figure 12.5 *Relative effect (RE) of an array of five equispaced in-line geophones. The apparent wavelength is equal to the actual wavelength divided by the sine of the angle between the wavefront and the ground surface. It is equal to the true wavelength for the direct wave and infinite for a wave rising vertically. The 100% RE is attained when there is zero spacing between the geophones, and also when the apparent wavelength is equal to the interelectrode spacing (*L = 1*).*

direct wave containing frequencies between 200 and 35 Hz would be strongly attenuated.

Non-linear arrays produce more complicated curves, and may perform better, but simple arrays are preferred since mistakes are less easily made in setting them out. The range of frequencies over which the direct wave is attenuated is proportional to array length, and it may be necessary to overlap the geophones in adjacent arrays. It would be unusual in a shallow survey to have more than five geophones in one array.

12.2.3   Shot arrays

Seismic cables for use with 12- or 24-channel seismographs are not designed with arrays in mind, and non-standard connectors may have to be fabricated to link the geophones to each other and to the cable. It may be easier to use arrays of shots instead.

A shot array using explosives usually involves simultaneous detonation of charges laid out in a pattern resembling that of a conventional geophone array. If an impact source is used with an enhancement instrument, the same effect can be obtained by adding together results obtained with the impact at different points. This is the simplest way of reducing the effects of surface waves when using a hammer.

12.2.4   Common mid-point shooting

Improving signal-to-noise ratios by adding together several traces (*stacking*) is fundamental to deep reflection surveys. In shallow surveys this technique was originally used only to stack (enhance) results obtained with identical source and detector positions. However, now that data are routinely recorded digitally, NMO corrections can be applied to traces produced with different source-receiver combinations. The technique normally used is to collect together traces that have the same source-receiver mid-point (*common mid-point* or CMP), apply the corrections and then stack.

The number of traces gathered together in a CMP stack defines the *fold* of coverage. Three traces forming a single synthetic zero-offset trace constitute a three-fold stack and are said to provide *300% cover*. The maximum fold obtainable, unless the shot-point and geophone line are moved together by fractions of a geophone interval (as can easily be done in marine surveys but not on land), is equal to half the number of data channels.

Figure 12.6 shows the successive geophone and source positions when a six-channel instrument is used to obtain 300% cover. Special cables and switching circuits are available for use in deep reflection surveys, but CMP fieldwork with the instruments used for shallow surveys can be very slow and laborious. Because traces from several different shots have to be combined, CMP processing can generally not be done in the field.
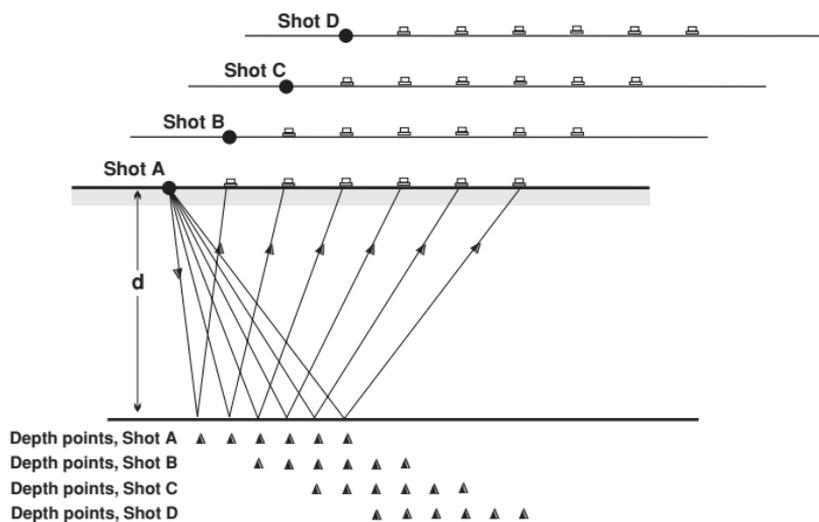
Figure 12.6 *Common mid-point (CMP) schematic, for three-fold cover with a six-channel system. Shot-points A, B, C and D are progressively one geophone group interval further to the right. Note that the distance between reflection points (depth points) on the interface is only half that between the geophone groups on the surface. Shots A and D have no depth points in common.*

The geometry of a CMP gather (Figure 12.7) differs from that for single-fold coverage since the shot-point positions as well as the geophone positions change, and the effect of dip is therefore different. The aim of stacking is to produce a noise-reduced seismic trace that approximates to the normal incidence trace, i.e. to the trace that would have been produced had the source and detector been coincident at the mid-point, and the 'depth' associated with this trace is, again, the slant distance d. However, in contrast to the situation shown in Figure 12.3, the minimum time is associated with the normal incidence ray, which travels the distance $2d$. The equations shown in Figure 12.7 replace the offset 'x' of the NMO equation with $x.\cos(\alpha)$, where the interface dips at an angle α, and the velocity deduced from a CMP stack is therefore equal to $V/\cos(\alpha)$. This is always greater than V, but by how much is generally not known, at least in the early stages of work in an area, because α is generally not known.

The initials CMP replaced an earlier acronym, CDP (*common depth point*), used for the same method. The newer term is preferable, because labelling the depth points (reflection points) as 'common' implies that all the reflections in a gather have come from the same point on the subsurface interface, which is true only for horizontal interfaces.

Figure 12.7 *Effect of dip in common mid-point (CMP) shooting. In contrast to single-fold shooting (see Figure 12.3), the shot-points as well as the geophone locations are different for the different traces. Shot-points and detector locations are interchangeable and the 'depth point', P, on the reflector moves updip as the offset increases. The moveout equation is most easily derived by noting that the path from source to detector is equal in length to the path SG' from the source to the detector 'image point', and that the geometric relationships between similar triangles imply the equality of all the lengths marked 'y'. Thus $AG' = 2$d. The Pythagoras relationship can be applied to the triangle $SG'A$ to obtain SG', and the times can be obtained by dividing the distances by* V. *Thus,* $T_0 = 2$d$/$V *and* $T = SG'/$V.

12.2.5 Depth conversion

Reflection events are recorded not in depth but in *two-way time* (TWT). Velocities are needed to convert times into depths, but the Dix velocities obtained from NMO curves (see Section 12.1.3) may be 10–20% in error, even for horizontal reflectors, and by still more where dip is involved. Interpretations should be calibrated against borehole data wherever possible,

Figure 12.8 *Geometric distortion on seismic sections. The image is of a small graben structure beneath an unconformity. The position of the true fault plane BB (indicated by the dashed line) can be estimated from the positions of the terminations of the sub-horizontal reflectors representing the sediment fill within the graben. The event AA is the seismic image of BB, but is displaced because the techniques used to display the data assume that reflections are generated from points vertically beneath the surface points, whereas they are actually generated by normal-incidence rays that are inclined to the vertical if reflected from dipping interfaces. The reflections from the fault and the opposite side of the graben cross over near the lower symbol 'A', forming a 'bow-tie'. Convex-upward reflections near point 'C' are diffraction patterns generated by the edges of fault blocks. See discussion in Section 10.3.2.*

and field crews should always be on the lookout for opportunities to measure vertical velocities directly.

12.2.6 Geometric distortion

Seismic reflection data are normally presented as sections prepared by playing out, next to each other and vertically down the sheet of paper, the traces from adjacent CMP gathers. Such sections are subject to geometric distortion. Artefacts such as displaced reflectors, diffraction patterns and 'bowties', described in Section 10.3.2 as affecting radar sections, also appear on seismic imagery, as shown in Figure 12.8.

Figure 12.9 *MicroVib shear-wave (SH) reflection source.*

12.2.7 Use of S-waves

As discussed in Section 11.1.2, S-waves (shear waves) travel more slowly than P-waves. In many materials the velocity ratio is about 1:2. This implies that the wavelength of an S-wave will be significantly less than that of the equivalent P-wave, and that higher resolution can be achieved using S-waves. This can be very useful in shallow reflection surveys. Also, S-waves are not as easily generated as P-waves, and are present to only a minor extent in traffic noise, making them ideally suited for surveys in urban environments. Another advantage of S-waves in some surveys is that they do not propagate in water. They therefore do not change velocity at the water table (i.e. V_{water} would be zero for S-waves in Example 11.1), and no S-wave reflections are produced there. This can be useful because the water table may be hard to penetrate using P-waves.

S-waves can be crudely generated by banging a spike into the ground and hitting it from the side with a hammer. This is rarely effective. The S-wave source shown in Figure 12.9 uses a swept frequency from 30 to 350 Hz over a period of 6 s. On soft ground, spikes are attached to the base of the vibrator to couple it to the ground. On tarmac and other hard surfaces, the spikes are removed and the vibrator is coupled through an array of studs protruding from its base. The coupling is further improved if a member of the survey team stands on it.

239

13

SEISMIC REFRACTION

Refraction surveys are widely used to study the water table and, for engineering purposes, the poorly consolidated layers near the ground surface, and also to determine corrections for the near-surface *low-velocity layer* (LVL) in deep reflection work. Travel times are usually only a few tens of milliseconds and there is little separation between arrivals of different types of wave or of waves that have travelled by different paths. Usually only the first-arrivals, which are always of a P-wave, can be *picked* with any confidence.

13.1 Refraction Surveys

Ideally the interfaces studied in small refraction surveys should be shallow, roughly planar and dip at less than $10°$. Provided that velocity increases with depth at each interface, the first-arrivals at the surface will come from successively deeper interfaces as distance from the shot-point increases. Survey results are displayed as plots of distance (on the horizontal axis) against arrival time (vertical axis). The gradient of any line on such a plot is the reciprocal of a velocity, i.e. steep slopes correspond to slow velocities.

13.1.1 The principal refractors

P-wave velocities for common rocks were shown schematically in Figure 11.2. In shallow refraction work it is often sufficient to consider the ground in terms of dry overburden, wet overburden and weathered and fresh bedrock. It is very difficult to deal with more than three interfaces.

The P-wave velocity of dry overburden is sometimes as low as $350\,\mathrm{m\,s^{-1}}$, the velocity of sound in air, and is seldom more than $800\,\mathrm{m\,s^{-1}}$. There is usually a slow increase with depth, which is almost impossible to measure, followed by an abrupt increase to $1500\text{–}1800\,\mathrm{m\,s^{-1}}$ at the water table.

Fresh bedrock generally has a P-wave velocity of more than $2500\,\mathrm{m\,s^{-1}}$ but is likely to be overlain by a transitional weathered layer where the velocity, which may be initially less than $2000\,\mathrm{m\,s^{-1}}$, increases with depth and the accompanying reduction in weathering.

13.1.2 Critical refraction and the head wave

Snell's Law (see Section 11.1.5) implies that if, in Figure 11.3, $\sin i = V_1/V_2$, which is possible only if V_2 is greater than V_1, then the refracted ray will travel parallel to the interface at velocity V_2.

After critical refraction, some energy will return to the ground surface as a *head wave,* represented by rays that leave the interface at the critical angle. This planar wavefront travels through the upper layer at velocity V_1 but, because of its inclination, appears to move across the ground at the V_2 velocity with which the wavefront expands below the interface. It will therefore eventually overtake the direct wave, despite the longer travel path. The *crossover* or *critical* distance for which the travel times of the direct and refracted waves are equal is:

$$x_c = 2d\sqrt{[(V_2 + V_1)/(V_2 - V_1)]}$$

This distance can be estimated from a plot of arrival time against distance (T-D plot). It is always more than twice the interface depth and is large if the depth is large or the difference in velocities is small. Simple methods of refraction interpretation use either this distance directly or the *crossover time,* which is equal to the crossover distance divided by the direct-wave velocity.

The term 'critical distance' is also sometimes used for the minimum distance at which refractions return to the surface, i.e. the distance from the shot-point at which energy arrives after reflection at the critical angle. This usage is not common amongst field crews because at this point, and for some distance beyond it, the refractions arrive after the direct wave and are difficult to observe. Use of the alternative 'crossover' terminology avoids this ambiguity.

If more than one interface is involved (Figure 13.1), multiple head waves are generated. The head wave from the nth interface makes an angle i_n with the ground surface given by:

$$\sin i_n = V_1/V_n$$

This angle, which is also the angle at which a ray must leave the source in order to be critically refracted at the nth interface, depends only on the velocities in the uppermost and lowermost layers and not on the velocities in between. For each interface, there is a corresponding crossover distance, but crossovers are difficult to locate precisely if several layers are involved and the Generalised Reciprocal (intercept-time) method discussed in Section 13.2 is preferred.
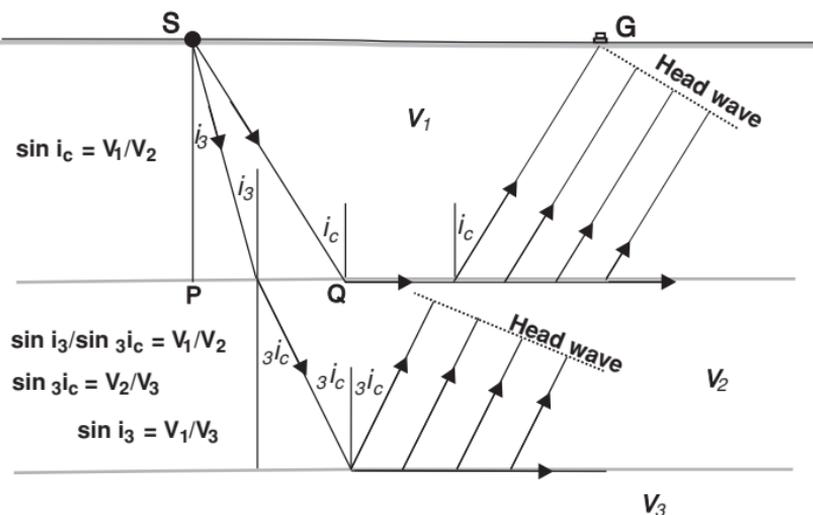
Figure 13.1 *Critical refraction at two interfaces. The relationships between the angles and the velocities all follow from Snell's Law.*

13.1.3 Lengths of refraction spreads

A line of geophones laid out for a refraction survey is known as a *spread*, the term *array* being reserved for geophones feeding a single recording channel. Arrays, which are common in reflection work, are almost unknown in refraction surveys, where the sharpest possible arrivals are needed.

Sufficient information on the direct wave and reasonable coverage of the refractor is obtained if the length of the spread is about three times the crossover distance. A simple but often inaccurate rule of thumb states that the spread length should be eight times the expected refractor depth.

13.1.4 Shot positioning

In most refraction surveys, *short shots* are fired very close to the ends of the spread. Interpretation is simplified if these shots are actually at the end-geophone positions so that travel times between shot-points are recorded directly. If this system is used, the geophone normally at the short-shot location should be moved half-way towards the next in line before the shot is actually fired (and must be replaced afterwards). Damage to the geophone is avoided and some extra information is obtained on the direct wave.

Long shots are placed sufficiently far from the spread for all the first-arrivals to have come via the deepest refractor, and short-shot data may therefore be needed to determine the minimum acceptable long-shot offsets. Distances to long shots need be measured accurately only if continuous

coverage is being obtained, on a system in which the long-shot location for one spread is a short- or centre-shot location for the next. If explosives are being used, it may be worth using very long offsets if these will allow firing in water (see Section 11.2.3).

13.1.5 Centre shots

The information provided by a conventional four-shot pattern may be supplemented by a centre shot. Centre shots are especially useful if there are considerable differences in interpretation at opposite ends of the spread, and especially if these seem to imply different numbers of refractors. They may make it possible to obtain a more reliable estimate of the velocity along an intermediate refractor or to monitor the thinning of an intermediate layer that is hidden, at one end of the spread, by refractions from greater depths. An additional reliable depth estimate is obtained that does not depend on assumptions about the ways in which the thicknesses of the various layers vary along the spread, and there will be extra data on the direct wave velocity. With long spreads, other shot positions may be occupied and are especially useful where, as in the example in Figure 13.2, there are rapid changes in subsurface conditions in a particular area (see discussion in caption to Figure 13.3).
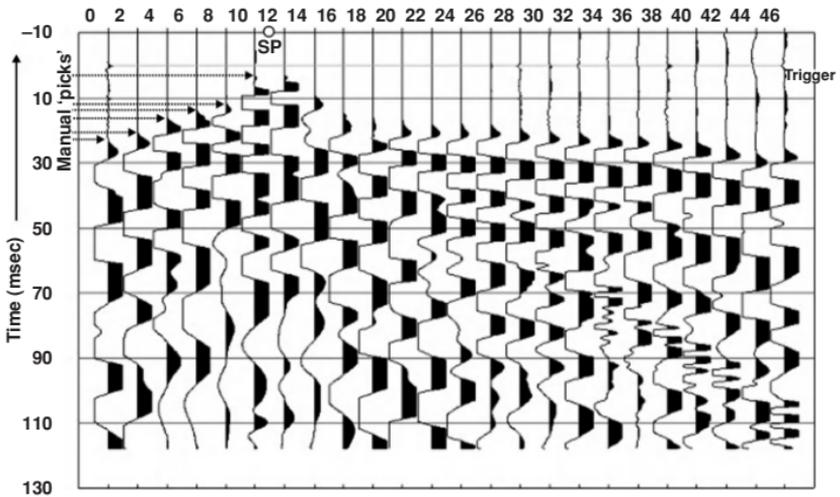
Figure 13.2 *Twenty-four-channel refraction record, with geophones spaced at 2 m intervals and the shot fired at a shot point (SP) between geophones 6 and 7. Manual first-break 'picks' at the start of a negative (right-going) pulse are identified by arrows, for the first six geophones only. This record would be considered good, and much more difficult decisions usually have to be made.*
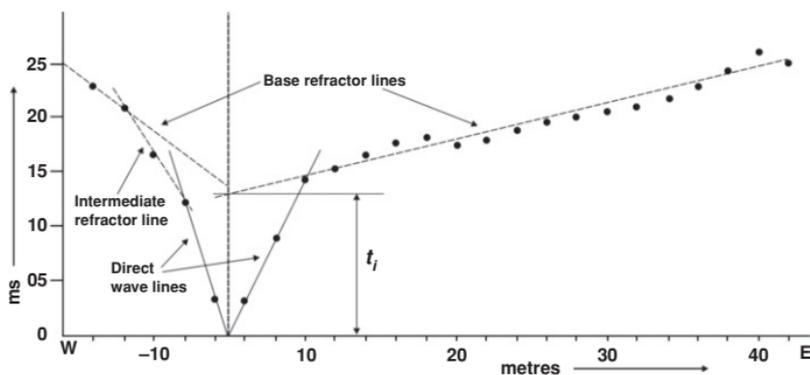
Figure 13.3 *Time-distance plot for the record shown in Figure 13.2. There appears to be an additional refractor to the west of of the shot-point, confirming this as a location at which there is a rapid change in subsurface conditions. The 'intercept time', t_i, for the refractor mapped in the east is given, to a first approximation, by the point where the 'best-fitting' straight line through the refracted arrival times cuts the vertical line through the shot-point. This point has then been used to decide which of the western arrivals have come via which refractor.*

Centre and other additional shots are used much less than they deserve. The extra effort is generally trivial compared to the work done in laying out the spread, and the additional and possibly vital information is cheaply obtained.

13.1.6 Annotation of field records

The several dozen records produced in a day's work that includes repeats, checks and tests as well as the completion of a number of different spreads must be carefully annotated if confusion is to be avoided. Annotations (*metadata*) should obviously include the date and the name of the observer-in-charge, along with the survey location and spread number. Orientation should be noted, and the position of Geophone 1 should be defined. Unless the geophone spacing is absolutely uniform, a sketch showing shot and geophone locations should be added. If the interval between timing lines on the records can be varied and/or variable time offsets can be applied, these settings must also be noted. Modern instruments store data, including the instrumental settings, digitally but hard-copy records can be (and should be) generated in the field by most of the seismographs now used for shallow refraction surveys.

Other items are optional. Amplifier gains and filter settings are not often recorded in notebooks but such information can be useful. The number of

shots or impacts combined in a single record can also be important. And, of course, features such as the use of S-wave geophones at some points or peculiarities in the locations of some of the geophones should always be noted.

Much of the metadata listed above will be printed on any hard-copy record generated, provided that they have first been entered into the machine. This is often a more tedious, and more error-prone, process than simply writing them on each record by hand, but is usually vital for identifying the individual records in digital memory.

13.1.7   Picking refraction arrivals

Picking first-arrivals on refraction records is still more reliably done by hand than by computer program, but may be difficult if the signal-to-noise ratio is poor. Some of the later peaks and troughs in the same wave train are likely to be stronger (Figure 13.2), and it is sometimes possible to work back from these to estimate the position of the first break. However, because high frequencies are selectively absorbed in the ground, the distance between the first break and any specific later peak gradually increases with increasing distance from the source. Furthermore, the trace beyond the first break is affected by many other arrivals as well as by later parts of the primary wave train, and these will modify peak and trough locations. Using later features to estimate first-arrival times is always a poor substitute for direct picking.

13.1.8   Time-distance plots

The first step in any refraction interpretation is to enter the arrival times (usually first-arrival times) on to a T-D plot (Figure 13.3). Times on these plots are measured along the vertical axes and distances along the horizontal axes, and the gradient of any line is equal to the reciprocal of a velocity.

If the arrival times lie on a number of clearly defined straight-line segments, best-fit lines may be drawn, defining velocities. These are not actually necessary if the Generalised Reciprocal interpretation method (see Section 13.2.5) is being used, and will be difficult to define if, as in Figure 13.3, the arrival times are irregularly distributed because of variations in refractor depth. It is often best to draw lines through only the direct-wave arrivals (which should plot on straight lines), leaving refracted arrivals either unjoined or linked only by faint lines between adjacent points.

In Generalised Reciprocal interpretation, all the data for a single spread are plotted on one sheet that has a working area covering only the ground where there are actually geophones (see Figure 13.9, accompanying Example 13.1). It is not necessary to extend the plot out to the long-shot positions. Since at least four sets of arrivals have to be plotted, as well as a set of time differences, different colours or symbols are needed to distinguish between data sets.
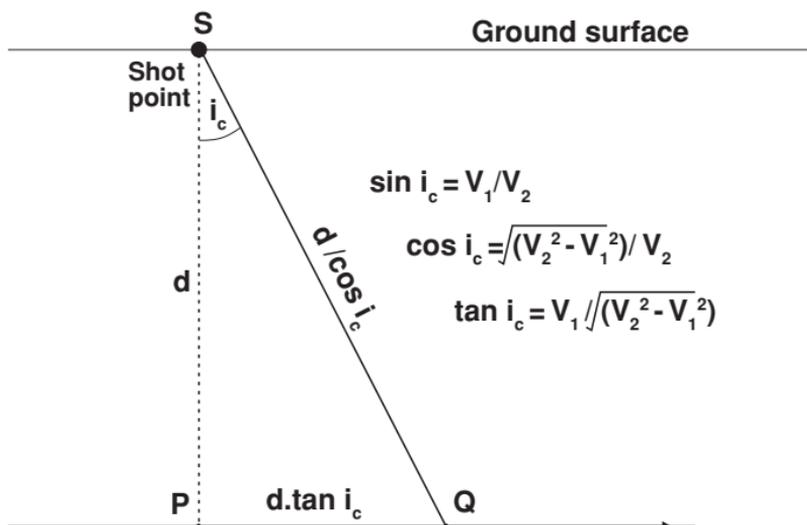
Figure 13.4 *The 'magic triangle' for calculating intercept times.* V_1 *and* V_2 *are the P-wave velocities in the upper and lower layers respectively, and* i_c *is the critical angle. For critical refraction below more than one layer, the same geometrical analysis can be applied repeatedly.*

13.2 Interpretation

Because the success of a refraction survey depends on parameters, such as line orientation, geophone spacing, shot positions and spread lengths, that can be varied almost at will, rapid 'first-pass' interpretation is essential. Only if analysis keeps pace with data collection will the right choices be made for the next day's work. Field interpretation has been made easier by computer programs that can be implemented on laptop PCs or on the seismographs themselves, but most are based on very simple models and are no substitute for actually thinking about the data.

13.2.1 Intercept times

The reason why the back-extrapolated refracted-arrival lines in Figure 13.3 do not pass through the t-x zero on the T-D plot is that it takes time for the energy to get down to the refractor and also to rise from the refractor to the detectors. For horizontal refractors these two delays are the same (Figure 13.1), and each makes up half of the total intercept time, because each is equal to the time taken, in Figure 13.4, for the energy to travel from S to Q at velocity V_1, minus the time it would have taken to travel from P to Q at velocity V_2. Simple trigonometry shows that this time is equal to:

$$[d/(V_1 . \cos i_c) - d . \tan i_c / V_2]$$

It is then necessary only to remember that $\sin i_c = V_1/V_2$ to show that the intercept time is:

$$t_i = 2d/V_{1,2} \quad \text{where} \quad V_{1,2} = V_1 V_2 / \sqrt{\left(V_2^2 - V_1^2\right)}$$

The quantity $V_{1,2}$ has the units of a velocity and is approximately equal to V_1 if V_2 is very much larger than V_1. The critical angle is then almost $90°$ and the delay suffered by the refracted ray in travelling between the surface and the refractor is roughly equal to twice the vertical travel time. If the difference between V_1 and V_2 is small, $V_{1,2}$ can be very large.

Intercept times can be estimated on T-D plots by drawing best-fit lines through the refracted arrival times (Figure 13.3), but even a very good fit is no guarantee that the depth of the refractor does not change in the region near the shot-point, from which no refractions are observed. If, however, a long shot is used as well, there should be a constant difference between long-shot and short-shot travel times at points towards the far end of the spread (Figure 13.5). An intercept time can then be obtained by subtracting this difference from the travel time between the long-shot and the short-shot locations, which can be done exactly if there is a geophone in the short-shot position when the long shot is fired. Otherwise, using the long-shot arrival at the geophone closest to the short-shot at least reduces the distance over which extrapolation must be made. Strictly speaking, the method will work only if the travel time from B to S_1 in Figure 13.5 is the same as the travel time from S_1 to D, but in practice the errors introduced by this assumption are usually small.

13.2.2 Multiple layers

The intercept-time equation can be extended to cases involving a number of layers. The intercept time associated with the nth refractor is:

$$t_n = 2d_1/V_{1,n+1} + 2d_2/V_{2,n+1} \cdots + \cdots 2d_n/V_{n,n+1}$$
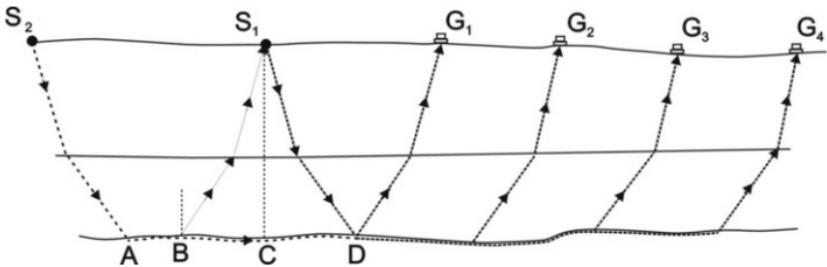
Figure 13.5 *Long-shot and short-shot travel paths for a three-layer case. The paths for energy travelling to any given geophone from S_1 and S_2 are identical once past the point D, and the long-shot/short-shot travel-time differences will therefore be the same at all detectors from G_1 onwards.*

where d_n is the thickness of the nth layer, which overlies the nth refracting interface, at which velocity increases from V_n to V_{n+1}. The definition of the various quantities $V_{m,n}$ is exactly analogous to the definition of $V_{1,2}$ cited above.

The presence of intermediate layers is often obvious but can sometimes only be recognised by comparing long- and short-shot data, since the constant long-shot–short-shot time difference exists only for head waves from the same refractor. However, at least two points on the time-distance plot are needed to define a velocity, and three for any confidence to be placed in the estimate so, at best, only four layers can be easily investigated with a 12-channel system.

Complicated field procedures can be devised to overcome this limitation; geophones may, for example, be moved one half-interval after a shot has been fired and the same shot-point can then be reused. Progress is extremely slow, and the problems presented by refractor topography, hidden layers and blind zones (see Section 13.3) still exist. In most circumstances, the firing of multiple shots into a modified spread is an attempt to extract more from the method than is really obtainable. A better approach is to use short geophone intervals with semi-continuous profiling in which shot-points are reused for short shots, long shots and sometimes centre shots, as the spread is moved on.

13.2.3 Effect of dip

Depths estimated from refraction surveys are referenced to geophone and shot-point elevations, which must therefore be measured to obtain a true picture of the subsurface refractor. Furthermore, the 'depths' determined are the perpendicular, not the vertical, distances to the refractors from the shot-points or geophones. With this proviso, 'horizontal' formulae can be applied without modification wherever the ground surface and the refractor are parallel. More usually their slopes will be different. Formulae are then most commonly quoted in terms of a horizontal ground and dipping refractors, but can equally well be applied if, for example, the ground slopes above a horizontal water table.

For the intercept-time equations to work, the true value of V_2 must be used. However, a wave that travels downdip not only has to travel further at velocity V_2 to reach more distant geophones, but also further at the slow velocity V_1 in the upper layer (Figure 13.6). It therefore arrives late; i.e. it has a low *apparent velocity*. The reverse is true of shooting updip, and in rare cases a wave may actually arrive earlier at further geophones than at nearer ones. The slope of the line through the refracted arrivals on a T-D plot depends on the dip angle, α, according to the equation:

$$V_{\text{app}} = V_2/(1 + \sin \alpha)$$
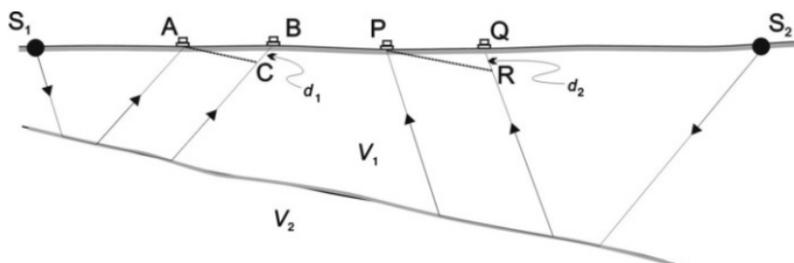
Figure 13.6 *Refraction at a dipping interface. The refracted energy from S_1 arrives later at B than at A not only because of the greater distance travelled along the refractor at velocity V_2, but also because of the extra distance d_1 (=BC) travelled at the lower velocity V_1. On the other hand, energy from S_2 arrives at P earlier than might have been expected, because it does not have to travel the distance d_2 (=QR) at velocity V_1 that is needed to reach Q from the refractor (to which the lines AC and PR are parallel).*

If shots are fired from both ends of the spread, different apparent velocities will be measured because the sign of the dip angle will differ. For dips of less than about $10°$, the true velocity is given by the dip-velocity equation:

$$2/V_2 = 1/V_{up} + 1/V_{down}$$

13.2.4 Refractor relief and true velocities

If, in the situation illustrated in Figure 13.6, and for each geophone, the travel time for energy from one long shot is subtracted from the travel time for energy from the other, the *difference line* that can be plotted will be straight and have a slope equal to the sum of the slopes of the individual lines, i.e. equal to $2/V_2$. This is a graphical expression of the dip-velocity equation.

Most refractors (apart from the water table) are irregular. If there were only a single local depression in an otherwise flat refractor, the refracted arrivals at geophones away from the depression would plot on straight lines of equal slope and, as discussed above, the differences between the two long-shot arrival times at each geophone would plot on a line with slope $2/V_2$. Both waves would arrive late at a geophone immediately above the depression, and, for small dips, the delays would be similar (Figure 13.7). The difference between the arrival times would thus be almost the same as if no depression existed, and would plot on the straight difference-line generated by the horizontal parts of the interface. Generalising from this result, it follows that, for an irregular refractor and provided that all dips are small, the time differences will still plot along a straight line with a slope corresponding to half the refractor velocity.
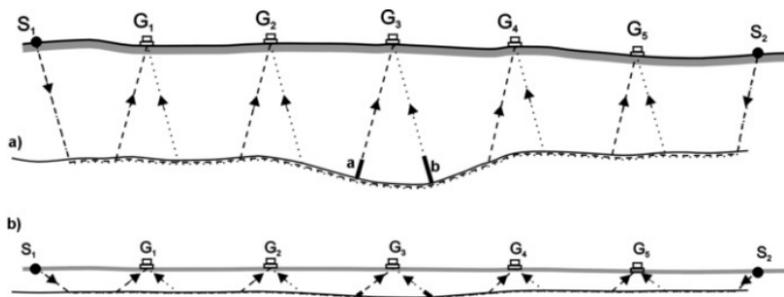
Figure 13.7 *(a) Effect on travel times of a bedrock depression. The arrivals at G_3 of energy from S_1 and S_2 are delayed by approximately the same amount (a/V_1 or b/V_1). As in all the diagrams in this chapter, vertical exaggeration has been used to clarify travel paths. (b) A more realistic, although still vertically exaggerated, picture of the likely relationships between geophone spacing and refractor depth and gradient.*

The approach described above generally works far better than the very qualitative 'proof' (and the rather contrived positioning of the geophones in Figure 13.7) might suggest. Changes in slopes of difference lines correspond to real changes in refractor velocity, so that zones of weak bedrock can be identified.

The importance of long shots is obvious. Short-shot difference lines can generally not be used, because the part of the spread over which the first-arrivals from both shots have come via the refractor will be limited and may not even exist. However, it is sometimes possible, especially when there are centre shots, for the differencing technique to be also applied to an intermediate refractor.

Difference lines are plotted using an arbitrary time-zero line placed where it will cause the least confusion with other data (see Figure 13.9). Computers can plot difference times as well as arrival times, but if graph paper is being used instead, differences are easily measured graphically and can then be transferred directly onto the T-D plot using dividers, or a pencil and a straight-edged piece of paper.

13.2.5 The Generalised Reciprocal method

In Example 13.1 the reciprocal-time method, or Generalised Reciprocal method, is used to interpret a four-shot refraction 'shoot'. This was the best method available before computers capable of tomographic modelling (see Section 13.2.6) became common in field camps. It still provides insights that are hard to obtain from a computer package.

The *reciprocal time*, t_R, is defined as the time taken for seismic energy to travel between long shots positioned beyond the opposite ends of a spread.
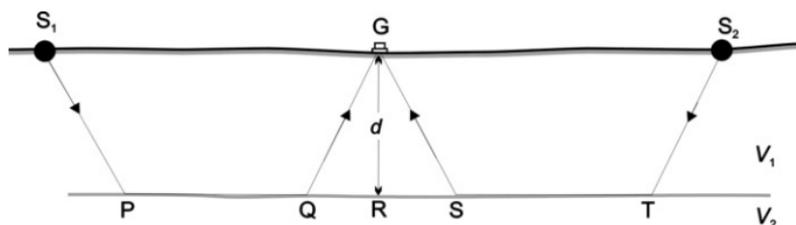
Figure 13.8 *Reciprocal time interpretation. The geometry of the triangles GQR and GRS is identical to that of the 'magic triangle' of Figure 13.4. The sum of the travel times from S_1 and S_2 to G differs from the reciprocal time, t_R, taken to travel from S_1 to S_2 by the difference between the times taken to travel QS at velocity V_2 and QGS at velocity V_1.*

In Figure 13.8, the difference between t_R and the sum of the travel times t_A and t_B from the two long shots to any geophone, G, is:

$$t_A + t_B - t_R = 2d/V_{1,2}$$

where d is the thickness of the upper layer. If there are multiple interfaces, d in this equation is replaced by D, the depth of the refractor beneath G, and $V_{1,2}$ is replaced by a *depth conversion factor*, **F**, which is a function of all the velocities involved, weighted according to the layer thicknesses. At a short shot $2D/\mathbf{F} = t_i$ (the intercept time) and the **F** value can be calculated. The ways in which **F** varies between short shots may be very complicated, but linear interpolation is usually adequate in the field (see Example 13.1).

Although t_R can be measured directly, it is more convenient to calculate it by applying the equations at the short-shot locations, where $2d/V_{1,2}$ (or its multi-layer equivalent) is the intercept time. Geophones are therefore located at the short-shot points when the long shots are fired, allowing t_A and t_B at these points to be measured. The two estimates of t_R should agree to within about 3 ms, and if they do not, the raw data and the calculations should be thoroughly checked to find the reason for the discrepancy.

Short-shot reciprocal times are measured directly if short-shots are fired from the end-geophone positions, and the fact that they should be equal may help in picking arrivals. However, they have little interpretational significance.

Example 13.1

Field interpretation of a four-shot refraction spread with long-shot (LS) and short-shot (SS) arrivals from west (W) and east (E) ends plotted on same set
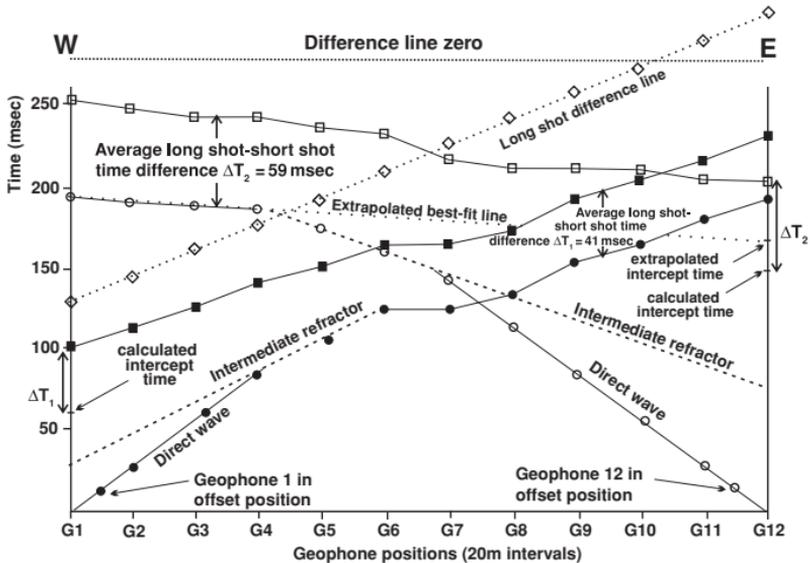
Figure 13.9 *Complete time-distance plot for a four-shot refraction spread. The long-shot difference times, indicated by diamonds, are referred to a zero line arbitrarily placed at* t = *280 ms. Note the difference at the G12 position between the intercept times obtained by extrapolation of short-shot data and by using the long-shot/short-shot difference times. Extrapolation of the western short-shot refracted arrival line back to the G1 position would be even more problematic and could lead to an even more erroneous estimate.*

of axes (Figure 13.9). After plotting the data, interpretation proceeds in the following stages (note that velocities are quoted only to the nearest $10\,\mathrm{ms}^{-1}$; even this overstates the attainable accuracy).

Stage 1 Base refractor intercept times

Measure LS(W)–SS(W) time differences: These are roughly constant and close to 41 ms from G6 to G12, indicating that in this region the SS(W) arrivals have come from the base refractor. Similarly, LS(E)–SS(E) time differences are close to 59 ms from G1 to G4.

Intercept times:
LS(W) time at W end = 101 ms. Intercept time = 101 − 41 = 60 ms.
LS(E) time at E end = 208 ms. Intercept time = 208 − 59 = 149 ms.
Note the difference of 21 ms between the LS(E) estimate and the corresponding extrapolated intercept time on Figure 13.9 of about 170 ms.

Stage 2 Velocities

Direct wave velocity: The straight line from the W origin through nearby SS(W) arrivals extends 60 m to G4.

Velocity $V_1 = 60/0.079 = 759\,\mathrm{m\,s^{-1}}$.

The straight line from the E origin through nearby SS(E) arrivals extends 100 m to G7.

Velocity $V_1 = 100/0.134 = 746\,\mathrm{m\,s^{-1}}$ Average V_1 value $= 750\,\mathrm{m\,s^{-1}}$

Intermediate refractor: Arrivals at G5 from SS(W) and at G5 and G6 from SS(E) do not belong to the 'base refractor' sets (see Stage 1) nor do they fall on the direct-wave arrival line, suggesting the presence of an intermediate, 'V_2', refractor. The V_2 velocity is poorly controlled but the arrivals lines should pass above all direct wave (V_1) and base refractor first-arrivals. For the most likely positions, as shown

$$\text{SS(W): } V_2 = 1470\,\mathrm{ms^{-1}} \quad \text{Intercept time} = 29\,\mathrm{ms}$$
$$\text{SS(E): } V_2 = 1560\,\mathrm{ms^{-1}} \quad \text{Intercept time} = 77\,\mathrm{ms}$$

These velocities suggest that the interface is probably the water table, with a velocity of about $1500\,\mathrm{m\,s^{-1}}$.

Base refractor velocity: Plot LS(W)–LS(E) time differences at each geophone, using a convenient (280-ms) line as time zero.

$$V_3 = 2/(\text{slope of difference line}) = 2 \times 220/0.182 = 2420\,\mathrm{ms^{-1}}.$$

Velocity functions:

$$V_{1,2} = V_1 \times V_2/\sqrt{(V_2^2 - V_1^2)} = 750 \times 1500/\sqrt{(1500^2 - 750^2)} = 870\,\mathrm{ms^{-1}}$$

$$V_{1,3} = V_1 \times V_3/\sqrt{(V_3^2 - V_1^2)} = 750 \times 2420/\sqrt{(2420^2 - 750^2)}$$
$$= 790\,\mathrm{ms^{-1}}$$

$$V_{2,3} = V_2 \times V_3/\sqrt{(V_2^2 - V_1^2)} = 1500 \times 2420/\sqrt{(2420^2 - 1500^2)}$$
$$= 1910\,\mathrm{ms^{-1}}$$

Stage 3 Depths at shot-points

Depths to intermediate refractor ($d_1 = \frac{1}{2}t_i\,V_{1,2}$):

$$\text{W end: } d_1 = \tfrac{1}{2} \times 0.029 \times 870 = 12.6\,\mathrm{m}$$

$$\text{E end: } d_1 = \tfrac{1}{2} \times 0.077 \times 870 = 33.5\,\mathrm{m}$$

Thickness of intermediate layer ($d_2 = \frac{1}{2}[t_i - 2d_1/V_{1,3}] \times V_{2,3}$):

$$W \text{ end: } d_2 = \frac{1}{2} \times \{0.060 - 25.2/790\} \times 1910 = 26.8 \text{ m}$$
$$D = 26.8 + 12.6 = 39.4 \text{ m}$$
$$E \text{ end: } d_2 = \frac{1}{2} \times \{0.149 - 67.0/790\} \times 1910 = 61.3 \text{ m}$$
$$D = 33.5 + 61.3 = 94.8 \text{m}$$

Stage 4 Reciprocal time interpretation (example using Geophone 8)

Reciprocal time ($t_A + t_B - t_i$):

$$W \text{ end: } t_R = 101 + 254 - 60 = 295 \text{ ms}$$
$$E \text{ end: } t_R = 233 + 208 - 149 = 292 \text{ ms}$$

Average $= 293$ ms

Depth conversion factors at short shots ($F = 2 \times D/t_i$):

$$W \text{ end: } F = 2 \times 39.4/0.060 = 1310 \text{ ms}^{-1}$$
$$E \text{ end: } F = 2 \times 94.8/0.149 = 1270 \text{ ms}^{-1}$$

Example: F at G8 (by interpolation) $= 1280 \text{ m s}^{-1}$

Depth at G8 ($D = t_A + t_B - t_R$):

$D = \frac{1}{2} \times (0.174 + 0.213 - 0.293) \times 1280 = 60.2 \text{ m}$

13.2.6 Seismic refraction imaging

Traditional methods of interpreting seismic refraction data assume a layered Earth where the layers are gently dipping ($<10°$) and the velocity structure is relatively simple. In more complex geological environments, characterised by steeply dipping and/or discontinuous refractors or strong lateral changes in velocity, modelling techniques that assume continuous velocity gradients can provide more realistic results. Software packages are now available that use finite element or finite difference approaches to automatically model varying velocity gradients, without any prior assumptions of subsurface structure. First-arrival picks are used to develop a best-fit velocity model by iteratively comparing the modelled first-arrivals from different velocity distributions with the observed data. This approach (Figure 13.10) is referred to by some workers as refraction tomography, and the interested reader is referred to an allied discussion on nomenclature in Section 6.5.1.

Figure 13.10 shows that perfect agreement between the observed and modelled T-D plots was not achieved. Moreover, and in contrast to traditional methods, which require some understanding of geology as well as
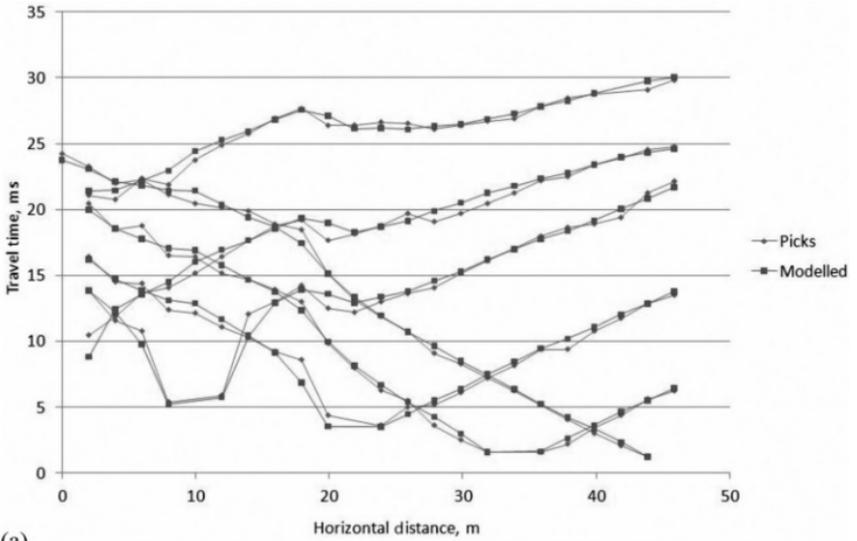
(a)
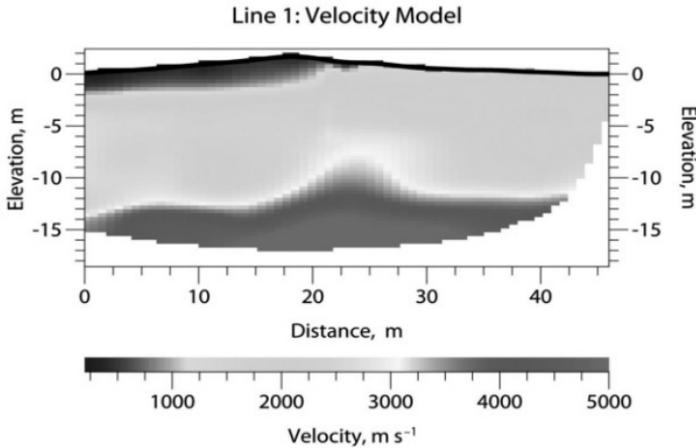
(b)

Figure 13.10 *Refraction modelling. (a) Comparison of first-arrival picks with modelled first-arrivals using a commercially available non-linear optimisation routine to obtain the optimal subsurface velocity distribution. (b) Continuous velocity model that gave rise to the modelled arrivals in (a). The interpreted bedrock depth of between 10 and 13 m was verified by a borehole.*

seismic principles, automatic imaging programs produce subsurface velocity models without human intervention and should be treated with healthy scepticism. Mis-picked arrivals will be treated as real by the computer, but would usually be recognised in manual interpretation. Moreover, if the actual velocity change at a boundary were to be greater than the maximum velocity change allowed by the software, the interface would be spread out, with overestimation of the velocities on the low-velocity side and underestimation of velocities on the high-velocity side. This could lead to an erroneous choice of ripping equipment (see Figure 11.2). Even the basic intercept-time method (see Section 13.2.1) would give a better result for shallow dipping layers in this case. It is always prudent to check the results of one type of modelling against another, to increase confidence in the interpretation, but imaging techniques have, generally speaking, significantly improved the interpretation of refraction data in complex geological settings and have radically reduced interpretation time and effort.

13.3   Limitations of the Refraction Method

First-arrival refraction work uses only a small proportion of the information contained in the seismic traces, and it is not surprising that interpretation is subject to severe limitations. These are especially important in engineering work; in low-velocity-layer studies only a time delay estimate is sought and short shots alone are often sufficient.

13.3.1   Direct waves

The *ground roll* consists of a complex of P and S body waves and Love and Rayleigh surface waves travelling with different but generally slow velocities. There is often some doubt as to which component actually produces the first break, since conventional geophones respond only poorly to the horizontal ground motions of direct P-waves. Close to the source, enough energy is associated with the P-waves for the response to be measurable, but at greater distances the first breaks may record the arrival of S-waves, surface waves or even the air wave.

The complex character of the direct wave may be amongst the reasons for the commonly observed failure of best-fit arrival lines to pass through the origin. Delays in the timing circuits may also play a part but can be measured directly using a detonator or a light hammer blow close to a geophone. A more important reason may be that the amplifier gains at geophones close to the shot-point are often set so low that the true first-arrivals are overlooked (Figure 13.11). Full digital storage of the incoming signals should allow the traces to be examined individually over a range of amplifications, but if this is not possible, then the most reliable velocity estimates may be those that do not treat the origin as a point on the line.
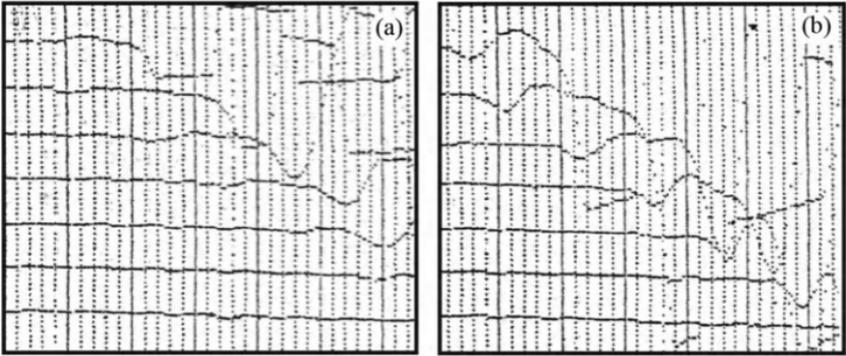
Figure 13.11 *Hard copy of a single stored data set played back at two different amplifications. First-arrivals that would probably be overlooked or dismissed as noise on (a) are clearly visible on (b). A direct-wave velocity based on the events clearly imaged on (a) would be roughly correct provided that the best-fit line was not forced through the origin. The crossover distance would also be wrong but the intercept time would not be affected, being based on the refracted arrivals only.*

13.3.2   Vertical velocities

However much care is taken to obtain valid direct-wave or refracted-wave velocities, the refraction method is fundamentally flawed in that the depth equations require vertical velocities but what are actually measured are horizontal velocities. If there is significant anisotropy, errors will be introduced. This is a problem for interpreters rather than field observers, but the latter should at least be aware of the importance of using any boreholes or recent excavations for calibration or to measure vertical velocities directly.

13.3.3   Hidden layers

A refractor that does not give rise to any first-arrivals, because the head wave from it is overtaken by the one from a deeper refractor before it itself overtakes the direct wave, is said to be *hidden*. A layer is likely to be hidden if it is much thinner than the layer above or has a much lower seismic velocity than the layer below. Weathered layers immediately above basement are often hidden. The presence of a hidden layer can sometimes be recognised from second-arrivals but this is only occasionally possible, in part because refracted waves are strongly attenuated in thin layers.

A layer may also be hidden, even if it produces a head wave that does arrive first over some part of the ground surface, if there are no appropriately located geophones. Concentrating geophones in the critical region can sometimes

be useful (although never convenient) but the need to do so will only be recognised if preliminary interpretations are being made on a daily basis.

13.3.4   Blind zones

If velocity decreases at an interface, critical refraction cannot occur and no refracted energy returns to the surface. Little can be done about these *blind* interfaces unless vertical velocities can be measured directly. Surface wave methods (see Chapter 14) may have to be used.

Thin high-velocity layers such as perched water tables and buried terraces often rest on blind interfaces. The refracted waves within thin layers lose energy rapidly with increasing distance from the source and ultimately become undetectable. Much later events may then be picked as first-arrivals, producing discontinuities in the time-distance plot. A similar effect is seen if a layer ends abruptly.

13.3.5   Limitations of drilling

Despite the limitations of refraction surveys, interpretations are not always wrong when they disagree with drill-hole data. Only a very small subsurface volume is sampled by the drill, and many drill tests of drift thickness have been terminated in isolated boulders some distance above the true top-of-bedrock. It is always important to find explanations for differences between drilling and seismic results.

14

SEISMIC SURFACE WAVE METHODS

The growth in the use of seismic surface waves in earthquake and foundation engineering over the past decade has been remarkable. Their main attraction is the ability to derive values of shear wave velocities, and hence shear moduli, at depths ranging from less than a metre to 100 m below the surface, as a practical alternative to drilling expensive boreholes.

14.1 Surface Wave Surveys

An important advantage of using surface waves in studies of the near subsurface is that the limitation of the refraction method to cases where velocities increase with depth does not apply.

14.1.1 Surface wave fundamentals

Rayleigh and Love surface waves (Section 11.1.1) travel more slowly than P and S body waves but carry considerable amounts of energy and cause the most damage to surface structures in earthquakes. In normal circumstances, more than two-thirds of the total seismic energy generated by a compression source will be converted to Rayleigh waves, which are the main component of the 'ground roll'. Particles follow retrograde elliptical paths in vertical planes that are aligned along the propagation paths (Figure 14.1). Love waves are horizontally polarised shear waves reflected between the surface and the base of a low-velocity layer. They are significant only if a layer with a much higher S-wave velocity underlies the low-velocity layer, but this is a very common situation.

Surface waves of both types decay exponentially in amplitude with depth, and are *dispersive*, i.e. different frequency components travel at different velocities. Of the two, the Rayleigh waves are the most important in engineering geophysics, as their velocities are related to those of the shear waves in the same elastic media. The exact relationship depends on the Poisson ratio, but for most geological materials, V_R, the Rayleigh wave velocity, is between 0.91 to 0.955 of V_S, the shear wave velocity. Approximating shear wave velocities by Rayleigh wave velocities, even without applying a correction factor, thus introduces an error of less than
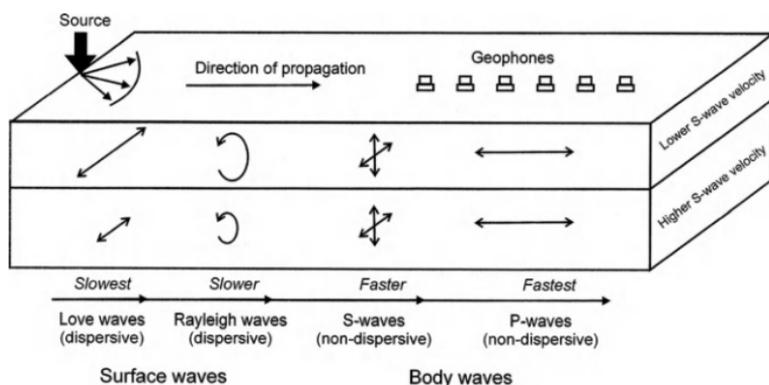
Figure 14.1 *Body and surface seismic waves generated by an impact or vibratory source and received by an array of geophones. The directions of particle motion in the ground are shown, in simplified form, by the arrowed lines. The relative ordering of transmission velocities is also indicated. Love and Rayleigh waves are dispersive, i.e. components with different frequencies travel at different velocities. The presence of the interface at which velocity increases is essential for the generation of Love waves.*

10% across a range of materials. Shear wave velocities (see Table 14.1) are in turn related to the maximum shear modulus or stiffness by the formula:

$$V_s = \sqrt{(G_{max}/\rho)}$$

G_{max} replaces μ in this version of the formula originally quoted in Section 11.1.2, to emphasise its role as the upper bound for the bulk stiffness

Table 14.1 *Typical shear-wave velocities for some common geological materials*

Material	Shear-wave velocity ($m\,s^{-1}$)
Soft muds	<200
Dry sand	300–600
Wet sand	700–900
Clays	500–800
Tills	1000–1200
Sandstone	1600–2600
Shale	2200–2400
Limestone	2500–3100
Granite	3200–3800
Basalt	3400–4000

of materials at the low strains associated with soil-structure interactions. Shear wave velocity profiles can be used to predict deformation in response to earthquakes (including earthquake-induced liquefaction analysis), in soil compaction control, in locating weak zones in embankments and in classifying soils according to the International Building Code (IBC).

Dispersion is the key to using Rayleigh waves for shear modulus depth profiling. Longer wavelengths penetrate deeper into the ground and generally have higher velocities. The velocities at specific frequencies are often referred to as the *phase velocities*, since they are manifest in the changing phase relationships between components of different frequencies. Velocities, and therefore wavelengths, can be estimated by measuring travel times between known points at selected frequencies, and the results can be used to build up wavelength-frequency or phase velocity-frequency dispersion curves. Depths can then be assigned to each velocity value using a *factored wavelength method* (typically by taking the depth of investigation to be equal to one-third of the wavelength). This may be adequate if only a few data points are available and there are no abrupt changes in shear wave velocities. Alternatively, full 1D inversion can be used to fit layered Earth models to the measured dispersion curves. In some cases, slowness-frequency curves are used (where slowness is the reciprocal of velocity).

There are three main steps in carrying out surface-wave seismic surveys:

1. Collect seismic data with a system designed to detect wavelengths (see Section 14.1.2).
2. Identify the Rayleigh wave fundamental mode, determine the variation in velocity as a function of frequency and present this as a dispersion curve.
3. Use the dispersion curve to determine the stiffness structure that matches the measured dispersion.

14.1.2 Types of survey

Seismic surface waves can be produced using impact sources such as sledgehammers, weight drops and vibrators (*active* sources) and are measured at frequencies typically between 3 and 30 Hz. Surface waves are also generated by random 'natural', 'passive' or 'cultural' sources such as wind, thunder, site construction activities and vehicular and pedestrian traffic. These can be merely noise but are also used in some methods as sources of signal. They can carry sufficient low-frequency energy for results to be obtained to depths of 100 m. Of the active sources, sledgehammers generate relatively high frequencies that allow exploration to a maximum depth of 10–30 m, depending on material type, whereas the Vibroseis trucks used for deep

seismic reflection work can produce usable surface wave data from depths of as much as 100 m.

Acronyms abound in surface wave studies and seem deliberately designed to confuse. The simplest possible (Common-Offset Rayleigh Wave) method uses a hammer source and a single geophone with a resonant frequency below 5 Hz. The two are moved together to carry out the survey. When the geophone or accelerometer is left in one place but a number of different source positions are used, the first acronym (MISW - Multiple Impact Surface Wave) is introduced. Add additional geophones in a linear array and the possibility of using a weight drop source, and the method becomes SASW (Spectral Analysis of Surface Waves). Switching to a controlled vibrator source, operating typically at between 5 and 600 Hz, may also involve a change of acronym to CSWS (Continuous Surface Wave Seismics). Use of random sources such as passing vehicles and site activities such as piling or drilling, with either linear or 2-D arrays of detectors, is implied by the MSM (Microtremor Survey Method) acronym. The recording times for MSM and other random-source methods should not be less than half a minute. Add a hammer wielded at random to increase the high-frequency content or a weight drop, also untimed, for deeper penetration, and the survey might be described as ReMi (Refraction Microtremor), although most users would restrict this term to surveys with linear detector arrays. Use almost every possible combination of random (untimed) and controlled (possibly timed) sources, and of linear and 2-D arrays and use also the acronym MASW (Multi-channel Analysis of Surface Waves). The names are not important, as long as the limitations are understood.

Surface wave arrivals are also unavoidably recorded in seismic reflection and refraction surveys, although usually attenuated in reflection surveys by the use of arrays (Section 12.2.2), and these can be processed and interpreted to produce shear-wave velocity models. Combination surveys are becoming increasingly popular.

All of the methods listed have in common the recording of time-domain seismic data, transformation into the frequency domain, conversion to dispersion plots of phase velocity (or wavelength) against frequency, and modelling to produce shear-wave velocity-depth profiles. Differences arise not only in the field parameters but in the methods used to derive the dispersion curves and in the modelling techniques used to produce the profiles.

14.1.3   Recording arrays

Linear arrays used for surface-wave studies are similar to those used in refraction and reflection surveys, but low-frequency geophones are usually required to reach the required target depths. For combined (ReMi)/refraction surveys it is usual to set out and switch between two rows of geophones.
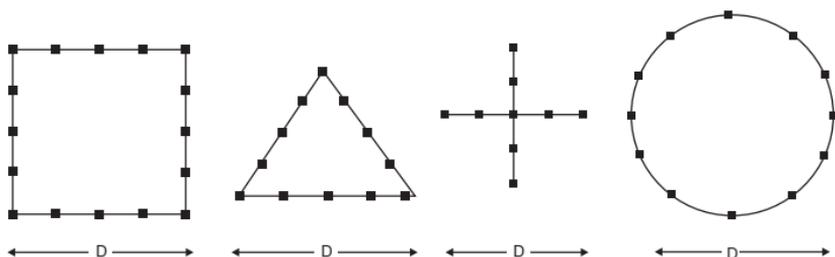
Figure 14.2 *Examples of symmetric 2D arrays for passive surface-wave surveys. D is the array maximum dimension and is related to the depth of investigation. The geophone spacing determines the minimum depth of investigation.*

A range of symmetric 2D arrays has been tested by the Kansas Geological Society and Figure 14.2 shows examples suitable for random-source MASW surveys.

14.1.4 Survey design

As a rule of thumb, the longest wavelength that can be detected, which in turn determines the maximum depth of investigation, is equal to the length of the geophone spread, D. The geophone spacing defines the shortest wavelength that can be detected and the shallowest depth of investigation. Near-field source effects are important and must be avoided, but there is currently little agreement concerning the minimum safe distance. Estimates vary between $D/5$ and $D/2$.

A typical record length for an active-source survey is 2 s, with a 0.5- or 1-ms sampling interval, but this may be increased over particularly slow materials. Records are stacked to maximise the signal-to-noise ratio (SNR). Passive surveys generally require multiple records, each of the order of 30 s long, sampled at intervals certainly no longer than 4 ms and preferably 2 ms. These records cannot be stacked.

The ReMi approach of combining random passive sources with untimed 'active' sources, typically sledgehammers, to increase the high-frequency content has proven very effective and can be preferable to carrying out separate active and passive MASW surveys, which might well require more time on site than is actually available. Geophone land streamers (see Section 11.3.1) are now being routinely used to maximise the productivity of linear-array surveys over relatively smooth surfaces, using both controlled and random sources. Geophones with resonant frequencies of 4.5 Hz or less are needed to get 50-m depth penetration, depending on the subsurface materials. For a 100 m depth, 1-Hz geophones are normally required.

The effect of topography is negligible for surfaces with constant slope but can be significant if elevations vary by more than 10% of the array length. 2D arrays are recommended by some workers for accurate evaluation and correction of Rayleigh wave directional effects. It has been claimed that velocity errors of up to 25% can result if these are not taken into account, but 2D deployment may not be possible in built-up areas.

14.1.5   Earthquake hazards

The parameter V_s is important in standards such as the International Building Code (IBC), where it is used for ground-motion amplification hazard ranking, and is also widely used in structure design in earthquake prone areas. Table 14.2 shows a soil hazard ranking scheme based on V_s values. Drilling and logging to the depths required for earthquake hazard investigations is expensive and often impractical in urban settings and these facts have been behind much of the development in seismic surface-wave methods during the past decade.

14.2   Data Processing

Fast Fourier Transforms (FFTs) are used to convert the time-domain field data into the frequency domain, and derive phase differences. The Rayleigh wave velocities and wavelengths are then computed using the distances and phase differences between receivers, via the relationship:

$$V_R(f) = 2\pi f \frac{\Delta x}{\Delta \phi}$$

where $\Delta \phi$ is the phase difference and Δx is the geophone spacing. The Rayleigh wavelength is then:

$$\lambda_R(f) = \frac{V_R}{f}$$

The f-k (frequency-wavenumber) and p-tau (slowness-intercept time) transforms are effective at isolating fundamental-mode Rayleigh wave energy from the higher harmonics, body waves and other forms of noise in the ground roll data.

14.2.1   Dispersion curves

A variety of techniques are available to calculate dispersion curves. The CSWS method utilises frequency information to plot dispersion curves of phase-difference against frequency, and wavelength against frequency,

Table 14.2 *Example of soil categorisation based on shear-wave velocity. Five soil types are defined on the basis of their shear-wave velocities (*V_s*) and susceptibility to earthquake shaking amplification. The definitions are those applicable in the Bay Area around San Francisco*

Category	Description
A: $V_s > 1500\,\mathrm{m\,s^{-1}}$	Soil type A includes unweathered intrusive igneous rock. Soil types A and B do not contribute greatly to shaking amplification
B: $1500\,\mathrm{m\,s^{-1}} > V_s > 750\,\mathrm{m\,s^{-1}}$	Soil type B includes volcanics, most Mesozoic bedrock, and some Franciscan bedrock. Does not contribute greatly to shaking amplification
C: $750\,\mathrm{m\,s^{-1}} > V_s > 350\,\mathrm{m\,s^{-1}}$	Soil type C includes some Quaternary sands, sandstones and mudstones, some Upper Tertiary sandstones, mudstones and limestone, some Lower Tertiary mudstones and sandstones, and Franciscan mélange and serpentinite
D: $350\,\mathrm{m\,s^{-1}} > V_s > 200\,\mathrm{m\,s^{-1}}$	Soil type D includes some Quaternary muds, sands, gravels and silts. **Significant amplification of shaking by these soils is generally expected**
E: $200\,\mathrm{m\,s^{-1}} > V_s$	Soil type E includes water-saturated mud and artificial fill. **The strongest amplification of shaking is expected for this soil type**

Source: National Earthquake Hazards Reduction Program (NEHRP), USA.

as shown in the example in Figure 14.3. The MASW method uses an (f-k) transform to calculate a phase velocity-frequency dispersion image (Figure 14.4). The ReMi method uses a p-tau (also known as or slant-stack) transformation to produce a slowness-frequency dispersion image (Figure 14.5). Because ReMi sources can be in any direction, this transform is applied for multiple directions through the geophone array and summed.

ReMi and MASW dispersion curvces are obtained in different ways. For ReMi data, there is uncertainty in the arrival azimuth of the Rayleigh wave
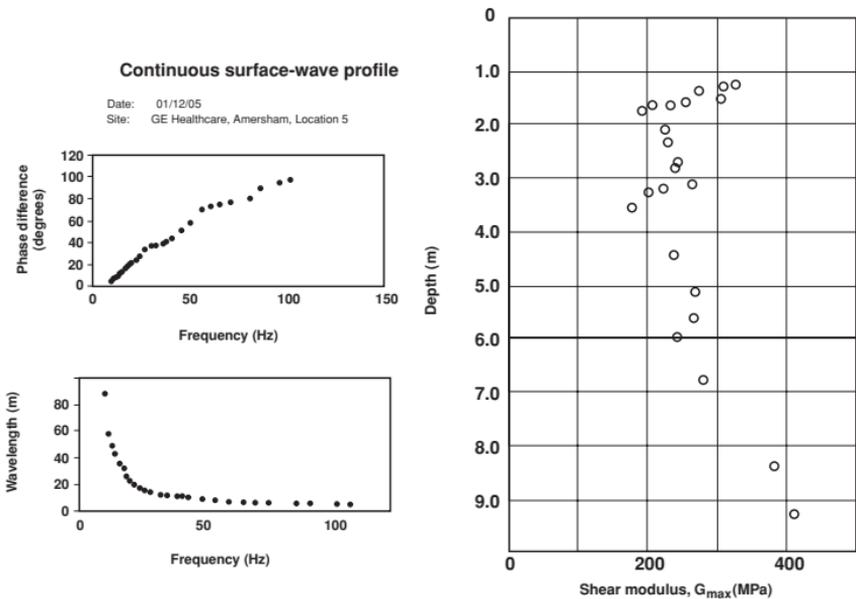
Figure 14.3 *Phase angle-frequency and wavelength-frequency dispersion curves with modelled shear modulus-depth profile. Data from a CSWS survey using a 70-kg weight ground vibrator with 489 N force (see Figure 11.5).*
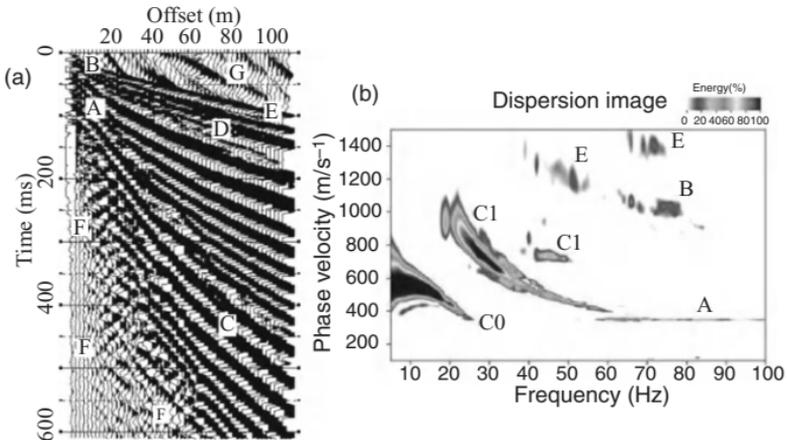
Figure 14.4 *(a) MASW seismic data obtained using active and passive sources. (b) Wavefield transformation into a dispersion image that separates out the Rayleigh wave fundamental (C0) from its higher harmonics (C1), an air wave (A) and reflections and refractions (E and B). A dispersion curve is manually or automatically picked by tracing the peak of C0 (and higher harmonic modes if required). Reproduced with kind permission of C. Park, Kansas Geological Survey.*
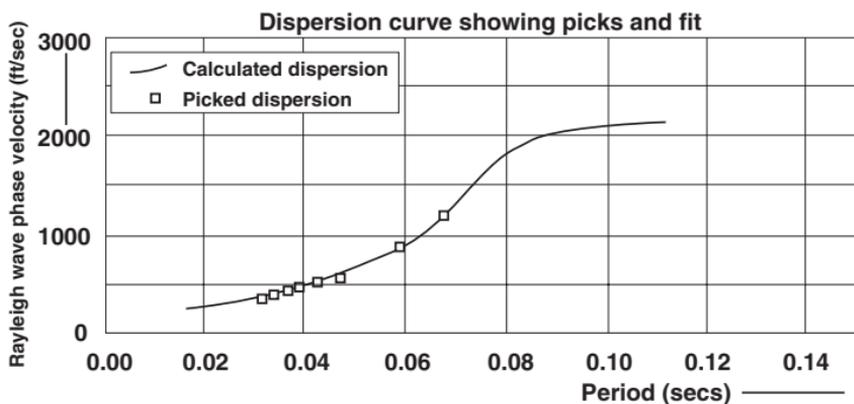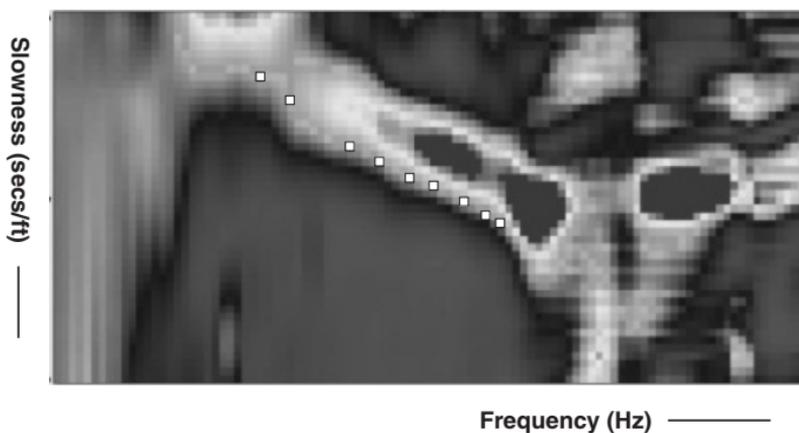
Figure 14.5 *Example of (ReMi) wavefield transformation to a slowness-frequency dispersion image (bottom) with dispersion picks tracing a dispersion curve along the leading edge of the Rayleigh wave fundamental mode. The modelled dispersion curve is also shown (top).*

energy, and picking is done along a lowest velocity envelope bounding the Rayleigh wave fundamental (Figure 14.5). The p-tau slowness-frequency dispersion image takes into account waves travelling in different directions. The MASW f-k approach is more sensitive to direction, but 2D receiver arrays can go a long way towards dealing with this. MASW dispersion curves are picked on the spectral peak of the Rayleigh wave fundamental.

269

Dispersion curves can also be extracted from multi-fold seismic reflection data and can be inverted using laterally constrained inversion techniques to obtain pseudo-2D models of the shear-wave velocity. This method gives better results than the individual inversion of single dispersion curves.

14.2.2 Modelling

Forward modelling or linear least-squares inversion of dispersion curves requires initial values for Poisson ratios and densities which can be based on the known properties of local materials.

Modelled shear-wave velocity profiles are insensitive to reasonable variations in density, but differences in the P-wave velocity, Vp, in saturated and unsaturated sediments (which cannot be obtained from the dispersion data) can imply differences of up to 10–20% in surface wave velocities, because of the dependence of the Rayleigh wave velocity on the Poisson ratio. The ratio can be obtained by a combination of P-wave refraction and MASW (or ReMi) surveys, at least to the maximum refractor depth.

Forward modelling is iterative and starts with an assumed shear-wave velocity profile from which a theoretical dispersion curve can be derived. The shear-wave model is adjusted until the measured dispersion curve and the theoretical dispersion curve match. Linear inversion depends heavily on the choice of initial model. If *a priori* information is unavailable, multiple inversion solutions with equally good data fits are possible.

Calculation methods for fully automated 1D inversion, including simulated annealing, are beyond the scope of this book. The starting points are multi-layer models with predefined thicknesses, shear-wave velocities, Poisson ratios and densities. More layers are used than are believed to be present.

Modelling by 1D inversion or even pseudo-2D inversion (Figure 14.6) is currently limited to flat-lying layers. The future development of full 2D and 3D modelling will take into account wavefield scattering and thus further improve the accuracy and robustness of the method.

14.3 Limitations of the Method

Surface-wave methods are at an early stage of development, and many problems remain. Some limitations are fundamental, but improvements can be expected in many areas. All methods currently suffer from near-field incompatibility effects. In some methods (e.g. SASW and CSWS) the Rayleigh wave fundamental cannot be easily separated from higher harmonics and body waves, and with these methods it has currently to be assumed that the fundamental mode is dominant. This can cause errors in defining the dispersion curves for modelling.
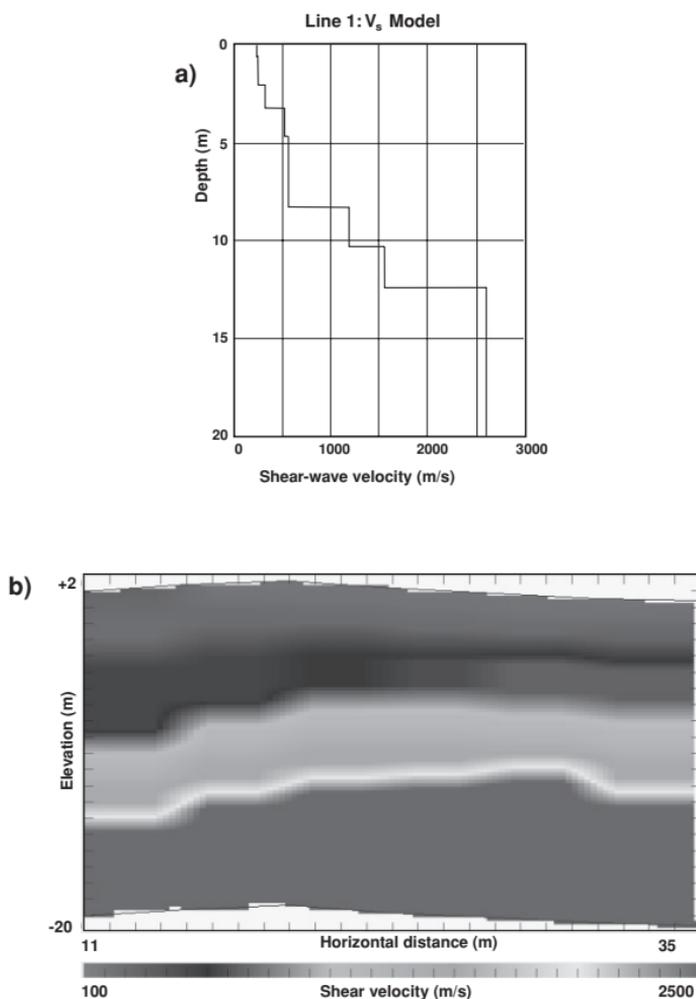
Figure 14.6 *Examples of ReMi V_s 1D model (a) and 2D model (b) based on the data in Figure 14.5.*

14.3.1 Depth of penetration and resolution

The depth of penetration in a surface-wave survey is determined by the longest wavelength that can be generated by the source, measured accurately in the field and resolved in modelling. Generally speaking, heavier sources generate longer wavelengths, but unfavourable site conditions, such as unconsolidated made ground and materials such as peat, may drastically attenuate the signals. Cultural noise, whilst necessary for passive methods,
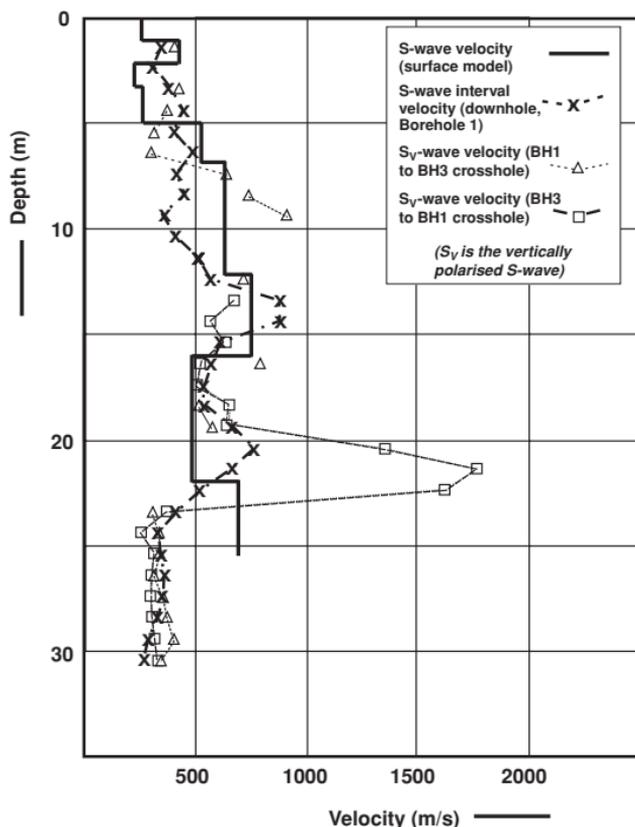
Figure 14.7 *Comparison of 1D shear-wave velocity model derived from a ReMi surface survey with downhole and crosshole measurements. The agreement is reasonable from surface to 20 m below ground level. Thereafter discrepancies become significant. Data collected at Zetica's test site.*

adversely affects the signal-to-noise ratio in active methods. Smoothing may be required, reducing resolution.

Layer-thickness resolution decreases with depth. As a general rule of thumb, the minimum thickness that can be resolved is of the order of one-fifth of the layer depth.

A comparison of borehole (crosshole and downhole) measurements of shear-wave velocity with surface wave (ReMi) estimates is shown in Figure 14.7. There is a reasonable match down to about 15–20 m below ground level, beyond which, at this site, the resolution of the surface wave method falls off quickly.

15

MAPS, MAPPING AND GPS

Advances in technology seldom make life easier. They make possible what was previously impossible, but can make life more complicated in doing so. This certainly seems to be true of the impact of the Global Positioning System (GPS) on geophysics. Even the terminology is complicated. GPS is just one type of Global Navigation Satellite System (GNSS) or Radio Navigation Satellite Service (RNSS), these being generic terms for any technology that uses satellites to determine position. GPS is the United States' contribution and was the first in the field, but other countries have developed or are developing alternatives.

15.1 Maps and Mapping

Before GPS became available, field crews seldom needed to worry about such things as map projections and ellipsoids, but now they have instruments that require that sort of information. Moreover, they have had to get to grips with a suite of new (and constantly changing) acronyms.

15.1.1 Map projections

Mapping systems are attempts to represent on a flat piece of paper a part of the roughly spherical surface of the Earth. Local topographic variations are usually far more obvious than Earth curvature, but because satellite positioning has made it possible for even small surveys to be referenced to national mapping systems, which do take curvature into account, it is now usually a contractual requirement that this be done. Failure to appreciate the issues involved can result in positioning errors of several hundred metres.

Latitudes and longitudes define positions on the Earth in a spherical polar co-ordinate system. Ideal for many purposes and essential for work covering large parts of the Earth's surface, these *geographic* co-ordinates are inconvenient for small areas, where a grid based on a standard unit of length, now almost universally the metre, is preferable. To use such a system it is necessary to treat the part of the Earth's surface under consideration as planar. This is the process of projection, and involves distortion that can be minimised but never eliminated. Projections incorporate scale factors for

transforming between geographic and grid co-ordinates and, because lines of longitude on the surface of the Earth converge but lines of latitude do not, at least one of the scale factors must change over the surface of any map, no matter what the projection. A projection in which both scale factors are adjusted in the same way, so that scale is independent of direction, is said to be *orthomorphic*. Most modern maps use orthomorphic projections.

The two main projection systems map the spherical Earth surface onto the surfaces of either cylinders or cones. In the orthomorphic Mercator projection, the mapping information is projected onto a cylindrical surface that is tangential to the Earth at the Equator. Distortion is minimal near the Equator but increases towards the poles (rendering Greenland enormous on Mercator maps).

The low distortion that characterises Mercator maps in equatorial regions can be achieved in other areas by redefining the Equator. For the Transverse Mercator (TM) projection, the regular Mercator projection is rotated through $90°$ so that one of the lines of longitude (*meridians*) effectively replaces the Equator, and distortion is small for points within about three longitude degrees of this line. The system is ideal for countries such as Chile, which cover a wide range of latitudes but only a small range of longitudes. Countries, or areas, that are elongated E–W rather than N–S are less well served, and mapping onto the surface of a cone, generally with its apex on an extension of the Earth's spin axis, may be preferred. The surface of the cone may either be tangential to the Earth's surface, giving zero distortion at a single latitude, or may cut the surface. The first option is the simpler but the second, with zero distortion at two latitudes, is preferred where the mapped area has significant N–S as well as the E–W extent.

Once a projection has been chosen, a linear Cartesian grid can be established, and this requires a point of origin. This point should be in a region of minimal distortion, e.g. on the central meridian for a transverse Mercator system or on the contact latitude in a conic system with one standard parallel. The origin defined in this way is not usually assigned zero co-ordinate values. Instead, values are chosen so that negative co-ordinates do not occur anywhere within the region of interest.

The Universal Transverse Mercator (UTM) system, which is centred on meridians spaced at 6-degree intervals, beginning at $3°$E and W of Greenwich, is now widely accepted as a global standard. The prime meridian through Greenwich is a zone boundary in this system, and UTM co-ordinates therefore change abruptly and discontinuously in east London. This would create enormous problems for mapping in the UK, and the central meridian of the British National Grid is therefore set at $2°$W – i.e. the grid uses a TM, but not the UTM, projection.

15.1.2 Ellipsoids

The Earth is, to a first approximation, spherical, but the equatorial radius is about 22 km longer than the polar radius. An ellipsoid of revolution, also known as a spheroid (in surveying, the terms are to all intents and purposes synonymous) is commonly used as a second-order approximation. Ellipsoids are characterised by the lengths of their major and minor axes, or by the length of the major axis and a flattening factor (eccentricity), and many different Earth ellipsoids have been used over the last 200 years. This has not only been because of the steady advances in knowledge of the true shape of the Earth over that period, but also because all ellipsoids are approximations, and some fit given parts of the Earth's surface better than others.

Latitudes and longitudes define points on an ellipsoid surface that approximates the Earth's surface. To use these coordinates accurately requires knowledge of which ellipsoid is being used. The currently most widely used system, the 1984 World Geodetic System (WGS84), with major and minor axes of 6378.137 and 6356.752 314 km respectively, is a poor fit in some parts of the world, where local and traditional systems continue to be used.

Choosing an ellipsoid does not, of itself, define a mapping system. It is also necessary to define the location of its centre with respect to the true centre of the Earth. This is now usually done in terms of its X, Y and Z coordinates in a Cartesian system, so that transformations between different systems using the same ellipsoid may be expressed (rather mysteriously to the uninitiated) in terms of a ΔZ (in the direction of the Earth's spin axis), a ΔX and a ΔY.

Deviations of the Earth's mean surface from the ellipsoid may affect basic surveying. Conventional surveying techniques depend on gravity, since this defines the orientation of a plumb bob or other levelling device. Surveys relying on such methods are inherently referenced not to an ellipsoid but to a surface over which the gravitational potential is constant. The most important such surface is the one that, at sea, coincides with average sea level. This *geoid* is in places as much as 100 m above or below the ellipsoid. Topographic elevations are normally referred to the geoid, since most users demand that their maps show local mean sea level as zero. GPS receivers, however, may show heights relative either to the geoid or to an idealised ellipsoid. It is not always easy to discover which!

With the increasing use of navigation satellites, these formerly rather esoteric considerations have become very important in field surveys. GPS receivers offer an enormous range of possible co-ordinate systems, and also of possible ellipsoids. It is all too easy to input the correct projection and co-ordinate system into the survey GPS and then relax, forgetting that the

ellipsoid must also be specified and that positional shifts of several hundred metres can be introduced by switching between commonly used versions. As an example of the importance of this effect, selecting WGS84 instead of the AGD66 ellipsoid on which the maps of the Bulolo goldfields of Papua New Guinea are based, would not merely move observers some 200 metres laterally but put them on the wrong side of the wide Bulolo River!

15.2   SATELLITE NAVIGATION

Satellite navigation is now being used for applications as diverse as personal navigation devices in smart phones, air and marine traffic control, and guiding ploughs and harvesters. The major users outside the United States are consequently reluctant to rely entirely on the goodwill of the US military, and some of the alternatives to GPS, the US system, are expected to become important during the lifetime of this edition.

15.2.1   Navigation satellites

As the uses of navigation satellites have multiplied, so have the acronyms. Currently four alternatives to GPS are in various stages of development. These are Russia's Global Navigation Satellite System (GLONASS), the European Union's Galileo, China's COMPASS and the Global Indian Navigation System (GINS). Apart from GPS, only GLONASS is currently operational. Galileo and COMPASS should be operational by 2013 and 2015 respectively. The systems are being designed to have some compatibility, so that it will in theory be possible for signals broadcast from any of the over 100 satellites to be received by all users, significantly improving both availability and accuracy.

15.2.2   Locational accuracies

The simplest GPS receivers use only *code-phase* measurements, with distances from the satellites (*pseudo-ranges*) deduced from the times it takes unique Pseudo-Random Codes (PRCs) generated by each selected satellite to reach the measuring point. At least three satellites are required to obtain a positional (XY) fix, and four if elevation is also required.

Small and reasonably cheap hand-held GPS receivers have been available since about 1990, but their accuracy was originally no better than a 100 m in position and even less in elevation because of deliberate signal degradation, known as Selective Availability (SA), implemented by the US military. On 1 May 2000 the US government turned off SA on all GPS satellites and the positional accuracy on stand-alone receivers improved to a few tens of metres. The readout precision, for both elevations and co-ordinates,

is generally to 1 m, or its rough equivalent ($0.000\,01°$) in latitude and longitude.

Accuracies of the order of 1–2 m can be obtained using differential GPS (dGPS or DGPS) methods. These apply distance corrections from external reference systems operating at surveyed locations and can eliminate errors that are common to both the base station and field receiver. The corrections can be applied by *post-processing*, but can also be carried out in real-time using corrections transmitted via satellite or from a local base station to a field receiver.

Even greater accuracies can be obtained using the phase of the carrier wave on which the PRC is superimposed. The GPS carrier frequency of 1.57 GHz implies, in principle, an ability to provide accuracies better by several orders of magnitude than the 1-MHz PRC modulation but, because each cycle of the carrier wave is almost identical to the previous one, *cycle-skipping* is possible, introducing errors that are multiples of 20 cm. Real-time kinetic (RTK) systems eliminate cycle-skipping by using a base station receiver and transmitter in a known location to re-broadcast the phase of the carrier to the mobile unit(s). Dual-frequency receivers that track a second carrier-wave phase signal (L2) as well as the standard code- and carrier-phase (L1) signal are used if even higher accuracy is required.

A technique called kinematic carrier-phase tracking (KCPT) utilises the L1 information to provide centimetre-scale locational accuracies but may require long (20–30-minute) initialisation periods to give an accurate position. RTK solutions can also be obtained using additional KCPT measurements on L2, providing the highest real-time accuracies currently available.

Multi-path errors (i.e. reflections from topography or buildings that provide alternative paths of different lengths) and variations in the properties of the atmosphere can significantly reduce accuracy. The main atmospheric effects occur in the ionosphere and depend on the magnitude and variability of the ionisation. They are thus most severe during periods of high solar activity, and particularly during magnetic 'storms' (see Section 3.2.4).

It is now often possible to obtain fixes through the canopy in rainforests, but buildings or solid rock between receiver and satellite are still insuperable obstacles. Interestingly, GPS receivers installed on smart phones can outperform even the most expensive RTK systems in the urban jungle, due in part to multi-path errors and weak direct satellite signals in the urban environment and thanks to technologies such as assisted-GPS (A-GPS), where cell phone transmitters also broadcast GPS satellite orbit information. Between 2007 and 2010, more GPS receivers were built into cell phones than were used in all the other applications put together.

Figure 15.1 *Example of a hand-held Personal Digital Assistant (PDA) and global positioning system (GPS) receiver that can receive real-time differential GPS (DGPS) corrections via satellite to provide 2–5 m locational accuracies. It is equipped with WiFi and Bluetooth for easy email and data exchange.*

15.2.3 Practical considerations in using DGPS

Differential GPS data can be obtained in a number of ways:

1. A dedicated base station may transmit corrections via a radio link for real-time corrections, or be used to post-process the field data if real-time corrections are not required.
2. A government-operated SBAS (*Satellite-Based Augmentation System*) such as WAAS (*Wide Area Augmentation Service*) in the United States or EGNOS (*European Geostationary Navigation Overlay Service*) in Europe can be used to provide real-time differential corrections. The GPS instrument shown in Figure 15.1 has a built-in radio receiver, allowing these corrections to be applied in real time.
3. A subscription can be bought for a commercial differential service such as Omnistar. Omnistar's dual-frequency, carrier-phase H-Star technology can achieve sub-30-cm accuracies in real time provided that at least 2 minutes of uninterrupted satellite signal (satellite lock) have been recorded. Lock must be maintained on a minimum of five satellites.
4. Carrier-wave RTK techniques provide centimetre accuracies in real time if carrier lock is maintained, but no position at all if it is lost. Post-processing can provide accuracies of 30 cm to 1 cm using data acquired

over intervals ranging from 10 to 45 minutes. This is the highest level of post-processed GPS accuracy. The base station, which must log dual-frequency GPS data at a frequency no lower than 0.03 Hz, must be located at a precisely surveyed point within 10 km if accuracies of a few centimetres are required.

5. In some countries there are networks of GPS reference stations (VRS or *Virtual Reference Stations*) that, if sufficiently dense, can provide the high-accuracy corrections required for RTK positioning. These are only available as subscription services.

There are, of course, advantages and disadvantages to each method. Dedicated base stations introduce complications into survey operations that field crews may well feel they can do without, and post-processing is therefore often favoured. Government systems are free but currently cover only limited areas. Use of a VRS relies on a cell-phone link to allow corrections to be continuously downloaded, and this may not be available throughout the survey area. The charges for network systems such as VRS and for commercial satellite services are significant, but can be well worth paying if productivity is improved by eliminating the need to set up base stations. Carrier-phase data are more difficult to collect in environments where satellite signals are subject to interruption, and under these conditions the use of code-phase data alone may be the more practical option. Remember that the theoretical locational accuracy of a given GPS system is not necessarily achievable at your site.

15.2.4 Ensuring a GPS is fit for purpose

The locational accuracy required for a geophysical survey obviously depends on its objectives. These should especially be borne in mind when using instruments with built-in GPS receivers. The fact that a receiver is integrated into a geophysical instrument should not be taken as evidence of its adequacy for all surveys using that instrument. Moreover, there is no point in generating very accurate target co-ordinates using sub-metre RTK positioning if a hand-held non-DGPS system will subsequently be used to dig or drill on those targets.

Most GPS receivers provide a quality-control parameter such as '%Precision' for positions recorded in real time, and can be programmed so that locations are not recorded if this parameter falls outside a predetermined range. Planning software provided by most instrument vendors can be used to predict when a factor called the DOP (dilution of precision), which is based on the number and positions of satellites in the sky and their relationships to each other, will fall outside an acceptable level. The lower the DOP, the better the signal quality. The DOP can be further broken down into

horizontal HDOP, vertical VDOP and time TDOP. VDOP may not be important if elevations are not critical, but where an accurate 3D position is required, as in the case of the void detection survey in Table 15.1, a minimum of four satellites are needed. Most receivers also record a signal-to-noise ratio (SNR) and the number of satellites actually used to determine position.

To ensure that recorded co-ordinates remain within acceptable tolerances, a reference point, which could also be a geophysical drift-measurement or nulling base station, should be revisited twice daily for comparison of the computed positions.

15.2.5 GPS as a timing device
Every GPS receiver in contact with a satellite is automatically synchronised to Greenwich Mean or Coordinated Universal Time (GMT or UTC). GPS receivers can therefore be used for synchronising base station and roving field magnetometers, or different instruments on multi-instrument platforms. So precise are the times that they can even be used to synchronise transmitters and receivers in electromagnetic surveys, removing the need for a cable link to provide a phase reference.

15.2.6 GPS-plus
No single navigation system can provide accurate positioning all the time and everywhere. The coming decade is likely to be the decade of GPS-plus rather than GPS-only, with integrated GPS + Inertial Navigation Systems (INS), GPS + GLONASS, Galileo and COMPASS and GPS + WiFi, to name but a few of the possibilities. GPS + INS combines the GPS receiver with a gyroscope, distance encoder, 3D accelerometers and a 2D digital compass to infill any gaps in GPS coverage. It can be particularly useful where satellites are obscured in parts of a survey area by topography or buildings. Multiple technologies are currently expensive, but during the lifetime of this edition will almost certainly fall in price and become more readily accessible.

15.2.7 GPS-minus
Reliance on new technology can be taken too far. In a micro-gravity survey on a flat site where elevations of closely-spaced stations are required to centimetre accuracy, the traditional optical level and graduated staff will often still provide the required information in the quickest, cheapest and most accurate fashion.

APPENDIX

Terrain corrections for Hammer zones B to M, in μGal, and per compartment (see Section 2.3.4). Row R lists the inner and outer radii of the zones, in metres up to and including Zone C and in kilometres thereafter. Row N lists the number of compartments into which each zone is divided.

Zone:	B	C	D	E	F	G
R (m):	2 166	53.3 170	390	895	1530
N	4	6	6	8	8	12
1	0.5	1.9	3.3	7.6	11.5	24.9
2	0.7	2.6	4.7	10.7	16.3	35.1
3	0.8	3.2	5.8	13.1	19.9	43.3
4	1.0	3.8	6.7	15.2	23.0	49.8
5	1.1	4.2	7.5	17.0	25.7	55.6
6	1.2	4.6	8.2	18.6	28.2	60.9
7	1.3	5.0	8.9	20.1	30.4	65.8
8	1.4	5.4	9.5	21.5	32.6	70.4
9	1.5	5.7	10.1	22.9	34.5	74.7
10	1.6	6.0	10.6	24.1	36.4	78.7
20	2.4	8.7	15.1	34.2	51.6	111.6
30	3.2	10.9	18.6	42.1	63.3	136.9
40	3.9	12.9	21.7	48.8	73.2	158.3
50	4.6	14.7	24.4	54.8	82.0	177.4
60	5.3	16.5	26.9	60.2	90.0	194.7
70	6.1	18.2	29.3	65.3	97.3	210.7
80	6.9	19.9	31.5	70.1	104.2	225.6
90	7.8	21.6	33.7	74.7	110.8	239.8
100	8.7	23.4	35.7	79.1	117.0	253.2

Effect (μGal):   Height differences (metres)

Zone:	H	I	J	K	L	M
R(km): 1.53	2.61	4.47	6.65	9.9	14.7	21.9
N:	12	12	16	16	16	16
1	32	42	72	88	101	125
2	46	60	101	124	148	182
3	56	74	125	153	186	225
4	65	85	144	176	213	262
5	73	95	161	197	239	291
6	80	104	176	216	261	319
7	86	112	191	233	282	346
8	92	120	204	249	303	370
9	96	127	216	264	322	391
10	103	134	228	278	338	413
20	146	190	322	394	479	586
30	179	233	396	483	587	717
40	206	269	457	557	679	828
50	231	301	511	624	759	926
60	253	330	561	683	832	1015
70	274	357	606	738	899	1097
80	293	382	648	790	962	1173
90	311	405	688	838	1020	1244
100	328	427	726	884	1076	1312

Effect (μGal): Height differences (metres)

These tables list the exact height differences that, assuming a density of $2.0\,\mathrm{Mg\,m^{-3}}$, will produce the tabulated terrain effects. Thus a height difference of 32 m between the gravity station elevation and the average topographic level in one compartment of Zone E (between 170 and 390 m from the gravity station) would be associated with a terrain effect of about 18 μGal. Most commercial gravity meters have sensitivities of only 10 μGal, but estimates to 1 μGal are used in the tabulation for small elevation differences to avoid the accumulation of 'rounding off' errors when summing the contributions from large numbers of compartments. For larger elevation differences, these levels of accuracy are not obtainable using Hammer charts.

INDEX