PART II: GEOMAGNETISM

1. INTRODUCTION

Geomagnetism, the study of Earth’s magnetic field, has a long history and has revealed much about the way the Earth works. As we shall see, the existence and characteristics of the field essentially demand that the fluid outer core be made of electrically conducting material that is convecting in such a way as to maintain a self-sustaining dynamo. The study of the field as it is recorded in rocks allows us to track the past motions of continents and leads directly to the idea of sea-floor spreading. Variations in the external part of the geomagnetic field induce secondary variations in Earth’s crust and mantle which are used to study the electrical properties of the Earth, giving insight into porosity, temperature, and composition in these regions.

The magnetic field was the first property attributed to the Earth as a whole, aside from its roundness. This was the finding of William Gilbert, physician to Queen Elizabeth I, who published his inference in 1600, predating Newton’s gravitational *Principia* by about 87 years. The magnetic compass had been in use, beginning with the Chinese, since about the second century B.C., but it did not find its way to Europe until much later, where it became an indispensable tool for maritime navigation. Petrus Peregrinus can be credited with producing the first scientific work devoted to magnetism, discovering magnetic meridians, the dipolar nature of the magnet, and describing two versions of the magnetic compass. The *Epistola de Magnete* was written in 1269, and subsequently widely circulated in Europe, but not actually published until the 16th century. Gilbert placed the source of magnetism within the Earth in 1600, but the temporal variations in the magnetic field were not well documented until the middle of the seventeenth century when Henry Gellibrand appreciated that the differences among repeated measurements were not just inaccuracy in the observations. In 1680 Edmund Halley published the first contour map of the geomagnetic variation as the declination was then known: he envisioned the secular variation of the field as being caused by a collection of magnetic dipoles deep within the earth drifting westward with time with about a 700 year period, a model not dissimilar to many put forward during the twentieth century, although he did not know of the existence of the fluid outer core. A formal separation of the geomagnetic field into parts of internal and external origin was first achieved by the German mathematician Karl Friedrich Gauss in the nineteenth century. Gauss invented spherical harmonics and deduced that by far the largest contributions to the magnetic field measured at Earth’s surface are generated by internal rather than external magnetic sources, thus confirming
Gilbert’s earlier speculation. He was also responsible for beginning the measurement of the geomagnetic field at globally distributed observatories, some of which are still running today.

Figure 1.1

The magnetic field is a vector quantity, possessing both magnitude and direction; at any point on Earth a free compass needle will point along the local direction of the field. Although we conventionally think of compass needles as pointing north, it is the horizontal component of the magnetic field that is directed approximately in the direction of the North Geographic Pole. The difference in azimuth between magnetic north and true or geographic north is known as declination (positive eastward). The field also has a vertical contribution; the angle between the horizontal and the magnetic field direction is known as inclination and is by convention positive downward (see Figure 1.1). Three parameters are required to describe the magnetic field at any point on the surface of the Earth, and the conventional choices vary according to subfields of geomagnetism and paleomagnetism. Traditionally, the vector $\mathbf{B}$ at Earth’s surface is referred to a right-handed coordinate system: north-east-down for $x$-$y$-$z$. But often instead of using the components in this system, three numbers used are: intensity, $B = |\mathbf{B}|$, declination, $D$, and inclination, $I$ as shown in the sketch or $D$, $H$ and $Z$; $H$, or equivalently $B_h$, is the projection of the field vector onto the horizontal plane and $Z$, or equivalently $B_z$, is the projection onto the vertical axis. $D$ is measured clockwise from North and ranges from $0 \rightarrow 360^\circ$ (sometimes $-180^\circ \rightarrow 180^\circ$). $I$ is measured positive down from the horizontal and ranges from $-90 \rightarrow + 90^\circ$ (because field lines can also point out of the Earth, indeed it is only in the northern hemisphere that they are predominantly downward). From the diagram we have

$$H = B\cos I; \quad B_z = B\sin I. \quad (1)$$

2
When components of $\mathbf{B}$ are used they are called $X$, $Y$, $Z$, and:

$$X = B_x = B \cos I \cos D; \quad Y = B_y = B \cos I \sin D; \quad Z = B_z = B \sin I.$$  \hspace{1cm} (2)

The field strength $B$ is sometimes measured in gammas ($\gamma$) where $1 \text{ gamma} = 10^{-5}$ Gauss; today you find SI units in use: $B$ is measured then in tesla (T); 1 T is a very large field. More commonly in geophysics the unit of choice is the submultiple nanotesla (nT); 1 nT = $10^{-9}$ T = 1 gamma, by pure coincidence; occasionally the $\mu$T is also used, with $1 \mu \text{T} = 10^{-6}$ T.

Figure 1.2: (a) Radial component of the magnetic field at Earth’s surface in $\mu$T and (b) its rate of change in nT/yr for the year 2000.

When the standard geocentric spherical coordinate system is used the magnetic field elements are usually designated $B_r$, $B_\theta$, and $B_\phi$, corresponding to locally radial, southward, and eastward unit vectors referred to a position vector $\mathbf{r}$ on a spherical surface $S(\alpha)$. It is generally important to account for the distinction between geocentric and geographic latitude, especially when combining surface and satellite observations.
Detailed maps of the present day field show that it is a complicated function of position on the surface of the Earth although it is dominantly dipolar, and can be approximated to first order by a dipole located at the center of the Earth, with its axis tilted about 11° relative to the geographic axis. The magnitude of the field, the magnetic flux density passing through Earth’s surface, is about twice as great at the poles (about 60 $\mu$T) as at the equator (about 30 $\mu$T).

Figure 1.3: (a) Scalar magnetic field at Earths’ surface in $\mu$T and (c) its rate of change in nT/yr for the year 2000. (b) and (d) are the same with the dipole part of the field subtracted out.

The present and historical magnetic field is measured at observatories, by surveys on land and at sea, and from aircraft. Since the late 1950s a number of satellites, each carrying a magnetometer in orbit around Earth for months at a time, have provided more uniform coverage than previously possible. Early satellites only measured the magnitude of the field: however, it was shown in the late 1960s that measurements of the field’s direction are also required to specify the field accurately. Prehistoric magnetic field records can also be obtained through paleomagnetic studies of remanent magnetism recorded in rocks and archeological materials. These can be useful for geomagnetic studies if there is an independent means of determining the timing that the magnetization was acquired.

Contours of the radial component of the geomagnetic field at Earth’s surface for 2000 are shown on Figure 1.2(a). The magnetic equator corresponds to the zero contour and differs significantly from the geographic equator. At the magnetic poles the field is vertical (inclination is $\pm 90^\circ$, and declination is undefined).
Earth’s magnetic field is generated in the liquid outer core, where fluid flow is influenced by Earth rotation and the inner core geometry (which defines the tangent cylinder). Core fluid flow produces a secular variation in the magnetic field, which propagates upward through the relatively electrically insulating mantle and crust. Above the insulating atmosphere is the electrically conductive ionosphere, which supports Sq currents as a result of dayside solar heating. Outside the solid Earth the magnetosphere, the manifestation of the core dynamo, is deformed and modulated by the solar wind, compressed on the sunside and elongated on the nightside. At a distance of about 3 Earth radii, the magnetospheric ring current acts to oppose the main field and is also modulated by solar activity. Magnetic fields generated in the magnetosphere and ionosphere propagate by induction into the conductive Earth, providing information on electrical conductivity variations in the crust and mantle. Magnetic satellites fly above the ionosphere, but below the magnetospheric induction sources. The ring current and solar wind are not drawn to scale.
that the magnetic poles are distinguished from the *geomagnetic poles* which correspond to the axis of the best fitting geomagnetic dipole. A different representation of the field is given in Figure 1.3 where the scalar field intensity is plotted along with its rate of change in the upper panel. In the lower part the dipole contribution to the field is omitted. From these figures we see that the field strength is lowest in the South Atlantic region, and at high latitudes it is dominated by pairs of approximately symmetric lobate structures. The non-dipole part of the field is weakest in the Pacific hemisphere. Looking at the secular variation or rate of change of the field, the largest changes are occurring over the Americas and the Atlantic and in the southern hemisphere over Africa and the Indian Ocean. Again the Pacific region shows relatively weak variations. The longevity of these features is an active area of research.

Figures 1.2 and 1.3 show the largest scale features of the internal magnetic field which originate in Earth’s core, but there are a number of different physical sources that contribute to the measured field. A whirlwind tour of spatial and temporal variations of both internal and external parts of the field is given in Chapter 1 of *Foundations of Geomagnetism*, by Backus, Parker and Constable (called *Foundations* henceforth). Figure
Figure 1.6: Power spectrum of paleomagnetic dipole moment variations as a function of frequency. At longest periods the spectrum is derived from the magnetostratigraphic time scale (black, gray lines), intermediate (red, blue, orange, green, brown) are from marine sediment paleomagnetic records, and shortest periods (purple) are from archeomagnetic and lake sediment data.

1.4 gives a simplified view of the parts of the magnetic field that are most important for our purposes: these can be roughly divided according to spatial scale and the frequency range in which they operate. The corresponding amplitude spectrum of variations as a function of frequency is given in Figure 1.5.

The bulk of Earth’s magnetic field is generated in the liquid outer core, where fluid flow is influenced by Earth rotation and the geometry of the inner core. Core fluid flow produces a secular variation in the magnetic field (see Figure 1.2(b); 1.3(c), (d)), which propagates upward through the relatively electrically insulating mantle and crust. Short term changes in core field are attenuated by their passage through the mantle so that at periods less than a few months most of the changes are of external origin. At Earth’s surface the crustal part is orders of magnitude weaker than that from the core, but remanent magnetization carried by crustal rocks has proved very important in establishing seafloor spreading and plate tectonics, as well as a global magnetostratigraphic timescale. The crust makes a small static contribution to the overall field, which only changes detectably on geological time-scales making an insignificant contribution to the long period spectrum. On very long timescales (about $10^6$ years) the field in the core reverses direction, so that a compass needle points south instead of north, and inclination reverses sign relative to today’s field.
The present orientation of the field is known as normal, the opposite polarity is reversed. The occurrence of reversals is unpredictable and the average rate varies with time. Figure 1.6 supplements the long period part of Figure 1.5, showing the power spectrum of dipole moment variations inferred from various kinds of paleomagnetic data.

Returning to Figure 1.4 we note that above the insulating atmosphere the electrically conductive ionosphere supports $S_q$ currents with a diurnal variation as a result of dayside solar heating. Lightning generates high frequency Schumann resonances in the Earth/ionsphere cavity. Outside the solid Earth the magnetosphere, the manifestation of the core dynamo, is deformed and modulated by the solar wind, compressed on the sunside and elongated on the nightside. At a distance of about 3 earth radii, the magnetospheric ring current acts to oppose the main field and is also modulated by solar activity. Although changes in solar activity probably occur on all time scales the associated magnetic variations are much smaller than the changes in the core field at long periods, and only make a very minor contribution to the power spectrum.

The Earth’s magnetosphere plays an important role in protecting us from cosmic ray particle radiation, because the incoming ionized particles can get trapped along magnetic field lines, preventing them from reaching Earth. One consequence of this is that rates of production of radiogenic nuclides such as $^{14}$C and $^{10}$Be are inversely correlated with fluctuations in geomagnetic field intensity. This means that knowledge of Earth’s dipole moment in the past plays an important role in paleoclimate studies that use cosmogenic nuclide production to infer solar insolation during prehistoric times.

**Supplemental Reading**

2. CLASSICAL ELECTRODYNAMICS IN GEOMAGNETISM

As we turn to the geomagnetic part of this course we will apply many of the same mathematical tools as are used in studying Earth’s gravitational potential. However, instead of Newton’s law, the fundamental physics are described by the equations of classical electrodynamics. This chapter starts with Helmholtz’s theorem, Maxwell’s equations, and Ohm’s law in a moving medium, and motivates the equations that are used in static geomagnetic field modeling. Much of the material covered here is to be found in Chapter 2 of Foundations; a less advanced treatment is given in Chapters 2 and 4 of Blakely’s book on Potential Theory.

2.1 Helmholtz’s Theorem and Maxwell’s Equations

The universe of classical electrodynamics begins with a vacuum containing matter solely in the form of electric charges, possibly in motion, and electric and magnetic fields. We can detect the presence of these fields by the forces they exert on a moving point charge $q$. If the charge $q$ is located at position $r$ at time $t$ and moves with velocity $v$ relative to an inertial frame, then

$$
 f = q[E(r, t) + v \times B(r, t)].
$$

This expression allows us in principle to measure the electric and magnetic fields using a moving charge as a detector in an inertial reference frame.

Maxwell’s equations provide the curl and divergence of the electric fields and magnetic fields in terms of other things. The reason this is useful is that Helmholtz’s Theorem tells us that if we know the curl and the divergence of a vector field, we can explicitly calculate the field itself, and furthermore, the curl and the divergence represent sources for the field, essentially creating the field. Here is Helmholtz’s theorem. A vector field $F$ in $\mathbb{R}^3$ which is continuously differentiable (except for jump discontinuities across certain surfaces) is uniquely determined by its divergence, its curl and jump discontinuities if it approaches 0 at infinity. The field can be written as the sum of two parts

$$
 F = -\nabla V + \nabla \times A
$$

where $V$ is called a scalar potential and $A$ a vector potential. These two potentials can be explicitly computed from the following two integrals:

$$
 V(r) = \frac{1}{4\pi} \int d^3s \, \frac{\nabla \cdot F(s)}{|r - s|}
$$

$$
 A(r) = \frac{1}{4\pi} \int d^3s \, \frac{\nabla \times F(s)}{|r - s|}.
$$
What these two equations state is that the field $F$ is generated by two kinds of sources: one is the divergence of $F$, the other its curl. Recall that in classical gravity, the gravitational potential $V$ is generated by matter density $\rho$ and we see this through:

$$V(r) = -G \int d^3s \frac{\rho(s)}{|r - s|}. \quad (6a)$$

But this is just equation (5), because of Poisson’s equation, $\nabla^2 V = 4\pi G \rho$ and the fact that here $F = g = -\nabla V$. Helmholtz tells us that if there is no divergence or curl anywhere in space, then $F$ must vanish, again confirming that $\nabla \cdot F$ and $\nabla \times F$ are the sources of the field. Now back to Maxwell’s equations in a vacuum.

Recall the universe we are operating in comprises an infinite vacuum containing electrical charges, represented by a local density $\rho$, which my be moving, and hence generating electric current, represented by a local current density $J$. Here are Maxwell’s equations:

\[ \nabla \times E = -\frac{\partial}{\partial t} B \]  
\[ \nabla \cdot E = \frac{\rho}{\epsilon_0} \]  
\[ \nabla \times B = \mu_0 (J + \epsilon_0 \frac{\partial}{\partial t} E) \]  
\[ \nabla \cdot B = 0 \]

where $\rho$ is charge density (in SI units coulombs/m$^3$), $J$ is current density (amperes/m$^2$), $\mu_0$ is permeability of vacuum ($4\pi \times 10^{-7}$ henries/m), $\epsilon_0$ is capacitivity of vacuum ($10^7/4\pi c^2$ farads/m); $B$ is in teslas, and $E$ is in volts/m. The quantities $\mu_0$ and $\epsilon_0$ are exactly defined constants of the SI measurement conventions; the quantity $c$ is the velocity of light in a vacuum, which is also an exact number in SI. Notice I have introduced the somewhat unconventional, streamlined notation for time derivative: $\partial/\partial t = \partial_t$.

Viewed from the perspective of Helmholtz’s Theorem we see that the Maxwell equations (7) and (8) tells how a magnetic field is generated (by changing magnetic fields — Faraday’s law) or by the presence of electric charges (Coulomb’s law); and equations (9) says we can generate a magnetic field by a combination of moving charges (Ampere’s law) and by changing the electric field in time (Maxwell’s discovery, which does not have the word law associated with it). Equation (10) says there are no isolated magnetic charges, that is, no magnetic monopoles.

2.2 The Static Case for Geomagnetic Field Modeling

For some purposes we can neglect time variation in geomagnetic processes and imagine a system of stationary charges and steady current flows. Many geomagnetic phenomena take place over long time
scales and certainly for the purposes of modeling the present geomagnetic field this seems like a reasonable approximation. In (7), the first of the Maxwell equations, we set $\partial_t B = 0$; then the curl of the electric field vanishes. Making use of this in (4) and (6) we find that the electric field may be written as the gradient of a scalar $\phi$ (the electric potential). Thus

$$E = -\nabla \phi. \quad (11)$$

Putting this together with (8) and (5) we get $\nabla^2 \phi = -\rho/\varepsilon_0$ (Poisson’s equation again, but notice the sign!) and

$$\phi(r) = \frac{1}{4\pi\varepsilon_0} \int d^3s \frac{\rho(s)}{|r - s|}. \quad (12)$$

For the magnetic field it follows from $\nabla \cdot B = 0$ and (5),(4) that we can always write $B = \nabla \times A$. The vector field $A$ is known as the magnetic vector potential. Now if we specialize to the static case with $\partial_t E = 0$, we find from (9), and (6) that

$$A(r) = \frac{\mu_0}{4\pi} \int d^3s \frac{J(s)}{|r - s|}. \quad (13)$$

Again from (9) we have that $J = \nabla \times B/\mu_0$ and taking the divergence yields

$$\nabla \cdot J = 0. \quad (14)$$

### 2.3 Constitutive Relations

Maxwell’s equations as written in (7)-(10) apply to a vacuum. When we need equations describing the behavior of electromagnetic fields inside a material we require some mechanism for spatial averaging of the charge and current distributions due to the atoms making up the material. This question is considered in most courses on electromagnetism, and in *Foundations*. These lead us to a form of Maxwell’s equations capable of describing field within various materials

$$\nabla \times E = -\partial_t B \quad (16)$$

$$\varepsilon_0 \nabla \cdot E = \rho - \nabla \cdot P \quad (17)$$

$$\nabla \times B/\mu_0 = (J + \partial_t P + \nabla \times M + \varepsilon_0 \partial_t E) \quad (18)$$

$$\nabla \cdot B = 0 \quad (19)$$

where $P$ and $M$ are the electric polarization per unit volume and the magnetization, or magnetic polarization per unit volume of the material. Physically what happens is that the presence of an electric field (for simplicity) polarizes the material, causing charge separation. This introduces a large number of tiny electric
dipoles into the medium, quantified by the term $P$ – this is simply the density of electric dipole moment present in the material. If the dipole density were precisely constant, there would be no effect on $E$, because the dipole fields would cancel on average (except at the ends of the specimen, where charges would accumulate). But variations in the dipole density do cause electric fields – this is seen in the fact that the term in the modified equations is $\nabla \cdot P$. The magnetic effect is similar, but more complicated because electrons’ intrinsic magnetic moments and their motions within atoms cause magnetic fields.

The solution of Maxwell’s equations for $E$ and $B$ in a material thus requires knowledge of $J$, $P$ and $M$ and these in turn depend in the way the material responds to the fields. These are called the constitutive relations for the material and are often determined by $E$ and $B$ themselves. They are not fundamental like Maxwell’s equations, but are the result of empirical observations and experiments done on different materials. The simplest possible behavior is linear. For example for many materials over a wide range of field values, we find

\begin{align*}
J &= \sigma E \quad (20) \\
P &= \epsilon_0 \chi_E E \quad (21) \\
M &= \beta B / \mu_0 \quad (22)
\end{align*}

where $\sigma$, $\chi_E$ and $\beta$ are constants. Of course, we recognize $\sigma$ as the electrical conductivity, $\chi_E$ as the electrical susceptibility, and $\beta$ as a kind of magnetic susceptibility.

We can simplify Maxwell’s equations by defining new fields $H$ and $D$,

\begin{align*}
D &= \epsilon_0 E + P \quad (23) \\
H &= B / \mu_0 - M. \quad (24)
\end{align*}

$D$ is called the electric displacement vector. $H$ has traditionally been called the magnetic field vector, while $B$ was called the flux intensity or magnetic induction. The two are often confused. In view of its primary place in the theory we shall call $B$ the magnetic field vector and by analogy with $D$, $H$ will be the magnetic displacement vector. You should be aware these names are not yet standard, but they ought to be. With these definitions in place we achieve a form of Maxwell’s equations for the second and third relations:

\begin{align*}
\nabla \cdot D &= \rho \quad (25) \\
\nabla \times H &= J + \partial_t D. \quad (26)
\end{align*}
The last term is called the \textit{displacement current} and is a way of generating magnetic fields without any charges having to move. Notice that the first and last of Maxwell’s equations remain unchanged from their vacuum forms, (7) and (10). In fact we rarely use $D$ in geomagnetism; one reason is that most Earth materials are not highly polarizable, and another is that we almost always drop the term involving $D$ in (26) as we shall see next.

\section*{2:4 Application to the Geomagnetic Field}

A reasonable approximation in geophysical problems is to neglect the displacement current $\partial_t D$ in (26). This can be shown by a crude dimensional analysis as follows. (For more details see \textit{Foundations}, Section 2.4) Take the time derivative of (26), and insert (20); for simplicity assume $\chi E$ and $\beta$ in (21)-(22) are negligible; then (26) becomes:

$$\nabla \times \partial_t B / \mu_0 = \sigma \partial_t E + \epsilon_0 \partial_t^2 E.$$ \hspace{1cm} (27)

Now use (16) and rearrange slightly

$$\nabla \times \nabla \times E + \mu_0 \sigma \partial_t E + \mu_0 \epsilon_0 \partial_t^2 E = 0.$$ \hspace{1cm} (28)

We would like to estimate the approximate size of each of the three terms in (28). If we assume length scales of variation are $L$ or larger and time scales $T$ or larger, very roughly we can replace space derivatives by $1/L$ and time derivatives by $1/T$; then

$$0 = [1 + \mu_0 \sigma (L^2 / T) + \mu_0 \epsilon_0 (L/T)^2] \|E\|.$$ \hspace{1cm} (29)

Because $\mu_0 \epsilon_0 = 1/c^2$, where $c$ is the speed of light, the last term represents the ratio of typical speeds in the system over $c$ squared. In geomagnetism scales are typically many thousands of kilometers, and time scales can be as low as minutes, but may be years: even for $L = 10^3$ km and $T = 10$ s, the last term in (29) is $10^{-5}$. The term with conductivity is much larger than this in the interior, say $\sigma \approx 10^{-3}$ S/m, then the second term is roughly $4\pi \times 10^{-7} \times 10^{-3} \times 10^{12}/10$ or about 120. So displacement current is unimportant and the balance is between the first two terms. The four equations (7)-(10), or the set valid within material (7), (10), (25), (26), in which the displacement current is neglected ($\partial_t E$ or $\partial_t D$)are sometimes referred to as the \textit{pre-Maxwell equations}.

But in the atmosphere $\sigma$ is so small, the second term is small too. When this happens we see the simple-minded analysis breaks down and the discover that the size of $\nabla \times \nabla \times E$ cannot be $\|E\|/L^2$ – terms in the spatial derivative cancel among themselves and the corresponding term in (27) vanishes by itself.
The magnetic field can always be written as the curl of a vector potential (because of Helmholtz’s Theorem, (4)-(6), and $\nabla \cdot \mathbf{B} = 0$). In certain circumstances there is an alternative representation in terms of a scalar potential for $\mathbf{B}$. In our application to the geomagnetic field we will make the approximation that Earth’s atmosphere is an insulator with no electrical currents (actually $\sigma \approx 10^{-13}$ S/m close to the ground so $\mathbf{J} = 0$ seems like a reasonable approximation). The atmosphere is also only very slightly polarizable magnetically so we can set $\mathbf{M} = 0$, thus within the atmospheric cavity we find the essential content of (26) is

$$\nabla \times \mathbf{B} = 0.$$  \hspace{1cm} (30)

(30) tells us that $\mathbf{B}$ can be written as the gradient of a scalar because when $\mathbf{B} = -\nabla \phi$, (30) is automatically satisfied (recall $\nabla \times \nabla = 0$). Since $\mathbf{B}$ is also solenoidal (divergence free) from (19), the scalar potential $\phi$ is harmonic: $\nabla^2 \phi = 0$. This why we can use so much of our gravity machinery in geomagnetism, for example, spherical harmonics.