3:4 Geomagnetic Field Models

In our discussions of the field so far we have not dealt with the issue of how to determine the magnetic field from a practical set of observations. In an earlier lecture it was stated that the spherical harmonic representation provides a unique description of any magnetic field represented by a harmonic potential, and in Section 3.1 we saw how in principle this allows unique separation of the internal and external field contributions to spherical harmonic expansions. Continuing with the specialization to the internal field coefficients begun in Section 3.2, we note that the field coefficients for \( \Omega(r, \theta, \phi) = rB_r(r, \theta, \phi) \) can, in principle be derived from equation (48):

\[
\beta^m_l = \int_{S(a)} B_r(a, \theta, \phi)Y^m_l(\theta, \phi) \frac{d^2r}{a^2} \tag{56}
\]

but this evidently requires knowledge of \( B_r \) all over \( S(a) \), which is not possible in practice. We will need a strategy for dealing with incomplete and noisy observations. But before we tackle the question of finding approximate models, (and here we will follow custom and say that a geomagnetic field model is a finite collection of SH coefficients), we discuss the mathematical question of what are sufficient measurements to define the field uniquely.

**Geomagnetic Elements**

Equation (56) relies on knowledge of the radial magnetic field, but these are not the only kind of observations that are made; typically when surface survey, observatory, and satellite data are involved we must consider all of the following kinds of observations:

- \( B_r, B_\theta, B_\phi \) – orthogonal components of the magnetic field in geocentric reference frame.

- \( X, Y, Z \) – orthogonal components of the geomagnetic field in local coordinate system, directed north, east, and downwards respectively. This is the geodetic coordinate system.

- \( H = (X^2 + Y^2)^{\frac{1}{2}} \) – horizontal magnetic field intensity

- \( B = (B_\theta^2 + B_\phi^2 + B_r^2)^{\frac{1}{2}} = (X^2 + Y^2 + Z^2)^{\frac{1}{2}} \) – total field intensity

- \( D = \tan^{-1}(Y/X) \) – declination

- \( I = \tan^{-1}(Z/H) \) - inclination.

If we are prepared to accept the approximation that the Earth is a sphere, then we can write

\[
X \approx -B_\theta; \quad Y \approx B_\phi; \quad Z \approx -B_r. \tag{57}
\]

But as we saw in Part I, the shape of the Earth is much better approximated by a spheroid or ellipsoid of revolution, with equatorial radius \( a \), polar radius \( b \), eccentricity, \( e \)

\[
e^2 = \frac{(a^2 - b^2)}{a^2} \tag{58}
\]

and flattening

\[
f = \frac{(a - b)}{a} = 1/298.257. \tag{59}
\]
In paleomagnetic field modeling it’s probably not necessary to take account of the difference between geodetic and geocentric latitude but in modern field modeling the data may be accurate enough for it to really matter.

Exercise:

What is the size of the error you make if you neglect to correct for the ellipsoidal shape of the earth and use geocentric latitude and longitude in calculating the field from a spherical harmonic model?

As we already noted, it is impossible to acquire complete knowledge of the radial magnetic field (or indeed any other component) on any spherical surface. Our data are always incomplete and noisy, and consequently there will always be ambiguity in the models derived from them. Because of this it might be argued that the issue of uniqueness in the case of complete and perfect data is of purely academic interest. However, experience in making magnetic field models based exclusively on one kind of observation has shown that this is not the case.

3.4.1 Limitations on Uniqueness

Non-Uniqueness and $|B|

The fact that $B$, on a surface $S(a)$ uniquely determines the field (because it is the solution of a Neumann boundary value problem for Laplace’s equation) might lead one to hope that complete measurements of $|B|$ (the magnitude of the field) would uniquely specify the internal magnetic field to within a sign. Most marine observations are of $|B|$ and so are many satellite measurements, and scalar observations are generally more accurate than their vector counterparts so there might be advantages to just using those data in modeling. For the case of intensity data alone, George Backus showed in 1968 (Q. J. Mech and Appl. Math 21, pp 195-221, 1968; and J. Geophys. Res. 75, pp 6339-41, 1970; see also Jacobs, Geomagnetism, Vol 1, p 347) by constructing a counter example that knowledge of $|B|$ was not sufficient. This provided an explanation of the poor quality of directional information predicted by models based on $B$ alone, especially in equatorial regions. Early satellites had only measured total field intensity, and this result led to the launching of the first vector field satellite in 1979, known as Magsat. Backus’ counter example is constructed in the following way: first find a magnetic field with potential $\Psi_M$ that is everywhere orthogonal to a dipole field; then $\Psi_D + \Psi_M$ and $\Psi_D - \Psi_M$ have the same $B$ everywhere.

One might reasonably ask whether this counter-example matters or is merely an artificial mathematical construct. If it does not one would surely be better off using scalar measurements (with typical accuracy of < 1nT) than vector components which are usually several nT, because of the need to orient a low earth orbiting measurement platform within an Earth centered reference frame. However the effect is real. Independent field models using exclusively vector or scalar Magsat data predict similar intensity at Earth’s surface, but show differences of up to 2000 nT in the vector components. This reflects the fact that optimization in the $|B|$ model is insensitive to errors perpendicular to the field, while the full vector is sensitive misfit in all directions. This perpendicular error effect is also referred to as the Backus effect.

Subsequently, Khokhlov, Hulot and Le Mouël (Geophys. J. Internat. 130, pp 701-3, 1997) showed that if knowledge of
the location of the dip equator is added to knowledge of $|\mathbf{B}|$ everywhere then uniqueness of the solution is guaranteed. The practical usefulness of this result is debatable, and full field $\mathbf{B}$ measurements are now routine.

**Non-Uniqueness and $\mathbf{B}$**

Another potentially interesting case is exact knowledge of the direction of the field all over Earth’s surface, $S(\alpha)$, which one might imagine could determine the field to within a multiplicative scaling constant. Paleomagnetic observations can determine ancient field directions much more reliably than field intensities, so one might hope to discover the shape of the geomagnetic field in those epochs when paleomagnetic data are plentiful. Similarly, (despite an earlier inaccurate argument to the contrary) it was shown in 1990 by Proctor and Gubbins, *(Geophys. J. Internat.* 100, pp 69-79, 1990) that $\mathbf{B}$ on a spherical surface does not determine the field to within a scalar multiple.

The counter examples are again questionable representatives of Earth-like fields, being axisymmetric, antisymmetric wrt the Equator, and octupolar in type. This result too has been investigated further and it has now been shown that if the magnetic field only has two poles on the surface $S(\alpha)$ then knowledge of its direction everywhere allows the geomagnetic field to be recovered, except for a constant multiplier (see Hulot, Khokhlov and Le Mouël, *Geophys. J. Internat.* 129, pp 347-54, 1997).

For much more discussion of uniqueness questions, including some extensions to the case where both external and internal field are considered see Sabaka et al. *(2010, Handbook of Geomathematics, doi: 10.1007/978-3-642-01546-5_17).*

The essential result of importance for geomagnetic field modeling using observations from satellites and geomagnetic observatories is summarized in their Figure 4 and reproduced in Figure 3.4.1.1 below. Despite intrinsic non-uniqueness one can recover sufficient information for field modeling to be a useful activity.

![Fig. 4](image)

Uniqueness of a magnetic field recovered from partial information within a current-carrying shell. In this special case relevant to geomagnetism, it is assumed that any source can lie below $r = a$ *(internal $J(r < a)$ sources), and above $r = c$ (external $J(r > c)$ sources), no sources can lie within the lower subshell ($a < r < b$, the neutral atmosphere), a spherical sheet current can lie at $r = b$ (the $E$-region $J_s(r = b)$ sources), and only poloidal sources can lie within the upper subshell ($b < r < c$, the $F$-region ionosphere). The knowledge of $\mathbf{B}$ on a sphere $r = R$ in the upper subshell (as provided by, e.g., a satellite) and of enough components of $\mathbf{B}$ on the sphere $r = a$ (as provided by, e.g., observatories at the Earth's surface), is then enough to recover the field produced by most sources in many places (see text for details)*
3:4.2 Construction of Field Models

We will suppose here that we have a finite collection of inaccurate observations of orthogonal components of the geomagnetic field \( B_i = \mathbf{B} \cdot \hat{r}_i \). Our goal is to derive from these observations the spherical harmonic coefficients that best represent the real geomagnetic field. Clearly, we cannot use (56) because our knowledge of \( B_i(a, \theta, \phi) \) is incomplete. Also it would not make use of our measurements of the tangential part of the field. The problem is analogous to (and in some respects identical with) that of interpolation to find a curve passing through a finite number of data. Many curves will do the job, and we cannot choose which one is more desirable without supplying additional information of some kind. One extra thing we know is that the magnetic field obeys a differential equation, but that turns out not to be enough information by itself. When the observations are not exact, but uncertain as in all real situations the issue is even more complicated. We obviously shouldn’t expect the interpolant to pass through all (perhaps even any) of the observations. In the case where we would like to downward continue a geomagnetic field model to the core-mantle boundary, we need to be especially careful about how we deal with noise in the data; if we fit models with small scale structure derived from this noise, then it may dominate the real signal after downward continuation.

3:4.3 Least Squares Estimation

The time-honored technique for the construction of geomagnetic field models (invented by Gauss for this very purpose!) used to be that of least squares estimation of the spherical harmonic coefficients in a truncated spherical harmonic expansion for the measured field components. But, instead of the exact expansion with infinitely many terms, we decide ahead of time to model the data with an expansion truncated to degree \( L \): so the scalar potential is

\[
\Psi(r, \theta, \phi) = a \sum_{l=1}^{L} \sum_{m=-l}^{l} b_l^m \left( \frac{a}{r} \right)^{l+1} Y_l^m(\theta, \phi)
\]

and we find

\[
\mathbf{B} = -\nabla \Psi.
\]

If we measure orthogonal components of the field \( B_r, B_\theta, B_\phi \) we have a set of observations at \( N \) sites designated \( \mathbf{r}_j \)

\[
d_j = \hat{s}_j \cdot \mathbf{B}(\mathbf{r}_j), \quad j = 1, \ldots, N
\]

\[
= \sum_{l=1}^{L} \sum_{m=-l}^{l} b_l^m a^{l+1} \hat{s}_j \cdot \nabla \left[ \frac{Y_l^m(\hat{r}_j)}{r_j^{l+1}} \right] + \epsilon_j.
\]

The data \( d_j \) are then linear functionals of the \( b_l^m \) specifying the field. With an appropriate indexing scheme for the \( b_l^m \) we can write a prediction for our observations \( d_j \) as a matrix equation.

\[
d = \mathbf{G} \mathbf{b} + \mathbf{n}
\]

with \( \mathbf{d}, \mathbf{n} \in \mathbb{R}^N, \mathbf{b} \in \mathbb{R}^K \). \( \mathbf{G} \) is an \( N \times K \) matrix, and for each \( d_j \) we can compute \( g_{ij} \), the contribution of the relevant spherical harmonic at that point; the vector \( \mathbf{n} \) contains the misfits between the model predictions and the actual measurements. The
total number of parameters, the length of the vector $\mathbf{b}$, is determined by the truncation level: $K = L(L + 2)$. Thus

$$\mathbf{b} = \begin{bmatrix} b_1^1 \\ b_1^0 \\ \vdots \\ b_L^1 \end{bmatrix}. \quad (64)$$

The value of $L$ is chosen so that $K$ is (much) less than $N$, the number of data, so there are fewer free parameters than data to be fit. This means that it is impossible to choose $\mathbf{b}$ to get an exact match to the data, and so $\mathbf{n}$ is not a vector of zeros. Least squares estimation involves finding the values for $\mathbf{b}$ that minimize $||\mathbf{n}||^2 = ||\mathbf{d} - \mathbf{G} \cdot \mathbf{b}||^2$, where the notation $|| \cdot ||$ is called a norm – in this case it is the ordinary length of the vector. The idea here is to do the best job possible with the available free variables and make the model predictions as close to the data as they can be measured by the length of the misfit vector.

Straightforward calculus can show that the LS vector can be written in terms of the solution to the normal equations:

$$\hat{\mathbf{b}} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{d}. \quad (65)$$

Note that (65) is for several reasons not a good way to find $\hat{\mathbf{b}}$ in a computer – first linear systems of equations ought never to be solved by calculating a matrix inverse (it wastes time and is inaccurate); second, there is a clever way of writing the LS equations that avoids a serious numerical precision problem arising in (65). A result from statistics, known as the Gauss-Markov Theorem, shows that provided the misfits are due to random, uncorrelated perturbations with zero mean, and have a common variance, then the least squares solution is the best linear unbiased estimate (BLUE) available, in the sense that they have the smallest variance amongst such estimates. Also

$$E[||\mathbf{n}||^2] = (N - K)\sigma^2 \quad (66)$$

where $\sigma^2$ is the variance of the noise process. If the noise has a known covariance structure, then the theory can readily be adjusted to take this into account.

A fundamental problem with this approach is that although we might have an idea of the size of the uncertainty in the various observations, and we could choose the truncation level of the spherical harmonic expansion so (66) is approximately satisfied, we do not know that the $K$ spherical harmonics we have chosen adequately describe the geomagnetic field. In other words, the misfit has two sources, not one: measurement error and an insufficiently detailed model. To guarantee a complete description of the real field $L = \infty$; in that case the Gauss-Markov theorem does not apply, nor does (66), and we have no uncertainty estimates for our model. Truncation at finite $L$ corresponds to an assumption about simplicity in the model that has no physical basis – the resulting field model may be biased by the truncation procedure.

### 3.4.4 Regularization – an Alternative to Least Squares

An alternative to LS fitting that has been very widely used since the mid 1980s and is the basis of almost all modern geomagnetic field modeling is to choose “simple” models for the field of an explicit kind. To illustrate the concept we return to our one-dimensional interpolation problem.
One widely used solution to the problem of interpolation is to use what is known as cubic spline interpolation. Data are connected by piecewise cubic polynomials, with continuous derivatives up to second order. The cubic spline interpolant has the property that it is the smoothest curve connecting the points, in the precise sense of minimizing RMS second derivative:

\[
\int_{x_1}^{x_N} \left( \frac{d^2 f}{dx^2} \right)^2 dx.
\]

(67)

The spline solution is the solid line in the Figure 3.4.4.1; any other curve, like the dashed ones, have a larger value for the integral (67).

**Figure 3.4.4.1**

![Figure 3.4.4.1](image1)

**Figure 3.4.4.2**

![Figure 3.4.4.2](image2)
If the data are noisy we can still model them by the same kind of curve, but we no longer require the curve to pass exactly through the model points. Instead, we ask for a satisfactory fit, usually defined in terms of norm of the misfit. So we seek the curve with the smallest RMS second derivative, that has a target misfit. This is another example of a constrained minimization, and like the one we saw in Part I, we can solve it with a Lagrange multiplier. The solution to the constrained problem is determined by finding the stationary points of an objective functional of the following kind:

\[ U = \sum_{j=1}^{N} \left( \frac{f(x_j) - y_j}{\sigma_j^2} \right)^2 + \lambda \int_{x_1}^{x_N} \left[ \frac{d^2 f}{dx^2} \right]^2 dx \]  

subject to

\[ \sum_{j=1}^{N} \left( \frac{f(x_j) - y_j}{\sigma_j^2} \right)^2 = T. \]  

The size of the Lagrange multiplier \( \lambda \) is dictated by the constraint that the data be fit to a reasonable tolerance level. If the uncertainty in the observations, \( \sigma_j \), is known one reasonable choice for \( T \) is the expected value of \( \chi_n^2 \), namely \( N \).

The same ideas are used in geomagnetic field modeling. However, the vector nature of the field makes the choice of penalty functional more complicated – we want to find some property of the magnetic field that can be used like (67) to minimize wiggleness or complexity in our models of the geomagnetic field, either at Earth’s surface or at the CMB (core-mantle boundary). The general idea of constructing models that minimize a penalty other than data misfit is called regularization.

One candidate penalty function is

\[ E = \int_{r>a} |B|^2 \, d^3r. \]  

Since \( |B(r)|^2/2\mu_0 \) is the energy density of the magnetic field at \( r \), \( E/2\mu_0 \) is the total energy stored in \( B \) outside the sphere of radius \( a \). We can reduce this integral to a manageable form in terms of spherical harmonics. We write \( B = -\nabla \Psi \), then

\[ E = \int_{r>a} \nabla \Psi \cdot \nabla \Psi \, d^3r. \]  

Next we make use of a familiar vector identity [number 4, in our list: \( \nabla \cdot (\phi A) = \nabla \phi \cdot A + \phi \nabla \cdot A \), and letting \( A = \nabla \Psi \)] followed by Laplace’s equation to write

\[ E = \int [\nabla \cdot (\Psi \nabla \Psi) - \nabla^2 \Psi] \, d^3r \]

\[ = \int \nabla \cdot (\Psi \nabla \Psi) \, d^3r. \]  

Using Gauss’ Divergence Theorem we can rewrite the volume integral in terms of a surface integral over \( S(a) \)

\[ E = \int_{S(a)} -\Psi \frac{\partial \Psi}{\partial r} \, d^2r. \]
Now we simply substitute the spherical harmonic expansion (60) for the potential \( \Psi \).

\[
E = \int_{S(a)} \left[ \alpha \sum_{l,m} b_{l,m}^{m} Y_{l}^{m}(\hat{\mathbf{r}}) \right] \left[ \sum_{l',m'} (l' + 1)b_{l',m'}^{m'} Y_{l'}^{m'}(\hat{\mathbf{r}}) \right]^{*} a^{2} d^{2}\mathbf{r} \]

\[
= \alpha^{3} \sum_{l,m} \sum_{l',m'} (l' + 1) b_{l,m}^{m}(b_{l,m}^{m})^{*} \int_{S(a)} Y_{l}^{m}(\hat{\mathbf{r}})Y_{l'}^{m'}(\hat{\mathbf{r}})^{*} d^{2}\mathbf{r} \]

\[
= \alpha^{3} \sum_{l=1}^{\infty} \sum_{m=-l}^{l} (l + 1)|b_{l,m}^{m}|^{2}. \quad (74)
\]

Hence \( E \) can be written as a positive weighted sum of the squared absolute values of the spherical harmonic coefficients. This sum is now in a form we can use in minimizing an objective function like (71) for the magnetic field. The weighting by increasing \( l \) means that higher degree (shorter wavelength) contributions to the model will be strongly penalized in minimizing the objective functional. (74) is one example of a set of norms (measures of size) of the kind

\[
||B||_{w}^{2} = \sum_{l=1}^{\infty} \sum_{m=-l}^{l} |B_{l,m}^{m}|^{2}, \quad w_{l} > 0. \quad (75)
\]

Many interesting properties corresponding to smoothness or small size of the field can be written in this form

\[
\int_{S(a)} B \cdot B \, d^{2}\mathbf{r} \quad w_{l} = (2l + 1)(l + 1) \quad (76)
\]

\[
\int_{S(a)} [\nabla_{l} \dot{\mathbf{r}} \cdot B]^{2} \, d^{2}\mathbf{r} \quad w_{l} = l(l + 1)^{2}(l + \frac{1}{2}) \quad (77)
\]

\[
\int_{S(a)} [\nabla^{2}_{l} \dot{\mathbf{r}} \cdot B]^{2} \, d^{2}\mathbf{r} \quad w_{l} = l^{2}(l + 1)^{4} \quad (78)
\]

\[
\int_{r<a} \mathbf{J}_{T} \cdot \mathbf{J}_{T} \, d^{3}\mathbf{r} \quad w_{l} = a^{3} \mu_{0}(l + 1)(2l + 1)^{2}(2l + 3). \quad (79)
\]

In (79) \( \mathbf{J}_{T} \) is the toroidal part of the current flow in Earth’s core, whose significance is discussed in section 3.5. These ideas were first set out in a paper by Shure, Parker, and Backus, *Phys. Earth Planet. Inter.* 28, pp 215-29, 1982.

3:4.5 Results - Gauss Coefficients

In Part I you may recall that the expansion of the Earth’s gravitational potential was called Stokes’ expansion. In geomagnetism the honor goes to Karl Friedrich Gauss: the expansion coefficients are invariably called Gauss coefficients. But there are some differences. First the fundamental one, that makes geomagnetism more interesting than gravity – the field is changing on times scales of a human lifetime. So the coefficients must always be dated. Secondly, and almost trivially, geomagnetists never use fully normalized spherical harmonics – instead they employ an awkward real (as opposed to complex) representation in which the basis functions (spherical harmonics) are normalized so that

\[
\int_{S(1)} (Y_{l}^{m})^{2} d^{2}\mathbf{r} = \frac{4\pi}{2l+1}. \quad (80)
\]
Then the scalar potential is written

$$\Psi(r, \theta, \phi) = a \sum_{l=1}^{\infty} \left( \frac{2}{\ell + 1} \right)^{l+1} \sum_{m=0}^{l} N_{lm} (g_{lm}^m \cos m \phi + h_{lm}^m \sin m \phi) P_l^m(\cos \theta)$$  \hspace{1cm} (81)

where $a = 6,371.2$ km, the mean Earth radius, the $P_l^m$ are the Associated Legendre functions from Part I, and

$$N_{lm} = 1, \quad m = 0$$

$$= \sqrt{\frac{2(l-m)!}{(l+m)!}}, \quad m > 0. \hspace{1cm} (82)$$

You will find the numbers $g_{lm}^m, h_{lm}^m$ tabulated in many places. There are official models designated IGRF for the International Geomagnetic Reference Field, agreed upon every five years by the International Association for Geomagnetism and Aeronomy as a good approximation to the field. The linearized rate of change of the field is also given for interpolation or extrapolation across the five year period. We are currently at Version 12 and the model coefficients can be found at http://www.ngdc.noaa.gov/IAGA/vmod/igrf.html. See Langel’s chapter in Jacob’s book Geomagnetism, an article rich in nuance, with a detailed description of how these are computed, and many tables. Below is the IGRF-9, 2000, onto degree and order 10. The values are in nanotesla. The most recent version of IGRF (2015-2020) can be found online at http://www.ngdc.noaa.gov/IAGA/vmod/igrf.html. Figure 3.4.5 shows various views of the radial component of the magnetic field from IGRF 2000: top is $B_r$ at Earth’s surface; middle the non-dipole contribution to $B_r$ at Earth’s surface and the bottom panel gives $B_r$ downward continued to the core-mantle boundary. IGRF models come with a Health Warning about their limitations. Recent versions extend to degree and order 13, but they do not include crustal anomalies.

### Gauss Coefficients for IGRF-2000

<table>
<thead>
<tr>
<th>$l$</th>
<th>$m$</th>
<th>$g_{lm}^m$</th>
<th>$h_{lm}^m$</th>
<th>$l$</th>
<th>$m$</th>
<th>$g_{lm}^m$</th>
<th>$h_{lm}^m$</th>
<th>$l$</th>
<th>$m$</th>
<th>$g_{lm}^m$</th>
<th>$h_{lm}^m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>-29615</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>74</td>
<td>64</td>
<td>9</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1728</td>
<td>5186</td>
<td>6</td>
<td>3</td>
<td>-161</td>
<td>65</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>-20</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>-2267</td>
<td>0</td>
<td>6</td>
<td>4</td>
<td>-5</td>
<td>-61</td>
<td>9</td>
<td>2</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>3072</td>
<td>-2478</td>
<td>6</td>
<td>5</td>
<td>17</td>
<td>1</td>
<td>9</td>
<td>3</td>
<td>-8</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1672</td>
<td>-458</td>
<td>6</td>
<td>6</td>
<td>-91</td>
<td>44</td>
<td>9</td>
<td>4</td>
<td>6</td>
<td>-6</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1341</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>79</td>
<td>0</td>
<td>9</td>
<td>5</td>
<td>-9</td>
<td>-8</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>-2290</td>
<td>-227</td>
<td>7</td>
<td>1</td>
<td>-74</td>
<td>-65</td>
<td>9</td>
<td>6</td>
<td>-2</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1253</td>
<td>296</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>-24</td>
<td>9</td>
<td>7</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>715</td>
<td>-492</td>
<td>7</td>
<td>3</td>
<td>33</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>-4</td>
<td>-8</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>935</td>
<td>0</td>
<td>7</td>
<td>4</td>
<td>9</td>
<td>24</td>
<td>9</td>
<td>9</td>
<td>-8</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>787</td>
<td>272</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>15</td>
<td>10</td>
<td>0</td>
<td>-2</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>251</td>
<td>-232</td>
<td>7</td>
<td>6</td>
<td>8</td>
<td>-25</td>
<td>10</td>
<td>1</td>
<td>-6</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-405</td>
<td>119</td>
<td>7</td>
<td>7</td>
<td>-2</td>
<td>-6</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>110</td>
<td>-304</td>
<td>8</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>10</td>
<td>3</td>
<td>-3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>-217</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>6</td>
<td>12</td>
<td>10</td>
<td>4</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>351</td>
<td>44</td>
<td>8</td>
<td>2</td>
<td>-9</td>
<td>-22</td>
<td>10</td>
<td>5</td>
<td>4</td>
<td>-6</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>222</td>
<td>172</td>
<td>8</td>
<td>3</td>
<td>-8</td>
<td>8</td>
<td>10</td>
<td>6</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>-131</td>
<td>-134</td>
<td>8</td>
<td>4</td>
<td>-17</td>
<td>-21</td>
<td>10</td>
<td>7</td>
<td>2</td>
<td>-3</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>-169</td>
<td>-40</td>
<td>8</td>
<td>5</td>
<td>9</td>
<td>15</td>
<td>10</td>
<td>8</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>-12</td>
<td>107</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>0</td>
<td>-2</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>72</td>
<td>0</td>
<td>8</td>
<td>7</td>
<td>-8</td>
<td>-16</td>
<td>10</td>
<td>10</td>
<td>-1</td>
<td>-8</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>68</td>
<td>-17</td>
<td>8</td>
<td>8</td>
<td>-7</td>
<td>-3</td>
<td>10</td>
<td>10</td>
<td>-1</td>
<td>-8</td>
</tr>
</tbody>
</table>
Field values in microteslas

**Figure 3.4.5**
3:4.6 The Geomagnetic Spectrum

In 1974 Frank Lowes (Geophys. J. Royal Astron. Soc. 36, pp 717-730, 1974) had the idea of plotting the magnetic energy averaged over the Earth’s surface as a function of wavelength. Of course on a sphere this really means as a function $SH$ degree, because by Jeans’ rule $\lambda = 2\pi a / (l + \frac{1}{2})$. Recall the energy density in a magnetic field is given by $B \cdot B / 2\mu_0$ so the energy integrated over a spherical surface is

$$E(a) = \int_{S(a)} \frac{B \cdot B}{2\mu_0} d^2s$$

$$= \frac{1}{2\mu_0} \int_{S(a)} \nabla \Psi \cdot \nabla \Psi \ d^2s$$  \hspace{1cm} (83)

$$= \frac{a^2}{2\mu_0} \int_{S(1)} [(\partial_i \Psi)^2 + |\nabla_i \Psi|^2] \ d^2\hat{s}$$  \hspace{1cm} (84)

where $\Psi$ is the usual scalar potential. Inserting the familiar SH expansion for $\Psi$ [equation (60) with $L = \infty$] we have:

$$\frac{2\mu_0 E(a)}{a^2} = \sum_{l,m} \sum_{l',m'} [(l+1)(l'+1) b^{m'}_l(b^{m'}_{l'})^* \int_{S(1)} Y^m_l(Y^m_{l'})^* d^2\hat{s} + b^m_l(b^m_l)^* \int_{S(1)} \nabla_1 Y^m_l \cdot (\nabla_1 Y^m_{l'})^* d^2\hat{s}].$$  \hspace{1cm} (85)

By the orthogonality relations of $Y^m_l$ and its surface gradient (3 and 12 in the table of lore) the double sums all collapse

$$\frac{2\mu_0 E(a)}{a^2} = \sum_{l,m} [(l+1)^2 + l(l+1)]|b^m_l|^2$$

$$= \sum_{l=1}^{\infty} (2l + 1)(l+1) \sum_{m=-l}^{l} |b^m_l|^2.$$  \hspace{1cm} (86)

This is essentially the result stated by (76). The idea is to see how the various harmonic degrees (equivalently, length scales) contribute to the total energy, and so traditionally one writes

$$R = \sum_{l=1}^{\infty} R_l$$  \hspace{1cm} (87)

where we define

$$R_l = \frac{(2l + 1)(l+1)}{4\pi} \sum_{m=-l}^{l} |b^m_l|^2 = (l+1) \sum_{m=0}^{l} [(g^m_l)^2 + (h^m_l)^2].$$  \hspace{1cm} (88)

Here $R_l$ is called the geomagnetic spectrum or the Lowes spectrum (or even the Mauersberger-Lowes spectrum, because Mauersberger mentioned $R_l$ some years before Lowes, though he didn’t do anything very original with it).

What Lowes discovered when he plotted the log of $R_l$ against $l$, and what we see with much more complete data, is that the spectrum breaks very obviously into two parts, the lower $l$ behavior well-fitted by a straight line on a linear-log plot. See Figure 3.4.6.1; the field model is due to Olsen et al., 2009 (Geophys. J. International, vol. 179, p. 1477-87, doi:
Sources of the Geomagnetic Field and the Modern Data That Enable Their Investigation

Fig. 3.4.6.1
Spatial power spectrum of the geomagnetic field at the Earth's surface. Black dots represent the spectrum of a recent field model (Olsen et al. 2009; Maus et al. 2008). Also shown are theoretical spectra (Voorhies et al., 2002) for the core (blue) and crustal (magenta) part of the field, as well as their superposition (red curve).

10.1111/j.1365-246X.2009.04386.x, We have exponential behavior, completely different from the gravity spectrum’s, which is essentially a power law (See Fig 7, Part I, though we plot a slightly different kind of spectrum, without the multiplying quadratic in $l$).

The natural interpretation of this result is that the two parts of the spectrum reflect different source regions – the long wavelength fields with $l \leq 14$ come almost entirely from the core, those with $l > 14$ from the crust. This idea is given considerable support from the observation that the equation of the best straight line through the core spectrum is $\alpha(r_1/a)^{2l}$ where $r_1 = 3,407$km, which is not very different from the core radius $c = 3,486$ km, according to seismologists. Why would one expect this equation for fields with their sources in the core? One argument runs as follows: consider the spectrum on spheres of different radii $r$. If it is permissible to downward continue the field through mantle, treating the mantle as a nonmagnetic insulator, then we find that

$$R_l(r) = (a/r)^{2l+4}R_l(a).$$

When we substitute $r = c$ and plot the spectrum that would be observed at the surface of the core we obtain the spectrum in Figure 3.4.6.2, which is almost flat. A flat spectrum is one in which there is equal energy at every scale, the sort of thing predicted in homogeneous turbulence. A plausible argument might be made that the fluid motions in the core cause a distribution of magnetic energy evenly into the different scales. This seems to work, except for the dipole term, which
clearly is unusually large, and doubtless this fact reflects the dominance of the Earth’s rotation in the geodynamo.

The large-$l$ spectrum makes no sense if the sources are the core; why should the energy apparently increase exponentially with $l$? Rather we can show (Jackson, *Geophys. J. Int.* 103, pp 657-74, 1990) that randomly distributed sources of magnetization in the crust would yield a mildly rising spectrum like the one seen in Figure 3.4.6.1. We suspect the core contribution continues on its exponential decline, but is completely obscured by the crustal field at shorter wavelengths.

Finally, Constable and Parker (*J. Geophys. Res.* 93, pp 11569-81, 1988) suggested that the flatness of $R_l(c)$ was most unlikely to be an accident of the 20th century, but instead it may be a persistent feature of the geomagnetic field common to the geodynamo throughout geologic history, with the possible exception of times near reversals. They also discovered that the coefficients $b_i^m$ when normalized to the radius of the core resembled a set of numbers drawn from a zero mean, Gaussian random process, a remarkable fact in itself. If this property holds for other geologic times, it becomes possible to predict what constitutes “normal” statistical behavior of the ancient geomagnetic field. Subsequent work has shown that the true field had some small but definite departures from the uniform Gaussian model, but it remains a useful approximation.

**3:5 Toroidal and Poloidal Fields**

Before we start, let us assemble two useful vector calculus identities, numbers 5 and 9 in our list:

\[
\nabla \times (\mathbf{U}s) = (\nabla \times \mathbf{U})s - \mathbf{U} \times \nabla s
\]

(90)

\[
\nabla \times \nabla \times \mathbf{U} = -\nabla^2 \mathbf{U} + \nabla(\nabla \cdot \mathbf{U}).
\]

(91)

In an earlier lecture we showed that the magnetic field can always be written as the curl of the magnetic vector potential $\mathbf{A}$. Now we consider a sphere of radius $c$ surrounded by an insulator and divide the vector potential into parts parallel to...
and perpendicular to \( \mathbf{r} \) by writing
\[
\mathbf{A} = T \mathbf{r} + \nabla P \times \mathbf{r} = T \mathbf{r} + \nabla \times (P \mathbf{r})
\]  
\tag{92}
where \( T \) and \( P \) are scalar functions of \( \mathbf{r} \), known as the defining scalars of the toroidal and poloidal fields. That this is always possible is shown in detail in Foundations, Chapter 5. To find \( \mathbf{B} \) we take the curl:
\[
\mathbf{B} = \nabla \times (T \mathbf{r}) + \nabla \times \nabla \times (P \mathbf{r}) = \mathbf{B}_T + \mathbf{B}_P
\]  
\tag{93}
and \( \mathbf{B}_T \) is called the \textit{toroidal part} of \( \mathbf{B} \), while \( \mathbf{B}_P \) is the \textit{poloidal part}. This decomposition for \( \mathbf{B} \) is unique and can always be done for all solenoidal vector fields (those with \( \nabla \cdot \mathbf{F} = 0 \)). We are very familiar with the idea of a potential for \( \mathbf{B} \) in an insulator; when \( \mathbf{J} \) does not vanish, we need two scalars, not one to describe \( \mathbf{B} \) completely. Conventionally, the scalars are always restricted to a class of functions whose average value over every sphere is zero, that is
\[
0 = \int_{S(c)} T(r \hat{\mathbf{r}}) \, d^2 \hat{\mathbf{r}} = \int_{S(c)} P(r \hat{\mathbf{r}}) \, d^2 \hat{\mathbf{r}}
\]  
\tag{94}
With this property, the scalars become unique in (92), which means that if \( \mathbf{B}_T \) vanishes, then \( T = 0 \), and similarly for \( \mathbf{B}_P \).

Using the vector identity (90) on the toroidal part of the field, we see
\[
\mathbf{B}_T = (\nabla \times \mathbf{r})T - \mathbf{r} \times \nabla T
\]  
\tag{95}
because \( \nabla \cdot \mathbf{r} = 0 \); hence \( \hat{\mathbf{r}} \cdot \mathbf{B}_T = 0 \) always; in other words, the toroidal magnetic field has no radial component – it is a tangent vector field on every concentric spherical surface. The lines of force lie on spherical surfaces and are thus confined to the interior of the conducting sphere. If we think of the sphere as Earth’s core, outside the core we have (for the sake of argument) \( \mathbf{J} = 0 \) and
\[
\mathbf{B} = -\nabla \Psi, \quad \text{with} \quad \nabla^2 \Psi = 0.
\]  
\tag{96}
Now \( \mathbf{B} \) is continuous at \( S(c) \) and since \( \hat{\mathbf{r}} \cdot \mathbf{B}_T = 0 \) just inside we conclude that \( \hat{\mathbf{r}} \cdot \mathbf{B}_T \) is also zero just outside the core. But we showed in Part I section 8, that any harmonic function with internal sources and vanishing radial component on \( S(c) \) is identically zero outside. Therefore \( \mathbf{B}_T \) vanishes outside the conducting sphere. Hence the toroidal part of \( \mathbf{B} \) in Earth’s core is invisible outside the core and only the poloidal part, \( \mathbf{B}_P \), has any detectable influence at the Earth’s surface.

When we neglect the displacement current term in Maxwell’s equations we can make a similar decomposition for the current flow in core, as we now show:
\[
\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \tag{97}
\]
\[
\nabla \cdot \mathbf{J} = 0 \tag{98}
\]
\[
\mathbf{J} = J_T + J_P. \tag{99}
\]
By substituting for $\mathbf{B}$ in terms of $\mathbf{B}_T$ and $\mathbf{B}_P$ in (93) we will show that $\mathbf{B}_P$ comes from $\mathbf{J}_T$ and $\mathbf{B}_T$ from $\mathbf{J}_P$, that is, toroidal fields are generated by poloidal current systems, and poloidal fields by toroidal currents.

Here is the proof: take the curl of (93)

$$\mu_0 \mathbf{J} = \nabla \times \mathbf{B} = \nabla \times \nabla \times (T \mathbf{r}) + \nabla \times \nabla \times (P \mathbf{r}).$$

(100)

It is obvious from the first term on the right that the toroidal part of $\mathbf{B}$ is associated with a poloidal current, whose scalar is just $T/\mu_0$. To show the last term, with three curls, is in fact toroidal, we analyze $S = \nabla \times \nabla \times (P \mathbf{r})$. From the identity (91)

$$S = -\nabla^2 (rP) + \nabla(\nabla \cdot (rP)).$$

(101)

We expand the first term in (101) with the Einstein notation:

$$\nabla^2 (rP) = \partial_j \partial_j (x_k P) = \partial_j ((\partial_j x_k) P + x_k \partial_j P) = \partial_j (\delta_{jk} P + x_k \partial_j P)$$

(102)

$$= \partial_k P + \partial_j (x_k \partial_j P) = \partial_k P + (\partial_j x_k) \partial_j P + x_k \partial_j \partial_j P$$

$$= \partial_k P + \delta_{jk} \partial_j P + x_k \partial_j \partial_j P = 2\partial_j P + x_k \partial_j \partial_j P$$

$$= 2\nabla P + r\nabla^2 P.$$  

(103)

Then (101) becomes

$$S = -r\nabla^2 P + \nabla(\nabla \cdot (rP) - 2P)$$

$$= -r\nabla^2 P + \nabla(r \partial_j P + P).$$

(104)

---

**Figure 3.5.1**
Finally, we take the curl of (104), noting the vanishing of the term involving a grad because $\nabla \times \nabla = 0$, so that

$$\nabla \times S = -\nabla \times (r\nabla^2 P)$$

and we recover (100)

$$\mu_0 J = -\nabla \times (r\nabla^2 P) + \nabla \times \nabla \times (Tr) = \mu_0 (J_T + J_P).$$

(105)

Thus the toroidal part of the current is derived from the toroidal scalar $-\nabla^2 P/\mu_0$.

The lengthy algebra may obscure what a toroidal or poloidal magnetic field might actually look like. The sketches in Figure 3.5.1 give an example of the simplest types of fields. You can easily see the toroidal system on the right as the currents generating the poloidal field on the left. But if you reverse the process, the field lines that stray outside the conductor (gray area) on the left are not allowed. Obviously not every toroidal or poloidal field has axial symmetry like these.

So we know that the geomagnetic field that we see is connected to a poloidal field in the core, which is generated by a toroidal current system $J_T$. The following nice result is also true (shown first by Gubbins 1975, and in more detail in Foundations, section 5.5, although there is an error of a factor $4\pi$ in the bounds given there):

$$\mu_0^2 \int_{r<c} J_T \cdot J_T d^3 r \geq c \sum_{l=1}^{\infty} \sum_{m=-l}^{m=l} \frac{(l+1)(2l+1)^2(2l+3)}{l} |b_{lm}|^2$$

(106)

where the coefficients $b_{lm}$ are those obtained on the surface of the core, whose radius is $c$. The amount of heat generated by ordinary Ohmic losses (also called Joule heating) at a point is given by $J \cdot E$. Thus (106) allows us to compute the minimum Joule heating from toroidal currents:

$$Q = \int_{r<c} \frac{J_T \cdot J_T}{\sigma} d^3 r$$

$$\geq \frac{c}{\mu_0^2 \sigma} \sum_{l=1}^{\infty} \sum_{m=-l}^{m=l} \frac{(l+1)(2l+1)^2(2l+3)}{l} |b_{lm}|^2.$$  

(107)

By using $Q$ as a regularizing norm, we can find the minimum amount of heating in the core associated with the poloidal magnetic field, the only part we can see. Of course the actual ohmic heating could be much larger than this because of the poloidal current term which remains invisible to us and because the estimate from (96) is a lower bound on toroidal current power in any case. It is controversial in dynamo theory whether the toroidal component of $B$ in the core is large compared with the poloidal part. Some models predict $|B_T| \sim 50 \times |B_P|$.

The Joule heating in the core is probably a small fraction of the heat budget (see the exercise), but it is believed that in certain neutron stars, called magnetars, the collapse of the magnetic field (with intensity $10^9$ T!) is the major source of energy in the body. See S. Kalkarni, Nature, v 419, pp 121-2, 2002.

Exercise:
(a) Make an estimate of the minimum rate of Joule heating generated in core. Compare this with terrestrial heat flow at the surface.

(b) A toroidal magnetic field \( \mathbf{B}_T \) fills a conducting sphere of radius \( a \); its scalar is the function \( T(r) \). Imagine creating a contour map of the values of \( T \) on the surface \( S(b) \) with \( b < a \). What connection does the contour map of the scalar \( T \) have with the magnetic field \( \mathbf{B}_T \)?