Introduction to Global Geophysics

GEO371/PHY371

Princeton University

December 23, 2021

Contents

1
Prelude

Geophysics is the physics of the Earth, the terrestrial planets, and moons. Physics is the science that studies relations between matter and energy, force fields and their observable effects, interactions between motion and momentum, action and reaction, transformations of measurable quantities, numbers, geometric objects. Physics deals with materials, their behaviors, and the constitutive laws that govern them. All the mathematical tools of *classical* physics (primarily mechanics and electromagnetism, and some thermodynamics), and, yes, even some borrowed from *modern* (quantum, statistical, computational) physics, are necessary to observe, describe, explain (and numerically reproduce) geophysical processes at work in and on the Earth. Hence this chapter.

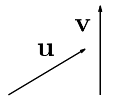
You will benefit from reading sections out of any of the many excellent textbooks aimed at budding geophysicists, e.g., Chapter 4 of [1], Chapter 1 of [2], but really, any introductory text on the mathematics of physics, especially [3], should be able to serve as a refresher. More advanced texts on mathematical (geo)physics that are on my bookshelf are, e.g., [4] and [5]. For an applied mathematics book that literally, has it all, I recommend [6]. Warning: you may just *have* to go to the library at some point. I know you've been postponing it.

Starting with some generality will simplify things greatly down the line. Many textbooks, none cited here, make a different choice, and they end up being often more confusing than enlightening. Hence, in this chapter, you will learn nothing about the Earth. Of course we know that the Earth is a *sphere* (a *ball*, if you will)—or just about; more about that later. When we do specify a coordinate system—and we *will* avoid this where we can—spherical polars will be handier than Cartesian coordinates. (Try finding the volume of the unit sphere by triple integration over x, y, and z! Call me when you have $4/3\pi$.)

1.1 Scalars, vectors, tensors, and their products

A *scalar* is a number. Negative six, $\frac{1-\sqrt{3}}{4\sqrt{2}}$, zero, one, two and three-quarters, one divided by 137.8, 1729, four-and-a-half billion, 2.718281828459046..., $(-1)^{1/2}$, π, ∞, you name it. Not `NaN`. Enough said.

A*vector* is an entity endowed with both magnitude and direction, which are both scalar. For this reason, vectors are most often represented geometrically by arrows (harpoons), with lengths proportional to their *magnitude*, and with arrowheads indicating their *direction*. You have, presumably, known this for longer than you can remember. A *unit vector* will have a hat on: $\hat{\mathbf{u}} = \mathbf{u}/||\mathbf{u}||$.

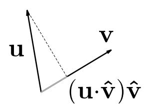


The **dot product** of vectors $\mathbf{u}$ and $\mathbf{v}$ is the scalar quantity given by

$\mathbf{u} \cdot \mathbf{v}$

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \Delta, \tag{1.1}$$

where Δ is the *acute* angle between the vectors $\mathbf{u}$ and $\mathbf{v}$, and $\|\mathbf{u}\|$ and $\|\mathbf{v}\|$ are their lengths. A formal definition of this notation is to come. It follows, however, immediately, that if $\mathbf{u}$ is perpendicular (*orthogonal*) to $\mathbf{v}$, their dot product vanishes.



Vectors living in geometrical space have *components*, which are scalars whose values ultimately will depend on the particular coordinate system chosen. The dot product (1.1) between two vectors is the sum of the component-by-component multiplication of both. Let us assume that there are three physical *dimensions*, thus three components u_i, v_i, $i = 1, 2, 3$. We will then have

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3, \tag{1.2}$$

and in a more compact notation, we write

$$\mathbf{u} \cdot \mathbf{v} = \sum_i u_i v_i = u_i v_i. \tag{1.3}$$

In the last equality the summation sign was skipped altogether. Indeed, in the *Einstein summation notation* a sum over repeated *indices* (i, in the multiplicative expression $u_i v_i$, but i in the sum $u_i + v_i$ would not count) is implied.

Only now do we arrive at the formal definition of the *length* of a vector used in eq. (1.1), as its *norm*. In the various notations developed so far, $\|\mathbf{u}\|$

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \Big(\sum_i u_i u_i\Big)^{1/2} = \sqrt{u_i^2}. \tag{1.4}$$

By combining eqs (1.1) and (1.4), the Euclidean distance *between* two points $\mathbf{r}$, for which $\|\mathbf{r}\| = r$, and $\mathbf{r}'$, for which $\|\mathbf{r}'\| = r'$, is also given by $\|\mathbf{r} - \mathbf{r}'\|$

$$\|\mathbf{r} - \mathbf{r}'\| = \left(r^2 + r'^2 - 2rr' \cos \Delta\right)^{1/2}, \tag{1.5}$$

known as the *law of cosines* in some circles, and as the *Pythagorean theorem* when the opening angle $\Delta = \pi/2$.

Eq. (1.4) is a “2-norm” since we end up with squares inside the square root. Sometimes you will see the explicit notation $\|\mathbf{u}\|_2$ for the above, and you will find its square, $\|\mathbf{u}\|_2^2 = \mathbf{u}\cdot\mathbf{u}$ being manipulated in other texts. With this elaboration we are ready for the usage of more general “p-norms”, defined as:

$$\|\mathbf{u}\|_p = \Big(\sum_i |u_i|^p\Big)^{1/p}. \tag{1.6}$$

In statistics and data analysis, vectors are simply *sets* of numbers. There, you will encounter the symbol $\|\mathbf{u}\|_0$ to mean *the number of nonzero elements* in a set $\mathbf{u}$, and $\|\mathbf{u}\|_2$ will be everywhere least-squares analysis for regression is being discussed. But with our definition, $\|\mathbf{u}\|_0$ is not a *proper* norm: eq. (1.6) is undefined when $p = 0$. On the contrary, $\|\mathbf{u}\|_1$ is well-defined, arising as the “1-norm”, the sum of the absolute values of the vector components or elements in the set. While we’re at it, $\|\mathbf{u}\|_\infty$ is the *largest element* of the component set.

$\mathbf{uv}$ Finally, there is another product by which to relate vector and tensor quantities. Let the dot product be known as the *inner product*, then the *outer product* shall be the **dyadic product** of vectors $\mathbf{u}$ and $\mathbf{v}$,

$$\mathbf{T} = \mathbf{u}\,\mathbf{v}, \tag{1.7}$$

a new quantity that we define to have the components

$$T_{ij} = u_i v_j. \tag{1.8}$$

A *tensor* is an object with more than one “index”. Tensors deserving the moniker are of *rank* two or more. By that token, *vectors* are “tensors” of rank one, and *scalars* of rank zero. A *matrix* is a popular way of representing tensors of rank two, written generically as $\mathbf{T}$, with scalar components T_{ij}, where both indices i and j range over the dimensions, as they did in eqs (1.2)–(1.3).

$\mathbf{R}(\Delta)$ An example is the *orthonormal* tensor $\mathbf{R}$ that rotates vectors clockwise in the plane over an angle Δ, preserving their lengths. Its rows and columns are of unit length, and its transpose is its inverse, which defines the *identity tensor*,

$$\mathbf{R}(\Delta) = \begin{pmatrix} \cos\Delta & \sin\Delta \\ -\sin\Delta & \cos\Delta \end{pmatrix}, \qquad \mathbf{R}^{-1} = \mathbf{R}^{\mathrm{T}}, \qquad \mathbf{R}\cdot\mathbf{R}^{\mathrm{T}} = \mathbf{I}. \tag{1.9}$$

$\mathbf{T}\cdot\mathbf{u}$ Tensors are *linear operators*: a “two-tensor” will act on a vector (or “one-tensor”) to produce another vector. To enact this property we form a dot product again, but now between a tensor and a vector. The result is a vector:

$$\mathbf{v} = \mathbf{T} \cdot \mathbf{u}. \tag{1.10}$$

The rule is simple: a dot product represents the *contraction*, or multiplication-and-summation-of-components between adjacent, dummy, indices. After contraction that index is gone: in component notation, the result of eq. (1.3) is

$$v_i = \sum_j T_{ij} u_j = T_{ij} u_j. \tag{1.11}$$

Now, look again upon eqs (1.7)–(1.8). There was no dot, indices did not repeat, they were not summed out, and instead of reducing the order, the result of the operation was a quantity of increased order!

The **cross product** of vectors $\mathbf{u}$ and $\mathbf{v}$ is the vector perpendicular to both $\mathbf{u}$ and $\mathbf{v}$ whose magnitude is given by $\mathbf{u} \times \mathbf{v}$

$$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \sin \Delta. \tag{1.12}$$

Its orientation is determined by the *right-hand rule*. Imagine positioning a cork screw perpendicularly onto the plane containing both $\mathbf{u}$ and $\mathbf{v}$, and then twisting it in the direction from $\mathbf{u}$ to $\mathbf{v}$. The direction of $\mathbf{u} \times \mathbf{v}$ will be given by the motion of the corkscrew, i.e. into ($\otimes$) or out of ($\odot$) this plane. Clearly, $\mathbf{u} \times \mathbf{v}$ and $\mathbf{v} \times \mathbf{u}$ will be different vectors, although they have the same magnitude: the cross product is not *commutative*, rather $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$. Also, $\mathbf{u} \times \mathbf{u} = \mathbf{0}$.

The vector that is the cross product

$$\mathbf{w} = \mathbf{u} \times \mathbf{v} \tag{1.13}$$

has elements that, in component notation, are given by

$$w_i = \sum_j \sum_k \epsilon_{ijk} u_j v_k = \epsilon_{ijk} u_j v_k. \tag{1.14}$$

Here, ϵ_{ijk} is the Levi-Cività *alternating symbol*. It takes on the value 1 if the ordered list $\{i, j, k\}$ is an even permutation of the numbers $\{1, 2, 3\}$, the value -1 if $\{i, j, k\}$ is an odd permutation of $\{1, 2, 3\}$, and 0 otherwise. At least four different identities [8] relate ϵ_{ijk} to δ_{ij}, the *Kronecker delta*, a symbol that evaluates to 1 when $i = j$ and 0 if $i \neq j$, and which we will encounter frequently. Easily verified is, for example, $\epsilon_{ijk}\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}$. ϵ_{ijk} δ_{ij}

Let us revisit the cross-product in a determinant notation that should give us the equivalent to eq. (1.14),

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix} \tag{1.15}$$

$$= (u_y v_z - u_z v_y)\,\hat{\mathbf{x}} + (u_z v_x - u_x v_z)\,\hat{\mathbf{y}} + (u_x v_y - u_y v_x)\,\hat{\mathbf{z}}. \tag{1.16}$$

The volume of the parallelepiped described by three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ is

$$V = |\mathbf{u} \cdot \mathbf{v} \times \mathbf{w}|. \tag{1.17}$$

1.2 From Cartesian to spherical coordinates, and back

(x, y, z) In a *Cartesian system*, the location of a point $\mathbf{r}$ is given in terms of the fixed unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ as

$$\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}, \tag{1.18}$$

whereby it is most convenient to think of $\mathbf{r}$ as the vector joining the origin of the coordinate system to the point of interest, with the set (x, y, z) containing its *Cartesian* coordinates, see Fig. 1.1.

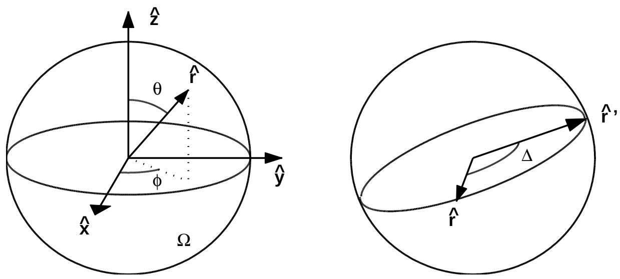


Fig. 1.1. Cartesian and spherical coordinates.

$\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ The vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ are mutually perpendicular and have unit length: they define an *orthogonal*, indeed, an *orthonormal* coordinate system. The components of $\mathbf{r}$ are "resolved" via the dot product:

$$x = \mathbf{r} \cdot \hat{\mathbf{x}}, \tag{1.19}$$

$$y = \mathbf{r} \cdot \hat{\mathbf{y}}, \tag{1.20}$$

$$z = \mathbf{r} \cdot \hat{\mathbf{z}}. \tag{1.21}$$

(r, θ, ϕ) *Spherical coordinates*, on the other hand, are the set (r, θ, ϕ) that describes the location of $\mathbf{r}$ in terms of its *distance* from to the origin, $r = \|\mathbf{r}\|$, its *colatitude* $0 \le \theta \le \pi$, and its *longitude* $0 \le \phi < 2\pi$. The transformation between Cartesian and spherical coordinates is achieved by the relations

$$\begin{aligned} x &= r\sin\theta\cos\phi, & r &= \sqrt{x^2+y^2+z^2}, \\ y &= r\sin\theta\sin\phi, & \theta &= \tan^{-1}(\sqrt{x^2+y^2}/z), \\ z &= r\cos\theta, & \phi &= \tan^{-1}(y/x). \end{aligned} \tag{1.22}$$

The *fixed* Cartesian **unit vectors** $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ have counterparts that *vary* with position: at every location $\mathbf{r}$ the unit vectors $\hat{\mathbf{r}}(\mathbf{r})$, $\hat{\boldsymbol{\theta}}(\mathbf{r})$ and $\hat{\boldsymbol{\phi}}(\mathbf{r})$ define a *local* coordinate system that can be related to the Cartesian axes $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ as $\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}$

$$\begin{pmatrix} \hat{\mathbf{r}} \\ \hat{\boldsymbol{\theta}} \\ \hat{\boldsymbol{\phi}} \end{pmatrix} = \boldsymbol{\Gamma}^{\mathrm{T}} \cdot \begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{y}} \\ \hat{\mathbf{z}} \end{pmatrix}, \qquad \begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{y}} \\ \hat{\mathbf{z}} \end{pmatrix} = \boldsymbol{\Gamma} \cdot \begin{pmatrix} \hat{\mathbf{r}} \\ \hat{\boldsymbol{\theta}} \\ \hat{\boldsymbol{\phi}} \end{pmatrix}, \tag{1.23}$$

whereby $\boldsymbol{\Gamma}$ defines the *orthogonal* (its transpose being its inverse, producing the identity upon multiplication) matrix transformation $\boldsymbol{\Gamma}$

$$\boldsymbol{\Gamma} = \begin{pmatrix} \sin\theta\cos\phi & \cos\theta\cos\phi & -\sin\phi \\ \sin\theta\sin\phi & \cos\theta\sin\phi & \cos\phi \\ \cos\theta & -\sin\theta & 0 \end{pmatrix}, \qquad \boldsymbol{\Gamma}\cdot\boldsymbol{\Gamma}^{\mathrm{T}} = \mathbf{I}. \tag{1.24}$$

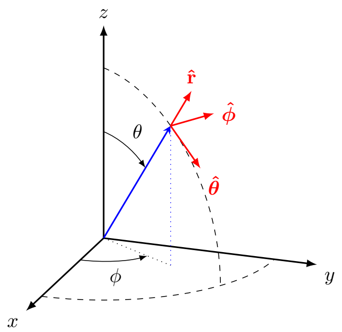


Fig. 1.2. Spherical unit vectors.

By definition, the Cartesian unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ point in the directions in which the position coordinates x, y and z, of $\mathbf{r}$ increase, thus

$$\partial_x \mathbf{r} = \hat{\mathbf{x}}, \quad \partial_y \mathbf{r} = \hat{\mathbf{y}}, \quad \partial_z \mathbf{r} = \hat{\mathbf{z}}. \tag{1.25}$$

In the same manner we define the spherical unit vectors $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$ to point in the direction of increasing r, θ and ϕ with the position vector $\mathbf{r}$. We write

$$\partial_r \mathbf{r} = \hat{\mathbf{r}}, \quad \partial_\theta \mathbf{r}/r = \hat{\boldsymbol{\theta}}, \quad \partial_\phi \mathbf{r}/(r\sin\theta) = \hat{\boldsymbol{\phi}}. \tag{1.26}$$

The *scale factors* 1, r and $r\sin\theta$ follow from the requirement that the right-hand sides are of length unity. Use eq. (1.22) to write the position vector $\mathbf{r} = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)$, take the derivatives, and then calculate the norm of the result. Use trigonometry to find the divisor that normalizes the result to one. After you've done that, you will have discovered that the Cartesian components of the spherical unit vectors are the *columns* of $\boldsymbol{\Gamma}$ in eq. (1.24) and the left-hand side of eq. (1.23) will verified. Inverting the relation will add the right-hand side once you notice the orthogonality of $\boldsymbol{\Gamma}$.

With the colatitude of a geographical point on the unit sphere $\hat{\mathbf{r}}$ denoted by $0 \leq \theta \leq \pi$ and the longitude by $0 \leq \phi < 2\pi$, the geodesic angular distance between two points $\hat{\mathbf{r}}$ and $\hat{\mathbf{r}}'$, as shown in Fig. 1.1, will be denoted by Δ, where

$$\cos\Delta = \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}' = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi - \phi'). \tag{1.27}$$

Finally, by differentiation of the columns of $\boldsymbol{\Gamma}$ of eq. (1.24), we find that the *partial derivatives* of the spherical unit vectors themselves are given by

$$\begin{aligned}
&\partial_r \hat{\mathbf{r}} = \mathbf{0}, \quad &&\partial_\theta \hat{\mathbf{r}} = \hat{\boldsymbol{\theta}}, \quad &&\partial_\phi \hat{\mathbf{r}} = \hat{\boldsymbol{\phi}}\,\sin\theta,\\
&\partial_r \hat{\boldsymbol{\theta}} = \mathbf{0}, \quad &&\partial_\theta \hat{\boldsymbol{\theta}} = -\hat{\mathbf{r}}, \quad &&\partial_\phi \hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\phi}}\,\cos\theta,\\
&\partial_r \hat{\boldsymbol{\phi}} = \mathbf{0}, \quad &&\partial_\theta \hat{\boldsymbol{\phi}} = \mathbf{0}, \quad &&\partial_\phi \hat{\boldsymbol{\phi}} = -\hat{\mathbf{r}}\sin\theta - \hat{\boldsymbol{\theta}}\cos\theta.
\end{aligned} \tag{1.28}$$

We will be using these soon enough.

1.3 Fields

$\mathbf{u}(\mathbf{r})$ A **vector field** $\mathbf{u}(\mathbf{r})$ is represented in a Cartesian system as

$$\mathbf{u} = u_x\hat{\mathbf{x}} + u_y\hat{\mathbf{y}} + u_z\hat{\mathbf{z}}, \tag{1.29}$$

and in spherical polar coordinates as

$$\mathbf{u} = u_r\hat{\mathbf{r}} + u_\theta\hat{\boldsymbol{\theta}} + u_\phi\hat{\boldsymbol{\phi}}. \tag{1.30}$$

As should follow immediately from eqs (1.23)–(1.30), the coordinate functions in both representations, which all vary with position $\mathbf{r}$, transform as

$$\begin{pmatrix} u_r \\ u_\theta \\ u_\phi \end{pmatrix} = \boldsymbol{\Gamma}^{\mathrm{T}} \cdot \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix}, \qquad \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} = \boldsymbol{\Gamma} \cdot \begin{pmatrix} u_r \\ u_\theta \\ u_\phi \end{pmatrix}. \tag{1.31}$$

The easiest route to this result is to rewrite eq. (1.30) using a vector product for the components, then using eq. (1.23), evaluating eq. (1.24), writing it all out, and collecting terms.

A **tensor field** $\mathbf{T}(\mathbf{r})$ is represented in a Cartesian system as $\mathbf{T}(\mathbf{r})$

$$\begin{aligned}\mathbf{T} &= T_{xx}\hat{\mathbf{x}}\hat{\mathbf{x}} + T_{xy}\hat{\mathbf{x}}\hat{\mathbf{y}} + T_{xz}\hat{\mathbf{x}}\hat{\mathbf{z}} \\ &\quad + T_{yx}\hat{\mathbf{y}}\hat{\mathbf{x}} + T_{yy}\hat{\mathbf{y}}\hat{\mathbf{y}} + T_{yz}\hat{\mathbf{y}}\hat{\mathbf{z}} \\ &\quad + T_{zx}\hat{\mathbf{y}}\hat{\mathbf{x}} + T_{zy}\hat{\mathbf{z}}\hat{\mathbf{y}} + T_{zz}\hat{\mathbf{z}}\hat{\mathbf{z}},\end{aligned} \tag{1.32}$$

using dyads of the coordinate vectors $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$, and, similarly, using $\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}$,

$$\begin{aligned}\mathbf{T} &= T_{rr}\hat{\mathbf{r}}\hat{\mathbf{r}} + T_{r\theta}\hat{\mathbf{r}}\hat{\boldsymbol{\theta}} + T_{r\phi}\hat{\mathbf{r}}\hat{\boldsymbol{\phi}} \\ &\quad + T_{\theta r}\hat{\boldsymbol{\theta}}\hat{\mathbf{r}} + T_{\theta\theta}\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}} + T_{\theta\phi}\hat{\boldsymbol{\phi}}\hat{\boldsymbol{\phi}} \\ &\quad + T_{\phi r}\hat{\boldsymbol{\phi}}\hat{\mathbf{r}} + T_{\phi\theta}\hat{\boldsymbol{\phi}}\hat{\boldsymbol{\theta}} + T_{\phi\phi}\hat{\boldsymbol{\phi}}\hat{\boldsymbol{\phi}},\end{aligned} \tag{1.33}$$

whose position-dependent elements we collect in a matrix as

$$\mathbf{T} = \begin{pmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz}. \end{pmatrix}, \qquad \mathbf{T} = \begin{pmatrix} T_{rr} & T_{r\theta} & T_{r\phi} \\ T_{\theta r} & T_{\theta\theta} & T_{\theta\phi} \\ T_{\phi r} & T_{\phi\theta} & T_{\phi\phi}. \end{pmatrix}. \tag{1.34}$$

Peeling off the individual entries is as simple as it was with eqs (1.19)–(1.21), by application of the dot product on either side of the tensor with the unit vectors in whichever coordinate system:

$$T_{rr} = \hat{\mathbf{r}} \cdot \mathbf{T} \cdot \hat{\mathbf{r}}, \qquad T_{xz} = \hat{\mathbf{x}} \cdot \mathbf{T} \cdot \hat{\mathbf{z}}, \qquad T_{r\phi} = \hat{\mathbf{r}} \cdot \mathbf{T} \cdot \hat{\boldsymbol{\phi}}, \tag{1.35}$$

and so on. Either way, if we *number* the coordinate vectors instead of *naming* them explicitly, the expressions generalizing eqs (1.29)–(1.33) simplify to

$$\mathbf{T} = T_{ij}\hat{\mathbf{x}}_i\hat{\mathbf{x}}_j \qquad \text{and} \qquad \mathbf{u} = u_i\hat{\mathbf{x}}_i. \tag{1.36}$$

1.4 Chain rule and the change-of-variables theorem

Remember Leibniz' **differentiation rule** for *composed* functions:

$$\frac{d}{dr} f(x(r)) = \left(\frac{dx}{dr}\right)\left(\frac{d}{dx} f\right). \tag{1.37}$$

Remember how to *integrate* as well as you can *differentiate*? Recall the fundamental theorem of the calculus, without worrying about the details for now:

$$\int_a^b f(x)\,dx = F(b) - F(a), \qquad \text{whereby} \qquad \frac{dF}{dx} = f(x) \tag{1.38}$$

defines F to be the *primitive function* or *antiderivative* of f. If we now also recall the *substitution rule* for a case where the variable x becomes identified with a function in terms of another variable, $x(r)$:

$$\int f(x)\,dx = \int f(x(r))\left(\frac{dx}{dr}\right)\,dr, \tag{1.39}$$

we can see that eq. (1.37) helps us find a primitive, in the sense of eq. (1.38), for the integrand on the right side of eq. (1.39).

Suffice to write eq. (1.38) as

$$\int f(x(r))\,dx(r) = F(x(r)) = \int g(r)\,dr, \qquad \text{if} \qquad \frac{dF}{dr} = g(r), \tag{1.40}$$

and thus if $F(x(r))$ is the composite function of eq. (1.37),

$$g(r) = \left(\frac{dx}{dr}\right)\left(\frac{d}{dx}F\right) = \left(\frac{dx}{dr}\right) f(x), \tag{1.41}$$

which is consistent with how we've done the substitution in eq. (1.39).

Integrate a function, differentiate again, and you get the same function back. Differentiate a function, integrate the result, and you also recover the original (up to a constant). When the upper limit of a *definite* integral is itself a function of the variable with respect to which you subsequently wish to differentiate the result, we can use the chain rule to prove Leibniz **integral rule** for differentiation under the integral sign:

$$\frac{d}{dx}\int_{f(x)}^{g(x)} h(t)\,dt = h[g(x)]g'(x) - h[f(x)]f'(x). \tag{1.42}$$

Consider the upper limit by positing $u = g(x)$ and using the chain rule

$$\frac{d}{dx}\int_{f(x)}^{g(x)} h(t)\,dt = \left(\frac{d}{du}\int_{f(x)}^{u} h(t)\,dt\right)\frac{du}{dx} = h(g(x))\frac{dg}{dx}. \tag{1.43}$$

J The multidimensional **chain rule** of differentiation is a generalization of eq. (1.37), applied for the other variables y, z, and for θ and ϕ. We express the effect on the partial derivatives of the transformation to and from Cartesian to spherical coordinates $(x, y, z) \leftrightarrow (r, \theta, \phi)$ concisely as:

$$\begin{pmatrix} \partial_r \\ \partial_\theta \\ \partial_\phi \end{pmatrix} = \mathbf{J}^{\mathrm{T}} \cdot \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix}, \qquad \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} = \mathbf{J}^{-\mathrm{T}} \cdot \begin{pmatrix} \partial_r \\ \partial_\theta \\ \partial_\phi \end{pmatrix}, \tag{1.44}$$

whereby $\mathbf{J}$ defines the *Jacobian* of the transformation $(r, \theta, \phi) \to (x, y, z)$ as

$$\mathbf{J} = \begin{pmatrix} \partial_r x & \partial_\theta x & \partial_\phi x \\ \partial_r y & \partial_\theta y & \partial_\phi y \\ \partial_r z & \partial_\theta z & \partial_\phi z \end{pmatrix} \tag{1.45}$$

$$= \begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}.$$

The multi-dimensional version of the integral **substitution rule** of eq. (1.39) is the *change-of-variables theorem*, written in the short-hand form

$$\iiint\limits_V u^{\mathrm{C}}(x, y, z)\, dxdydz = \iiint\limits_{V'} u^{\mathrm{S}}(r, \theta, \phi) \left| \frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} \right| drd\theta d\phi. \tag{1.46}$$

For the Cartesian-to-spherical case we end up with the well-known

$$\iiint\limits_V u^{\mathrm{C}}(x, y, z)\, dxdydz = \iiint\limits_{V'} u^{\mathrm{S}}(r, \theta, \phi)\, r^2 \sin\theta\, drd\theta d\phi. \tag{1.47}$$

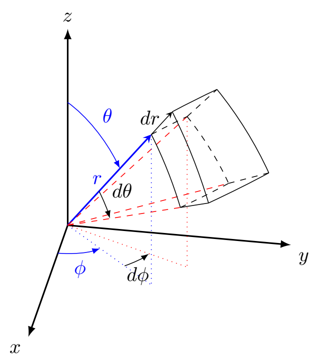


Fig. 1.3. Spherical volume element.

The Cartesian **volume element**

$$dV^C = dxdydz, \tag{1.48}$$

becomes, in spherical polar coordinates,

$$dV^S = r^2 \sin\theta\, dr\, d\theta\, d\phi. \tag{1.49}$$

The ratio of the volume elements in the two coordinate systems is the determinant of the Jacobian, $|\mathbf{J}| = \det \mathbf{J} = r^2 \sin\theta$. It is also precisely the product of the scale factors as can be derived purely geometrically from Fig. 1.3.

2
Gravity

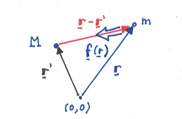
What do we want? To answer the age-old question: what's inside the Earth? The prime variable of interest is the **mass density**, an intrinsic property that is meaningfully correlated with *temperature* (rocks typically expand with temperature), *pressure* (rocks typically get denser with increased pressure), and *chemical composition* (rocks are polymineralic aggregates whose density depends on the types of atoms and molecules that compose them, and on how these are assembled). If we can come up with models for the gross density structure of the Earth, from measurements *at* (made by humans with instruments) or *outside* (made by satellites) its surface, we then need to find a way to study **density anomalies** inside the Earth from measurements of gravitational potential, gravitational acceleration, moments of inertia, however obtained. In short, we need to solve the *gravimetric inverse problem*.

We start with **Newton** and his point masses, and make our way to observing the Earth in the satellite age. Our derivations rely on an increasingly complete description of the structure of the Earth as a planet—a *ball*, or rather, a flattened *ellipsoid*, and one that *rotates*, at that. As we strip off, one by one, the "known", or most easily explained "big" effects, what will be left are "anomalies" that will reveal geological structure: chemical differentiation, pressure effects, and temperature perturbations.

Most of our interest in this chapter will be kept by the decription and understanding of the largely unchanging, static gravity field of the Earth. We will set up a formalism that is not time-dependent. This means we will be ignoring Earth and oceanic tides, and we will only mention any secular time-dependence at the end—the inevitable result of ice cap melting and sea-level change under global warming.

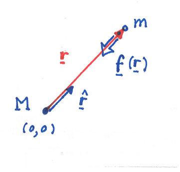
The geophysical literature on gravity that is on my shelf is dominated by an exceptionally well-written book [9], which I heartily recommend to all.

2.1 Force and acceleration



A point mass M at $\mathbf{r}'$ exerts on a point mass m at $\mathbf{r}$ a gravitational pull $\mathbf{f}$ according to, with the gravitational constant $G = 6.67408 \times 10^{-11}\ \mathrm{m^3 kg^{-1} s^{-2}}$,

$$\mathbf{f}(\mathbf{r}) = -G\frac{Mm}{\|\mathbf{r}-\mathbf{r}'\|^3}\,(\mathbf{r}-\mathbf{r}'). \tag{2.1}$$



The name *inverse-square law* is all the more intuitively justified if we let the origin of the coordinate system coincide with the location of the mass M, as

$$\mathbf{f}(\mathbf{r}) = -G\frac{Mm}{\|\mathbf{r}\|^2}\,\hat{\mathbf{r}} = -G\frac{Mm}{r^2}\,\hat{\mathbf{r}}. \tag{2.2}$$

Be careful about the sign in the previous expression, and think about the dimensions, [MLT^{-2}], and units of force, N—the *Newton*.

Of course we will be thinking of M as the mass of the "big" object (the Earth!), and of m as the mass of the "test" object (Newton's apple!). Normalized by the mass of the test object (force equals mass times acceleration), the *gravitational acceleration* is the vector quantity

M

$\mathbf{g}$

$$\mathbf{g}(\mathbf{r}) = -G\frac{M}{r^2}\,\hat{\mathbf{r}}. \tag{2.3}$$

Again, watch the sign, and observe the dimensions, [LT^{-2}] and units, ms^{-2}. From the ballpark figure for the magnitude of the gravitational acceleration at the surface of the Earth, derive the mean density of the Earth—and note that it is much larger than expected from the density of rocks that you will find at the Earth's surface. Conclusion: the Earth's density must increase with depth.

2.2 Potential

The gravitational *point-mass potential* $U(\mathbf{r})$ is the *energy* it takes to bring a unit test mass from infinity, where its potential is defined to be zero, to the position $\mathbf{r}$ in the gravitational field of the *point mass* M. We thus have

U

$$U(\mathbf{r}) = \int_\infty^{\mathbf{r}} \mathbf{g}(\mathbf{r}')\cdot d\mathbf{r}' = G\int_\infty^{r}\frac{M}{r'^2}\,dr' = -G\frac{M}{r}. \tag{2.4}$$

Note the spherical symmetry: only the mass and the distance count. The sign is established by convention: *at* the location of the point mass, the potential is negative infinity.

2.3 The gradient (of a scalar field)

dU From eq. (2.4), we heuristically deduce that, in Cartesian coordinates,

$$dU = \mathbf{g} \cdot d\mathbf{r} = -g_x dx - g_y dy - g_z dz. \tag{2.5}$$

∂_i Another way of describing the *total variation* of the potential $U(\mathbf{r})$ is as

$$dU = \left(\frac{\partial U}{\partial x}\right) dx + \left(\frac{\partial U}{\partial x}\right) dy + \left(\frac{\partial U}{\partial x}\right) dz. \tag{2.6}$$

$\boldsymbol{\nabla} U$ From eq. (2.6) we see that the Cartesian components of the gravitational acceleration are the negatives of the *partial derivatives* of the gravitational potential. Note that when the potential is constant, there is *no* acceleration. We write

$$\boxed{\mathbf{g}(\mathbf{r}) = -\boldsymbol{\nabla} U(\mathbf{r})}, \tag{2.7}$$

and have thereby introduced the *gradient*, which, in Cartesian coordinates, is
$\boldsymbol{\nabla}$ the differential *operator*

$$\boldsymbol{\nabla} = \hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}, \qquad \text{or, equivalently,} \qquad \boldsymbol{\nabla} = \hat{\mathbf{x}}_i \partial_i. \tag{2.8}$$

It is to be understood that operators act on *something*—in the case of eq. (2.7), on the scalar function of position $U(\mathbf{r})$. In spherical coordinates,

$$\boldsymbol{\nabla} = \hat{\mathbf{r}}\frac{\partial}{\partial r} + \hat{\boldsymbol{\theta}}\frac{1}{r}\frac{\partial}{\partial \theta} + \hat{\boldsymbol{\phi}}\frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}, \tag{2.9}$$

which can be derived by using the expressions (1.23) and the chain rule (1.44), collecting and equating terms.

In the next chapter we'll be needing, from eq. (1.28), that

$$\boldsymbol{\nabla}\hat{\mathbf{r}} = \frac{1}{r}(\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}} + \hat{\boldsymbol{\phi}}\hat{\boldsymbol{\phi}}) = \frac{1}{r}(\mathbf{I} - \hat{\mathbf{r}}\hat{\mathbf{r}}). \tag{2.10}$$

2.4 Point masses, no more

Let us consider the potential $dU(\mathbf{r})$, at some location $\mathbf{r}$, that is due to the presence, at some other location $\mathbf{r}'$, of an infinitesimal amount of mass $dM(\mathbf{r}')$ inside some infinitesimal volume dV'. This increment of potential is given by

$$dU(\mathbf{r}) = -G\,\frac{1}{\|\mathbf{r}-\mathbf{r}'\|}\,dM(\mathbf{r}') = -G\,\frac{\rho(\mathbf{r}')}{\|\mathbf{r}-\mathbf{r}'\|}\,dV', \tag{2.11}$$

where we have introduced the *volumetric mass density*, $\rho(\mathbf{r})$. If $\mathbf{r}' = \mathbf{0}$, we recover the differential form of eq. (2.4). Upon integration over the total volume, we obtain the universal expression ρ V

$$\boxed{U(\mathbf{r}) = -G \int_V \frac{\rho(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|} \, dV'} . \tag{2.12}$$

The gravitational field, with that same generality of expression, is

$$\boxed{\mathbf{g}(\mathbf{r}) = -G \int_V \frac{\rho(\mathbf{r}') \, (\mathbf{r} - \mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|^3} \, dV'} , \tag{2.13}$$

which harks back all the way to eq. (2.1) and of course recovers eqs (2.3) and (2.4) for a point mass at $\mathbf{r}' = \mathbf{0}$. Per eq. (1.49), $dV' = r'^2 \sin\theta' \, d\theta' \, d\phi'$, and eq. (1.5) in our coordinate system is $\|\mathbf{r} - \mathbf{r}'\| = \left(r^2 + r'^2 - 2rr' \cos\theta\right)^{1/2}$.

2.5 A *spherical* Earth

The gravitational potential and the acceleration *outside* a *spherically symmetric* "Earth", where $\rho(\mathbf{r}) = \rho(||\mathbf{r}||) = \rho(r)$, from eqs (2.12) and (2.13), are identical to eqs (2.3) and (2.4). In other words, such an idealized situation is identical to one in which all of the mass were to be concentrated at the origin of the coordinate system. Newton's law not only applies to point masses, but also to spherically symmetric mass distributions—in the right coordinate system.

What about potential and acceleration *inside* a *spherically symmetric* body of radius a? The expression (2.12) breaks down into an integral over the radial coordinate r' in two parts: (I) in which the queried position r is on the *outside* of the mass that it encloses, $(r'/r) < 1$, and (II) in which r is on the *inside* of the mass that still surrounds it, $(r/r') < 1$. Hence, factoring r or r' out from eq. (1.5), and developing the result $(1+x)^{1/2}$, where x is contains either (r'/r) or (r/r'), in a Taylor series around $x = 0$, as appropriate, and disregarding the colatitudinal dependence on the grounds of symmetry considerations, leaves a

$$U(\mathbf{r}) = \underbrace{-\,G\frac{4\pi}{r}\int_0^r \rho(r')\, r'^2 \, dr'}_{\text{(I)}} \underbrace{-\,G\,4\pi \int_r^a \rho(r')\, r' \, dr'}_{\text{(II)}} \tag{2.14}$$

when $\|\mathbf{r}\| = r \leq a$. The term (I) simply picks up $-GM(r)/r$ as in eq (2.4).

The gravitational acceleration at a distance $r \leq a$ from the center is only due to the portion of mass that is enclosed in the sphere of radius r as follows:

$$\mathbf{g}(\mathbf{r}) = -\hat{\mathbf{r}}\, G \frac{4\pi}{r^2} \int_0^r \rho(r')\, r'^2\, dr' = -G\frac{M(r)}{r^2}\,\hat{\mathbf{r}}. \tag{2.15}$$

The expressions (2.14)–(2.15) contain the results that the gravitational potential inside a *homogeneous* sphere grows quadratically with the distance from the center, whereas the accceleration grows linearly. Indeed, if the density were to be *constant*, $\rho(\mathbf{r}) = \rho$, we can evaluate eqs (2.14)–(2.15) to yield

$$U(\mathbf{r}) = -G\frac{2}{3}\pi\rho\left(3a^2 - r^2\right), \qquad \text{for} \quad r \le a, \tag{2.16}$$

$$\mathbf{g}(\mathbf{r}) = -G\frac{4}{3}\pi\rho\, r\,\hat{\mathbf{r}}, \qquad \text{for} \quad r \le a. \tag{2.17}$$

The results in this section comprise the notions that the gravitational potential everywhere inside a spherically symmetric hollow *shell* is constant, and thus the acceleration zero. Newton knew all of that, and so do we, now.

It is this first-order picture of a non-rotating, spherically symmetric Earth, that we shall be refining in subsequent sections.

2.6 Potential *outside* an Earth-like body

We shall now reevaluate eq. (2.12) outside a mass distribution with some more generality, i.e., without recourse to spherical symmetry in the density nor the shape of the volume of interest. Take a look at Figure 2.1 for what follows.

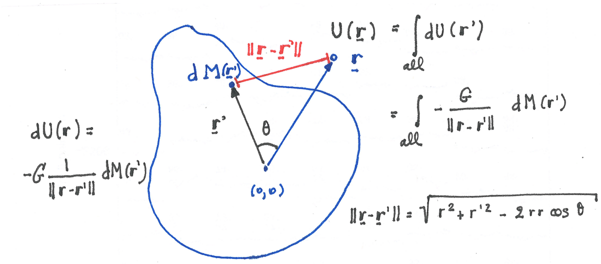


Fig. 2.1. Derivation of eq. (2.20), which expresses the gravitational potential *outside* a matter-filled portion of space.

We rewrite the distance $\|\mathbf{r} - \mathbf{r}'\|$ of eq. (1.5), by factoring out the query point r, and expand the result in a Taylor (Maclaurin) series in $(r'/r) < 1$ that we truncate after the *third* term, to obtain, accurate to *second* order,

$$\frac{r}{\|\mathbf{r}-\mathbf{r}'\|} = \left[1+\left(\frac{r'}{r}\right)^2 - 2\left(\frac{r'}{r}\right)\cos\theta\right]^{-\frac{1}{2}} \tag{2.18}$$

$$\approx \left[1+\left(\frac{r'}{r}\right)\cos\theta + \frac{1}{2}\left(\frac{r'}{r}\right)^2\left(3\cos^2\theta - 1\right)\right], \tag{2.19}$$

which is accurate at some distance away from the object, when $(r'/r) < 1$. After integration, this turns the potential due to, but sufficiently *away* from, the *entire* mass assemblage, correct to *second* order, into

$$U(\mathbf{r}) = \underbrace{-\frac{G}{r}\int dM}_{\text{(I)}} \underbrace{-\frac{G}{r^2}\int r'\cos\theta\, dM}_{\text{(II)}} \underbrace{-\frac{G}{2r^3}\int r'^2\left(3\cos^2\theta-1\right) dM}_{\text{(III)}}. \tag{2.20}$$

The first term in this sum is simply the point-mass potential eq. (2.4) again. It equals the potential at some distance r due to *all* of the mass concentrated at the origin—the *zeroth moment* of the mass distribution. The second term is a *first moment* of the mass density. It vanishes upon choosing the origin of our coordinate system to coincide with the center-of-mass of the object under consideration. The third term is a *second moment* of the mass density distribution.

We denote the nth *moment* of a distribution $\rho(\mathbf{r})$ about a certain point $\mathbf{r}$ as its integral weighted by the shifted coordinate raised to the nth power:

$$\int ||\mathbf{r}'-\mathbf{r}||^n \rho(\mathbf{r}')\, d^3\mathbf{r}'. \tag{2.21}$$

If, as in statistics, ρ were a *probability density function*, the zeroth, first (about the origin) and second (about the first) moments would be the total *mass of probability*, its *expectation*, and its *variance*. In physics, with ρ the mass density, the *center of mass* is the location about which the first moment of mass density vanishes. The second moment, as you will recognize from your study of mechanics, is a measure of the *rotational inertia* in the mass distribution. Do make the connection with eq. (1.6). But for now, we recognize in eq. (2.20) the succession of powers in the spherical distance r'. Next, we will relate those explicitly to the rotational moments of inertia of our planetary system.

2.6.1 *Factoring out moments of inertia*

Let us rewrite the integral in the third term of eq (2.20) using the trigonometric identity $\sin^2\theta + \cos^2\theta = 1$ as follows:

$$\frac{1}{2}\int r'^2\left(3\cos^2\theta - 1\right)\,dM = \underbrace{\int r'^2\,dM}_{\text{(I)}} - \frac{3}{2}\underbrace{\int r'^2\sin^2\theta\,dM}_{\text{(II)}}. \tag{2.22}$$

The first term of the right hand side of eq. (2.22) is proportional to the sum of the three *moments-of-inertia* for rotation of the mass point at $\mathbf{r}' = (x', y', z')$ around the arbitrarily oriented axes $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$, and the second term to its moment-of-inertia with respect to rotation around the axis $\hat{\mathbf{r}}$:

$$\text{(I)}\begin{cases} I(\hat{\mathbf{x}}) = \int (y^2+z^2)\,dM, & \text{around } \hat{\mathbf{x}},\\ I(\hat{\mathbf{y}}) = \int (x^2+z^2)\,dM, & \text{around } \hat{\mathbf{y}},\\ I(\hat{\mathbf{z}}) = \int (x^2+y^2)\,dM, & \text{around } \hat{\mathbf{z}}, \end{cases} \tag{2.23}$$

$$\text{(II)}\;\left\{\; I(\hat{\mathbf{r}}) = \int (r'\sin\theta)^2\,dM, \quad \text{around } \hat{\mathbf{r}}. \right. \tag{2.24}$$

This can be seen in Fig. 2.2 by rewriting the squared distance to the origin system as $r'^2 = [(y'^2+z'^2)+(x'^2+z'^2)+(x'^2+y'^2)]/2$, and recognizing the perpendicular distance of the integration point $\mathbf{r}'$ to $\hat{\mathbf{r}}$ as $r'\sin\theta$. Hence

$$\frac{1}{2}\int r'^2\left(3\cos^2\theta - 1\right)\,dM = \frac{1}{2}\left[I(\hat{\mathbf{z}}) + I(\hat{\mathbf{y}}) + I(\hat{\mathbf{z}}) - 3I(\mathbf{r})\right]. \tag{2.25}$$

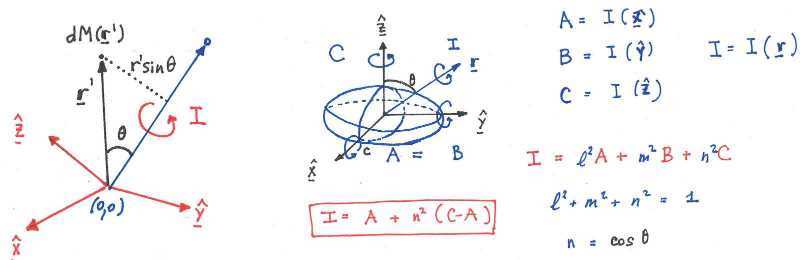


Fig. 2.2. Derivation of eqs (2.25) and (2.26).

2.6.2 Picking a principal-moments coordinate system

After *anchoring* our coordinate system $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ to center of mass of our planet, we now *orient* it so that the *principal axes* of the moment-of-inertia *tensor* are $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$, such that A, B, and C are the *principal* moments-of-inertia, see Fig. 2.2. In doing so we transform eq. (2.20) through eq. (2.25) into a result known as **MacCullagh's formula**

$$U(\mathbf{r}) = -\frac{GM}{r} - \frac{G}{2r^3}\left[A + B + C - 3I(\mathbf{r})\right]. \tag{2.26}$$

The potential at a point $\mathbf{r}$ far enough outside a mass distribution, expressed in a coordinate system centered around its center-of-mass, is now written in terms of its three principal (A, B, C), and one generic (around $\hat{\mathbf{r}}$) moment-of-inertia.

The first term in eq. (2.26), the point-mass contribution, dominates at large distances away from the body. Furthermore, eq. (2.26) contains the intuitive result that the potential for *spherically symmetric* bodies is identical to that due to a point with the same total mass located at the center of mass. Indeed, in that case, $A = B = C = I$, which annihilates the second term in eq. (2.26).

2.6.3 An Earth-like oblate ellipsoid

We finally assume an Earth-like situation in which we define C to be the *polar* moment around the rotation axis $\hat{\mathbf{z}}$, and in which, due to rotational symmetry, we assume the equality of the *equatorial moments*, $A = B$. In other words: we consider the Earth to be an *oblate ellipsoid*. Note that rotation is the *ultimate* cause of the flattening, but the Earth's ability to deform its *proximal* cause. In that case, and referring again to Fig. 2.2, the moment $I(\mathbf{r})$ around the axis $\mathbf{r}$ making the colatitudinal angle θ with the north polar axis is given by

$$I(\mathbf{r}) = A + (C - A)\cos^2\theta. \tag{2.27}$$

The potential of such an Earth-shaped body ultimately becomes

$$U(\mathbf{r}) = -\frac{GM}{r} + \frac{G}{r^3}(C - A)\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right) \tag{2.28}$$

$$= -\frac{GM}{r} + \frac{GM}{r^3} J_2\, a^2 \left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right). \tag{2.29}$$

In its second form, the expression (2.29) contains the Earth model's *equatorial radius* a, and introduces the important *dynamical flattening* parameter a J_2

$$J_2 = \frac{C - A}{Ma^2}, \tag{2.30}$$

which depends on the Earth's shape and its internal density distribution.

2.6.4 *The effect of rotation*

The above applied to a realistically ellipsoidal but non-rotating Earth, where we'll let the irony stand that it is the rotation that is responsible for the fact that the Earth is a flattened ellipsoid. Let the Earth be rotating with a vector $\boldsymbol{\omega} = \Omega\,\hat{\mathbf{z}}$. We obtain an expression for the potential due to rotation as in the derivation of eq. (2.4), but instead of $\mathbf{g}$ we will use the *centrifugal acceleration*,

$$U_{\mathrm{rot}}(\mathbf{r}) = -\int_{\infty}^{\mathbf{r}} \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}')\cdot d\mathbf{r}' = \int_{\infty}^{r} \Omega^2 r' \sin^2\theta\, dr'$$
$$= -\frac{1}{2}\,\Omega^2 r^2 \sin^2\theta = -\frac{1}{2}\|\boldsymbol{\omega}\times\mathbf{r}\|^2. \tag{2.31}$$

2.6.5 *A four-parameter Earth model*

Only now do we pick some actual values for the parameters introducted up until this point: the reference Earth. The values for the international WGS-84 reference ellipsoid are $a = 6378137$ m, $GM = 3986004.418 \times 10^8$ m^3 s^{-2} and geodetically determined C and A that fix $J_2 = 108263 \times 10^{-8}$. Only one additional parameter is necessary to completely define a reference ellipsoid, namely the Earth's angular velocity, $\Omega = 7292115 \times 10^{-11}$ s^{-1}.

2.7 The gravitational potential *on* the reference ellipsoid

The potential outside of a *rotating* Earth-like reference ellipsoid becomes

$$U(\mathbf{r}) = -\frac{GM}{r} + \frac{GM}{r^3} J_2\, a^2 \left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right) - \frac{1}{2}\,\Omega^2 r^2 \sin^2\theta. \tag{2.32}$$

We now call c the polar radius of the reference ellipsoid with equatorial radius a. We require that the reference ellipsoid be an *equipotential surface* with value U_{ref}. At its north pole (where $\theta = 0$) and on its equator (where $\theta = \pi/2$),

$$U_{\mathrm{ref}} = \underbrace{-\frac{GM}{c} + \frac{G}{c^3}J_2 M a^2}_{\text{at the north pole}} = \underbrace{-\frac{GM}{a} - \frac{G}{2a^3}J_2 M a^2 - \frac{1}{2}a^2\Omega^2}_{\text{at the equator}}. \tag{2.33}$$

f

From these equations we derive the normalized difference between the polar and equatorial radii of the reference equipotential surface, equal to

$$f = \frac{a-c}{a} = J_2\left(\frac{c}{2a} + \frac{a^2}{c^2}\right) + \frac{1}{2}\frac{a^2 c\,\Omega^2}{GM} \approx \frac{1}{2}\left(3J_2 + \frac{a^3\Omega^2}{GM}\right). \tag{2.34}$$

This quantity, termed the *geometrical flattening* is approximately

$$f \approx \frac{1}{2}\,(3J_2 + m)\,, \tag{2.35}$$

isolating a term from within the brackets in eq. (2.34) that we define as m

$$m = \frac{a\Omega^2}{GM/a^2}. \tag{2.36}$$

With the values for the WGS-84 reference ellipsoid, we thus have $U_{\mathrm{ref}} = 62636860.8497\,\mathrm{m}^2\,\mathrm{s}^{-2}$, $f = 1/298.257223563$ and also $c = 6356752.3142$ m and $m = 0.003461391393112$. At the North Pole the equipotential surface is about 21 km closer to the center of the Earth that at the equator.

2.8 The gravitational acceleration *on* the reference ellipsoid

Let us now calculate the gravitational acceleration at the surface of such a "model Earth" as defined by the reference ellipsoid. We use eq. (2.7) and eq. (2.9) and the potential given by eq. (2.32). Owing to the small flattening of the reference ellipsoid, we will be justified in neglecting all but the radial component of the acceleration. We thus write:

$$g(\mathbf{r}) \approx \frac{GM}{r^2} - 3\frac{GM}{r^4} J_2\, a^2 \left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right) - \Omega^2 r \sin^2\theta. \tag{2.37}$$

Now we see that m in eq. (2.36) is the ratio between the magnitudes of the centrifugal and the gravitational acceleration at the equator, as we shall see,

A useful approximation for the radius of the reference ellipsoid is

$$r_{\mathrm{ref}} \approx a\,(1 - f\cos^2\theta), \tag{2.38}$$

and for its inverse square we can write the approximation

$$r_{\mathrm{ref}}^{-2} \approx a^{-2}\,(1 + 2f\cos^2\theta), \tag{2.39}$$

which we use to reduce the first term of eq. (2.37). Its second and third terms are already small enough for us to use the approximation $r_{\mathrm{ref}} \approx a$, and thus the gravitational acceleration *on* the reference ellipsoid will be

$$\begin{aligned} g_{\mathrm{ref}}(\theta) &\approx \frac{GM}{a^2}\left[1 + 2f\cos^2\theta - 3J_2\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right) - m\sin^2\theta\right] \\ &\approx \frac{GM}{a^2}\left(1 + \frac{3}{2}J_2\sin^2\theta - m + 2m\cos^2\theta\right). \end{aligned} \tag{2.40}$$

If we now define an analog to eq. (2.34), namely the *gravity flattening*, f^*

$$f^* = \frac{g_{\mathrm{np}} - g_{\mathrm{eq}}}{g_{\mathrm{eq}}} = \frac{-3J_2/2 + 2m}{1 + 3J_2/2 - m}, \tag{2.41}$$

we can rewrite eq. (2.40) in the form of, once again, an ellipse:

$$\boxed{g_{\rm ref}(\theta) = g_{\rm eq}\left(1 + f^* \cos^2\theta\right).} \tag{2.42}$$

So this is the magnitude of the gravitational acceleration *of* the reference ellipsoid *on* the reference ellipsoid, at the point uniquely defined by its colatitude, and expressed relative to the gravitational acceleration at the equator. To the same order of the approximation as the above equations,

$$f^* = 5\,m/2 - f. \tag{2.43}$$

Eqs (2.42)–(2.43) express the amazing results obtained by **Clairaut** in the mid-eighteenth century: the *geometrical* shape of the Earth can be obtained from a combination of purely *dynamical* quantities, i.e., by measuring gravity.

2.9 Practical formulas and international reference values

Maybe here a short first section on gravity "anomalies".

After subtracting reference values from the measurements, the resulting *anomalies* are those variations of gravity that cannot be accounted for by simply approximating the Earth as a rotating oblate ellipsoid. The insights we get from such anomalies about the interior of the Earth are of fundamental importance for our understanding of our planet.

Note the Fischer geoid.

To conclude, the defining parameters for WGS-84 are given in Table 2.1, and from this, all of the above can be derived.

Table 2.1. *Defining parameters of the WGS-84.*

a	6378137	m
f	1/298.257223563	
GM	3986004.418×10^8	$\mathrm{m^3\,s^{-2}}$
Ω	7292115×10^{-11}	$\mathrm{s^{-1}}$

Picture of the geoid.

2.10 Flux

Is what, exactly? Treat flux graphically, and derive from this the expression for divergence in Cartesian coordinates. Which you only name in the next section. The gate analogy.

2.10.1 The divergence (of a vector field)

The *divergence* $\boldsymbol{\nabla}\cdot$ can be thought of as the *dot product* between the gradient and its argument, which must be *at least* a vector. In that case, the result is scalar. Physically, it is the outward *flux* of the vector field per unit volume. In Cartesian coordinates, the divergence of a vector function $\mathbf{u}(\mathbf{r})$ equals

$$\boldsymbol{\nabla}\cdot\mathbf{u} = \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z}, \tag{2.44}$$

which is easily derived from eqs (2.8) and (1.29). In spherical coordinates, after a rather more lengthy calculation via eqs (2.9), (1.30) and (1.28), we get

$$\boldsymbol{\nabla}\cdot\mathbf{u} = \frac{\partial u_r}{\partial r} + \frac{2}{r}u_r + \frac{1}{r}\left(\frac{\partial u_\theta}{\partial \theta} + u_\theta \cot\theta + \frac{1}{\sin\theta}\frac{\partial u_\phi}{\partial \phi}\right). \tag{2.45}$$

2.10.2 The Gauss theorem

One of the most fundamental results underlying the potential theory of gravity is the *divergence theorem* due to Gauss, which states, for vector fields, that the *volume* integral of the divergence of the quantity $\mathbf{u}$ is equal to the flux, or *surface* integral of its surface-parallel component, with unit normal vector $\hat{\mathbf{n}}$:

$$\boxed{\int_V \boldsymbol{\nabla}\cdot\mathbf{u}\, dV = \int_{\partial V} \hat{\mathbf{n}}\cdot\mathbf{u}\, \mathrm{d}\Sigma.} \tag{2.46}$$

For good measure, we state the version for surfaces in two dimensions:

$$\int_\Sigma \boldsymbol{\nabla}\cdot\mathbf{u}\, \mathrm{d}\Sigma = \int_{\partial\Sigma} \hat{\mathbf{n}}\cdot\mathbf{u}\, dl, \tag{2.47}$$

with normal unit vector $\hat{\mathbf{n}}$, and for lines in one dimension, we get the well-known result

$$\int_a^b \frac{\partial u}{\partial l}\, dl = u(b) - u(a). \tag{2.48}$$

2.11 Poisson's and Laplace's equations

Need Kellogg, brief treatment by Blakely. Bottom line: once again give the perspective that we need to write potentials and their derivatives both inside and outside of a mass distribution, and that it depends on the niceties of densities to be able to pull the derivatives out of the integral.

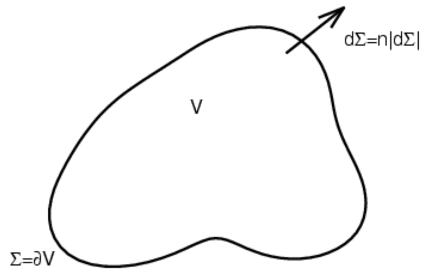


Fig. 2.3. Diagram illustrating Gauss' theorem, eq. (2.46).

So now we apply eq. (2.46) to the flux of gravity through a small sphere containing a little bit of mass. Heuristic derivation.

$$\int_V \boldsymbol{\nabla} \cdot \mathbf{g}\ dV = \int_{\partial V} \hat{\mathbf{n}} \cdot \mathbf{g}\ d\Sigma. \tag{2.49}$$

We use eq. (2.3) and eq. (2.7), pretending we're in some tiny volume of radius r with surface area $4\pi r^2$, with a homogeneous bit of enclosed mass, so $\hat{\mathbf{r}} = \hat{\mathbf{n}}$,

$$\int_V \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} U\ dV = \frac{GM}{r^2} \int_{\partial V} d\Sigma. \tag{2.50}$$

And then

$$\int_V \nabla^2 U(\mathbf{r})\ dV = 4\pi G \int_V \rho(\mathbf{r})\ dV \tag{2.51}$$

and then the biggest thing of all

$$\boxed{\nabla^2 U(\mathbf{r}) = 4\pi G \rho(\mathbf{r})}. \tag{2.52}$$

This is called *Poisson's* or *Laplace's equation*—depending on whether the equation has a right hand sign or not. The right-hand side of eq. (2.51) follows from the right-hand side of eq. (2.50) for the surface, and by writing out the mass as a density integral.

2.12 The Laplacian (of a scalar field)

We've identified $\nabla^2 = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}$, the *Laplacian*, a measure of the curvature of a scalar field. In Cartesian coordinates:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{2.53}$$

And in spherical coordinates:

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\left(\frac{\partial^2}{\partial_\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right). \tag{2.54}$$

Use this expression to verify eq. (2.16).

2.13 Solutions to Laplace's equation—I

Some generalities of solving differential equations: boundary conditions, methods.

Let's say we're outside (or in the limit, on the surface) of some mass distribution, i.e. in a density-free region of space. There, Laplace's equation holds:

$$\nabla^2 U(\mathbf{r}) = 0. \tag{2.55}$$

We want to find the function U that represents the potential due to the mass distribution—this function is non-zero. It is, however, *harmonic*, that's the word. From looking at any of the foregoing expressions for $U(\mathbf{r})$ that we derived via more pedestrian means, we can see that sums and products of constants, powers of r and trigonometric functions etc somehow will be good candidate functions to describe the *general* shape of U in spherical coordinates. So we're not worrying wat causes the potential, we're just going to find the most general form for it.

We take the bold if uninspired move to test a solution by *separation of variables*, i.e. we propose that $R\,\Theta\,\Phi$

$$U(r,\theta,\phi) = R(r)\,\Theta(\theta)\,\Phi(\phi), \tag{2.56}$$

and thereby reduce this partial differential equation to a set of three ordinary ones.

2.13.1 Radial behavior

Focus on derivation, then reduction, then solution, without the details.

We we find two possible solutions: $R(r)$

$$R(r) = \begin{cases} r^l \\ r^{-l-1} \end{cases}. \tag{2.57}$$

Maybe here say we suspect they'll be useful for internal and external fields separatly, depending on their decay.

2.13.2 Azimuthal behavior

We plug eq. (2.54) into eq. (2.56), we chug, and find for the *radial dependence* an equation

$$\frac{1}{R}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}R\right). \tag{2.58}$$

This gives us the $l(l+1)$ behavior.

$\Phi(\phi)$ We plug, we chug, we find for the *azimuthal* or *longitudinal dependence* the equation

$$-\frac{1}{\Phi}\frac{d^2}{d\phi^2}\Phi = m^2, \tag{2.59}$$

which is solved for by a complex exponential modulo any constant,

$$\Phi(\phi) = C_{lm}\exp(im\phi), \tag{2.60}$$

$\Theta(\theta)$ for certain C_{lm} and S_{lm}, to be determined. We are prudently adding both indices to the coefficients multiplying the sine and cosine terms.

2.13.3 Colatitudinal behavior

We plug and chug and get for the *colatitudinal dependence*

$$\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{d}{d\theta}\Theta(\theta)\right) + [l(l+1)\sin^2\theta - m^2]\Theta(\theta) = 0. \tag{2.61}$$

P_{lm} Now we're in luck, as eq. (2.61) is an equation that is known and loved in the mathematics community and goes by the name of *associated Legendre equation*. The solutions are known as *associated Legendre functions*... and go by $P_{lm}(\theta)$ or, often $P_{lm}(\cos\theta)$, and often, in the notation $\mu = \cos\theta$. Connect up with them being eigenfunctions of the surface Laplacian.

This is the only one that triggers its own subsection.

2.13.4 Solutions to Legendre's equation

First some business about the Legendre functions. Here they are. The *Rodrigues' formula* for $\mu = \cos\theta$:

$$P_{lm}(\mu) = \frac{1}{2^l l!}(1-\mu^2)^{m/2}\left(\frac{d}{d\mu}\right)^{l+m}(\mu^2-1)^l. \tag{2.62}$$

As implicit and analytical this formula is, its numerical evaluation gets cumbersome very fast. Calculating factorials is not for the faint of heart. Luckily, the Legendre functions follow a host of three-term recursion relations, starting from some very simple building blocks, calculating the next order from the previous two, or the next degree from the previous two. Smart normalization will prevent numerical inaccuracies from building up. Table 2.2 lists the low-degree functions explicitly, and Fig. 2.4 offers a graphical rendition.

Table 2.2. *Associated Legendre functions of degree and order zero to two.*

l	m	P_{lm}
0	0	1
1	0	$\cos\theta$
1	1	$\sin\theta$
2	0	$\frac{3}{2}\cos^2\theta - \frac{1}{2}$
2	1	$3\sin\theta\cos\theta$
2	2	$3\sin^2\theta$

2.13.5 Back to the narrative

And the bottom line is that by combining eqs (2.57), (2.60) and (2.62) we have found the complete solution in the form of eq (2.56), and thus, we have found the general solutions of the Laplace equation, which we shall label U_{lm},

$$U_{lm}(r,\theta,\phi) = \left\{ \begin{array}{c} r^l \\ r^{-l-1} \end{array} \right\} P_{lm}(\cos\theta)\exp(im\phi), \tag{2.63}$$

call them *solid spherical harmonics*, if you will. People do.

Now we need to define the full solution for a harmonic potential in spherical coordinates, the main point being that we now write the potential for an *external field* as

$$U(\mathbf{r}) = \sum_{l=0}^{\infty}\sum_{m=0}^{l} r^l\,[C'_{lm}\cos m\varphi + S'_{lm}\sin m\varphi]P_{lm}(\cos\theta). \tag{2.64}$$

and the potential *outside the mass distribution* (i.e. for the *internal field*) as

$$U(\mathbf{r}) = \sum_{l=0}^{\infty}\sum_{m=0}^{l}\left(\frac{1}{r}\right)^{l+1}[C''_{lm}\cos m\varphi + S''_{lm}\sin m\varphi]P_{lm}(\cos\theta). \tag{2.65}$$

Earth models, clearly of the latter kind, are often given as normalized *band-limited* expansions

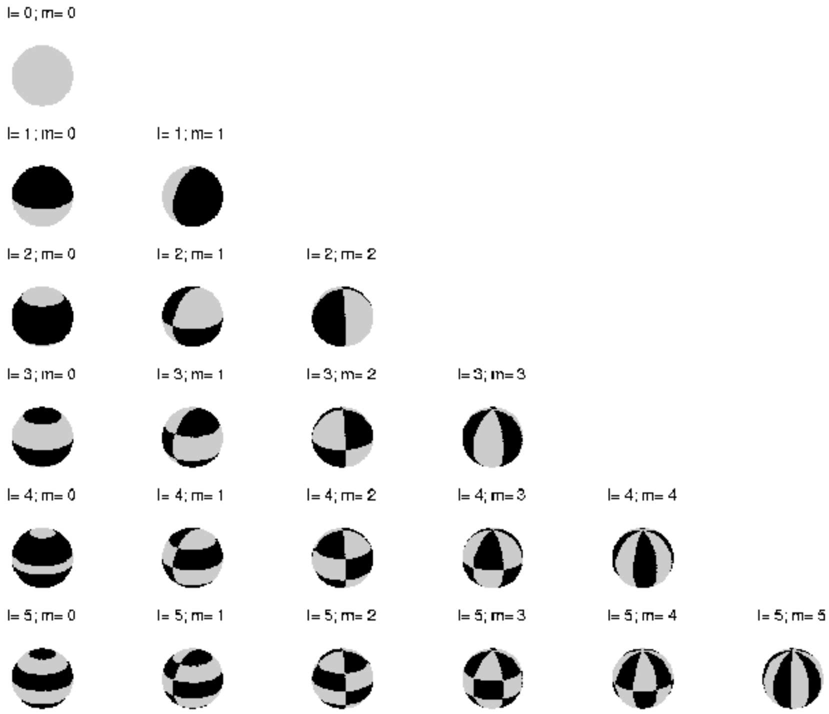


Fig. 2.4. A rendition of spherical harmonics of various degrees l and orders m, reduced to their essence: whether they are positive or negative, hence in two tones, and with the number of *nodal* crossings clearly visible. When $m = 0$, the spherical harmonic is said to be *zonal*, and when $l = m$ it is *sectoral*. The general case is called *tesseral*.

$$U(\mathbf{r}) = -\frac{GM}{a}\sum_{l=0}^{L}\sum_{m=0}^{l}\left(\frac{a}{r}\right)^{l+1}[C_{lm}\cos m\varphi + S_{lm}\sin m\varphi]P_{lm}(\cos\theta), \tag{2.66}$$

as one form of the sought-after solution to eq. (2.55). Take a good look at the canonical relation eq. (2.66). All the information about the planetary gravity contained is contained in the spherical harmonic expansion coefficients of the gravitational potential, C_{lm} and S_{lm}. Armed with these, and a knowledge of the normalizing constants, GM and a, the equatorial radius, it becomes easy to calculate the potential at any other value outside the mass distribution. Operations such as *upward* and *downward continuation* from one height up or down to another are a simple scaling by a degree-dependent term $(a/r)^{l+1}$.

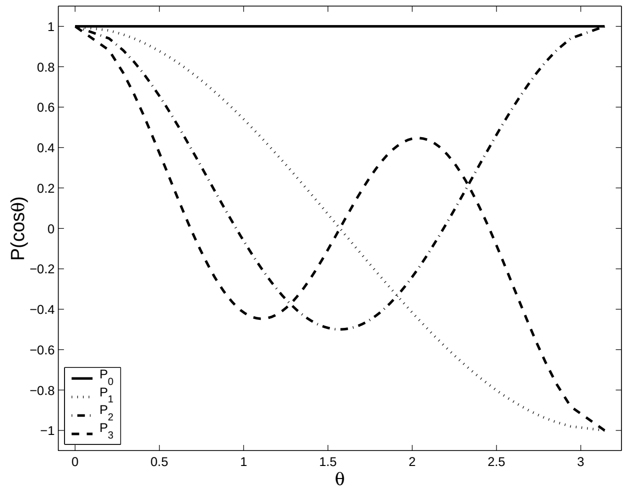


Fig. 2.5. Legendre functions of order $m = 0$, for the construction of *zonal* spherical harmonics, for various degrees $l = 0, 1, 2, 3$. The number of colatitudinal *nodal crossings* equals the degree l. As to the *associated Legendre functions*, the order m modifies the number of colatitudinal crossings. The spherical harmonics, which contain the longitudinal term $\exp(m\phi)$, will distribute the l nodal lines on the Earth's *surface* across additional *longitudinal nodal lines*.

2.14 Spherical harmonics

The *surface spherical harmonics* occupy a special place:

$$Y_{lm}(\theta,\phi) = e^{im\phi}X_{lm}(\theta), \tag{2.67}$$

$$X_{lm}(\theta) = (-1)^m\left(\frac{2l+1}{4\pi}\right)^{1/2}\left[\frac{(l-m)!}{(l+m)!}\right]^{1/2}P_{lm}(\cos\theta), \tag{2.68}$$

What their deal is with the surface Laplacian. What their orthogonality is.

Make sure to also list what Y_{00} is as we will be needing this later.

The least-squares approximation properties etc. that show us it's more than just an oddity. In fact, restricting our attention to those functions restricted to the unit sphere, *any* function, whether it is harmonic or not, can be represented as a sum of spherical harmonics as long as it is *square-integrable*, which is nothing more than the rather mild condition that

$$\int_\Omega |f(\mathbf{r})|^2\, d\Omega < \infty. \tag{2.69}$$

This is easily shown by the following argument which I have written down somewhere. Make the connection to Fourier analysis.

Table 2.3. *EGM96 coefficients in values of* 10^{-6}. *The scaling values are* $a = 6378136.3$ m *and* $GM = 3986004.415 \times 10^8$ m^3 s^{-1}.

l	m	C_{lm}	S_{lm}
0	0	1	0
1	0	0	0
1	1	0	0
2	0	-484.1654	0
2	1	-0.0002	0.0012
2	2	2.4391	-1.4002

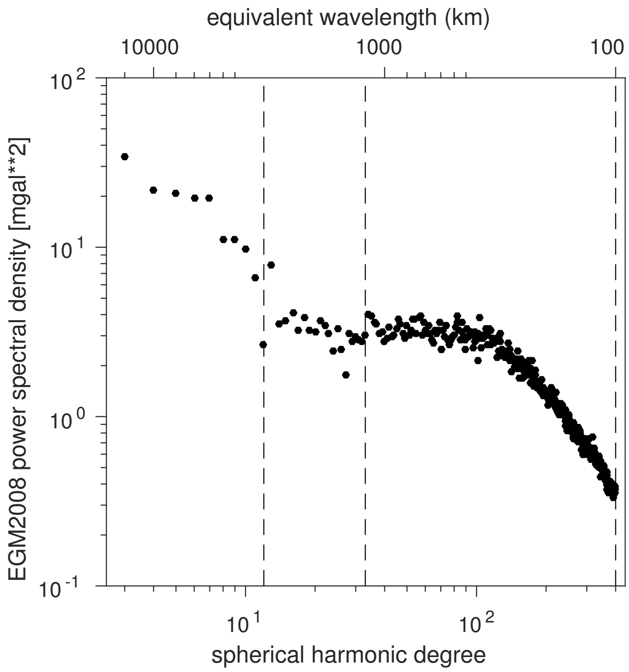


Fig. 2.6. Power in the coefficients for the gravitational field of Earth model EGM2008. Starting from $l = 0$, note the missing $l = 1$ term, and the steep drop-off with increasing spherical harmonic degree. Local slopes of *power spectral densities* like these are quite diagnostic of the planet under consideration.

2.15 Solutions to Laplace's equation—II

Rewriting the potential equation (2.11) in integral form, we have for the potential due the mass in the Earth at a point *outside* of the Earth,

$$U(\mathbf{r}) = -G \int_{\oplus} \frac{\rho(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|} \, dV'. \tag{2.70}$$

Going back all the way to Gauss is the result that, when $r' < r$,

$$\frac{1}{\|\mathbf{r}-\mathbf{r}'\|} = \frac{1}{r}\sum_{l=0}^{\infty}\left(\frac{r'}{r}\right)^{l} P_l(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'), \tag{2.71}$$

$$= \frac{1}{r}\sum_{l=0}^{\infty}\left(\frac{r'}{r}\right)^{l}\sum_{m=-l}^{l}\left(\frac{4\pi}{2l+1}\right) Y_{lm}(\hat{\mathbf{r}})Y_{lm}(\hat{\mathbf{r}}'), \tag{2.72}$$

where we recall eq. (1.27), $\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}' = \cos\Delta$. In some sense we've already seen this, as eq. (2.18). From this we conclude that

$$U(\mathbf{r}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left(\frac{1}{r}\right)^{l+1} U_{lm}Y_{lm}(\hat{\mathbf{r}}), \tag{2.73}$$

an expansion in solid harmonics, where the **Stokes' coefficients** are given by

$$\boxed{U_{lm} = -\frac{4\pi G}{2l+1}\int_{\oplus}\rho(\mathbf{r}')\, r'^{l}\, Y_{lm}(\hat{\mathbf{r}}')\, dV'}. \tag{2.74}$$

In a spherically symmetric Earth, we have the special case:

$$U_{lm} = -\frac{4\pi G}{2l+1}\int_{0}^{a}\rho(r')\, r'^{l+2}\, dr' \int_{0}^{\pi}\!\!\int_{0}^{2\pi} Y_{lm}(\theta',\phi')\sin\theta'\, d\theta'\, d\phi', \tag{2.75}$$

which we had seen in various guises before.

2.16 Potential, gravitational and geoidal anomalies

The real utility of using spherical harmonic expansions for the potential lies in the ease with which derived quantities can be computed by simple manipulations of the expansion coefficients. We begin by introducing and defining some fundamental quantities, which we will then restate in terms of some rather easily obtained functions of the spherical harmonic coefficients of the root quantity that is the gravitational potential. Spherical harmonic expansion coefficients of the gravitational potential of Earth and various planets and Moons are a prime target for geodetic satellite missions, and their tables are widely distributed by (international) space agencies.

The actual geopotential $U(\mathbf{s})$ and the potential $U_{\text{ref}}(\mathbf{s})$ of the reference Earth differ by a quantity called *a potential anomaly*, ΔU

$$\Delta U(\mathbf{s}) = U(\mathbf{s}) - U_{\text{ref}}(\mathbf{s}). \tag{2.76}$$

In eq. (2.76), $\mathbf{s}$ is a *generic* position coordinate. Henceforth, we introduce two *specific* coordinates: one, $\mathbf{r}$ that *defines* an equipotential surface of the *actual*

Earth, and another, $\mathbf{r}'$ that *defines* an equipotential surface of the reference, i.e. ellipsoidal, Earth. Both equipotential surfaces maintain the same constant value of the potential— a choice of convenience which we call $U_\oplus$.

$U_\oplus$ $\mathbf{r}'$ The *reference geoid* is the equipotential surface at the reference value,

$$\{\mathbf{r}' \,:\, U_{\mathrm{ref}}(\mathbf{r}') = U_\oplus\}, \tag{2.77}$$

where $\|\mathbf{r}'\| = r_{\mathrm{ref}}$, which we have previously approximated by eq. (2.38).

$\mathbf{r}$ Given that the reference Earth is at best an approximation of the real thing, the *actual geoid*, at the same potential, is the locus of points $\mathbf{r}$ defined by

$$\{\mathbf{r} \,:\, U(\mathbf{r}) = U_\oplus\}. \tag{2.78}$$

In what follows we will continue to speak of "the" actual and reference geoids as equipotential surfaces with the same potential $U_\oplus$, continuing to denote a point on the actual geoid by $\mathbf{r}$ and on the reference geoid by $\mathbf{r}'$. Hence, a specific case of eq. (2.76) becomes *the potential anomaly*,

$$\Delta U(\mathbf{r}) = U_\oplus - U_{\mathrm{ref}}(\mathbf{r}). \tag{2.79}$$

N The, to first order radial, difference between these two equipotential surfaces is the *geoid height* or *geoid undulation*,

$$N(\mathbf{r}) = \|\mathbf{r}\| - r_{\mathrm{ref}}(\theta). \tag{2.80}$$

Refer to Fig. 2.7. Over regions of mass *excess* the equipotential surface will be warped up compared to the reference, whereas it will be drawn down over regions of mass deficit (e.g., low-density areas).

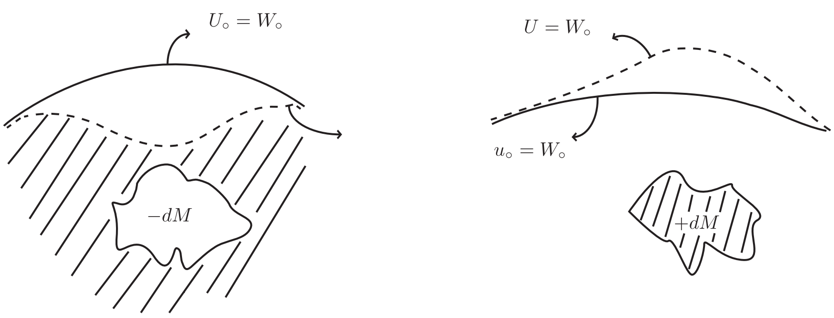


Fig. 2.7. Mass deficiencies cause geoid *lows*, mass excesses cause geoid *highs*.

How does the potential anomaly relate to the geoid height? We compute the value of the reference potential at $\mathbf{r}$ (i.e, on the *actual* geoid) from its value at $\mathbf{r}'$ (i.e., on the *reference* geoid), where $U_{\rm ref}(\mathbf{r}') = U_\oplus$. We make a first-order Taylor expansion in the radial direction using eq. (2.7) to write

$$U_{\rm ref}(\mathbf{r}) = U_\oplus + \left.\frac{\partial U_{\rm ref}(\mathbf{r})}{\partial r}\right|_{\mathbf{r}'} N(\mathbf{r}) = U_\oplus + g_{\rm ref}(\mathbf{r}')N(\mathbf{r}). \tag{2.81}$$

Now, the potential does *not* vary on the equipotential surface, but the gravity *does*. And thus, combining eqs (2.79) and (2.81) we use a reference of the kind of eq. (2.42) to derive that

$$\boxed{\Delta U(\mathbf{r}) = -g_{\rm ref}(\theta)N(\mathbf{r}).} \tag{2.82}$$

This result is known as **Bruns' formula**. It allows us to calculate the radial distance to the reference geoid from a measurement of the potential and a comparison with the theoretical reference gravity. Take a good look at Fig. 2.8.

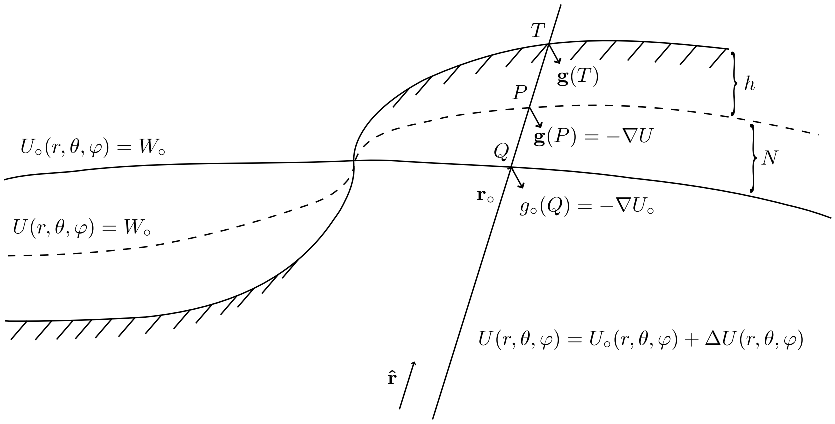


Fig. 2.8. Diagram illustrating the relation between the various *anomalies* (potential, geoid, free-air gravity) and the gravity *disturbance*.

The *vector gravity disturbance* is the difference between the measured gravity and the reference gravity at the same point in space: $\delta\mathbf{g}$

$$\delta\mathbf{g}(\mathbf{r}) = \mathbf{g}(\mathbf{r}) - \mathbf{g}_{\rm ref}(\mathbf{r}), \tag{2.83}$$

which, using eqs (2.7) and (2.76) we can easily reduce to

$$\delta\mathbf{g}(\mathbf{r}) = -\boldsymbol{\nabla}[\Delta U(\mathbf{r})]. \tag{2.84}$$

To first order, we obtain for the *scalar gravity disturbance* δg

$$\delta g(\mathbf{r}) = g(\mathbf{r}) - g_{\rm ref}(\mathbf{r}) = \frac{\partial \Delta U(\mathbf{r})}{\partial r}. \tag{2.85}$$

The above quantities relate actual *measurements* to their reference or *theoretical* values *at the same point in space*. For historical reasons, however, it has been more popular to relate the value measured on (or, as we see later, referred to) the actual geoid to the theoretical values on the reference surface. In other words, there, the comparison between gravity values is to points *at the same potential*, which, however, are at different locations spatially.

Δg The *free-air gravity anomaly* is defined as the gravity $g(\mathbf{r})$ at some point $\mathbf{r}$ on the Earth's actual geoid, minus the reference gravity, $g_{\rm ref}(\mathbf{r}')$, at the projection $\mathbf{r}'$ of this point onto the reference geoid. Neglecting the small differences in direction, the scalar free-air anomaly is given by:

$$\Delta g(\mathbf{r}) = g(\mathbf{r}) - g_{\rm ref}(\mathbf{r}'). \tag{2.86}$$

To reduce this further, we expand $g_{\rm ref}(\mathbf{r}')$ in a Taylor series to first order:

$$g_{\rm ref}(\mathbf{r}) = g_{\rm ref}(\mathbf{r}') + \left.\frac{\partial g_{\rm ref}(\mathbf{r})}{\partial r}\right|_{\mathbf{r}'} N(\mathbf{r}). \tag{2.87}$$

Using this and eq. (2.85) we rewrite eq. (2.86) as

$$\Delta g(\mathbf{r}) = \frac{\partial \Delta U(\mathbf{r})}{\partial r} + \left.\frac{\partial g_{\rm ref}(\mathbf{r})}{\partial r}\right|_{\mathbf{r}'} N(\mathbf{r}), \tag{2.88}$$

$\partial_r g_{\rm ref}(\mathbf{r})$ rewrite the second term of the above by approximating, starting from eq. (2.37),

$$\left.\frac{\partial g_{\rm ref}(\mathbf{r})}{\partial r}\right|_{\mathbf{r}'} \approx -\frac{2}{r_{\rm ref}}\, g_{\rm ref}(\theta). \tag{2.89}$$

Eq. (2.89) is a *free-air correction*: the formula by which we can relate, e.g., the gravitational acceleration measured at some point on the actual Earth (like, in Princeton) to the value that the gravitational acceleration has at some other point a certain distance away. This rewrites eq. (2.88), with Bruns' equation (2.82), as

$$\Delta g(\mathbf{r}) = \frac{\partial \Delta U(\mathbf{r})}{\partial r} + \frac{2}{r_{\rm ref}} \Delta U(\mathbf{r})\,. \tag{2.90}$$

This is the *fundamental equation of geodesy*. It relates the measurement quantity $\Delta g(\mathbf{r})$ to the unknown disturbing potential, $\Delta U(\mathbf{r})$. Since our measurements are inevitably confined to the surface of the Earth, however, eq. (2.88), embodies a boundary-value problem condition to Laplace's equation (2.55).

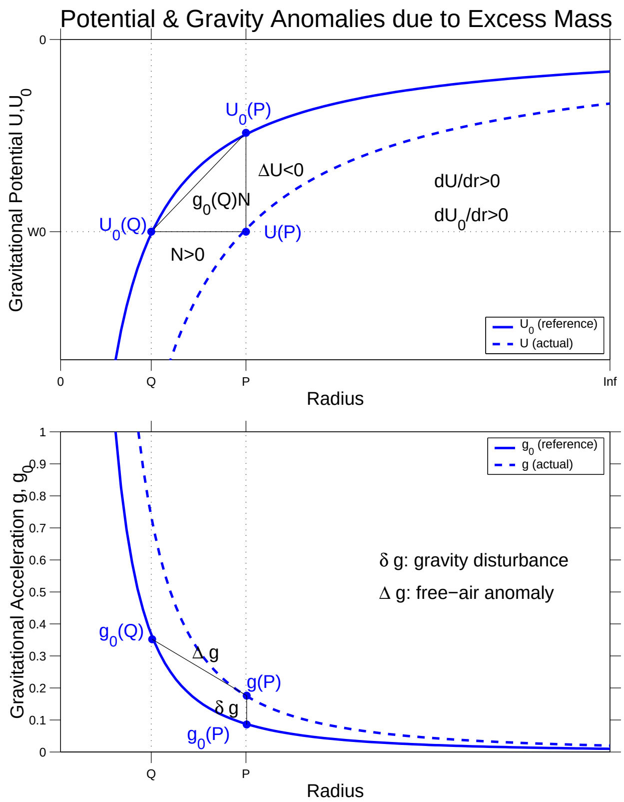


Fig. 2.9. Diagram illustrating the relation between the various *anomalies* (potential, geoid, gravity) and the gravity *disturbance*.

2.17 Spherical harmonic formulation of the above

How can we relate, e.g., gravity anomalies to geoid heights, from the spherical harmonic expansions coefficients of the potential? Simple transformations will do the trick.

For convenience, we can represent the *set* of coefficients of eq. (2.66) by using the subscripts A and B to label the coefficients of the $\cos m\varphi$ and $\sin m\varphi$ terms in the expansion as follows:

$$\left\{ \begin{array}{c} U_A \\ U_B \end{array} \right\} = -\frac{GM}{a} \left\{ \begin{array}{c} C_{lm} \\ S_{lm} \end{array} \right\}. \tag{2.91}$$

The reference potential is by definition *zonal*, $U_{\mathrm{ref}}(r,\theta,\phi) = U_{\mathrm{ref}}(r,\theta)$. We can represent its coefficients as follows:

$$\left\{ \begin{array}{c} U_{\mathrm{ref},A} \\ U_{\mathrm{ref},B} \end{array} \right\} = -\frac{GM}{a} \left\{ \begin{array}{c} C'_{lm} \\ 0 \end{array} \right\} \tag{2.92}$$

Note that we did not drop the m , even though $m = 0$ for the zonal harmonics used for the reference spheroid. We just require the coefficient $A_l^{'m}$ to be zero for $m \neq 0$. By doing this we can keep the equations simple.

The coefficients of the anomalous potential $\Delta U(r, \theta, \varphi)$ are obtained simply by subtracting the coefficients of the reference:

$$\left\{ \begin{array}{c} \Delta U_A \\ \Delta U_B \end{array} \right\} = -\frac{GM}{a} \left\{ \begin{array}{c} C_{lm} - C'_{lm} \\ S_{lm} \end{array} \right\} \tag{2.93}$$

Let us find the equations for the radial derivative of the potential anomaly, which represent the gravitational attraction due to an anomalous mass. Note that these are the equations for the gravity disturbance:

$$\left\{ \begin{array}{c} \delta g_A \\ \\ \delta g_B \end{array} \right\} = \left\{ \begin{array}{c} \frac{d\Delta U_A}{dr} \\ \\ \frac{d\Delta U_B}{dr} \end{array} \right\} = -\frac{GM}{a} \left(\frac{-(l+1)}{a} \right) \left\{ \begin{array}{c} C_{lm} - C'_{lm} \\ \\ S_{lm} \end{array} \right\} \tag{2.94}$$

We can find the expression for the coefficients of $\Delta g(r, \theta, \varphi)$ using eqs. (**??**) and (2.94):

$$\left\{ \begin{array}{c} \Delta g_A \\ \Delta g_B \end{array} \right\} = \frac{GM}{a} \left(\frac{l-1}{a} \right) \left\{ \begin{array}{c} C_{lm} - C'_{lm} \\ S_{lm} \end{array} \right\} \tag{2.95}$$

$$= g_{\mathrm{ref}}(a)(l-1) \left\{ \begin{array}{c} C_{lm} - C'_{lm} \\ S_{lm} \end{array} \right\} \tag{2.96}$$

Note that these expressions are only valid at the reference radius a since the degree-dependent attenuation factors are not represented here. Actually, should put them in here. The proportionality with $(l-1)g_{\mathrm{ref}}(a)$ means that the higher degree terms are magnified in the gravity field relative to those in the potential field. This leads to the important result that gravity maps typically contain much more detail than geoid maps because the spatial attenuation of the higher degree components is suppressed.

Using eq. (2.82) we can express the coefficients of the expansion of $N(r, \theta, \varphi)$ in terms of either the coefficients of the expanded anomalous potential:

$$g_{\mathrm{ref}}(a) \left\{ \begin{array}{c} N_A \\ N_B \end{array} \right\} = \frac{GM}{a} \left\{ \begin{array}{c} C_{lm} - C'_{lm} \\ S_{lm} \end{array} \right\}, \tag{2.97}$$

which then leads to the following form

$$\left\{ \begin{array}{c} N_A \\ N_B \end{array} \right\} = a \left\{ \begin{array}{c} C_{lm} - C'_{lm} \\ S_{lm} \end{array} \right\}, \tag{2.98}$$

or, in terms of the coefficients of the gravity anomalies (eqns 2.96 and 2.98)

$$\left\{ \begin{array}{c} N_A \\ N_B \end{array} \right\} = \frac{a}{(l-1)g_{\mathrm{ref}}(a)} \left\{ \begin{array}{c} \Delta g_A \\ \Delta g_B \end{array} \right\}. \tag{2.99}$$

The geoid heights can thus be synthesized from the expansions of either the gravity anomalies (2.99) or the anomalous potential (2.98). Geoid anomalies have been constructed from both surface measurements of gravity (2.99) and from satellite observations (2.98). Eq. (2.99) indicates that, relative to those of the gravity anomalies, the coefficients of the geoid height at radius $r_{\mathrm{ref}} = a \; N(r, \theta, \varphi)$ are suppressed by a factor of $1/(l-1)$. As a result, shorter-wavelength features are much more prominent on gravity maps. In other words, geoid (and geoid height) maps essentially depict the low harmonics of the gravitational field.

A last note concerns the spectral representation of the free-air gravity anomaly with respect to the gravity disturbance. From the foregoing,

$$\left\{ \begin{array}{c} \Delta g_A \\ \Delta g_B \end{array} \right\} = \frac{l-1}{l+1} \left\{ \begin{array}{c} \delta g_A \\ \delta g_B \end{array} \right\} \tag{2.100}$$

As we have seen, the gravity disturbance reveals much more of the heterogeneous structure of the Earth than the free-air anomaly. On the other hand, the free-air anomalies are only filtered coefficients of the gravity disturbances. However, this filter is very non-linear. Low harmonics are severely attenuated. If most of the gravity disturbances are in the long-wavelength part of the spectrum, the free-air anomaly will make those less visible.

2.18 Gravity due to (buried) bodies

Periodic loads. Infinite slabs. Buried loads. Derivation is best not in TS but rather in Snieder. At least one worked example. And then later return to it with flexure?

2.19 Gravity measurement and interpretation in practice

To reduce a gravity measurement $g(\mathbf{r}'')$, made at some altitude $h(\mathbf{r}'')$ above the actual geoid $\mathbf{r}$, to the actual geoid (so it can later be referred to the values on the reference geoid $\mathbf{r}'$ in order to compute the free-air anomaly), we need an adjustment called the *free-air correction*. We follow the derivation of eqs (2.87) and (2.89) and write, to first order,

$$g(\mathbf{r}) = g(\mathbf{r}'') + \frac{2}{r} g(\mathbf{r})\, h(\mathbf{r}''). \tag{2.101}$$

We *add* a bit of gravitational attraction to the measurement made above the geoid, to compensate for the fact that we are further removed from the mass. The free-air correction explains that part of the observed anomaly with respect to the reference gravity that is due to the altitude of the measurement (without any intervening mass, as for measurements made from an airplane, hence the name, free-air), and what is left is the free-air anomaly. Shipboard measurements minus the reference gravity are at once free-air anomalies.

Many assumptions have led us to finally writing down eq. (2.101), but practical implementations proceed in an even more cavalier fashion, by approximating the free-air gravity correction as

$$\frac{2}{r} g(\mathbf{r}) \approx 0.3086 \times 10^{-5} \mathrm{s}^{-2}. \tag{2.102}$$

In other words, 0.3086 mgal is to be added to the measurement per meter of observation altitude in order to reduce the measurement to the actual geoid — at least, approximately.

By the way, tell them what a mgal is, and an Eötvós is 10^{-9} per second squared, a unit of gravity gradient.

A *terrain correction* is needed to account for what we have been neglecting thus, namely the mass in-between the measurement point and the actual geoid (imagine a measurement taken on a mountain top).

Full problem done in Fourier space, or in spherical harmonics, gets complicated rather quickly. Quick and dirty, leads to the *Bouguer correction*. Then, there's the fact that you need to do a similar correction for an interface at depth. Perhaps should derive the Bouguer correction twice—once the easy way, and the second time as a special case for the compensation due to an interface. Via direct integration, or via Fourier-domain modelling.

Talk briefly about **isostasy**, it returns twice.

2.20 Gravity and topography

Flexure etc.

2.21 Time-variable gravity

GRACE, and what it does for us.

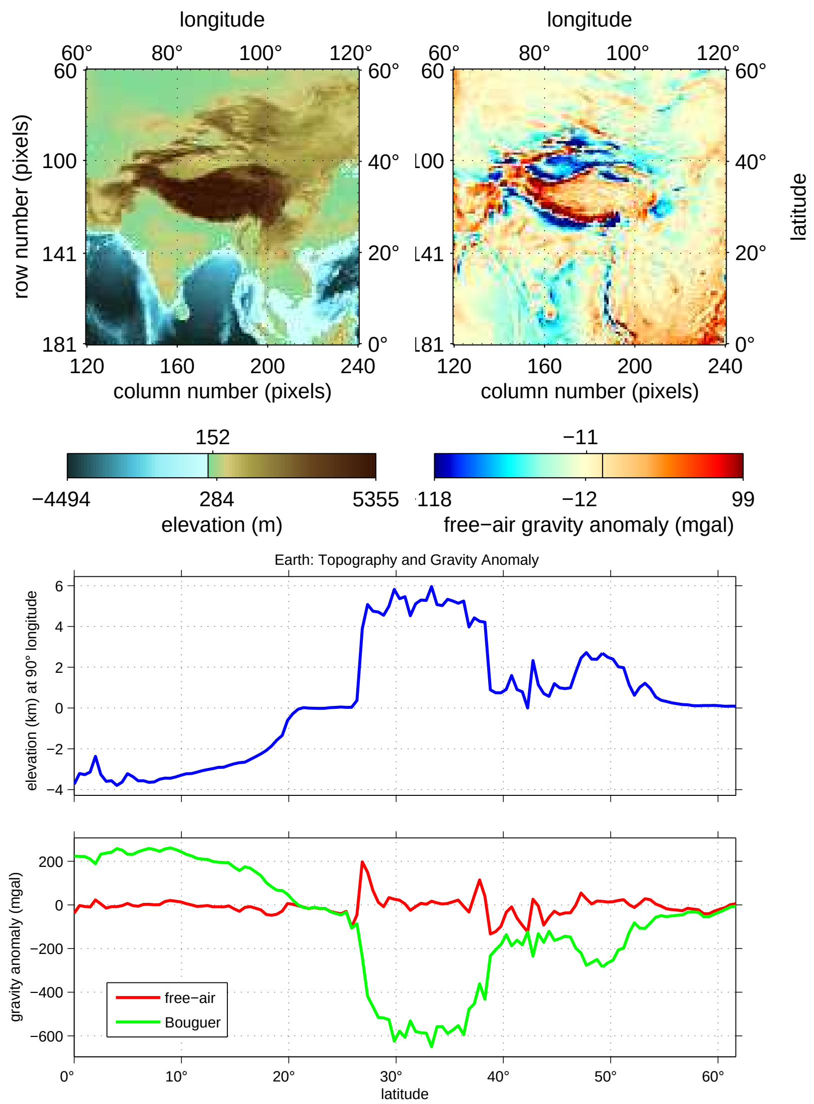


Fig. 2.10. Bouguer and free-air anomalies.

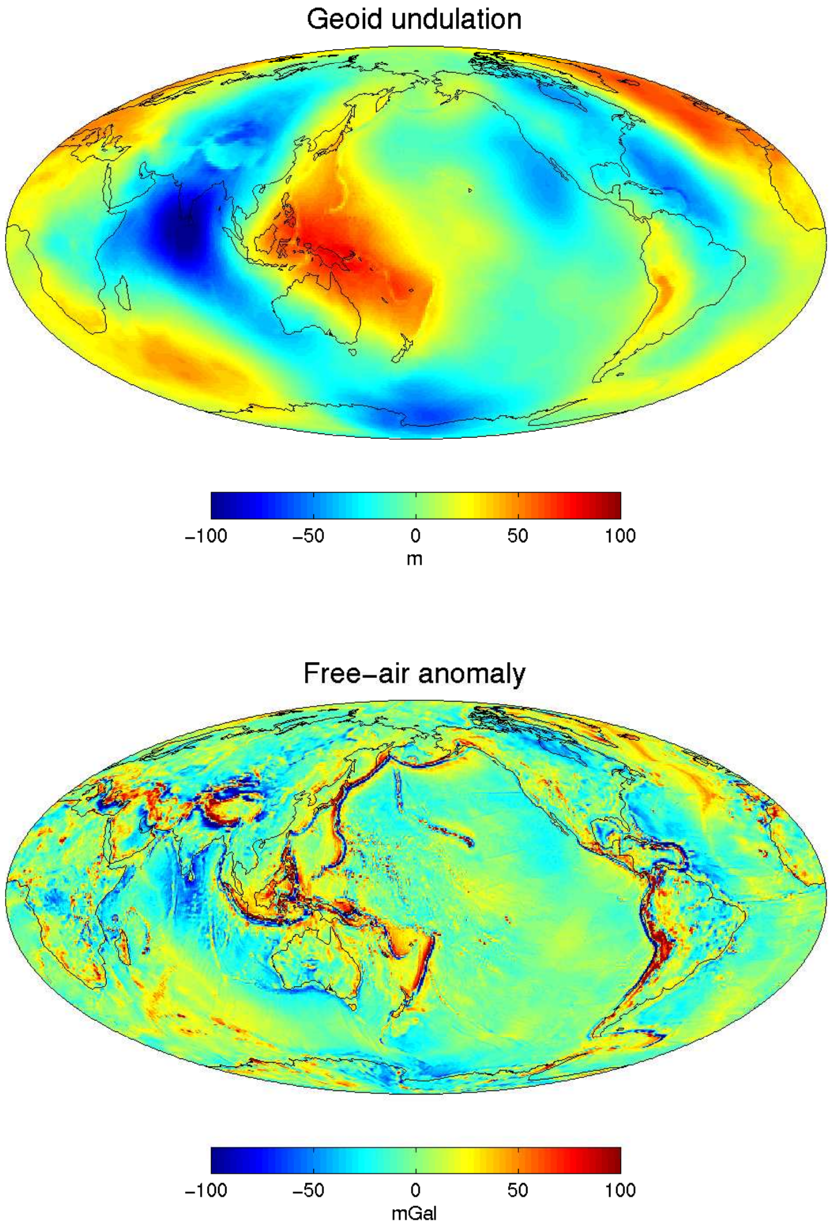


Fig. 2.11. The EGM96 Earth geopotential reference model.

3
Magnetism

All the world is a magnet. *“Magnus magnes ipse est globus terrestris”*, said Alexander von Humboldt, paraphrasing William Gilbert of Colchester [10]. Geomagnetism may well be the oldest *natural science* of all. Certainly older than any other area of “modern” geophyiscs–knowledge of the study of Earth’s magnetic field has been vital to humans ever since we learned how to sail.

Ancient Chinese civilizations knew how to navigate on the compass for ages. They knew the magnetic needle did not point precisely to geographic pole. That knowledge is said to have been passed on through Italy and Flanders to the ages of Western “discovery”, and played a role in the voyages of Columbus [11]. Formal inquiry had to wait for another few centuries, but as Newton was to gravity in the 18th century, so was **Gauss** to magnetism in the 19th.

Again we are able to start with nineteenth-century physics and make our way into the satellite age, with, as of 2019, the thirteenth-generation *International Geomagnetic Reference Field*. Like its predecessors, IRGF-13 is a model valid for a period of just five years, and it comes packaged with derivatives that attempt to capture its time dependence—the *secular variation*.

This fact alone should give us pause. Surely the geodetic *International Terrestrial Reference System* in the previous chapter wasn’t changing all that fast? After all we are still well served by the WGS84 geoid, are we not? But yes indeed, the study of geomagnetism introduces a temporal component that we will attempt to explain and understand.

The Earth’s *core* or *main* field is ever-changing, actively (re)generated by convective currents in the Earth’s outer core, a dynamic system referred to as the **geodynamo** [12]. In contrast, the *crustal* and *lithospheric* fields are remnant. Quite literally they are the fossils of plate tectonics [13]. Their interaction produces an *induced* field. And that’s just what happens in the solid Earth system—about the oceans and atmosphere, we will be largely silent.

3.1 Force, field and induction

Gravity was about the mutual *gravitational* attraction of two point masses, and up from there. Next we will be considering the mutual *magnetic* attraction of two small loops of electric current: "magnetic point masses", if you will.

$\mathbf{E}$ But first, introduce the **Lorentz force** on an electrical charge q moving with a velocity **v** in an electric field **E** and a magnetic field **B**:

$$\mathbf{f}(\mathbf{r}) = q\,[\mathbf{E}(\mathbf{r}) + \mathbf{v}(\mathbf{r}) \times \mathbf{B}(\mathbf{r})]. \tag{3.1}$$

$\mathbf{B}$ Electric currents generate a force on a moving charge which is proportonial to the cross product of the **magnetic induction field B** and the velocity of the charge. Through the force we simultaneously *define* and find a way to *measure* the electric and magnetic fields. This also fixes the units of **B** to be Tesla (T), where the dimensions are clearly given as [T]=MT^{-2}I^{-1}.

Let us take a quick look at the **electric force**, and thereby define the electric field, via **Coulomb's "law"**, for a point charge Q, namely

$$\mathbf{f}_E = q\,\mathbf{E} = \frac{1}{4\pi\epsilon_0}\frac{qQ}{r^2}\hat{\mathbf{r}}, \tag{3.2}$$

where ϵ_0 is the *electric constant*, and the **magnetic force** is

$$\mathbf{f}_B = q\,\mathbf{v} \times \mathbf{B}. \tag{3.3}$$

And thus much like with gravity in eq. (2.3), we obtain

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{Q}{r^2}\,\hat{\mathbf{r}}. \tag{3.4}$$

The electrical potential due to such a point charge (an electrical **monopole**) will be

$$V_E^P = \frac{1}{4\pi\epsilon_0}\frac{Q}{r}, \tag{3.5}$$

and for a **dipole** we will expect a term in r^{-2}.

3.2 Electrical and magnetic dipoles

Two equal but opposite charges separated by a distance **d** (in the direction from negative to positive) define an **electrical dipole** moment,

$$\mathbf{m}_E = Q\mathbf{d}. \tag{3.6}$$

From symmetry considerations, the expansion of $1/r$ etc., and using eq. (3.5), exploiting symmetry and to first order,

$$V_E^D = \frac{1}{4\pi\epsilon_0}\frac{\mathbf{m}_E \cdot \hat{\mathbf{r}}}{r^2}, \tag{3.7}$$

and for the strength of the dipole we find after a short amount of work

$$E^D(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{m_E}{r^3}(3\cos^2\theta + 1)^{1/2}. \tag{3.8}$$

Like for gravity, the electric dipole generates an electrical potential at a point a distance r removed from the center of the dipole that falls off as r^{-3}, modulated by the angle θ that the observation point makes with the axis of the dipole. Hence, (r, θ) are *dipole coordinates*.

Now let us anticipate that there is also something like a **magnetic dipole**, and let's pretend it is due to two elusive magnetic monopoles spaced *very* close together. The analogy to eq. (3.7) is apparent from the result

$$V^D(\mathbf{r}) = \frac{\mu_0}{4\pi}\frac{\mathbf{m}\cdot\hat{\mathbf{r}}}{r^2}. \tag{3.9}$$

3.3 Magnetic potential

As long as there are no electrical currents, outside of any magnetic materials, a scalar potential is sufficient,

$$\boxed{\mathbf{B} = -\boldsymbol{\nabla} V_M,} \tag{3.10}$$

Apply eq. (3.10) to eq. (3.9), using the product rule, and eq. (2.10), to obtain the vector expression

$$\mathbf{B}^D(\hat{\mathbf{r}}) = \frac{\mu_0}{4\pi}\frac{m}{r^3}\left[3(\hat{\mathbf{m}}\cdot\hat{\mathbf{r}})\,\hat{\mathbf{r}} - \hat{\mathbf{m}}\right], \tag{3.11}$$

an equation that we can actually use in practice, and we note that

$$B^D(\mathbf{r}) = \frac{\mu_0}{4\pi}\frac{m}{r^3}(3\cos^2\theta + 1)^{1/2}, \tag{3.12}$$

exactly as in eq. (3.8). In dipole coordinates of course.

3.4 Potential *outside* an Earth-like body

The above contained important simplifications. We want to get at Laplace's law for the magnetic potential but need to show the conditions under which this holds. How do we do that? We first write the full solution for the potential of an *external field* as eq. (2.64) and the potential for the *internal field* as eq. (2.65).

Earth models for an *internal field* are often given as normalized *bandlimited* expansions—no longer just a dipole, but a collection of *multipoles*:

$$V(\mathbf{r}) = a\sum_{l=1}^{L}\sum_{m=0}^{l}\left(\frac{a}{r}\right)^{l+1}[g_{lm}\cos m\phi + h_{lm}\sin m\phi]P_{lm}(\cos\theta). \tag{3.13}$$

The dipole terms alone are

$$V^D(\mathbf{r}) = \frac{a^3}{r^2}(g_{10}\cos\theta + g_{11}\cos\phi\sin\theta + h_{11}\sin\phi\sin\theta). \tag{3.14}$$

Writing out the dot product n eq. (3.9) after transforming the unit vector $\hat{\mathbf{r}}$ via eqs (1.23)–(1.24) yields

$$V^D(\mathbf{r}) = \frac{\mu_0}{4\pi}\frac{1}{r^2}(m_x\sin\theta\cos\phi + m_y\sin\theta\sin\phi + m_z\cos\theta), \tag{3.15}$$

which allows for the identification of the components of the dipole vector with the Gauss coefficients as follows:

$$\begin{pmatrix} m_x \\ m_y \\ m_z \end{pmatrix} = a^3\frac{4\pi}{\mu_0}\begin{pmatrix} g_{11} \\ h_{11} \\ g_{10} \end{pmatrix}. \tag{3.16}$$

Similarly, we may obtain the components of the dipole field by combining eq. (3.14) with eqs (3.10) and (2.9). This yields a field vector in a way that can be thought of as a coordinate transformation from the dipole components to the field components, if we reuse the matrix $\mathbf{\Gamma}$ from eq. (1.24):

$$\begin{pmatrix} \frac{1}{2}B^D_r \\ -B^D_\theta \\ -B^D_\phi \end{pmatrix} = \frac{a^3}{r^3}\mathbf{\Gamma}^{\mathrm{T}}\cdot\begin{pmatrix} g_{11} \\ h_{11} \\ g_{10} \end{pmatrix}. \tag{3.17}$$

In a frame of reference (r', θ', ϕ') with one axis aligned with the dipole itself, so that $\mathbf{m} = \hat{\mathbf{z}}\, m = \hat{\mathbf{z}}\, a^3 4\pi\, g_{10}/\mu_0$, the field components are given by

$$\begin{pmatrix} B^D_{r'} \\ B^D_{\theta'} \\ B^D_{\phi'} \end{pmatrix} = \frac{m}{r'^3}\frac{\mu_0}{4\pi}\begin{pmatrix} 2\cos\theta' \\ \sin\theta' \\ 0 \end{pmatrix}, \tag{3.18}$$

which validates the observation that the field strength at the magnetic north pole ($\theta' = 0$) is about twice that at the magnetic equator ($\theta' = \pi/2$). The dipolar equatorial field strength at the Earth's surface is

$$B^D_0 = \frac{m}{a^3}\frac{\mu_0}{4\pi} = (g^2_{11} + h^2_{11} + g^2_{10})^{1/2}. \tag{3.19}$$

3.5 Inclination and declination

The angle between $\mathbf{B} = B_r\hat{\mathbf{r}} + B_\theta\hat{\boldsymbol{\theta}} + B_\phi\hat{\boldsymbol{\phi}}$ and the tangent plane to the Earth's surface at the same location (i.e. the dip measured positively downward from the local horizontal) is called *inclination* and given in general by:

$$I = \operatorname{atan} \frac{B_r}{\sqrt{B_\theta^2 + B_\phi^2}}, \tag{3.20}$$

which, for simple dipoles and in a dipole reference frame, simplifies to

$$I^D = \operatorname{atan} \frac{B_{r'}^D}{B_{\theta'}^D} = \operatorname{atan}\left(2 \cot \theta'\right). \tag{3.21}$$

See Figure 3.5. The *declination* is the azimuth of the field vector, i.e. its clockwise-positive angle with the geographic meridian:

$$D = \operatorname{atan} \frac{B_\phi}{B_\theta}. \tag{3.22}$$

Dipoles expressed in a dipole reference frame have no declination, $D^D = 0$. Important for navigation, *isoclinal* maps show contours of equal inclination, as in Fig. 3.3; *isogonal* maps show contours of equal declination, as in Fig. 3.5.

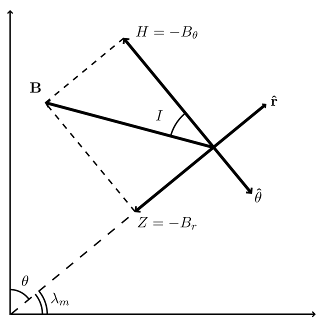


Fig. 3.1. Magnetic inclination.

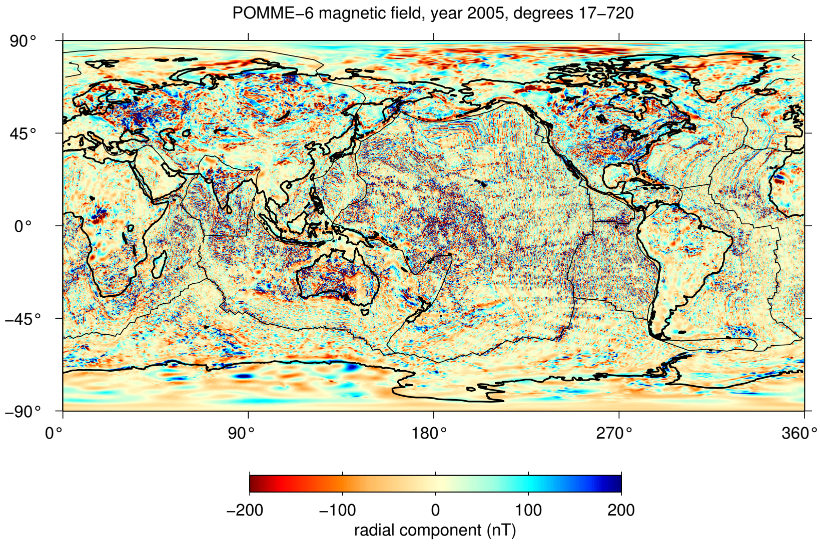


Fig. 3.2. POMME-6 for 2005!

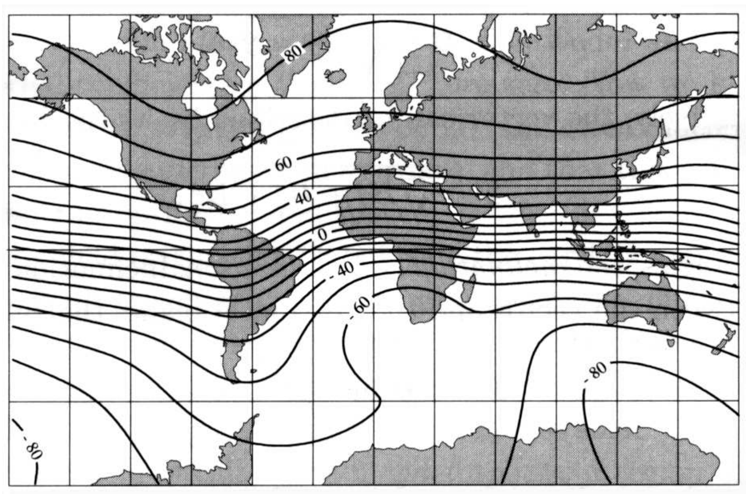


Fig. 3.3. Isoclinic map: constant inclination in the IGRF 1990 geomagnetic field.

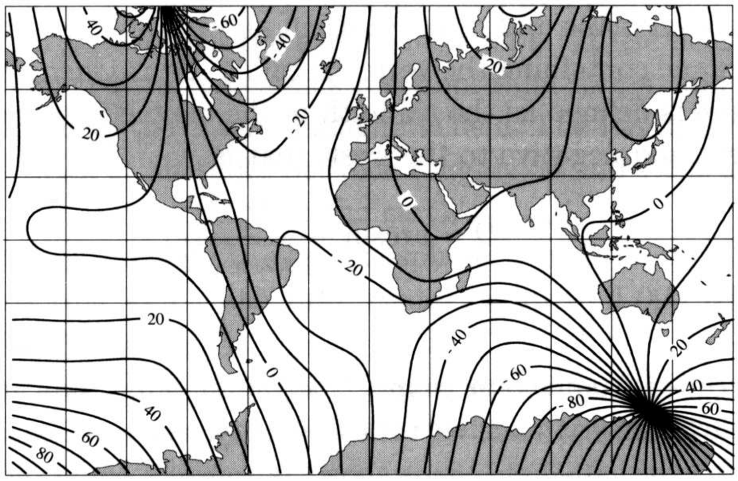


Fig. 3.4. Isogonic map: constant declination of the IGRF 1990 geomagnetic field .

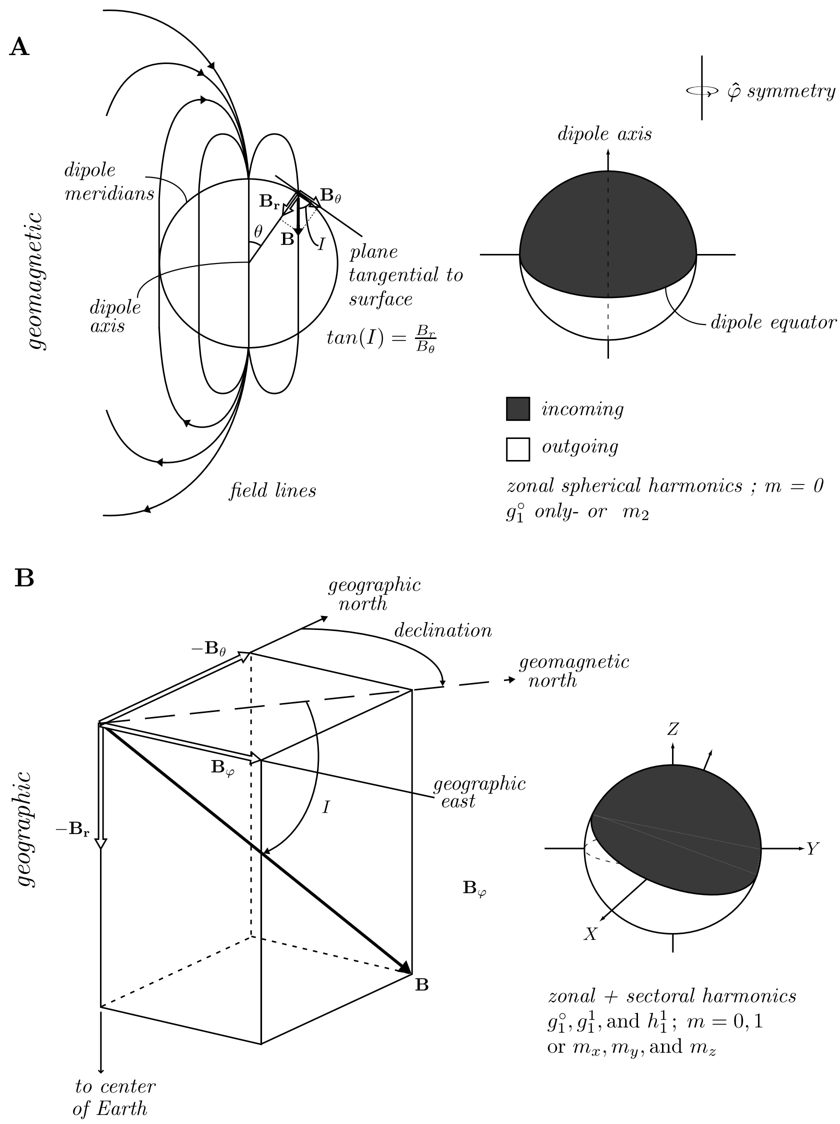


Fig. 3.5. Definitions and such.

3.6 Power spectrum and analytical continuation

Should write something about this. Simply about the description, causes later.
(1) My figure from SPIE would fit in great here.

3.7 International reference values

See the table. Make maps.

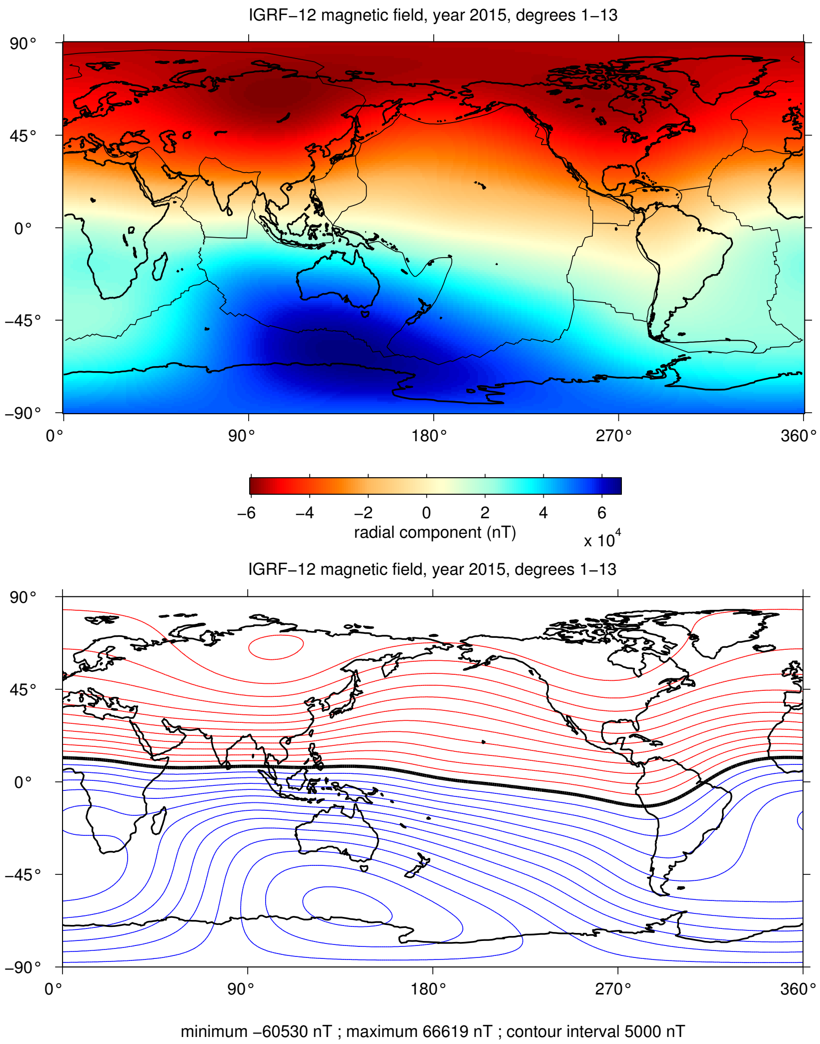


Fig. 3.6. The twelfth International Geomagnetic Reference Field model [14]. Shown is the radial component

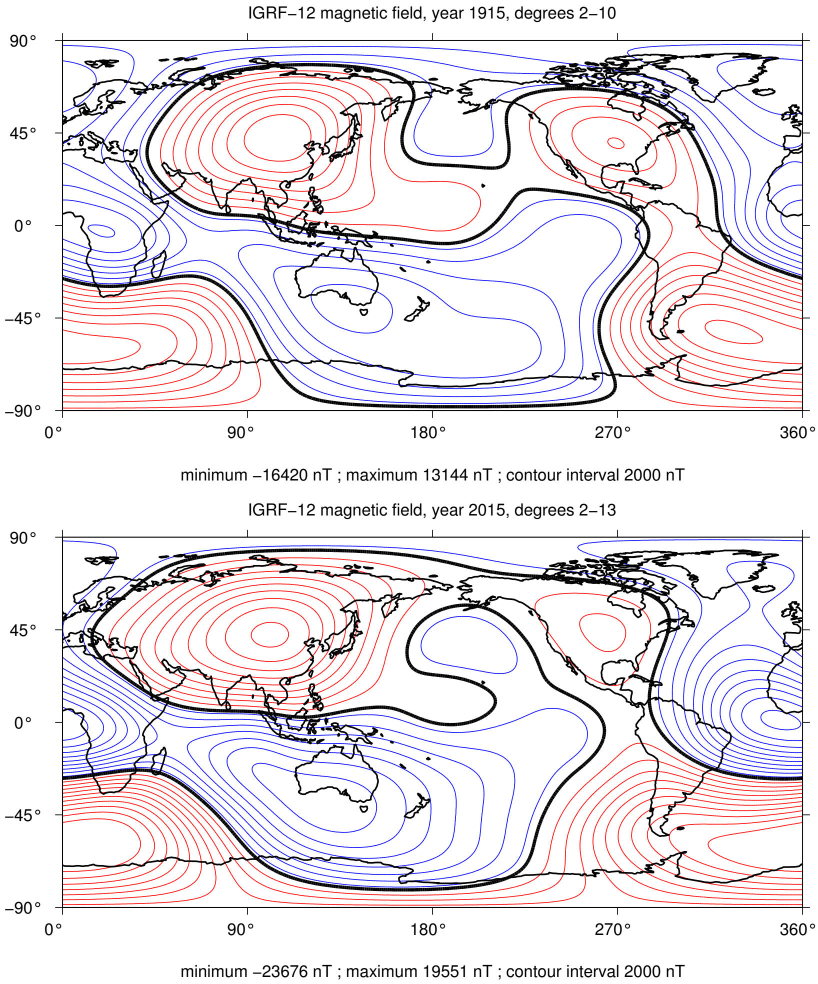


Fig. 3.7. The twelfth International Geomagnetic Reference Field model [14]. (*Top*) The complete field for 2015. (*Middle*) The non-dipolar field for 2015. (*Bottom*) The non-dipolar field for 1915. Note the clearly visible westward drift of the flux patches.

3.8 The curl (of a vector field)

In Cartesian coordinates. The **curl** or rotation operator $\nabla\times$, the *cross product* of the gradient with the argument, leaves the rank intact — for our vector

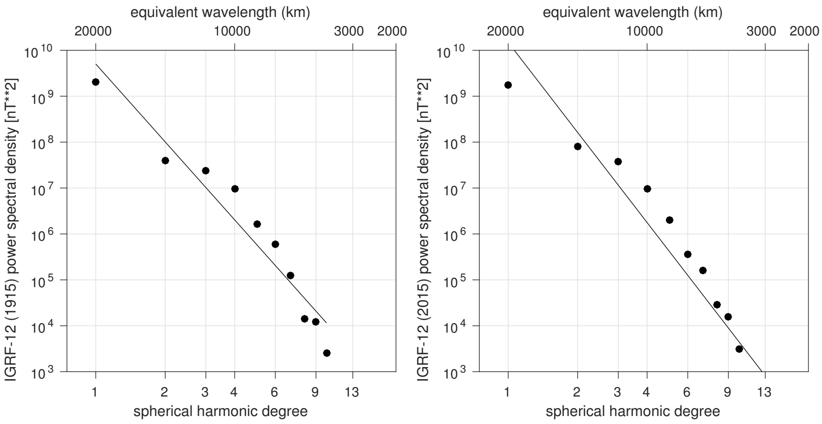


Fig. 3.8. Power-spectral density of the geomagnetic field model IGRF-12, for 1915 and for 2015 [14]. Fits to the power in the range $l = 1$–10 for 1915 and $l = 1$–10 for 2015 show the spectral decay of the core field, and how little, on balance, has changed in terms of the energy content of the main field.

l	m	g_{lm}	h_{lm}	$\dot{g}_{lm}$	$\dot{h}_{lm}$
0	0	0	0	0	0
1	0	-29556.8	0	8.8	0
1	1	-1671.8	5080.0	10.8	-21.3
2	0	-2340.5	0	-15.0	
2	1	3047.0	-2594.9	-6.9	-23.3
2	2	1656.9	-516.7	-1.0	-14.0

Table 3.1. *Gauss coefficients* (nT) *of the Earth's internal field and their temporal derivatives* (nT/yr)*, in 2005, according to the 10th International Geomagnetic Reference Field (IGRF-10), as in eq. (3.13) with* $a = 6371.2$ m. *Source:* `http://www.ngdc.noaa.gov/IAGA/vmod.`

function $\mathbf{u}$, the easiest representation is in *determinant* form:

$$\nabla \times \mathbf{u} = \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ u_x & u_y & u_z \\ \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \end{vmatrix}, \tag{3.23}$$

which is as much as

$$\left(\frac{\partial u_z}{\partial y} - \frac{\partial u_y}{\partial z}\right)\hat{\mathbf{x}} + \left(\frac{\partial u_x}{\partial z} - \frac{\partial u_z}{\partial x}\right)\hat{\mathbf{y}} + \left(\frac{\partial u_y}{\partial x} - \frac{\partial u_x}{\partial y}\right)\hat{\mathbf{z}}, \tag{3.24}$$

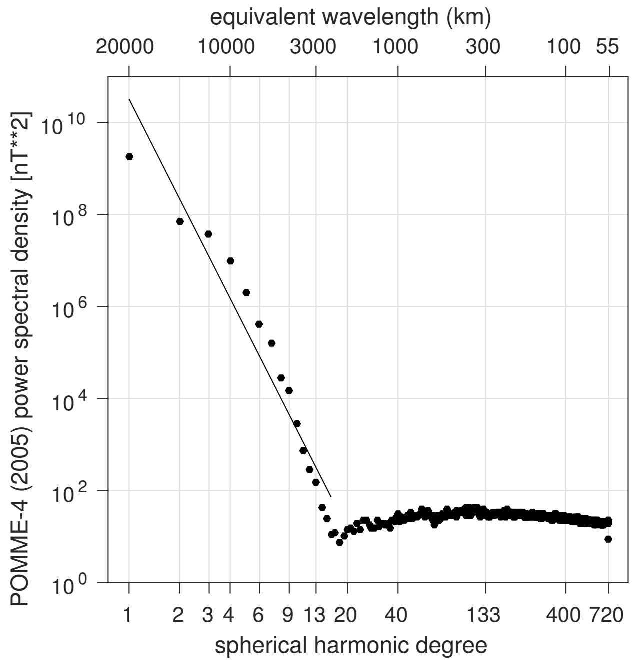


Fig. 3.9. Power-spectral density of the geomagnetic field model POMME-6, valid for 2005 [15]. A fit to the spectral power in the range $l = 1$–13 is shown to represent the decay of the core field.

which defines a vector field normal to $\mathbf{u}$ and its gradient. It measures the *rotation* or *vorticity* of the $\mathbf{u}$-field. We show by example. In spherical coordinates the expressions are a bit more complicated, but still:

$$\nabla \times \mathbf{u} = \frac{1}{r}\left(\frac{\partial u_\phi}{\partial \theta} + u_\phi \cot\theta - \frac{1}{\sin\theta}\frac{\partial u_\theta}{\partial \phi}\right)\hat{\mathbf{r}} + \left(\frac{1}{r\sin\theta}\frac{\partial u_r}{\partial_\phi} - \frac{\partial u_\phi}{\partial r} - \frac{1}{r}u_\phi\right)\hat{\boldsymbol{\theta}} + \left(\frac{\partial u_\theta}{\partial r} + \frac{1}{r}u_\theta - \frac{1}{r}\frac{\partial u_r}{\partial \theta}\right)\hat{\boldsymbol{\phi}}. \tag{3.25}$$

Another shorthand is, in component notation, valid for both

$$\nabla \times \mathbf{u} = \sum_i \sum_j \sum_k \epsilon_{ijk} \partial_j u_k \, \hat{\mathbf{x}}_i, \tag{3.26}$$

where ϵ_{ijk} is the Levi-Cività alternating symbol.

3.9 Maxwell's equations *in vacuo*

Maxwells equations are all there is to know about the production and interrelation of electric and magnetic fields. A few of them we've already seen (in various forms). In this section, we will give Maxwell's Equations in vector form but derive them from the integral forms which were based on experiments.

3.9.1 The Gauss, Stokes and Green theorems

Two results from vector calculus will be used here. The first we already know: it is Gauss' *divergence theorem*. It relates the integral of the divergence of the field, over some closed volume V, to the flux through the surface $\partial V = \Sigma$ that bounds the volume. The divergence measures the sources and sinks within the volume. If nothing is lost or created within the volume, there will be no flux through its surface. Let's repeat eq. (2.46) here:

$$\int_V \boldsymbol{\nabla} \cdot \mathbf{u}\, dV = \int_{\partial V} \hat{\mathbf{n}} \cdot \mathbf{u}\, \mathrm{d}\Sigma. \tag{3.27}$$

And remember its two-dimensional version, eq. (2.47).

$$\int_\Sigma \boldsymbol{\nabla} \cdot \mathbf{u}\, \mathrm{d}\Sigma = \int_{\partial \Sigma} \hat{\mathbf{n}} \cdot \mathbf{u}\, dl. \tag{3.28}$$

A second important law is Stokes' *curl theorem*. This law relates the curl of a vector field, integrated over some surface, to the line integral of the field over the curve that bounds the surface, with tangent unit vector $\hat{\mathbf{t}}$:

$$\boxed{\int_\Sigma \hat{\mathbf{n}} \cdot (\boldsymbol{\nabla} \times \mathbf{u})\, \mathrm{d}\Sigma = \int_{\partial\Sigma} \hat{\mathbf{t}} \cdot \mathbf{u}\, dl.} \tag{3.29}$$

See the derivation from Gauss' in Snieder, which is very enlightening. For good measure, let us write eq. (3.29) in two planar Cartesian dimensions, i.e. for a surface Σ whose normal is given by $\hat{\mathbf{z}}\, d\Sigma$. Using eq. (1.29) to write $\mathbf{u} = u_x\hat{\mathbf{x}} + u_y\hat{\mathbf{y}}$ for the vector field, and $\hat{\mathbf{t}}\, dl = dx\, \hat{\mathbf{x}} + dy\, \hat{\mathbf{y}}$ for the field tangent to the curve $\partial\Sigma$, and with the definition of the Cartesian curl, eqs (3.23)–(3.24), we have

$$\boxed{\int_\Sigma \left(\frac{\partial u_y}{\partial x} - \frac{\partial u_x}{\partial y} \right) \mathrm{d}\Sigma = \int_{\partial\Sigma} \left(u_x\, dx + u_y\, dy \right).} \tag{3.30}$$

In this form, eq. (3.30) is known as Green's theorem.

This is a very handy theorem, e.g. to compute the area of a certain (parametric) surface. For instance, let the curve be defined by the planar equations inspired by eq. (1.22),

$$x = r\cos\phi \quad , \quad y = r\sin\phi \qquad \text{and} \qquad 0 \le \phi < 2\pi, \tag{3.31}$$

describing, in other words, a circle of radius r. If we now let $u_x = -y$ and $u_y = x$, we notice that the left hand side of eq. (3.30) defines the area of the enclosed disk, which is thus equal to

$$2\pi r^2 = \frac{1}{2}\int_{\partial\Sigma} (x\,dy - y\,dx) = \frac{1}{2}\int_{\phi=0}^{\phi=2\pi} \left(x\,\frac{\partial y}{\partial\phi} - y\,\frac{\partial x}{\partial\phi} \right) d\phi, \tag{3.32}$$

as can be easily verified by substitution of eq. (3.31).

Now to an exercise in spherical geometry:

$$\int_\Sigma d\Omega = \int_\theta\int_\phi \sin\theta\,d\theta\,d\phi = \int_\theta\int_\phi \left(\frac{\partial u_\phi}{\partial\theta} - \frac{\partial u_\theta}{\partial\phi} \right) d\theta\,d\phi, \tag{3.33}$$

and thus what goes is

$$\int_{\partial\Sigma} (u_\theta\,d\theta + u_\phi\,d\phi)\,. \tag{3.34}$$

But can't now u_θ be anything non-ϕ dependent, and $u_\phi = -\cos\theta$? And note that the integral of ϕ is *along* the curve. This is the solution—whatever the constant, it vanishes on a closed curve....

Fig. 3.10. Diagram illustrating the geometry of Stokes' theorem.

3.9.2 The magnetic field is solenoidal

We have already seen that magnetic field lines begin and end at the magnetic dipole. Magnetic "charges" or "monopoles" are not generally thought to exist except in very exotic materials and under special circumstances. Hence, all field lines leaving a surface enclosing a dipole, reenter that same surface. There is no magnetic flux through a closed surface:

$$\int_{\partial V} \mathbf{B}\cdot \mathrm{d}\mathbf{\Sigma} = 0. \tag{3.35}$$

solenoidal

Rewriting this with Gauss' theorem (4.17) gives a first law of Maxwell's:

$$\boxed{\boldsymbol{\nabla}\cdot\mathbf{B} = 0.} \tag{3.36}$$

3.9.3 Electromagnetic induction

An empirical law due to Faraday says that changes in the magnetic flux through a surface induce a current in a wire loop that encloses the surface:

$$\frac{d}{dt}\int_{\Sigma} \mathbf{B}\cdot \mathrm{d}\boldsymbol{\Sigma} = -\int_{\partial\Sigma} \mathbf{E}\cdot \mathrm{dl}, \tag{3.37}$$

which can be rewritten using Stokes' theorem (3.29) to give a second law of Maxwell's:

$$\boxed{\boldsymbol{\nabla}\times\mathbf{E} = -\frac{\partial}{\partial t}\mathbf{B}.} \tag{3.38}$$

3.9.4 Displacement current

We've seen that a time-dependent magnetic flux induces an electric field. The reverse is true: a time-dependent electric flux induces a magnetic field. But a current by itself was also responsible for a magnetic field, as per Biot-Savart's law of eq. (**??**). Both effects can be combined into one equation as follows:

$$\mu_0\left(i + \epsilon_0 \frac{d}{dt}\int_{\Sigma} \mathbf{E}\cdot \mathrm{d}\boldsymbol{\Sigma}\right) = \int_{\partial\Sigma} \mathbf{B}\cdot \mathbf{dl} \tag{3.39}$$

The term i is the conductive "regular" current that was studied experimentally by Ampère. The second term within the brackets also has the dimensions of a current and is termed "displacement" current. Instead of current i we will now introduce the current *density* vector $\mathbf{J}$ (per unit of surface and perpendicular to the surface) so that $\int_{\Sigma}\mathbf{J}\cdot\mathrm{d}\boldsymbol{\Sigma} = i$. We can then use the curl theorem again and write the third of Maxwell's equations:

$$\boxed{\boldsymbol{\nabla}\times\mathbf{B} = \mu_0\left(\mathbf{J} + \epsilon_0\frac{\partial}{\partial t}\mathbf{E}\right).} \tag{3.40}$$

The curl of $\mathbf{B}$ is the vector sum of all forms of charge through the region. In the absence of moving charge, $\mathbf{B}$ is irrotational. irrotational

In the absence of conduction or displacement currents, the magnetic field is irrotational, $\boldsymbol{\nabla}\times\mathbf{B} = 0$, as per eq. (3.40), in additional to being solenoidal, $\boldsymbol{\nabla}\cdot\mathbf{B} = 0$, which we knew from eq. (3.36). In that case, Maxwell's equations imply the harmonicity of the magnetic potential, $\nabla^2 V = 0$.

3.9.5 Electric flux in terms of charge density

Remember how we obtained the flux of the gravity field in terms of the mass density. In contrast, the flux of the magnetic field was for a closed surface enclosing a dipole. For a closed surface enclosing a charge distribution, the flux through that surface will be related to the electrical charge density contained in the volume! This is a manifestation of the potential (rather than solenoidal) nature of the electric field. We write

potential

$$\int_{\partial V} \mathbf{E} \cdot \mathrm{d}\boldsymbol{\Sigma} = \frac{q}{\epsilon_0}, \tag{3.41}$$

which, with the help of the divergence theorem and introducing the volumetric electrical charge density ρ_E so $\int_V \rho_E \, dV = q$, transforms easily to the fourth of Maxwell's laws:

$$\boxed{\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\rho_E}{\epsilon_0}.} \tag{3.42}$$

Make connection with Helmholtz' theorem a la Backus, 2nd chapter

3.10 Maxwell's equations *in an Earth-like body*

The Earth ain't a vacuum. Got to do something about it. Basic point is that we need terms for the *electric polarization* and the *magnetic polarization* or *magnetization* of continuous media such as rocks. The main thing being that we consider Maxwells' equations to apply not just at points in a vacuum, but also *on average*, over some small volume including a number of atoms, whose average properties vary smoothly when the center of the volume is moved about a little. Here's the complete set again

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0, \tag{3.43}$$
$$\boldsymbol{\nabla} \times \mathbf{E} = -\partial_t \mathbf{B}, \tag{3.44}$$
$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \left(\mathbf{J} + \epsilon_0 \partial_t \mathbf{E} \right), \tag{3.45}$$
$$\boldsymbol{\nabla} \cdot \mathbf{E} = \rho_E / \epsilon_0. \tag{3.46}$$

And let's now have them hold *on average*, as in

$$\boldsymbol{\nabla} \cdot \langle \mathbf{B} \rangle = 0, \tag{3.47}$$
$$\boldsymbol{\nabla} \times \langle \mathbf{E} \rangle = -\partial_t \langle \mathbf{B} \rangle, \tag{3.48}$$
$$\boldsymbol{\nabla} \times \langle \mathbf{B} \rangle = \mu_0 \left(\langle \mathbf{J} \rangle + \epsilon_0 \partial_t \langle \mathbf{E} \rangle \right), \tag{3.49}$$
$$\boldsymbol{\nabla} \cdot \langle \mathbf{E} \rangle = \langle \rho_E \rangle / \epsilon_0. \tag{3.50}$$

At macroscopic distances an atom looks like a point charge plus an electric and a magnetic dipole. We take this into account as follows. Define a small volume $\Delta V \neq 0$ for which

$$\mathbf{P} = \frac{1}{\Delta V} \sum_i \mathbf{p}_i, \tag{3.51}$$

for the electric polarization per unit volume and likewise for the **magnetic polarization** per unit volume

$$\mathbf{m} = \frac{1}{\Delta V} \sum_i \mathbf{m}_i. \tag{3.52}$$

Otherwise called the **magnetization vector** it is the vector sum of all of the individual magnetic moments per unit of volume.

It can then be shown that the average charge density is

$$\langle \rho_E \rangle = \rho_F - \boldsymbol{\nabla} \cdot \mathbf{P}, \tag{3.53}$$

and that the average current density

$$\langle \mathbf{J} \rangle = \mathbf{J}_F + \boldsymbol{\nabla} \times \mathbf{m} + \partial_t \mathbf{P}, \tag{3.54}$$

where the subscript F refers to *free* charges and currents (the *macroscopic* ones), and the correction terms represent the *bound* charges and currents (the *microscopic* ones), those intrinsic to the atoms and molecules contained in the small volume of material over which is averaged.

So we rewrite Maxwell's equations for the last time in this section in a *real* medium as:

$$\boldsymbol{\nabla} \cdot \langle \mathbf{B} \rangle = 0, \tag{3.55}$$

$$\boldsymbol{\nabla} \times \langle \mathbf{E} \rangle = -\partial_t \langle \mathbf{B} \rangle, \tag{3.56}$$

$$\boldsymbol{\nabla} \times \langle \mathbf{B} \rangle = \mu_0 \left(\mathbf{J}_F + \boldsymbol{\nabla} \times \mathbf{m} + \partial_t \mathbf{P} + \epsilon_0 \partial_t \langle \mathbf{E} \rangle \right), \tag{3.57}$$

$$\boldsymbol{\nabla} \cdot \langle \mathbf{E} \rangle = (\rho_F - \boldsymbol{\nabla} \cdot \mathbf{P})/\epsilon_0. \tag{3.58}$$

3.11 The electric and the magnetic displacement vectors

Materials respond to an applied electric field $\mathbf{E}$ and an applied magnetic field $\mathbf{B}$ by producing their own internal *bound* charges and current distributions, which in turn contribute to $\mathbf{E}$ and $\mathbf{B}$. Since those properties are difficult to calculate, it is customary in the literature to define an *electric displacement vector*

$$\mathbf{D} = \epsilon_0 \langle \mathbf{E} \rangle + \mathbf{P}, \tag{3.59}$$

and a *magnetic displacement vector* or *magnetic field intensity*

$$\mathbf{H} = \langle \mathbf{B} \rangle / \mu_0 - \mathbf{m}, \tag{3.60}$$

which is seen to be the magnetic induction minus the effects of magnetization.

These auxiliary fields $\mathbf{H}$ qand $\mathbf{D}$ allow us to rerwite two of Maxwell's equations again in terms of the *free* current and charge densities

$$\nabla \times \mathbf{H} = \mathbf{J}_{_F} + \partial_t \mathbf{D}, \tag{3.61}$$

$$\nabla \cdot \mathbf{D} = \rho_F. \tag{3.62}$$

or, finally, assuming spatial variability on the length scale of averaging, all four of them in their most often quoted form

$$\nabla \cdot \mathbf{B} = 0, \tag{3.63}$$

$$\nabla \times \mathbf{E} = -\partial_t \mathbf{B}, \tag{3.64}$$

$$\nabla \times \mathbf{H} = \mathbf{J}_{_F} + \partial_t \mathbf{D}, \tag{3.65}$$

$$\nabla \cdot \mathbf{D} = \rho_F. \tag{3.66}$$

Eqs. (3.63)–(3.66) are not "better" or more general than the original equations (3.55)–(3.63). But it is possible to avoid needing to calculate the bound charges and currents if you can supply **constitutive relations** between $\mathbf{B}$ and $\mathbf{H}$ and between $\mathbf{E}$ and $\mathbf{D}$.

3.12 Constitutive relations

Now we're getting to real *materials*. We need *constitutive relations* that express how we can find the terms $\mathbf{J}_{_F}$, $\mathbf{P}$, and $\mathbf{m}$, which we need in order to solve Maxwell's equations. We also need ρ_F, but that one is easier: adding the time derivative of eq. (3.58) to the divergence of eq. (3.57), we obtain

$$\nabla \cdot \mathbf{J}_{_F} = -\partial_t \rho_F, \tag{3.67}$$

which is a continuity equation expressing the fact that free charges are conserved. So we know how they relate, and with an initial condition for $\rho_F(\mathbf{r}, 0)$ we can construct $\rho_F(\mathbf{r}, t)$ at all times $t > 0$.

It is the domain of experiment or *ab initio* calculations to obtain the material parameters

$$\mathbf{J}_{_F} = \boldsymbol{\sigma} \cdot \langle \mathbf{E} \rangle, \tag{3.68}$$

$$\mathbf{P} = \epsilon_0 \, \boldsymbol{\chi}_{_E} \cdot \langle \mathbf{E} \rangle, \tag{3.69}$$

$$\mathbf{m} = \boldsymbol{\chi}_{_B} \cdot \langle \mathbf{B} \rangle / \mu_0, \tag{3.70}$$

where $\boldsymbol{\sigma}$ is the electric conductivity, $\boldsymbol{\chi}_{_E}$ the electric susceptibility and $\boldsymbol{\chi}_{_B}$ the magnetic susceptibility. All of these quantities are tensors; in the absence of information about the anisotropy of the materials in question, we can think of them as scalars, and write

$$\mathbf{J}_F = \sigma\langle\mathbf{E}\rangle, \tag{3.71}$$
$$\mathbf{P} = \epsilon_0\,\chi_E\langle\mathbf{E}\rangle, \tag{3.72}$$
$$\mathbf{m} = \chi_B\langle\mathbf{B}\rangle/\mu_0, \tag{3.73}$$

Eq. (3.71) is known as **Ohm's "law"**. In a stationary reference frame, it becomes identifiable with the effect of the Lorentz force of eq. (3.1):

$$\mathbf{J}_F = \sigma(\langle\mathbf{E}\rangle + \mathbf{v}\times\langle\mathbf{B}\rangle). \tag{3.74}$$

3.13 Rock magnetism

The *relative magnetic susceptibility* is now usually defined not via $\mathbf{B}$ as in eq. (3.73) but rather via $\mathbf{H}$, as

$$\mathbf{m} = \chi\mathbf{H}, \tag{3.75}$$

The quantity $\mathbf{m}$ is the magnetization *induced* by the *ambient* or *applied* external field $\mathbf{H}$. We are now writing eq. (3.60) with eq. (3.75) as

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{m} = \frac{\mathbf{B}}{\mu_0(1+\chi)} = \frac{\mathbf{B}}{\mu}, \tag{3.76}$$

where μ is the *magnetic permeability* of the material. So $\mathbf{B}$ is the resulting field. If we write perhaps that

$$\mathbf{B} = \mu_0(\mathbf{H}+\mathbf{m}) = \mu_0(1+\chi)(\mathbf{H}) = \mu\mathbf{H} \tag{3.77}$$

it is easier to read this equation... the total magnetic field is given by the sum of the ambient field and the resulting magnetic field induced in the magnetizable substance. The materials science aspect of these relations is very complex. Not only can the relationships be tensorial, but they need not even be linear, and the are strongly temperature dependent.

3.13.1 Diamagnetism

Small effect. Opposing the applied field. This when $\chi < 0$.

3.13.2 Paramagnetism

Small effect. Reinforcing the applied field. This when $\chi > 0$.

3.13.3 Ferromagnetism

Strong effects with hysteresis. This when $\chi \gg 0$. The subclasses are ferromagnetic proper, antiferromagnetic, and ferrimagnetic.

Ferromagnetic materials will acquire an **induced magnetization**, as seen in eq. (3.75), which will disappear when the applied field $\mathbf{H}$ disappears. But they might also retain an **remanent magnetization** as a function of their geologic history. The sum total of both effects is

$$\mathbf{m}_\mathrm{T} = \chi\mathbf{H} + \mathbf{m}_\mathrm{R}, \tag{3.78}$$

and the relative importance of remanent to induced magnetization is the **Königsberger ratio**:

$$Q = \frac{\|\mathbf{m}_\mathrm{R}\|}{\chi\|\mathbf{H}\|}. \tag{3.79}$$

Ferromagnetism disappears above the material-dependent **Curie temperature**. While paramagnetic and diamagnetic effects persist at these temperatures, they are so small that above the Curie temperatures we may consider rocks and minerals to be nonmagnetic.

3.14 The Laplacian (of a vector field)

We'll need this

$$\nabla^2\mathbf{u} = \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{u}) - \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{u}). \tag{3.80}$$

3.15 The geodynamo

Under the **magnetohydrodynamic approximation**, we assume the displacement currents are small and that the material is Ohmic. Furthermore, in the Earth's core, it is assumed that no magnetization exists, so $\mathbf{H} = \mathbf{B}/\mu_0$ in that case. Eq. (3.74) holds, and eq. (3.65) becomes

$$\boldsymbol{\nabla}\times\mathbf{H} = \sigma(\mathbf{E} + \mathbf{v}\times\mathbf{B}). \tag{3.81}$$

Then it gets messy, because in order to *solve* the problem, we need to also solve for the velocity of the material... One ingredient in the required set of equations needed to solve for $\mathbf{v}$ is definitely Poisson's equation (2.52), but we also need to keep track of diffusion and advection of heat, and of supplying a good set of consitutive equations for the mass density. We're not nearly there, but we can get some insight quickly by looking at eq (3.81), taking the curl, taking its time derivative, using the vector rule eq. (3.80), and Maxwell's eqs (3.63) and (3.64) to get

$$\boxed{\frac{\partial \mathbf{H}}{\partial t} = \boldsymbol{\nabla} \times (\mathbf{v} \times \mathbf{H}) + \frac{1}{\mu_0 \sigma} \nabla^2 \mathbf{H}.} \tag{3.82}$$

which is known as the *magnetic induction equation*. The time variability of the field represents the balance between the competition between advective regeneration by core convection and Ohmic diffusional decay.

Define the ratio of the two terms as the *magnetic Reynolds' number*:

$$R_m = \mu_0 \sigma \, \frac{\|\nabla \times (\mathbf{v} \times \mathbf{H})\|}{\|\nabla^2 \mathbf{H}\|}. \tag{3.83}$$

And thus $R_m > 1$ is a necessary but not sufficient condition for the sustenance of the field, i.e. for the geodynamo to operate.

Full solutions require numerical, analog, mathematical, modelling, and coupling Maxwell's equations to those from fludi flow. It gets very complicated quickly. So we do two end-member cases to get some physical insight.

3.15.1 Diffusion-dominated regime

At low R_m, the right hand term in eq. (3.82) dominates. The field decays according to a *diffusion* equation:

$$\boxed{\partial_t \mathbf{H} = (\mu_0 \sigma)^{-1} \nabla^2 \mathbf{H}.} \tag{3.84}$$

3.15.2 Advection-dominated regime

At high R_m, the left-hand term dominates in eq. (3.82). We get an *advection* equation

$$\boxed{\partial_t \mathbf{H} = \boldsymbol{\nabla} \times (\mathbf{v} \times \mathbf{H}).} \tag{3.85}$$

We still won't be able to solve it without knowing what the fluid does, but we will just look at some implications for a generic surface through which we study the magnetic *flux*:

$$\int_\Sigma \hat{\mathbf{n}} \cdot (\partial_t \mathbf{H}) \, \mathrm{d}\Sigma = \int_\Sigma \hat{\mathbf{n}} \cdot \boldsymbol{\nabla} \times (\mathbf{v} \times \mathbf{H}) \, \mathrm{d}\Sigma. \tag{3.86}$$

Use Stokes' theorem (3.29) on the right hand side:

$$\int_\Sigma \hat{\mathbf{n}} \cdot (\partial_t \mathbf{H}) \, \mathrm{d}\Sigma = \int_{\partial\Sigma} \hat{\mathbf{t}} \cdot (\mathbf{v} \times \mathbf{H}) \, dl. \tag{3.87}$$

Noting that $\hat{\mathbf{t}} \cdot (\mathbf{v} \times \mathbf{H}) = -(\mathbf{v} \times \hat{\mathbf{t}}) \cdot \mathbf{H}$ we rewrite

$$\int_\Sigma \partial_t(\hat{\mathbf{n}} \cdot \mathbf{H})\, \mathrm{d}\Sigma + \int_{\partial\Sigma} (\mathbf{v} \times \hat{\mathbf{t}}) \cdot \mathbf{H}\; dl = 0. \tag{3.88}$$

The vector $\mathbf{v} \times \hat{\mathbf{t}}$ is the normal to the area swept out by the moving line segment in a unit time interval, and thus the second term represents the flux change experienced by the material enclosed within $\partial\Sigma$ as it moves about with the fluid. Rephrasing eq. (3.88) as as:

$$\frac{d}{dt} \int_\Sigma \hat{\mathbf{n}} \cdot \mathbf{H}\; \mathrm{d}\Sigma = 0, \tag{3.89}$$

we have arrived at the *frozen-flux principle*: field lines move with the flow in a perfect conductor.

3.16 Magnetic measurements and interpretation in practice

See what the magnetometer does.

3.17 Magnetism to (buried) bodies

A bit of an application. A worked example.

3.18 Magnetism and plate tectonics

In the context of plate tectonics, should tie it all together nicely.

3.19 Time-variable magnetism

Causes. Jerks. Reversals. Etc. Extension of the "secular variation bit". Check out the sign indifference of eq. (**??**).

3.20 Secular variation

Should write something about this. Simply about the description, causes later.

(1) Taylor series, jerks.

$$g_e(t) = g_e(t_0) + \left.\frac{\partial g}{\partial t}\right|_{t_0} (t - t_0) + \frac{1}{2} \left.\frac{\partial^2 g}{\partial t^2}\right|_{t_0} (t - t_0)^2 \tag{3.90}$$

(2) Relaxation time.

$$\tau = \left(\frac{\sum_{lm} (g_{lm}^2 + h_{lm}^2)}{\sum_{lm} (\dot{g}_{lm}^2 + \dot{h}_{lm}^2)} \right)^{1/2} \tag{3.91}$$

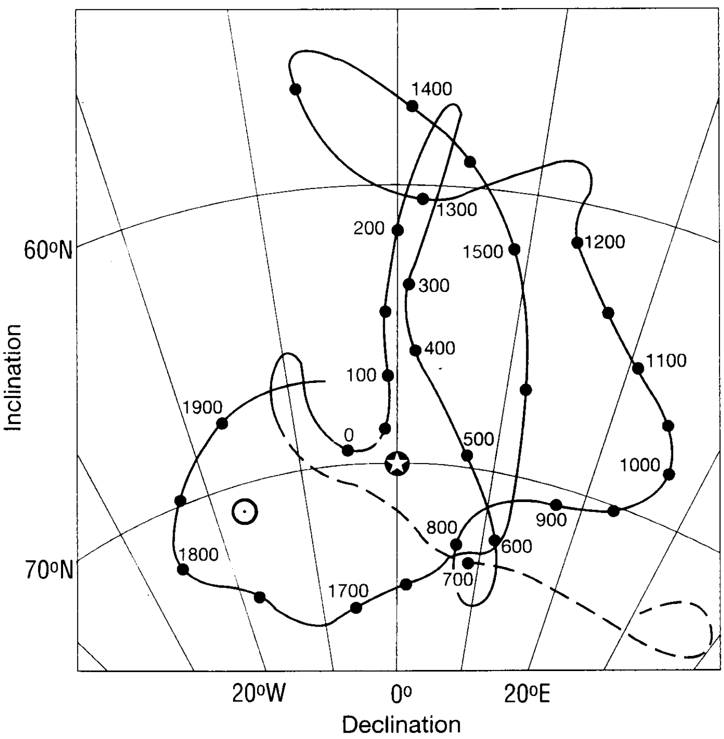


Figure 7.5. Secular variation of the magnetic field in Britain, plotted as inclination vs. declination as a function of time. The record of direct observations from about 1600 is extended back by archeomagnetic data. Numbers on the curve give dates. Archeomagnetic data were taken from a plot by Tarling (1989). The direction of the field due to a geocentric axial dipole is shown as the star and the direction corresponding to the present, inclined dipole is represented by the open circle.

Fig. 3.11. Bauer plot. Should really remake this myself.

3.20.1 Change in dipole strength

3.20.2 Change in orientation

3.20.3 Westward drift

4
Seismology

Not so long ago, none of us had literally *any* idea what earthquakes are, or what seismic waves are, for that matter. Seismology, before the word existed, was a branch of mathematics. Slawinski waxes philosophical about it.

Check out the historical account of Prince Galitzin at the Fifth Mathematical Congress.

Lord Rayleight knew stuff, as did Love, and Lamb.

Should also explicitly do the acoustic wave equation to make the link with exploration seismology and undergraduate physics. Think again also about my FRS2017 lecture on waves.

4.1 Force and traction

Continuum mechanics describes how materials (solids, liquids, gases) behave under the influence of external and internal forces. For the most part, it disregards the molecular structure of *continuous media*, and assumes that the mathematical functions that describe their properties are *continuous*—except at a finite number of interior surfaces separating regions of continuity.

A *force* $\mathbf{f}$ is a push or pull experienced by a mass m that is accelerated:

$$\mathbf{f} = m\,\mathbf{a}, \tag{4.1}$$

where $\mathbf{a}$ is the acceleration vector. This is known as *Newton's second law:* force equals mass times acceleration. The force of gravity or *weight* depends on the gravitational acceleration $\mathbf{g}$,

$$\mathbf{f} = m\,\mathbf{g}, \tag{4.2}$$

which we have described in detail in Chapter 2. *Body* ("action-at-a-distance") forces are often reckoned per unit mass. *Surface* ("contact") forces are usually defined as *tractions*, and reckoned per unit area.

Talk about inertia and linear momentum.

4.2 Torque

Need to connect that here, to moment of inertia in previous chapter! And angular momentum.

4.3 Cauchy's stress tensor

The ("Cauchy") *stress tensor* can be thought of as a matrix whose *columns* are filled with the components of the vector tractions $\mathbf{t}$ acting on three perpendicular faces of an imaginary infinitesimally small cube inside a solid object, as in Figure 4.1. Those three faces are denoted by their unit vectors. In a Cartesian coordinate system, these are $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$, and the stress tensor is written as:

$\mathbf{T}$

$$\mathbf{T} = \begin{pmatrix} t_x(\hat{\mathbf{x}}) & t_x(\hat{\mathbf{y}}) & t_x(\hat{\mathbf{z}}) \\ t_y(\hat{\mathbf{x}}) & t_y(\hat{\mathbf{y}}) & t_y(\hat{\mathbf{z}}) \\ t_z(\hat{\mathbf{x}}) & t_z(\hat{\mathbf{y}}) & t_z(\hat{\mathbf{z}}) \end{pmatrix} = \begin{pmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{pmatrix}. \tag{4.3}$$

Switching to a pure index notation, we write the elements of $\mathbf{T}$ as T_{ij}, in the three physical dimensions $i, j = 1, ..., 3$. The tractions T_{ij}, $i = j$, are perpendicular to the face on which they act and are called *normal*; the tractions T_{ij} for $i \neq j$ are called *shear*.

4.3.1 Symmetry of the stress tensor

Also known as: conservation of angular momentum.

We might have defined $\mathbf{T}$ in terms of *rows* of traction components. It doesn't matter. The stress tensor must be *symmetric*. This is a first law due to Cauchy:

$\mathbf{T}^{\mathrm{T}}$

$$\boxed{\mathbf{T} = \mathbf{T}^{\mathrm{T}}.} \tag{4.4}$$

or in index notation,

$$T_{ij} = T_{ji}. \tag{4.5}$$

Eq. (4.4) is a statement of the principle of *conservation of angular momentum*.

Proof here that is seat-of-the-pants. Maybe see Stein and Wysession. If we are to keep the infinitesimally small faces from rotating... Full treatment see Malvern or DT.

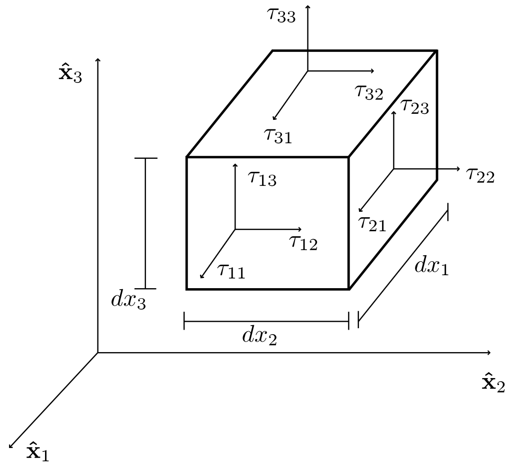


Fig. 4.1. Free-body diagram to illustrate the equilibrium condition and equations of motion. See GEO424 notes 9-30-02. In 2019 I used a single picture with three coordinate planes, an arbitrary cut, and a principal plane. Emphasizing the thread of choosing coordinate systems that runs through the course.

4.3.2 Tractions, from the stress tensor

It is to be understood that all the quantities defined so far will be considered to vary as a function of position $\mathbf{r}$ within the continuous medium (a strictly *Eulerian* viewpoint). By writing $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$, we've specified a *particular* Cartesian coordinate system—but why should our choice be special? What will the traction be on a plane that does *not* coincide with these coordinate axes? Let's label such a plane of some random orientation by its unit normal $\hat{\mathbf{n}}$.

Applying eq. (4.1) on the four sides of an imaginary ("Cauchy") tetrahedron show in Fig. (4.2), formed by cutting the coordinate axes by the plane $\hat{\mathbf{n}}$, we may show that $\mathbf{t}(\hat{\mathbf{n}})$

$$\boxed{\mathbf{t}(\hat{\mathbf{n}}) = \hat{\mathbf{n}} \cdot \mathbf{T}.} \tag{4.6}$$

or in index notation with the Einstein summation convention:

$$t_i(\hat{\mathbf{n}}) = \hat{n}_j T_{ji}. \tag{4.7}$$

This formula is also due to Cauchy. Knowing the stress tensor, we thus know precisely the tractions acting on *any* given plane inside the continuous medium.

Cauchy's relation (**??**) applies whether or not the medium is in equilibrium. It is true for an infinitesimally small volume element inside a medium; thus in the limit, it is true for a zero-volume volume! Thus, eq. (**??**) is true in fluid dynamics as well as solid mechanics. Incidentally, so is (4.4).

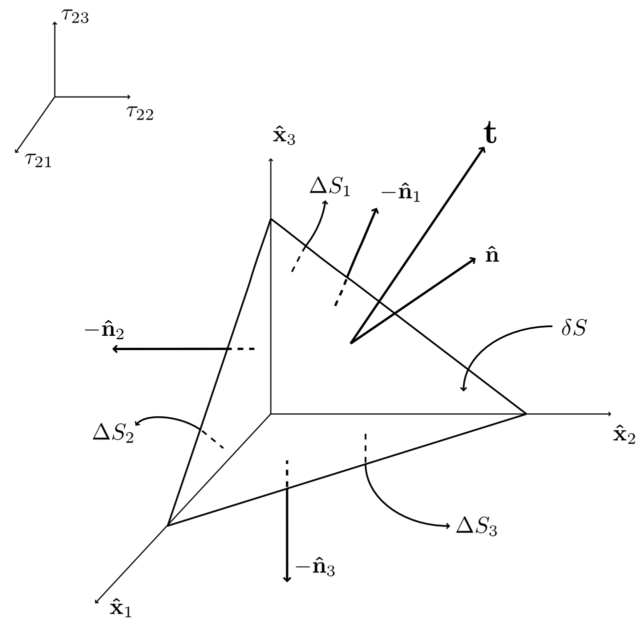


Fig. 4.2. The Cauchy tetrahedron. How to resolve the tractions on an arbitrary plane from the known tractions on the coordinate planes as given by the Cauchy stress tensor in a chosen reference frame.

The equation relating the stress tensor to the traction on an arbitrary plane appears in the work of A. L. Cauchy around 1823, and thus bears his name in continuum mechanics. Note that Cauchy wrote close to eight hundred research papers, and many, many relationships are known as "Cauchy's formula".

4.3.3 Principal axes and principal stresses

Instead of specifying all six components of the stress tensor in the "usual" coordinate system $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$, we may opt to specify three *principal stresses* (T_1, T_2, T_3) in a new special coordinate system defined by the three *principal stress directions*, $\hat{\mathbf{n}}_1$, $\hat{\mathbf{n}}_2$, and $\hat{\mathbf{n}}_3$. What are the special planes inside of a medium, those upon which all stresses are *normal* with out *shear* components? Another way of asking the questions is: what are the special planes whose unit normals coincide (as in: aligned in the same direction but with a scaling factor to allow for a different magnitude) with the tractions acting on them. The principal stresses are the *eigenvalues* and the principal axes the *eigenvectors* of the stress tensor,

$$\mathbf{T} \cdot \mathbf{N} = \mathbf{N} \cdot \mathbf{\Lambda}, \tag{4.8}$$

or in index notation,

$$\mathbf{t}(\hat{\mathbf{n}}_i) = \mathbf{t} \cdot \hat{\mathbf{n}}_i = \lambda_i \hat{\mathbf{n}}_i. \tag{4.9}$$

This representation is not more economical, but it is certainly more convenient: it allows us to distinguish quickly and easily between different states of stress: *hydrostatic* stress ($\mathbf{T}^{\mathrm{H}}$), *uniaxial compression* and *tension* ($\mathbf{t}^{1\mathrm{D}}$), *pure shear* ($\mathbf{T}^{\mathrm{PS}}$), *deviatoric* stress tensor ($\mathbf{T}^{\mathrm{D}}$), and so on.

$$\mathbf{T}^{\mathrm{H}} = \begin{pmatrix} P & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & P \end{pmatrix} \qquad \mathbf{T}^{1\mathrm{D}} = \begin{pmatrix} \pm T & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.10}$$

$$\mathbf{T}^{\mathrm{PS}} = \begin{pmatrix} \pm T & 0 & 0 \\ 0 & \mp T & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \mathbf{T}^{\mathrm{D}} = \begin{pmatrix} \pm T_1 - P & 0 & 0 \\ 0 & T_2 - P & 0 \\ 0 & 0 & T_2 - P \end{pmatrix} \tag{4.11}$$

The mean normal stress, a *tensor invariant*,

$$P = \tfrac{1}{3}(T_1 + T_2 + T_3) \tag{4.12}$$

is sometimes *called* pressure. However, it is usually preferred to reserve the word *pressure* for the thermodynamic quantity that is only realized in a state of *hydrostatic stress*, when $\mathbf{T} = P\mathbf{I}$.

We introduce the **trace** of a vector field as the scalar

$$\mathrm{trace}(\mathbf{t}) = t_{ii}. \tag{4.13}$$

Hence, by our definitions, we have decomposed any type of stress as a *hydrostatic* stress plus a *deviatoric* stress.

4.3.4 Equilibrium conditions and equations of motion

Or: *conservation of linear momentum.*

We can be more explicit and define what it really means for a solid to be in static equilibrium. Must vanish:

(i) the resultant of body and surface forces, and
(ii) the resultant moment about any axis.

We have looked at the second condition before, and it led to the symmetry of the stress tensor (4.4). Detailing the first condition is very similar to the setup with Cauchy's tetrahedron, but for an infinitesimal cube this time. See Figure 4.1 again.

Newton's second law (4.1) gives us an *equation of motion*; setting $\mathbf{a} = \mathbf{0}$ gives us an *equilibrium condition*:

$$\frac{\partial T_{ji}}{\partial x_j} + \rho f_i = 0 \quad , \quad i = 1, ..., 3. \tag{4.14}$$

The very presence of body forces implies that there is heterogeneity of stress in a body in equilibrium. The equation(s) of motion are more general and known as the *linear momentum conservation law(s)*:.

$$\frac{\partial T_{ji}}{\partial x_j} + \rho f_i = \rho \frac{du_i}{dt} \quad , \quad i = 1, ..., 3. \tag{4.15}$$

The stress tensor $\mathbf{t}$ collects the surface tractions; $\mathbf{f}$ is a body force per unit mass; ρ the mass density; $d\mathbf{u}/dt$ the acceleration. All physical quantities ($\mathbf{f}$, ρ, $\mathbf{t}$ and $\mathbf{u}$) are functions of specific points $\mathbf{r}$ in space—the *Eulerian* viewpoint. However, the derivation assumed we were dealing with a specific set of particles whose positions vary with time—the *Lagrangian* viewpoint.

The spatial derivatives are with respect to the location in the *deformed* medium, and d/dt represents a temporal rate of change at those evolving locations. The latter is thus a *material* derivative, in that it describes what happens to the physical qualities of a parcel of matter while journeying through space. In Section 4.6, we compare and contrast both viewpoints.

4.4 The divergence (of a tensor field)

We rewrite eq. (**??**) as a single vector equation and thereby define the divergence of a tensor field:

$\boldsymbol{\nabla} \cdot \mathbf{T}$

$$\boxed{\boldsymbol{\nabla} \cdot \mathbf{T} + \rho\, \mathbf{f} = \rho\, \frac{d\mathbf{u}}{dt}.} \tag{4.16}$$

Eq. (4.16) is known as Cauchy's law of motion. Once again, all variables are Eulerian and the derivative is material. This is exact (not linearized, stronger than the previous derivation would have had us believe) and applicable to a non-rotating earth. The body force could be gravity, but we'll ignore it in what's next.

Requote the divergence theorem (2.46) but now for a tensor field, namely:

$$\int_V \boldsymbol{\nabla} \cdot \mathbf{T}\, dV = \int_{\partial V} \hat{\mathbf{n}} \cdot \mathbf{T}\, d\Sigma. \tag{4.17}$$

Let us restate Newton's second law (DT. 2.59) for what it is, the conservation of (linear) momentum. In a comoving volume (which we neglect for now), the change in total momentum is the sum of all the surface tractions on the volume, which we know from eq. (**??**) to be given by the first term, and the sum of all of the body forces, which we express as the second term in

$$\frac{d}{dt}\int_V \rho\,\mathbf{u}\,dV = \int_{\partial V} \hat{\mathbf{n}}\cdot\mathbf{T}\,d\Sigma + \int_V \rho\,\mathbf{f}\,dV. \tag{4.18}$$

All we have to do next is use eq. (4.17) to express all the integrals in terms of the volume, and then consider that our (moving) volume wasn't special at all, once again directly implying eq. (4.16). To switch the derivatives inside the integral I used *Reynold's transport theorem* and the *continuity equation* without being very upfront about it. Maybe we should discuss this theorem in more detail. The bottom line is that after we have done that, the derivative is *material*.

4.5 The gradient (of a vector field)

We're going to need this before going on. Perhaps in the context of the multi-variable chain rule? The **gradient** of a vector field $\mathbf{u}$ is a tensor whose elements are given by $\boldsymbol{\nabla}\mathbf{u}$

$$(\boldsymbol{\nabla}\mathbf{u})_{ij} = \frac{\partial u_j}{\partial x_i}, \tag{4.19}$$

which, in three Cartesian dimensions amounts to a quantity that can be written in the convenient matrix form

$$\boldsymbol{\nabla}\mathbf{u} = \begin{pmatrix} \frac{\partial u_x}{\partial x} & \frac{\partial u_y}{\partial x} & \frac{\partial u_z}{\partial x} \\ \frac{\partial u_x}{\partial y} & \frac{\partial u_y}{\partial y} & \frac{\partial u_z}{\partial y} \\ \frac{\partial u_x}{\partial z} & \frac{\partial u_y}{\partial z} & \frac{\partial u_z}{\partial z} \end{pmatrix}. \tag{4.20}$$

As before, the gradient operator increases the rank of its argument by one: the gradient of a *vector* field (rank 1) becomes a *tensor* (rank 2). As it turns out, the gradient of a vector field is the dyadic product of the "gradient vector" of eq. (2.8) and its vector argument, $\mathbf{u}$:

$$\boldsymbol{\nabla}\mathbf{u} = \partial_i u_j\,\hat{\mathbf{x}}_i\hat{\mathbf{x}}_j. \tag{4.21}$$

Note that the divergence, e.g., eq (2.44), is the **trace** of the gradient.

If we kept going we'd define the gradient of a tensor field before its trace, the divergence of a tensor field, but we won't need it in here.

4.6 Eulerian and Lagrangian viewpoints

Let $\mathbf{r}$ describe a *position* inside a continuous body. As it deforms, a point initially at $\mathbf{r}_0$ in the undeformed medium, ends up in different places over time: $\mathbf{r} = \mathbf{r}(\mathbf{r}_0, t)$. Thus, $\mathbf{r}_0 = \mathbf{r}(\mathbf{r}_0, 0)$ identifies a *parcel* of matter, a *particle*, inside the undeformed medium. A *Lagrangian* description involves tracking **particles**; the *Eulerian* viewpoint, which we will continue to prefer unless otherwise noted, describes properties at fixed spatial **positions**. The ("Eulerian") *velocity* $\mathbf{u}(\mathbf{r}, t)$ specifies the velocity of whatever particle occupies the spatial position $\mathbf{r}$ at a certain time t. It is thus given by

$$\mathbf{u}(\mathbf{r}, t) = \frac{\partial \mathbf{r}}{\partial t}. \tag{4.22}$$

To compute the *acceleration* of a certain particle $\mathbf{r}_0$ we need to take the *partial* derivative of the *Lagrangian* velocity $\mathbf{u}_L(\mathbf{r}_0, t)$, the latter being equal to the *Eulerian* velocity $\mathbf{u}(\mathbf{r}(\mathbf{r}_0, t), t)$. Using the multivariable chain rule we identify a *total*, *substantial* or *material* derivative of the Eulerian velocity $\mathbf{u}(\mathbf{r}, t)$, according to which

$$\frac{d\mathbf{u}(\mathbf{r}, t)}{dt} = \frac{\partial^2 \mathbf{r}}{\partial t^2} + \mathbf{u}(\mathbf{r}, t) \cdot \boldsymbol{\nabla} \mathbf{u}(\mathbf{r}, t) \frac{\partial \mathbf{u}_L(\mathbf{r}_0, t)}{\partial t}. \tag{4.23}$$

You can by writing a "total derivative" in terms of $\mathbf{r}$ and t and then divide by dt. The whole thing is a dot product if you think about it—a four-dimensional derivative using the four-dimensional gradient which you simply split into two terms. In very general terms, changes in a Eulerian physical quantity q can be expressed in terms of its Lagrangian description q_L and its Eulerian velocity $\mathbf{u}$ by the relation

$$dq/dt = \partial_t q + \mathbf{u} \cdot \boldsymbol{\nabla} q = \partial_t q_L. \tag{4.24}$$

The first term is the first partial derivative of the property with respect to time; the second term is an *advective* term: physical quantities that refer to moving particles can change in time because the fields in which they move change in time, but also, and even in the absence of such changes, because the particles move through spatial gradients of those fields.

For infinitesimal deformation, we can drop the distinction between Eulerian and Lagrangian viewpoints, and confuse

$$d/dt \leftrightarrow \partial_t \tag{4.25}$$

It is important to note that *for infinitesimal deformation* means: to first order; small velocities and small displacements away from the undeformed state in which the initial values of velocity and displacement are zero. We will not in general be able to interchange partial and total derivatives for quantities such as density, gravity and stress, which have zeroth-order initial values and non-negligible initial spatial gradients.

4.7 Infinitesimal deformation and linearized theory

Displacements in seismology, as opposed to tectonics and geology, are small. Hence we shall be using *linearized* theory. The approach we take is to just rewrite the equation (4.16) not in terms of density ρ, stress $\mathbf{T}$ and velocity $\mathbf{u}$, but in terms of small *perturbations* from their reference state. We drop the body force and neglect rotation for the moment. $\boldsymbol{\tau}$

For the stress perturbation, we write $\boldsymbol{\tau} = \mathbf{T} - \mathbf{T}_0$; for the infinitesimal displacement, we write $\mathbf{s} = \mathbf{r} - \mathbf{r}_0$; and the unperturbed density is ρ_0. Neglecting gravity, from eq. (4.16), the Earth at rest satisfies $\boldsymbol{\nabla} \cdot \mathbf{T}_0 = 0$. $\mathbf{s}$ ρ_0 $\mathbf{T}_0$

Substituting the perturbation equations for stress and density into eq. (4.16), using the small displacement, and subtracting the equations for the Earth at rest, we then get to solve: $\boldsymbol{\nabla} \cdot \boldsymbol{\tau}$

$$\boxed{\boldsymbol{\nabla} \cdot \boldsymbol{\tau} = \rho_0 \, \frac{\partial^2 \mathbf{s}}{\partial t^2},} \tag{4.26}$$

which is the *linearized wave equation* valid in a non-rotating, non-gravitating Earth. Euler and Lagrange are now confused for good. Now the derivatives are evaluated in the undeformed Earth... Three equations, six stress unknowns.

Write in terms of pressure and deviatoric stress? And then lead on to the acoustic wave equation? See "Acoustic Theory".

Before solving eq. (4.26), we need a proper definition strain so we can get ready for the constitutive equations.

4.8 Displacement, strain and rotation

Consider two points at initial positions $\mathbf{r}_0$ and $\mathbf{r}_0'$ and their separation vector $\mathrm{d}\mathbf{r}_0 = \mathbf{r}_0' - \mathbf{r}_0$. Some time later, they end up at $\mathbf{r} = \mathbf{r}(\mathbf{r}_0, t)$ and $\mathbf{r}' = \mathbf{r}(\mathbf{r}_0', t)$, respectively, and now the vector joining them is $\mathrm{d}\mathbf{r} = \mathbf{r}' - \mathbf{r}$. We adopt a Lagrangian viewpoint (but don't make a big deal out of it since we've already confused the two scenarios) in writing that the first particle is experiencing the small displacement

$$\mathbf{s}(\mathbf{r}_0, t) = \mathbf{r} - \mathbf{r}_0 = \mathbf{s}. \tag{4.27}$$

and the second particle experiences the small displacement

$$\mathbf{s}(\mathbf{r}_0', t) = \mathbf{r}' - \mathbf{r}_0' = \mathbf{s}'. \tag{4.28}$$

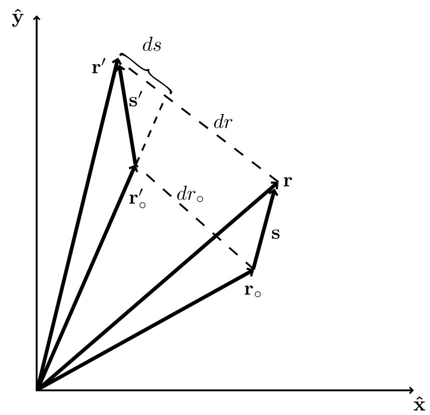


Fig. 4.3. Two diagrams. On the left, two points $\mathbf{r}_0$ and $\mathbf{r}'_0$, and their separation vector $\mathrm{d}\mathbf{r}_0 = \mathbf{r}'_0 - \mathbf{r}_0$. After a while, the new points are $\mathbf{r}$ and $\mathbf{r}'$, and $\mathrm{d}\mathbf{r} = \mathbf{r}' - \mathbf{r}$. In a different color, $\mathbf{s}$ and $\mathbf{s}'$. On the right, we show those same guys in gray but with now the moved vectors to illustrate the relations that exist between $\mathrm{d}\mathbf{s}$ and $\mathrm{d}\mathbf{r}$.

4.8.1 The displacement-gradient and deformation tensors

We are interested in how $\mathrm{d}\mathbf{r}$ relates to $\mathrm{d}\mathbf{r}_0$, or, alternatively, how the initial separation $\mathrm{d}\mathbf{r}_0$ maps into a difference in the respective displacements of the two points, $\mathrm{d}\mathbf{s} = \mathbf{s}' - \mathbf{s}$. We define $\mathbf{r}_0 = (x_1, x_2, x_3)$ and $\mathrm{d}\mathbf{r}_0 = (dx_1, dx_2, dx_3)$, or indeed $\mathbf{r}_0 = x_i\hat{\mathbf{x}}_i$ and $\mathrm{d}\mathbf{r}_0 = dx_i\hat{\mathbf{x}}_i$, summation convention implied.

As to the latter relation, we write, to first order, that individual components

$$s_i(\mathbf{r}'_0, t) = s_i(\mathbf{r}_0, t) + \frac{\partial s_i}{\partial x_j} dx_j, \tag{4.29}$$

$\boldsymbol{\nabla}\mathbf{s}$

which, in vector form, amounts to

$$\mathrm{d}\mathbf{s} = (\boldsymbol{\nabla}\mathbf{s})^{\mathrm{T}} \cdot \mathrm{d}\mathbf{r}_0 = \mathbf{J} \cdot \mathrm{d}\mathbf{r}_0. \tag{4.30}$$

On the other hand,

$\mathbf{J}$

$$\mathrm{d}\mathbf{r} = \mathrm{d}\mathbf{r}_0 + \mathrm{d}\mathbf{s} = [\mathbf{I} + (\boldsymbol{\nabla}\mathbf{s})^{\mathrm{T}}] \cdot \mathrm{d}\mathbf{r}_0 = \mathbf{F} \cdot \mathrm{d}\mathbf{r}_0. \tag{4.31}$$

Thus, the *displacement-gradient tensor*,

$$\mathbf{J} = (\boldsymbol{\nabla}\mathbf{s})^{\mathrm{T}}, \tag{4.32}$$

$\mathbf{F}$

relates the change in displacement experienced by two particles to their initial separation, and the *deformation tensor*,

$$\mathbf{F} = \mathbf{I} + (\boldsymbol{\nabla}\mathbf{s})^{\mathrm{T}} = \mathbf{I} + \mathbf{J} \tag{4.33}$$

describes how the separation between two particles evolves over time.

As with all tensors, we decompose $\mathbf{J}$ into a symmetric and an antisymmetric part: $\boldsymbol{\varepsilon}$

$$\mathbf{J} = \tfrac{1}{2}\left[(\boldsymbol{\nabla}\mathbf{s})^{\mathrm{T}} + \boldsymbol{\nabla}\mathbf{s}\right] + \tfrac{1}{2}\left[(\boldsymbol{\nabla}\mathbf{s})^{\mathrm{T}} - \boldsymbol{\nabla}\mathbf{s}\right] \tag{4.34}$$

$$= \boldsymbol{\varepsilon} + \boldsymbol{\omega}, \tag{4.35}$$

where we note that, by design, $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^T$ and $\boldsymbol{\omega} = -\boldsymbol{\omega}^T$. See the next two subsections for formal definitions of those two quantities. $\boldsymbol{\omega}$

4.8.2 The infinitesimal strain tensor

We have defined

$$\boldsymbol{\varepsilon} = \tfrac{1}{2}\left[\boldsymbol{\nabla}\mathbf{s} + (\boldsymbol{\nabla}\mathbf{s})^{\mathrm{T}}\right]. \tag{4.36}$$

Example of an infinitesimal cube pegged at the $(0, 0)$ corner being deformed with

$$\boldsymbol{\varepsilon} = \begin{pmatrix} 0 & \theta \\ \theta & 0 \end{pmatrix}, \qquad \boldsymbol{\omega} = \mathbf{0}. \tag{4.37}$$

$\mathbf{d}$ Infinitesimal shear strain. No volume change. Equivoluminal. $\boldsymbol{\nabla}\cdot\mathbf{s} = 0$.

Define the *deviatoric strain*:

$$\mathbf{d} = \boldsymbol{\varepsilon} - \tfrac{1}{3}(\mathrm{tr}\,\boldsymbol{\varepsilon})\mathbf{I} = \boldsymbol{\varepsilon} - \tfrac{1}{3}(\boldsymbol{\nabla}\cdot\mathbf{s})\mathbf{I} \tag{4.38}$$

4.8.3 The infinitesimal rotation tensor

We now have

$$\boldsymbol{\omega} = -\tfrac{1}{2}\left[\boldsymbol{\nabla}\mathbf{s} - (\boldsymbol{\nabla}\mathbf{s})^{\mathrm{T}}\right]. \tag{4.39}$$

Example of a cube being deformed with

$$\boldsymbol{\omega} = \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}, \qquad \boldsymbol{\varepsilon} = \mathbf{0}. \tag{4.40}$$

Infinitesimal rotation over clockwise θ. No volume change. $\boldsymbol{\nabla}\cdot\mathbf{s} = 0$. Let us say

$$\mathbf{s} = (\theta y, -\theta x) \tag{4.41}$$

and write down $\mathbf{J}$ and $\boldsymbol{\omega}$ and $\boldsymbol{\nabla}\times\mathbf{s} = -2\theta\hat{\mathbf{z}}$ is the rotation vector!

Write a little arc diagram with triangle with sides dx and dx and hypothenusa $-\theta dx$ and then $\tan(-\theta) \approx -\theta dx/dx$ and indeed, as $\theta \to 0$ we have $\sin(-\theta)/\cos(-\theta) = -\theta$.

Thus the rotation or curl returns a vector whose magntitude is the angular velocity and whose direction follows the movemement of a xcrew tightened by the rotation, clockwise or anticlockwise.

4.8.4 Cubic dilation and volume change

Somehow this comes out very easily. Let us think of $\mathbf{F}$ in the expression

$$\mathrm{d}\mathbf{r} = \mathbf{F} \cdot \mathrm{d}\mathbf{r}_0. \tag{4.42}$$

as the *Jacobian* of the transformation from the undeformed to the deformed coordinate system $\mathbf{r}_0 \rightarrow \mathbf{r}$, and from this conclude that the deformed and undeformed volume elements dV and dV_0 are related by

$$dV = \det(\mathbf{F})\, dV_0, \tag{4.43}$$

for which, to first order

$$\det(\mathbf{F}) \approx 1 + \boldsymbol{\nabla} \cdot \mathbf{s} = 1 + \mathrm{tr}(\boldsymbol{\varepsilon}) \tag{4.44}$$

represents the relative change in volume. This is easily verified heuristically by considering the infinitesimal cubic volume

$$dV_0 = dx_1 dx_2 dx_3 \tag{4.45}$$

deformed using eq. (4.42) to the new volume

$$dV = dx_1 \left(1 + \frac{\partial s_1}{\partial x_1}\right) dx_2 \left(1 + \frac{\partial s_2}{\partial x_2}\right) dx_3 \left(1 + \frac{\partial s_3}{\partial x_3}\right), \tag{4.46}$$

$$\approx dx_1 dx_2 dx_3 \left(1 + \frac{\partial s_1}{\partial x_1} + \frac{\partial s_2}{\partial x_2} + \frac{\partial s_3}{\partial x_3}\right), \tag{4.47}$$

$$d \approx V_0\,(1 + \boldsymbol{\nabla} \cdot \mathbf{s}). \tag{4.48}$$

to first order in the displacement. The quantity

$$\boldsymbol{\nabla} \cdot \mathbf{s} = (dV - dV_0)/dV_0 \tag{4.49}$$

$\boldsymbol{\nabla} \cdot \mathrm{s}$ is called the *cubic dilation*, or indeed the relative volume change per unit volume.

Example:

$$\mathbf{s} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} \tag{4.50}$$

$$\mathbf{s} = -x\hat{\mathbf{x}} + y\hat{\mathbf{y}} \tag{4.51}$$

No volume change.

4.8.5 Rotation and vorticity

$\boldsymbol{\nabla} \times \mathbf{s}$

From the definition of the wedge... in Cartesian coordinates, in components... and the antisymmetry... the rotation rate

$$\boldsymbol{\nabla} \times \mathbf{s} = -\epsilon_{ijk}\omega_{jk}\,\hat{\mathbf{x}}_i \tag{4.52}$$

Now the summation is implicit as compared to eq. (3.26).

Example:

$$\mathbf{s} = -y\omega\hat{\mathbf{x}} + x\omega\hat{\mathbf{y}} \tag{4.53}$$

No volume change.

Wrap-up here?

$$\boldsymbol{\tau} = \boldsymbol{\tau}^0 + \boldsymbol{\tau}^D \qquad \text{and} \qquad \boldsymbol{\tau}^D = \boldsymbol{\tau} - p\mathbf{I} \tag{4.54}$$

$$\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^D \qquad \text{and} \qquad \boldsymbol{\epsilon}^D = \boldsymbol{\epsilon} - (\Delta/3)\mathbf{I}. \tag{4.55}$$

Note that $\mathrm{tr}(\boldsymbol{\tau}^D) = 0$ and $\mathrm{tr}(\boldsymbol{\epsilon}^D) = 0$. Note that $\boldsymbol{\epsilon}^0$ is that part of $\boldsymbol{\epsilon}$ that results in volume changes. Note that $\boldsymbol{\epsilon}^D$ is that part of $\boldsymbol{\epsilon}$ that results in shape changes. Note that $\epsilon_{ij}^D = \epsilon_{ij}$ for $i \neq j$. Shear components of total strain and the strain deviator are identical. Off-diagonal terms relate to angular deformations. If $\boldsymbol{\epsilon}^D = \mathbf{0}$ then purely volumetric deformation. If $\boldsymbol{\epsilon}^0 = \mathbf{0}$ then purefly shape-changing deformation, but note that if $\boldsymbol{\epsilon} = \mathrm{diag}a, b, c$ if $a \neq b$, $a \neq c$ and $b \neq c$ then there is a shape change! Thus $\boldsymbol{\nabla} \cdot \mathbf{s}$ does not preclude shape changes, but requires null volume change.

Talk again about principal strains and principal strain directions. Normal strains, shear strains.

4.9 Hooke's law of linear elasticity

Ut tensio sic vis. Symmetry of stress, of strain, of thermodynamical Maxwell relations and strain energy function. $\mathbf{C} : \varepsilon$

Hooke's law relates stress and strain as follows:

$$\boldsymbol{\tau} = \mathbf{C} : \varepsilon \quad \text{or} \quad \tau_{ij} = C_{ijkl}\varepsilon_{kl} \tag{4.56}$$

Many, many elastic constants. Symmetries:

$$C_{ijkl} = C_{jikl} = C_{ijlk} = C_{klij} \tag{4.57}$$

Reduce all the way down to isotropy.

λ, μ

4.9.1 Isotropy

Now, in an isotropic medium, it turns out there is only one fourth-order rank tensor with all the required symmetries, and that one depends on two free parameters, λ and μ:

$$C_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}). \tag{4.58}$$

These are called Lamé's parameters, one of which, μ, is the **shear modulus**. The infinitesimal stress tensor in such a medium takes the form:

$$\tau_{ij} = \lambda\delta_{ij}\delta_{kl}\epsilon_{kl} + \mu(\delta_{ik}\delta_{jl}\epsilon_{kl} + \delta_{il}\delta_{jk}\epsilon_{kl}), \tag{4.59}$$

written in tensor notation

$$\boldsymbol{\tau} = \lambda(\boldsymbol{\nabla}\cdot\mathbf{s})\mathbf{I} + 2\mu\boldsymbol{\epsilon}. \tag{4.60}$$

Give an example. τ_{12} leads to $\epsilon_{12} = \tau_{12}/2/\mu$. Zero μ for a liquid. Using eq. (4.38) yields an alternative expression

$$\boldsymbol{\tau} = \kappa(\boldsymbol{\nabla}\cdot\mathbf{s})\mathbf{I} + 2\mu\mathbf{d} \tag{4.61}$$

whereby now we have introduced κ, the **bulk modulus**,

$$\kappa = \lambda + \frac{2}{3}\mu. \tag{4.62}$$

And then, $d\boldsymbol{\tau} = -dP\mathbf{I}$, using eq. (4.61) directly explains its alternative name,

κ the **incompressibility**,

$$\kappa = \frac{-dP}{d(\boldsymbol{\nabla}\cdot\mathbf{s})}. \tag{4.63}$$

Finally then we rewrite the alternative expression to eq. (4.58):

$$C_{ijkl} = (\kappa - \tfrac{2}{3}\mu)\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{4.64}$$

Write in tensor notation for fun...

This equation by itself shows that the principal stress axes and principal strain axes coincide in an isotropic medium! Proof. Let $\mathbf{n}^{(1)}$ be an eigenvector of $\boldsymbol{\epsilon}$ with eigenvalue e_1. Then

$$\boldsymbol{\epsilon}\cdot\mathbf{n}^{(1)} = e_1\mathbf{n}^{(1)} \tag{4.65}$$

$$\boldsymbol{\tau}\cdot\mathbf{n}^{(1)} = (\lambda(\boldsymbol{\nabla}\cdot\mathbf{s})\mathbf{I} + 2\mu\varepsilon)\cdot\mathbf{n}^{(1)} \tag{4.66}$$

$$= \lambda(\boldsymbol{\nabla}\cdot\mathbf{s})\mathbf{n}^{(1)} + 2\mu\varepsilon\cdot\mathbf{n}^{(1)} \tag{4.67}$$

$$= \left[\lambda(\boldsymbol{\nabla}\cdot\mathbf{s}) + 2\mu e_1\right]\mathbf{n}^{(1)} \tag{4.68}$$

Now, and because that's true let us write a relation between the eigenvalues

$$\begin{pmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \begin{pmatrix} \lambda+2\mu & \lambda & \lambda \\ \lambda & \lambda+2\mu & \lambda \\ \lambda & \lambda & \lambda+2\mu \end{pmatrix} \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix} \tag{4.69}$$

Shear, torsion, equivoluminal.

Things we might know from rock mechanics. Uniaxial, $\tau_1 \neq 0$ but $\tau_2 = \tau_3 = 0$. Then $e_2 = e_3 = \frac{-\lambda}{2(\lambda+\mu)} e_1$ which defines the Poisson's ratio aka DO NOT USE α or β yet ν

$$\nu = \frac{\alpha^2 - 2\beta^2}{2(\alpha^2 - \beta^2)}. \tag{4.70}$$

Can be negative "anti-rubber" but must be larger than negative 1. And E

$$E = \frac{\tau_1}{e_1} = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu} \tag{4.71}$$

defines the Young's modulus.

4.9.2 Anisotropy

Maybe just mention Love's numbers?

4.9.3 Anelasticity

Maybe just mention some key points. Get it right! Karato papers?

4.10 The wave equation in a homogeneous medium

Let us rewrite eq. (4.26) in a homogeneous medium. SNREI. The divergence of such a stress tensor is given by:

$$\boldsymbol{\nabla} \cdot \boldsymbol{\tau} = (\lambda + \mu)\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{s}) + \mu \nabla^2 \mathbf{s} \tag{4.72}$$

and thus the wave equation (4.26) becomes:

$$\boxed{\rho_0 \frac{\partial^2 \mathbf{s}}{\partial t^2} = (\lambda + 2\mu)\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{s}) - \mu \boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{s})} \tag{4.73}$$

after using eq. (3.80). These are called Navier's equations (1821, 1827) given, corrected by Cauchy, in 1822. So a balance between dilatational and torsional motion! Had we done this for incompressible fluid flow, we would have obtained the "Navier-Stokes" equations.

Note that in here, we have neglected any and all gradients of the elastic properties: $\boldsymbol{\nabla}\lambda = \boldsymbol{\nabla}\mu = \mathbf{0}$. This is a huge simplification (maybe should give the Jordan forms?), but as we'll see, "layering the Earth" allows us to treat it as locally homogeneous, and later patch up the solutions. Is this ever more than mathematically useful? Yes, it's a physical approximation which holds when the wavelengths of the waves are much too small to be aware of boundary effects!

4.10.1 A wave equation for the volume change

Remember the magnetic induction equation, and how we made sense of the two terms. This is a purely formal "finger exercise" to motivate what's coming next.

We remind ourselves that for any vector field $\mathbf{u}$ and for any scalar field ϕ, the divergence of the curl and the curl of the gradient, respectively, vanish: $\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\times\mathbf{u} = 0$ and $\boldsymbol{\nabla}\times\boldsymbol{\nabla}q = \mathbf{0}$.

Take the divergence of (4.73), rewrite as:

$$\partial_t^2(\boldsymbol{\nabla}\cdot\mathbf{s}) = \alpha^2\nabla^2(\boldsymbol{\nabla}\cdot\mathbf{s}), \tag{4.74}$$

α and we get a scalar wave equation where a first *phase speed* can be identified as

$$\alpha = \sqrt{\frac{\lambda + 2\mu}{\rho_0}} = \sqrt{\frac{\kappa + \frac{4}{3}\mu}{\rho_0}}. \tag{4.75}$$

Now it's a wave equation, i.e., a hyperbolic differential equation.

4.10.2 A wave equation for the rotation

and the right hand side of (4.72), using $\boldsymbol{\nabla}\times\nabla^2 = \nabla^2\boldsymbol{\nabla}\times$, rewrite as:

$$\partial_t^2(\boldsymbol{\nabla}\times\mathbf{s}) = \beta^2\nabla^2(\boldsymbol{\nabla}\times\mathbf{s}) \tag{4.76}$$

β and we get a vector wave equation where a second phase speed

$$\beta = \sqrt{\frac{\mu}{\rho_0}}. \tag{4.77}$$

And so this, too, is a wave equation. Write $\frac{\alpha^2}{\beta^2} =$

λ can be negative! You might wonder about κ, how about $\sqrt{\frac{\kappa}{\rho}} = \alpha^2 - \frac{4}{3}\beta^2$ the bulk sound speed.

4.10.3 Solution by separation of Cartesian variables

Let's solve eqs (4.74) and (4.76). Both of them are of the same form, though one is a scalar and the other one a vector equation. If we define

$$\Phi' = \boldsymbol{\nabla} \cdot \mathbf{s} \qquad \text{and} \qquad \boldsymbol{\Psi}' = \boldsymbol{\nabla} \times \mathbf{s}, \tag{4.78}$$

we see immediately that

$$\partial_t^2 \Phi' = \alpha^2 \nabla^2 \Phi', \tag{4.79}$$

$$\partial_t^2 \boldsymbol{\Psi}' = \beta^2 \nabla^2 \boldsymbol{\Psi}'. \tag{4.80}$$

These are known as the *Helmholtz* equations. Both the volume change and the vorticity travel through the medium as waves with speeds α and β.

Do the solution here for Φ for a simple scalar wave equation. Separation of Cartesian variables. Big story is Cartesian, could have done it in spherical also. Get integral at the end. Do it ON the ball, get sums, and spherical harmonics!

$$\Phi(x, y, z, t) = X(x)Y(y)Z(z)T(t) \tag{4.81}$$

and $\nabla^2\Phi = \partial^2\Phi/\partial_x^2 + \partial^2\Phi/\partial_y^2 + \partial^2\Phi/\partial_z^2$ hence

$$XYZ\frac{d^2T}{dt^2} = \alpha^2 \left(YZT\frac{d^2X}{dx^2} + XZT\frac{d^2Y}{dy^2} + XYT\frac{d^2Z}{dz^2} \right) \tag{4.82}$$

and thus, rather

$$\frac{1}{\alpha^2 T}\frac{d^2T}{dt^2} - \frac{1}{X}\frac{d^2X}{dx^2} - \frac{1}{Y}\frac{d^2Y}{dy^2} - \frac{1}{Z}\frac{d^2Z}{dz^2} = 0 \tag{4.83}$$

So each of the terms must be a constant and hence we pick numbers

$$\frac{1}{2T}\frac{d^2T}{dt^2} = \omega^2 \qquad \Rightarrow \qquad T = e^{\pm i\omega t} \tag{4.84}$$

$$\frac{1}{X}\frac{d^2X}{dx^2} = k_x^2 \qquad \Rightarrow \qquad X = e^{\pm i k_x x} \tag{4.85}$$

$$\frac{1}{Y}\frac{d^2Y}{dy^2} = k_y^2 \qquad \Rightarrow \qquad Y = e^{\pm i k_y y} \tag{4.86}$$

$$\frac{1}{Z}\frac{d^2Z}{dz^2} = k_z^2 \qquad \Rightarrow \qquad Z = e^{\pm i k_z z} \tag{4.87}$$

with clearly

$$k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{\alpha^2}. \tag{4.88}$$

And now do a simple suite of test functions, e.g. for $T = \sin(\omega t)$, $T =$

$\cos(\omega t)$, $T = \exp(i\omega t) = \cos(\omega t) + i\sin(\omega t)$, and so on, which proves our point. And then the general solution is

$$\Phi(x, y, z, t) = e^{\pm ik_x x} e^{\pm ik_y y} e^{\pm ik_z z} e^{\pm i\omega t} \tag{4.89}$$

That's a wave! A plane wave. Signs and amplitudes to depend on boundary conditions and initial conditions. Characteristics. Then discuss the general concepts such as wavefronts, period, phase, etc. And the dispersion relation. And the fact that we need to sum over all wavevectors, so in fact we've discovered the Fourier transform.

It's like we've replaced the original equation with something like

$$\frac{d^2\mathbf{s}}{dt^2} = \alpha^2 \boldsymbol{\nabla}\Phi - \beta^2 \boldsymbol{\nabla} \times \boldsymbol{\Psi} \tag{4.90}$$

and then taken the temporal Fourier transform to get (check the signs)!

$$\mathbf{s}(\omega) = -\frac{\alpha^2}{\omega^2}\boldsymbol{\nabla}\Phi + \frac{\beta^2}{\omega^2}\boldsymbol{\nabla} \times \boldsymbol{\Psi} \tag{4.91}$$

which we then turn into what's next.

4.11 Whole-space solutions to the wave equation

As Feynman is reported to have said, "the same equations have the same solutions". There is another way at arriving at equations that look similar to eqs (4.79)–(4.80), and one which is ultimately more useful in expressing the displacement field. If we look for the displacement field in terms of scalar and vector *potentials* by the *Helmholtz decomposition theorem*:

$$\mathbf{s} = \boldsymbol{\nabla}\Phi + \boldsymbol{\nabla} \times \boldsymbol{\Psi} \quad \text{with} \quad \boldsymbol{\nabla} \cdot \boldsymbol{\Psi} = 0. \tag{4.92}$$

Now Φ and $\boldsymbol{\Psi}$ are the "inverse Laplacians" of the divergence and curl of the displacement field, respectively:

$$\nabla^2\Phi = \boldsymbol{\nabla} \cdot \mathbf{s} \qquad \text{and} \qquad \nabla^2\boldsymbol{\Psi} = -(\boldsymbol{\nabla} \times \mathbf{s}). \tag{4.93}$$

Substituting eqs (4.92)–(4.93) into eq. (4.73) we need to satisfy:

$$\boldsymbol{\nabla}\left(\partial_t^2\Phi - \alpha^2\nabla^2\Phi\right) + \boldsymbol{\nabla} \times \left(\partial_t^2\boldsymbol{\Psi} - \beta^2\nabla^2\boldsymbol{\Psi}\right) = \mathbf{0}, \tag{4.94}$$

which is possible by requiring that the terms within the brackets equal zero. This leads to equations for Φ and $\boldsymbol{\Psi}$ of the exact same form as eqs (4.79)–(4.80). solutions to those equations — by separation of variables or Fourier methods — are *waves*. We obtain that, whether the phase speed $c = \alpha$ or $c = \beta$,

$$k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{c^2}. \tag{4.95}$$

$$\Phi(\mathbf{r}, t) = \Phi_0 \exp(i[\mathbf{k}\cdot\mathbf{r} - \omega t]) \quad \text{with} \quad |\mathbf{k}| = |\omega/\alpha|, \tag{4.96}$$

$$\boldsymbol{\Psi}(\mathbf{r}, t) = \boldsymbol{\Psi_0} \exp(i[\mathbf{k}\cdot\mathbf{r} - \omega t]) \quad \text{with} \quad |\mathbf{k}| = |\omega/\beta|, \tag{4.97}$$

in other words, oscillations in space and time linked together by the *dispersion relation* which is of the form $\mathbf{k}$

$$\boxed{|\mathbf{k}| = \left|\frac{\omega}{c}\right|,} \tag{4.98}$$

with $\mathbf{k}$ the *wave vector*, $|\mathbf{k}| = k$ the *wave number*, c the *phase speed* and ω the *angular frequency* of the wave. means... that the complete solutions are all possible frequencies and all possible wave vectors in the appropriate weightings. Plane waves. Wave fronts. Propagation direction.

Why P waves travel faster than S waves (back to the bulk modulus).

Can I give an equation for the density as a function of space and time in a passing wavefield? Should I give an expression for the energy in the P versus the S wave, which is a question I was asked and which is apparent in the seismogram.

4.11.1 Dispersion and plane waves

Eqs (4.96) and (4.97) were general. The complete wavefield will be a superposition. The complete solution to the wave equation is thus given by inverse transformation of $\Phi(\mathbf{r}, \omega)$ as follows:

$$\Phi(\mathbf{r}, t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \Phi(k_x, k_y, \omega, z) e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}\, dk_x\, dk_y\, d\omega. \tag{4.99}$$

Remember the spherical harmonics! They also were superposed. But by sums, not integrations, since they are in a quantized medium.

There are three independent quantities involved here (not four): k_x, k_y and ω, and their relationship is given by the dispersion equation. In other words,

$$\mathbf{k}\cdot\mathbf{r} = k_x x + k_y y + z\left(\frac{\omega^2}{c^2} - k_x^2 - k_y^2\right)^{1/2} \tag{4.100}$$

It's important to see Eq. 4.99 as what it is: a superposition (integral) of plane waves with a certain wave vector and frequency, each with its own amplitude.

The amplitude is a coefficient which will have to be determined from the initial or boundary conditions. We will describe plane waves in more detail, more particularly their general form $\exp(i(\mathbf{k}\cdot\mathbf{r}-\omega t))$. In the following section we will show an alternative way of solving the scalar potential equations—by separation of variables.

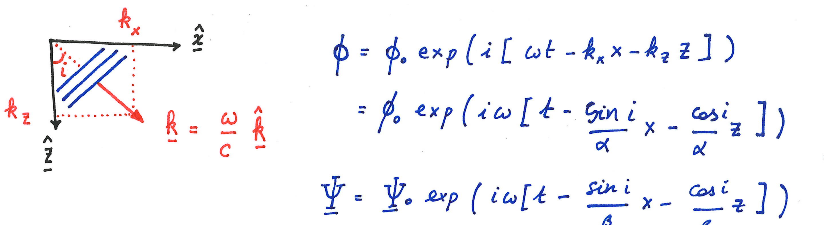


Fig. 4.4. *P* and *S* waves in an unbounded medium.

4.11.2 P and S waves

Going back to eq. (4.92), the displacement field is decomposed into

$$\mathbf{s} = \mathbf{s}_P + \mathbf{s}_S, \tag{4.101}$$

whereby the scalar *P-wave potential* Φ is of the form eq. (4.96) and the irrotational *P-wave displacement field* is given by

$$\boxed{\mathbf{s}_P = \boldsymbol{\nabla}\Phi,} \tag{4.102}$$

x and the divergence-free vector *S-wave potential* $\boldsymbol{\Psi}$ is of the form eq. (4.97) and the *S-wave displacement field* is given by

$$\boxed{\mathbf{s}_S = \boldsymbol{\nabla}\times\boldsymbol{\Psi}, \qquad \boldsymbol{\nabla}\cdot\boldsymbol{\Psi} = 0.} \tag{4.103}$$

It's only logical that *P* waves arrive first and *S* second; this is contained in the wave speeds eqs (4.75) and (4.77).

Note the similarity between eqs. (4.102), (2.7), (**??**), and (3.10): in each of the three chapters on gravity, electricity and magnetism, and now, seismology, we've managed to express an observable in terms of the gradient of a potential functions.

This is where you talk about particle motion and the relation between $\mathbf{k}$ and $\mathbf{s}$, namely $\mathbf{k}^P\cdot\mathbf{s}^P \neq 0$ but $\mathbf{k}^S\cdot\mathbf{s}^S = 0$.

Relate this directly to particle motion. Here the figure on particle motion vs wave vector. Aki Richards p 127. $\boldsymbol{\nabla}(\mathbf{k}\cdot\mathbf{r}) = \mathbf{k}$ is the direction in which the phase varies the most. $\boldsymbol{\nabla} \exp(i\omega\xi) = \exp(i\omega\xi)\boldsymbol{\nabla}\xi$. And $\boldsymbol{\nabla} \times (\boldsymbol{\Psi} \exp(i\omega\xi)) = \boldsymbol{\nabla} \exp(i\omega xi) \times \boldsymbol{\Psi} + \exp(i\omega\xi)\boldsymbol{\nabla} \times \boldsymbol{\Psi}$. And $\boldsymbol{\nabla} \times \xi\Psi = \boldsymbol{\nabla}\xi \times \boldsymbol{\Psi} + \xi\boldsymbol{\nabla} \times \boldsymbol{\Psi}$. Thus $\exp(i\omega\xi)\boldsymbol{\nabla}\xi \times \boldsymbol{\Psi} + \exp(i\omega\xi\boldsymbol{\nabla} \times \boldsymbol{\Psi})$ thus $\mathbf{k} \times \Psi$ both perpendicular.

So the Cartesian components of the complete displacement field in an infinite medium are given by (refer to the explicit equations for grad and curl)

$$s_x = \frac{\partial \Phi}{\partial x} + \left(\frac{\partial \Psi_z}{\partial y} - \frac{\partial \Psi_y}{\partial z} \right), \tag{4.104}$$

$$s_y = \frac{\partial \Phi}{\partial y} + \left(\frac{\partial \Psi_x}{\partial z} - \frac{\partial \Psi_z}{\partial x} \right), \tag{4.105}$$

$$s_z = \frac{\partial \Phi}{\partial z} + \left(\frac{\partial \Psi_y}{\partial x} - \frac{\partial \Psi_x}{\partial y} \right). \tag{4.106}$$

4.11.3 Displacements and tractions in plane-layered media

Plane-strain case.

We'll use the above to do something important in planar media. The above was for infinite media, but it's not hard to image we could solve the equations in homogeneous "layers" and determine the required multiplicative constants by patching up the solutions via the appropriate boundary conditions.

In a halfspace, plane-layered medium, write $\boldsymbol{\Psi} = \mathbf{B}\Psi$, and $k_y = 0$, nothing propagates in the y direction,

$$\mathbf{s}^P = \begin{pmatrix} -ik_x \\ 0 \\ -ik_z \end{pmatrix} \Phi \tag{4.107}$$

$$\mathbf{s}^S = \begin{pmatrix} ik_z B_y \\ -B_x ik_z + B_z ik_x \\ -iB_y k_x \end{pmatrix} \Psi \tag{4.108}$$

i.e., with $\mathbf{s} = \mathbf{s}^{SV} + \mathbf{s}^{SH}$ in the (x, z) plane and out of the (x, z) plane.

Since no variations or presence of anything occur in the y direction, we can write:

$$\mathbf{s} = \left(\frac{\partial \Phi}{\partial x} - \frac{\partial \Psi_y}{\partial z}, \frac{\partial \Psi_x}{\partial z} - \frac{\partial \Psi_z}{\partial x}, \frac{\partial \Phi}{\partial z} + \frac{\partial \Psi_y}{\partial x} \right), \tag{4.109}$$

in order to notice that the $(s_x, 0, s_z)$ or the *P-SV system* is completely decoupled from the $(0, s_y, 0)$ or the *SH system*. And then we might as well write

$$\mathbf{s} = \left(\frac{\partial \Phi}{\partial x} - \frac{\partial \Psi}{\partial z}, s_y, \frac{\partial \Phi}{\partial z} + \frac{\partial \Psi}{\partial x} \right), \tag{4.110}$$

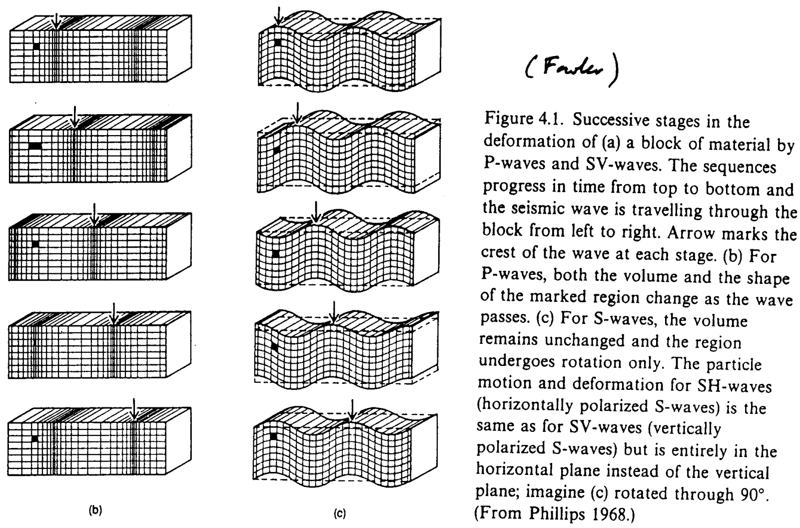


Fig. 4.5.

as if it resulted from the alternative plane-layered decomposition

$$\mathbf{s} = \boldsymbol{\nabla}\Phi + \boldsymbol{\nabla} \times (\Psi\hat{\mathbf{y}}) + s_y\hat{\mathbf{y}}, \tag{4.111}$$

where now the three *scalar* potentials Φ, s_y and Ψ each satisfy a Helmholtz equation of the type (4.79).

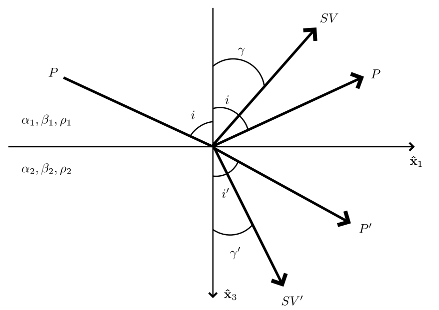


Fig. 4.6. Ray conversions in the *P-SV* system, should make sense they all couple together.

Let's consider a stress tensor of the kind (4.60) and write it in component notation as

$$\tau_{ij} = \lambda(\boldsymbol{\nabla} \cdot \mathbf{s})\delta_{ij} + \mu\left(\frac{\partial s_i}{\partial x_j} + \frac{\partial s_j}{\partial x_i}\right) \tag{4.112}$$

and let us consider the tractions on a horizontal plane (with unit normal $\hat{\mathbf{z}}$); from eq. (**??**) these would be given by

$$\mathbf{t}(\hat{\mathbf{z}}) = (\tau_{zx}, \tau_{zy}, \tau_{zz}). \tag{4.113}$$

We're here just going to derive what they look like. Later on, we'll do Snell's law for reflection and refraction etc... and then in the next chapter we'll actually put the tractions to zero at the free surface. And derive the Rayleigh and Love waves with them.

Let us now consider a medium in which three types of waves might exist: the *P*, the *SV* and the *SH* systems. Let the wave vectors be in the (x, z) plane... we can immediately position the zeroes in the following three expressions.

Consider tractions on a horizontal plane in this medium.

The tractions due to the *P* wave are:

$$\mathbf{t}_P = \left(2\mu\frac{\partial^2\Phi}{\partial x\partial z},\, 0\,, \lambda\nabla^2\Phi + 2\mu\frac{\partial^2\Phi}{\partial z^2}\right) \tag{4.114}$$

The tractions due to the *SV* wave are:

$$\mathbf{t}_{SV} = \left(\mu\left[\frac{\partial^2\Psi}{\partial x^2} - \frac{\partial^2\Psi}{\partial z^2}\right],\, 0\,, 2\mu\frac{\partial^2\Psi}{\partial x\partial z}\right) \tag{4.115}$$

The tractions due to the *SH* wave are:

$$\mathbf{t}_{SH} = \left(0\,,\, \mu\frac{\partial s_y}{\partial z}\,, 0\right) \tag{4.116}$$

4.12 Boundary conditions

Boundary conditions Various types, give a general intro also suitable for heat flow and computation? Whether bottom of the ocean is no-slip or free-slip? Dicussion in Vienna 2017 with Markus Jochum! Connect to geomagnetism by Nature NV of geodynamo simulations. Physical. Kinematic, dynamic. Mathematical. Dirichlet and Neumann and Cauchy. Computational. Absorbing, PMLStacey.

For *solid-solid* (i.e. *welded*) transitions, we have continuity of traction and continuity of displacement:

$$\mathbf{t}^+(\hat{\mathbf{n}}) = \mathbf{t}^-(\hat{\mathbf{n}}), \tag{4.117}$$

$$\mathbf{s}^+(\hat{\mathbf{n}}) = \mathbf{s}^-(\hat{\mathbf{n}}). \tag{4.118}$$

For *fluid-solid* transitions, we still have continuity of traction, but we require zero *tangential* traction, i.e. "stress-free", "free-slip" as opposed to "no-slip" commonly used for viscous fluids. In addition, we have continuity of the non-tangential displacements.

$$\mathbf{t}^{+}(\hat{\mathbf{n}}) = \mathbf{t}^{-}(\hat{\mathbf{n}}) = (\hat{\mathbf{n}} \cdot \mathbf{t})\hat{\mathbf{n}}, \tag{4.119}$$

$$\hat{\mathbf{n}} \cdot \mathbf{s}^{+} = \hat{\mathbf{n}} \cdot \mathbf{s}^{-} \tag{4.120}$$

For *solid-vacuum* (i.e. the *free surface*) transitions, we have zero traction:

$$\mathbf{t}(\hat{\mathbf{n}}) = \mathbf{0}. \tag{4.121}$$

So from here we now infer intuitively that rays get *reflected* and transmitted — but *refracted*.

4.12.1 Reflected and transmitted waves

Now more specifically, let us consider a situation in which we only have a *P* and an *SV* (and thus coupled) system... and no coupling to *SH* so leave it out of consideration. Let's say it's a case of reflection. Or transmission. So what. The point is to illustrate Snell's law, and later apply the general principle to the Earth.

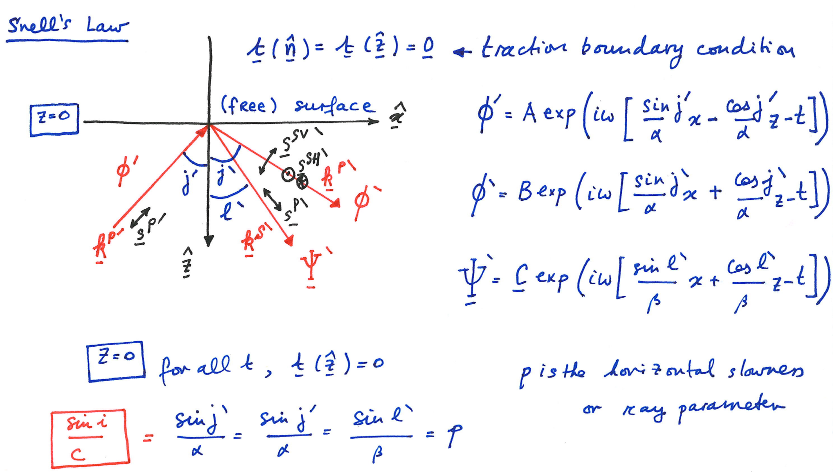


Fig. 4.7. Snell's law.

The wave vector is in the direction of propagation (must have proved this before). Schematic shows particle motion (must have shown this before). At any rate, make sense, from the dispersion relation (4.98), that the three potentials would be given by:

$$\Phi' = A \exp\left(i\omega\left[\frac{\sin j'}{\alpha}x - \frac{\cos j'}{\alpha}z - t\right]\right) \tag{4.122}$$

$$\Phi^{\backprime} = B \exp\left(i\omega\left[\frac{\sin j^{\backprime}}{\alpha}x + \frac{\cos j^{\backprime}}{\alpha}z - t\right]\right) \tag{4.123}$$

$$\Psi^{\backprime} = C \exp\left(i\omega\left[\frac{\sin l^{\backprime}}{\beta}x + \frac{\cos l^{\backprime}}{\beta}z - t\right]\right) \tag{4.124}$$

Boundary conditions? Involving first and second derivatives of those potentials. Treat the free-surface reflection. Each of the traction components is still some constant times the exponentials of above. Since the boundary conditions hold on $z = 0$ for all x and t, the phase factors must be unchanged. Whatever they are, will lead to needing

$$\frac{\sin j'}{\alpha} = \frac{\sin j^{\backprime}}{\alpha} = \frac{\sin l^{\backprime}}{\beta} \equiv p \tag{4.125}$$

Thus, for plane waves in plane-layered media, the whole system of rays is characterized by a common *horizontal slowness*. This is true for the whole wave field of reflected and transmitted (refracted) waves. Eq. (4.125) is known as *Snell's law* and p is called the *ray parameter*. Horizontal slowness is preserved upon *reflection* and also upon *conversion*. And *transmission*.

Frederiksen and Bostock put it well. By Snell's law, the component of phase slowness parallel to an interface is preserved across the interface.

It should be clear we can now, from the boundary conditions, define the ratios between the multiplicative factors B/A and C/A... thereby defining some sort of *reflection* and *transmission* coefficients that tell you how much of the potential/displacement/energy etc... is reflected/converted/transmitted. This is fairly pedestrian, but the subject of seismology proper.

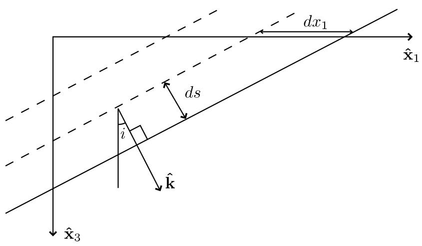


Fig. 4.8. Snell's law, need to make another point about this.

4.12.2 Snell's law via the principle of stationary time

Fermat's principle. Connect to Fresnel zone, easy way of saying it is as Yomogida.

An important principle in optics is Fermat's principle, which governs the geometry of ray paths. This principle states that a wave propagating from position A to position B follows a path of **stationary time**. The principle of stationary time plays a fundamental role in high frequency seismology. Note that stationary time does not necessarily mean *minimum* time; it can also be a maximum time.

See SIAM review article on thin lenses and the spring connections!

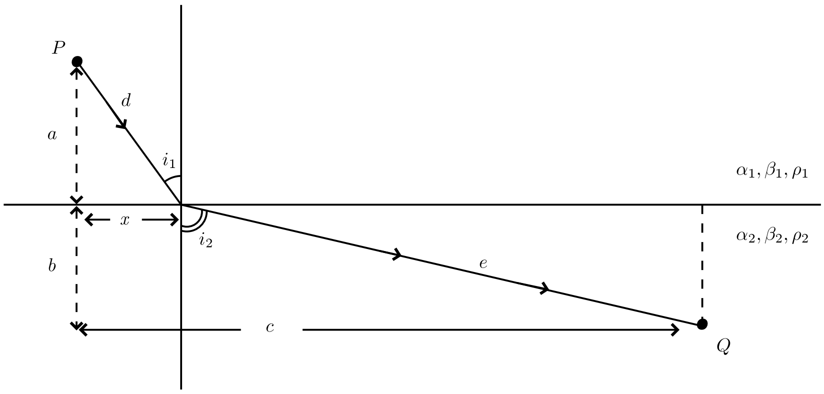


Fig. 4.9. The principle of stationary time.

Consider Fig. 4.9. A ray leaves point *P* that is in a medium with wave speed c_1 and travels to point *Q* in a medium with wave speed c_2. What path will the ray take to Q? Since the wave speeds in the media are constant the ray path in each medium is a straight line, so that in this simple case the geometry is completely defined by the positions of P, Q, and the point x where the ray crosses the interface.

The travel time on an arbitrary path between *P* and *Q* is given by

$$T_{P-Q} = \frac{d}{c_1} + \frac{e}{c_2} = \frac{\sqrt{a^2+x^2}}{c_1} + \frac{\sqrt{b^2+(c-x)^2}}{c_2} \tag{4.126}$$

For the path to be a stationary time path (i.e. time is maximum or minimum) we simply set the spatial derivative of the travel time to zero:

$$\frac{dT}{dx} = 0 = \frac{x}{c_1\sqrt{a^2+x^2}} - \frac{c-x}{c_2\sqrt{b^2+(c-x)^2}} \tag{4.127}$$

and note that

$$\frac{x}{\sqrt{a^2+x^2}} = \sin i_1 \qquad \text{and} \qquad \frac{c-x}{\sqrt{b^2+(c-x)^2}} = \sin i_2 \tag{4.128}$$

This gives Snell's law again:

$$\boxed{\frac{\sin i_2}{c_2} = \frac{\sin i_1}{c_1} \equiv p} \tag{4.129}$$

p is called the **ray parameter**.

The ray parameter p is constant along the entire system. As a ray enters material of increasing velocity, it is is deflected toward the horizontal; if it enters material with lower velocity, the ray is deflected to the vertical. The angle between the ray and the vertical is referred to as the **angle of incidence** or **take-off angle**.

Snell's law — in spherical geometry. Add in the scale factor. Bullen and Bolt is quick. Aster is decent.

4.12.3 Ray tracing

General equations. Then Crewes/Slotnick for linear media. Examples?

Let's take the standard derivation in two dimensions whereby an infinitesimal ray path segment ds in the xz plane defines the geometry

Maybe move slownesses from below up.

$dx/ds = \sin\theta = c$

$dx/dz = \cos\theta = sqrt1 - p^2c^2$

$dx/dz = \frac{p}{c^{-2}-p^2}$

which you then integrate. Maybe go straigh to the piecewise discretization.

$dt/dz = \frac{c^{-1}}{1-p^2c^2}$

which you then integrate.

Then rewrite these equations for spherical geometry.

Then make a piecewise linear approximation and you can do a whole lot of calculations a la Slotnick.

4.12.4 Propagating and evanescent waves

Consider a plane wave in two dimensions:

$$\Phi = e^{-i\omega t} e^{ik_x x + ik_z z} \tag{4.130}$$

The horizontal wavenumbers k_x and k_z reflect how fast a wave of a particular frequency travels in the direction of (given the sign, positive) x and z, respectively. For an **incidence angle** (the angle between the wave vector and the surface normal) i, we have

$$k_x = +\left|\frac{\omega}{c}\right|\sin i \tag{4.131}$$

$$k_z = -\left|\frac{\omega}{c}\right|\cos i \tag{4.132}$$

p We have defined the **ray parameter** p as:

$$p = \frac{\sin i}{c} \tag{4.133}$$

and noticed, with **Snell**, that in plane-layered media, the ray parameter is a constant throughout the plane wave system (including all reflections and transmissions). The ray parameter is related to the horizontal slowness. Let's define
η a similar quantity related to the *vertical slowness*:

$$\eta = \frac{\cos i}{c} = \sqrt{\frac{1}{c^2} - p^2}, \tag{4.134}$$

and let's rewrite the wave equation in this format:

$$\Phi = e^{-i\omega(t-px-\eta z)}. \tag{4.135}$$

In the case discussed above, p and η are real numbers and the wave is a propagating plane wave. When $\eta = 0$, the wave propagates horizontally along the surface since it has no vertical component of slowness.

Also, η could be a complex number; this happens when

$$p > \frac{1}{c} \tag{4.136}$$

In the latter case, we can represent $\eta = i\hat{\eta}$, and we get

$$\Phi = e^{-\hat{\eta}\omega z} e^{-i\omega(t-px)}. \tag{4.137}$$

In the positive $z-$direction (with increasing depth), the amplitude of the displacement dies out until it is zero at infinity: the wave is **evanescent**.

Fig **??** illustrates Snell's law for free surface interactions for a particular example. For an incident SV-wave to generate a critically reflected P wave (in situation b), the SV-wave needs to have an incidence angle given by

$$p = \frac{1}{\alpha} = \frac{\sin j_1}{\beta}. \tag{4.138}$$

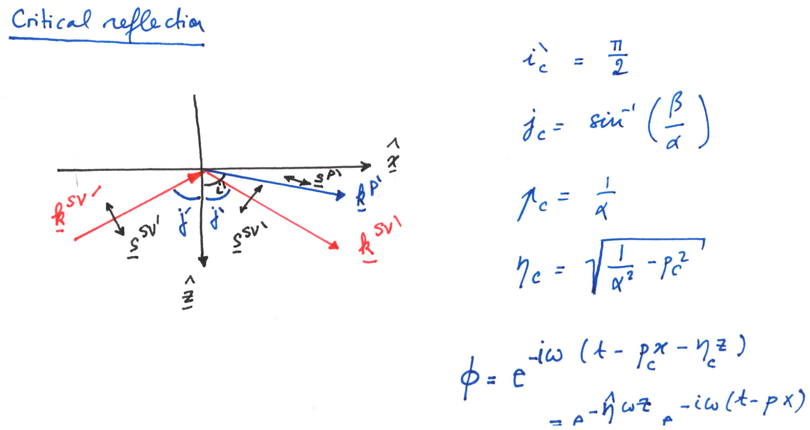


Fig. 4.10. Critical reflection.

The critical incidence angle, ray parameter and vertical slowness are given by:

$$j_c = \sin^{-1}\left(\frac{\beta}{\alpha}\right), \tag{4.139}$$

$$p_c = \frac{1}{\alpha}, \tag{4.140}$$

$$\eta_c = \sqrt{\frac{1}{\alpha^2} - p_c^2}. \tag{4.141}$$

If, as in Figure 4.11, the incoming SV wace comes in at an even steeper angle than the critical *reflection* angle, $j_1 > j_c$, its ray parameter will be bigger than the critical one $p > p_c > 1/\alpha$ and η will indeed be complex. This is how surface reflections can become evanescent. (If you like to turn this situation upside down, you can study headwaves in this manner. They, too, become evanescent for waves inciding above the critical *refraction* angle.)

It better be evanescent in an infinite halfspace! To conserve energy the amplitude of the horizontally propagating *P* wave must decrease with depth and vanish at some point, i.e., a critically refracted *P* wave is an *evanescent* wave.

4.12.5 Body-wave nomenclature; travel-time curves

Body-wave nomenclature. Classical τ-p curves and the first-order information derived from them.

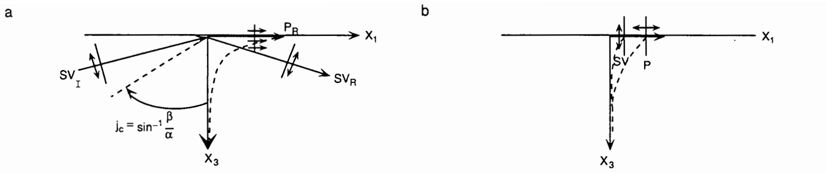


FIGURE 4.3 (a) Postcritical *SV* wave incident on a free surface gives rise to an evanescent *P* wave propagating along the boundary as well as to a phase-shifted *SV* reflection. (b) Simultaneous existence of evanescent *P*- and *SV*-wave energy traveling horizontally along a free surface produces the interference surface wave called a Rayleigh wave.

Fig. 4.11. Evanescent waves.

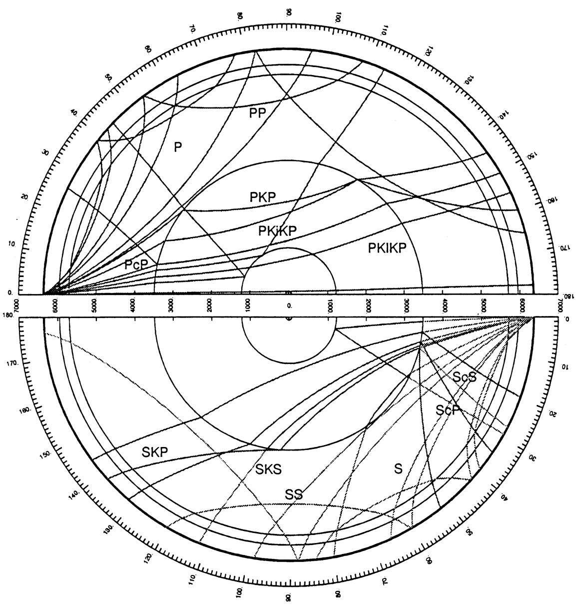


Figure 5.15. Paths of seismic rays through the Earth, illustrating their nomenclature. Figure by courtesy of B.L.N. Kennett.

Fig. 4.12.

4.13 Half-space solutions to the wave equation

Surface waves. Rayleigh and Love. First time we bring in a bounding surface, and the attendant zero-traction boundary conditions.

Rewrite the next sentence which is from Essentials: The existence of *P* and S-waves was first demonstrated by Poisson (in 1828). He also showed that *P* and *S*-type waves are, in fact, the only solutions of the wave equations for an

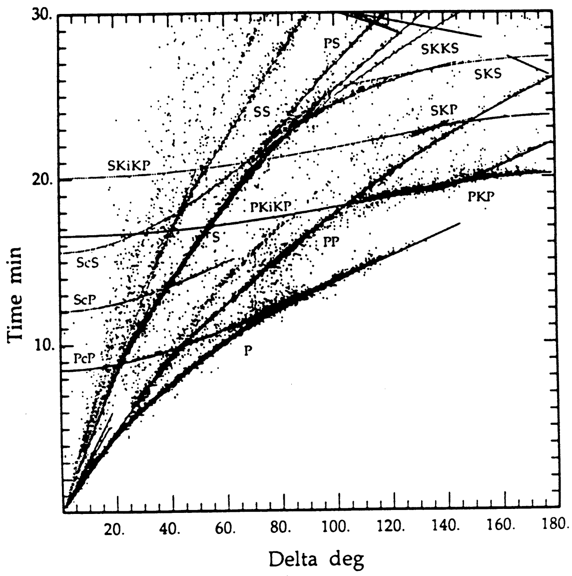


FIGURE 6.2 Six thousand travel times picked from phases of select shallow earthquakes and explosions with known or particularly well-determined locations. Superimposed on the travel times are the interpretation of the phases and the curves showing predicted arrival times based on the iasp91 Earth model. The phases are named using a convention that describes the wave's path through the Earth. For example, *PcP* is the *P* wave reflected from the Earth's core. Some of the arrivals continue to be observed beyond 180°, and they "wrap" around onto this plot. (From Kennett and Engdahl, 1991.)

Fig. 4.13.

unbounded medium (a "whole" space), so that eq. (4.101) provides the complete solution for the displacement in an elastic, isotropic and homogeneous medium. Or in the sense that spatial variations in elastic properties occur over much larger distances than the wavelength of the waves involved. If the medium is not unbounded, as a half-space with perhaps some stratification, there are more solutions to the general equation of motions. Those solutions are the surface (Rayleigh and Love) waves.

A simple treatment is by [16].

4.13.1 Rayleigh waves

As a conclusion to Lord Rayleigh's paper *On waves propagated along the plane surface of an elastic solid*, he writes: "It is not improbable that the surface waves here investigated play an important part in earthquakes, and in the

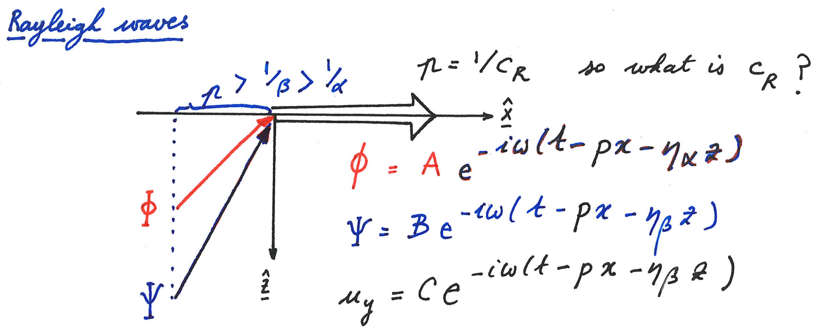


Fig. 4.14. Rayleigh waves sketch: a P and and SV wave that are *both* critically reflected at an interface

collision of elastic solids. Diverging in two dimensions only, they must acquire at a great distance from the source a continually increasing preponderance."

We turn our attention to a homogeneous **half-space**. In this half-space, we want to study the three kinds of scalar potentials we know could be present, and we know they're all plane waves with an unknown amplitude:

$$\Phi = A\, e^{-i\omega(t-px-\eta_\alpha z)}, \tag{4.142}$$

$$\Psi = B\, e^{-i\omega(t-px-\eta_\beta z)}, \tag{4.143}$$

$$u_y = C\, e^{-i\omega(t-px-\eta_\beta z)}. \tag{4.144}$$

We understand that the vertical wavenumbers are given by

$$\eta_\alpha = \sqrt{\frac{1}{\alpha^2} - p^2} \qquad \text{and} \qquad \eta_\beta = \sqrt{\frac{1}{\beta^2} - p^2}. \tag{4.145}$$

The tractions on the free surface vanish, hence

$$\mathbf{t}(\hat{\mathbf{z}}) = (\tau_{xz}, \tau_{yz}, \tau_{zz}) = 0. \tag{4.146}$$

This leads to the following three conditions (after some rearranging):

$$\boxed{\tau_{yz} = 0} \Rightarrow 0 = \mu C e^{-i\omega(t-px-\eta_\beta z)}(i\omega\eta_\beta) \tag{4.147}$$

$$\boxed{\tau_{zz} = 0} \Rightarrow 0 = A\left[(\lambda+2\mu)\eta_\alpha^2 + \lambda p^2\right] + B\left[2\mu\eta_\beta p\right] \tag{4.148}$$

$$\boxed{\tau_{xz} = 0} \Rightarrow 0 = A\left[2p\eta_\alpha\right] + B\left[p^2 - \eta_\beta^2\right]. \tag{4.149}$$

The first condition immediately dictates that $C = 0$. In other words, in a homogeneous halfspace no component of displacement can be perpendicular to the plane of propagation. SH waves in a halfspace do not exist and we shouldn't have bothered with them in the first place. The $P - SV$ system is what counts here. This leaves the last two conditions.

Given a particular system of plane waves in a homogeneous halfspace with compressional wave speed α and shear wave speed β, and a particular ray parameter p of an incoming SV-wave, one can calculate the amplitudes of the reflected SV and the converted reflected P-wave. You could start from an upcoming SV-wave and calculate how its kinetic energy is fractionated into a downgoing SV and a downgoing P-wave and hence determine the reflection coefficient in function of the incidence angle, etcetera.

Indeed, for subcritical incidence angles, when $p < 1/\alpha < 1/\beta$, this is nothing but a simple system of reflected waves.

But when

$$p > \frac{1}{\beta} > \frac{1}{\alpha}, \tag{4.150}$$

and both η_α and η_β are imaginary:

$$\eta_\alpha = i\sqrt{p^2 - \frac{1}{\alpha^2}} \quad \text{and} \quad \eta_\beta = i\sqrt{p^2 - \frac{1}{\beta^2}}. \tag{4.151}$$

we will get a resulting wave which is evanescent: the **Rayleigh** wave, which is formed by the interaction of P and SV at the free surface. Because it propagates horizontally, its phase speed will be given by $c_R = 1/p$.

What is the phase speed of the Rayleigh wave? In Rayleigh's paper, the remaining two boundary conditions are recognized as a matrix equation for the unknown amplitudes A and B, and the discriminant of the coefficient matrix is used to derive an explicit equation for c/β. For a Poisson solid, $\lambda = \mu$ and $\alpha^2 = 3\beta^2$, the ratio of the speed of the Rayleigh wave to the shear wave speed of the half-space is about 0.92.

1) Amplitudes evanescent. 2) speed is cR=1/p. 3) cR=rβ. Then go.

I've found a much easier way to show this for the Poisson solid. Let's indeed look for the ratio

$$r = \frac{c}{\beta}. \tag{4.152}$$

Going back to the boundary conditions, we can derive the ratio of the amplitudes of the P and the SV wave which interfere to give the Rayleigh wave.

By subsituting the definitions for $p = 1/c_R$, $c_R = r\beta$ and the complex valued functions η_α and η_β into eq. 4.148, that ratio is:

$$\frac{A}{B} = -\frac{2\mu\eta_\beta p}{(\lambda + 2\mu)\eta_\alpha^2 + \lambda p^2} = +\frac{1}{2i}\frac{r^2 - 2}{\sqrt{1 - r^2/3}}. \tag{4.153}$$

Similarly, we can use eq. 4.149 and write

$$\frac{A}{B} = -\frac{p^2 - \eta_\beta^2}{2p\eta_\alpha} = -\frac{1}{i}\frac{\sqrt{1-r^2}}{(1-r^2/2)}. \tag{4.154}$$

Note that as always $\lambda = (\alpha^2 - 2\beta^2)$ and $\eta_\alpha^2 = \frac{1}{\alpha^2} - p^2$ and $\eta_\beta^2 = \frac{1}{\beta^2} - p^2$, we can rewrite

$$\frac{(\lambda + 2\mu)}{\mu}\eta_\alpha^2 + \frac{\lambda}{\mu}p^2 = \frac{\rho\alpha^2}{\mu}\eta_\alpha^2 + \frac{\rho}{\mu}(\alpha^2 - 2\beta^2)p^2 \tag{4.155}$$

$$= \frac{\rho\alpha^2}{\mu}\left(\eta_\alpha^2 + p^2\right) - 2\frac{\rho}{\mu}\beta^2 p^2 \tag{4.156}$$

$$= \frac{\rho\alpha^2}{\mu}\left(\eta_\alpha^2 + p^2\right) - 2p^2 \tag{4.157}$$

$$= \frac{\rho}{\mu} - 2p^2 \tag{4.158}$$

$$= \frac{1}{\beta^2} - 2p^2 \tag{4.159}$$

$$= \eta_\beta^2 - p^2 \tag{4.160}$$

and now we have the same terms on the left of the right, and we can write the Rayleigh equation:

$$\frac{2\eta_\beta p}{\eta_\beta^2 - p^2} = \frac{p^2 - \eta_\beta^2}{2p\eta_\alpha}, \tag{4.161}$$

or indeed

$$\left(p^2 - \eta_\beta^2\right)^2 + 4p^2\eta_\alpha\eta_\beta = 0. \tag{4.162}$$

Not just Poisson. This is when a and b can be whatever. But when it IS Poisson, when a2=3b2, then As can be quickly verified graphically, in order to equate both expressions, besides the trivial solution at $r = 0$, the only other solution is indeed $r \approx 0.92$. This particular case implies that

$$B = -1.47iA \qquad \text{and} \qquad \eta_\alpha = 0.85\frac{i}{c} \qquad \text{and} \qquad \eta_\beta = 0.39\frac{i}{c}. \tag{4.163}$$

4.13.2 Particle motion of Rayleigh waves

With the expressions for the displacement from eq. **??**, and after filling in our findings from eq. 4.163 into eqs 4.142 and 4.143, we get, after some rearrang-

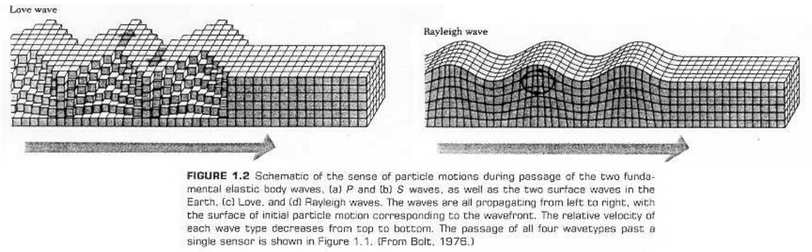


Fig. 4.15.

ing, that

$$s_x = e^{-i\omega(t-px)}\frac{\omega}{c}Ai\left(e^{-0.85z\frac{\omega}{c}} - 0.58e^{-0.39z\frac{\omega}{c}}\right) \tag{4.164}$$

$$s_z = e^{-i\omega(t-px)}\frac{\omega}{c}Ai\left(-0.85e^{-0.85z\frac{\omega}{c}} + 1.47e^{-0.39z\frac{\omega}{c}}\right) \tag{4.165}$$

The **particle motion** is given by the real part of the above expressions, and therefore we get:

$$\mathcal{R}e[s_x] = -\frac{\omega}{c}A\sin\left(\frac{\omega}{c}x - \omega t\right)\left[e^{-0.85z\frac{\omega}{c}} - 0.58e^{-0.39z\frac{\omega}{c}}\right], \tag{4.166}$$

$$\mathcal{R}e[s_z] = +\frac{\omega}{c}A\cos\left(\frac{\omega}{c}x - \omega t\right)\left(-0.85e^{-0.85z\frac{\omega}{c}} + 1.47e^{-0.39z\frac{\omega}{c}}\right). \tag{4.167}$$

Along the interface the critically refracted *P*-wave exists simultaneously with the incident *SV*-wave; in fact, the evanescent *P*-waves alone do not satisfy the stress-free boundary conditions and they *cannot* propagate along the interface without coupling to *SV*. The interference of *P* and *SV*-wave produces a particle motion in the $x - z$ plane that is *retrograde* at shallow depth, but changes to *prograde* at larger depth (see Fig. 4.16). This is similar to the particle motion in ocean waves (but there gravity is a restoring force, here we have ignored gravity).

The Rayleigh wave can thus be observed at both the vertical (in the direction of z) and horizontal (radial, i.e., in the direction of x) components of the displacement field (see also Fig. **??**).

It's useful to plot the particle motion as a function of depth. The particle motion in the $x-z$ plane is *retrograde* at shallow depth, but changes to *prograde* at larger depth (see Fig. 4.16). At the surface, we get $u_x = -0.42Ak\sin(kz-\omega t)$ and $u_z = 0.62Ak\cos(kz-\omega t)$. The change occurs at a depth $z = \lambda/a$, where $a = -2\pi(0.85 - 0.39)/\ln(0.58) \approx 5$. Hence, at a depth of about a fifth of $\lambda/5$

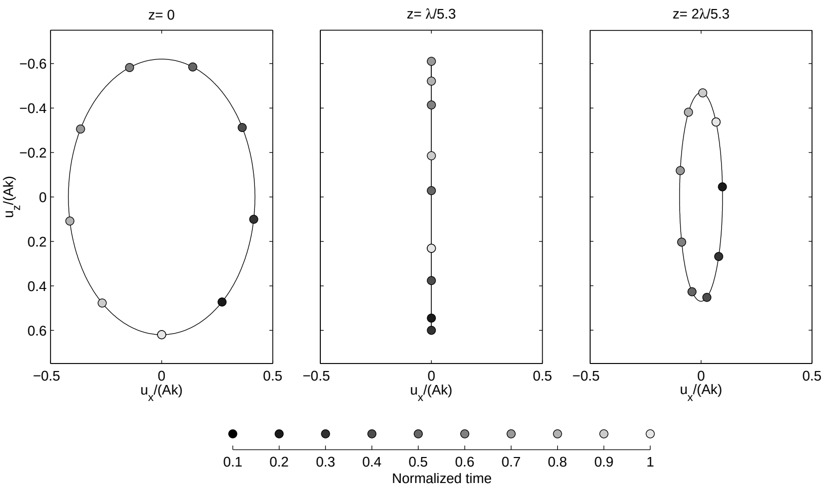


Fig. 4.16. Normalized Rayleigh-wave particle motion at different depths. The motion changes from retrograde to prograde elliptical at a depth of about one fifth of the wavelength.

the wavelength of the Rayleigh wave, the displacement is purely in the vertical direction, and the particle motion changes from retrograde to prograde elliptical.

Another important observation is that long-wavelength waves have larger displacements at greater depths than short-wavelength waves. As a rule of thumb, their sensitivity drops significantly below a depth of about half their wavelength. This can be appreciated qualitatively from Figure 4.17, where the magnitude of the displacement (defined as $\sqrt{u_x^2 + u_z^2}$) is plotted in function of the period of Rayleigh waves, for a homogenous Poisson halfspace with a

$\lambda/2$ Rayleigh wave phase speed of 4500 m/s. Below $\lambda/2$, the magnitude of the partical motion drops to about half of the magnitude at the surface.

4.14 Layered half-space solutions to the wave equation

Love: layer-over-a halfspace. Derive the important

$$\beta_1 < c_L < \beta_2 \tag{4.168}$$

The observation of earthquake waves shows that they generate more than just Rayleigh waves, and that the Earth is not a homogeneous halfspace. Among them are the fact that seismograms usually have a lot of energy on the transverse component, the fact that most motion is actually observed on the hori-

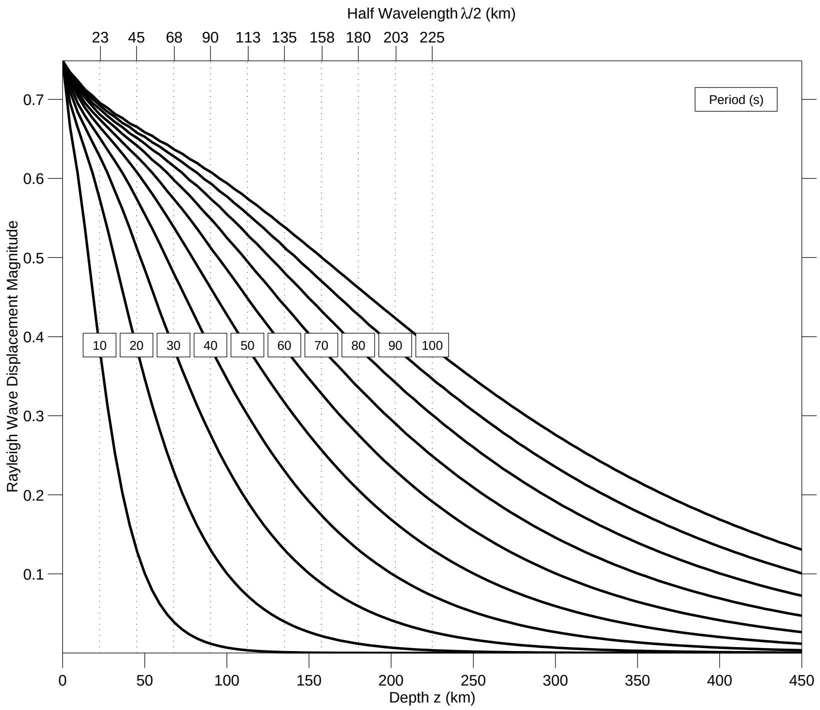


Fig. 4.17. Magnitude of displacement at depth for varying periods of Rayleigh waves in a homogeneous Poisson halfspace with constant (non-dispersive) phase speed of 4500 m/s.

zontal component, and that the wavetrains are usually dispersed, which implies that the phase speed shows a variation with frequency, unlike for the Poisson halfspace, where $c_R \approx 0.92\beta$, regardless of ω.

4.14.1 Love waves

Let us study the case where a layer (1) of thickness h and elastic properties μ', α', β' and ρ' overlies a halfspace (2) characterized by μ, α, β and ρ. Since we've already studied the $P - SV$ interactions and Rayleigh waves in the homogeneous halfspace (where we've shown that there could be no u_y), we will now study the behavior of the $u_y(x, z)$ wavefield which is non-zero in the layer-over-halfspace case. The fields u_x and u_z will be a function of P and SV waves and require the treatment with displacement potential functions as in the previous sections on Rayleigh waves. Now, u_y will always involve only

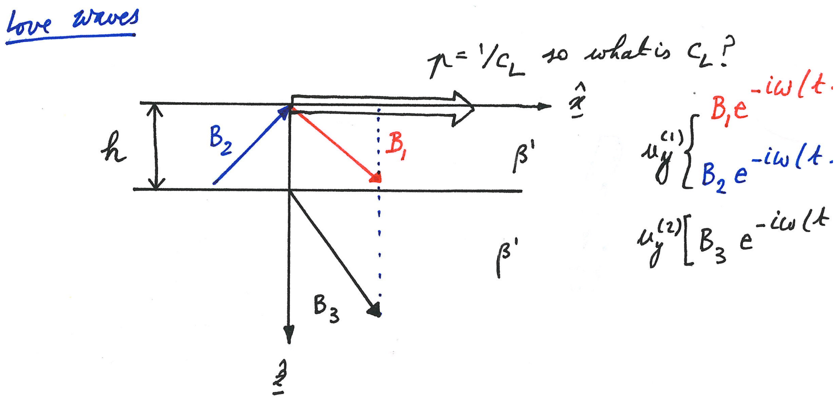


Fig. 4.18. Love waves sketch.

S-waves and can be solved for separately. There is no need for displacement potential functions, as u_y solves the wave equation directly.

Let's investigate the situation of a downgoing wave (with amplitude B_1) and an upgoing (with amplitude B_2) in the layer, and a downgoing wave (amplitude B_3) in the half-space, and see how they interact. The wavefields are given by:

$$u_y^{(1)} = B_1\, e^{-i\omega(t-px-\eta_{\beta'} z)} + B_2\, e^{-i\omega(t-px+\eta_{\beta'} z)} \tag{4.169}$$

$$u_y^{(2)} = B_3\, e^{-i\omega(t-px-\eta_{\beta} z)}, \tag{4.170}$$

where, as before, the vertical wavenumbers are

$$\eta_{\beta'} = \sqrt{\frac{1}{\beta'^2} - p^2} \qquad \text{and} \qquad \eta_{\beta} = \sqrt{\frac{1}{\beta^2} - p^2}. \tag{4.171}$$

Since we're studying the free surface, the tractions at the surface need to be zero. Hence, as before, $(\tau_{xz}, \tau_{yz}, \tau_{zz}) = (0, 0, 0)$. Using eq. **??**, we can write:

$$\boxed{\tau_{yz} = 0} \Rightarrow 0 = \mu B_1 e^{-i\omega(t-px-\eta_{\beta'} z)}(i\omega\eta_{\beta'}) - \mu B_2 e^{-i\omega(t-px+\eta_{\beta'} z)}(i\omega\eta_{\beta'}), \tag{4.172}$$

and this leads to $B_1 = B_2$. In other words, besides the fact that the upgoing SH wave is never coupled to a P-conversion (we already knew that), it is always totally reflected from the surface.

What are the boundary conditions on the interface? For the interface to be *welded*, the displacements as well as the stresses must be continuous across it.

Hence,

$$u_y^{(1)} = u_y^{(2)} \tag{4.173}$$

$$\tau_{yz}^{(1)} = \tau_{yz}^{(2)} \tag{4.174}$$

Let's use eqs 4.173 and 4.174 by equation eqs 4.169 and 4.170 at a depth $z = h$ for $B_1 = B_2$ and by using the expression of eq. **??** again. This leads to:

$$\boxed{u_y^{(1)} = u_y^{(2)}} \Rightarrow B_1 \left[e^{i\eta_{\beta'} h\omega} + e^{-i\eta_{\beta'} h\omega}\right] = B_3 e^{i\eta_\beta h\omega} \tag{4.175}$$

$$\boxed{\tau_{yz}^{(1)} = \tau_{yz}^{(2)}} \Rightarrow (i\omega\eta_{\beta'})\mu' B_1 \left[e^{i\eta_{\beta'} h\omega} - e^{-i\eta_{\beta'} h\omega}\right] = (i\omega\eta_\beta)\mu B_3 e^{i\eta_\beta h\omega}. \tag{4.176}$$

Recognizing that $\cos u = (e^u + e^{-u})/2$ and $\sin u = (e^u - e^{-u})/(2i)$, we can easily rewrite eq. 4.176 to yield the following quotient:

$$\tan(\eta_{\beta'} h\omega) = -i\frac{\eta_\beta}{\eta_\beta'}\frac{\mu}{\mu'}. \tag{4.177}$$

It's easy to see how a solution can only be found in case η_β is complex (hence, the waves in the halfspace are evanescent), and if η_β' is real (hence, oscillatory waves in the layer). In other words, the horizontal propagation speed c of what we'll call **Love waves** must obey

$$\beta > c > \beta'. \tag{4.178}$$

As before, we shall replace

$$\eta_\beta = i\sqrt{p^2 - \frac{1}{\beta^2}} \qquad \text{and} \qquad \eta_{\beta'} = \sqrt{\frac{1}{\beta'^2} - p^2}. \tag{4.179}$$

Hence, Love waves always require at least a low-velocity layer over a half-space. Because, as for Rayleigh waves, the magnitude of displacement at depth of Love waves is dependent on the wavelength $\lambda = 2\pi/k = 2\pi c/\omega$ of the wave (longer wavelength waves "feel" deeper than shorter wavelengths), a radial increase of $\beta > \beta'$ must mean that the phase speed c will depend on the frequency of the wave: Love waves are naturally **dispersive**.

Plugging eqs 4.179 and **??** into eq. 4.177, with $p = 1/c_L$, and realizing that $\tan$ is a periodic function of period, some rearranging quickly leads to

$$\boxed{\frac{\omega h}{c}\sqrt{\left(\frac{c}{\beta'}\right)^2 - 1} = \tan^{-1}\left[\frac{\mu}{\mu'}\frac{\sqrt{1 - \left(\frac{c}{\beta}\right)^2}}{\sqrt{\left(\frac{c}{\beta'}\right)^2 - 1}}\right] + n\pi} \tag{4.180}$$

So now

$$c_n(\omega) = \frac{\omega}{k_n(\omega)} \tag{4.181}$$

This naturally introduces the different **modes** of the Love wave system (of

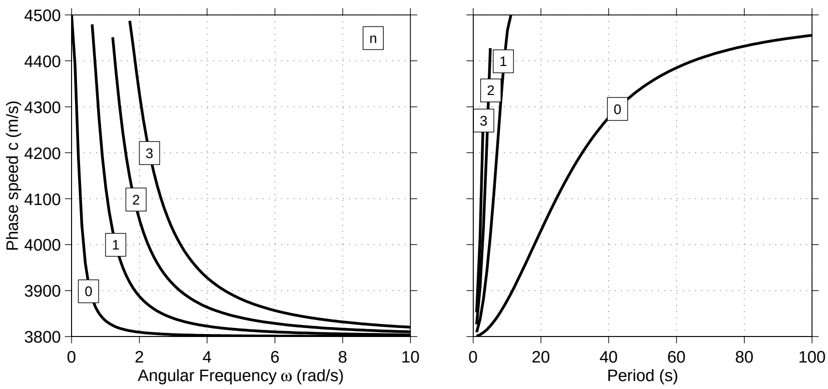


Fig. 4.19. Graphical representation of the solutions to eq. 4.180. A plot of phase speed versus frequency is usually called a *dispersion diagram.*

course, Rayleigh waves generated in a layer over a halfspace will also exhibit this phenomenon): for increasing frequency ω, more and more solutions to this equation exist. In the low-frequency limit, as $\omega \to 0$, the Love wave propagates with a phase speed of $c \approx \beta$, i.e. it is most sensitive to the halfspace. In the high limit, $\omega \to \infty$, its speed is $c \approx \beta'$ because it is mostly confined to the layer. The modes (labeled **fundamental mode** for $n = 0$ and **higher modes** or **overtones** for $n > 0$ can be thought of as corresponding to how many different times you can fit the oscillatory wave inside a layer of finite thickness h.

4.14.2 Particle motion of Love waves

It's easy to write expressions for the displacement functions once we know they are oscillatory in the layer and evanescent in the halfspace. Then, by taking the real part of u_y, you can get the particle motion.

In the layer:

$$s_y(0 < z < h) = 2B_1 e^{-\omega(t-x/c)} \cos\left(\frac{\omega}{c} z \sqrt{\left(\frac{c}{\beta'}\right)^2 - 1}\right) \tag{4.182}$$

In the halfspace:

$$s_y(z > h) = B_3 e^{-\omega(t-x/c)} e^{-z\frac{\omega}{c}\sqrt{1-\left(\frac{c}{\beta}\right)^2}} \tag{4.183}$$

In Figure 4.20, the magnitude of the displacement of a fundamental mode Love wave is plotted as a function of period. For this layered system, the phase velocities of the Love wave were computed using eq. 4.180 and the displacement functions u_y were computed using eqs 4.182 and 4.183. As indicated before, waves of infinitely long period $\omega \to 0$ will be affected by the entire halfspace and propagate with its shear-wave speed β, and in the short-period limit $\omega \to \infty$, the Love wave is almost entirely confined to the low-velocity layer, and propagates with the shear wave speed of the layer, β.

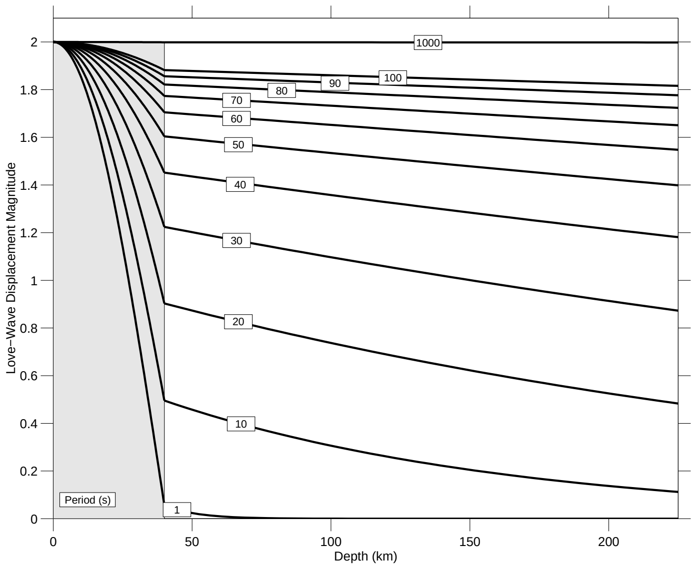


Fig. 4.20. Fundamental mode Love wave displacement in a layer over a halfspace with $\beta' = 3800$ m/s, $\beta = 4500$ m/s, $\rho' = 3000$ g/cm^3, $\rho = 3360$ g/cm^3.

As we have said before, the mode number equals the number of times the magnitude of the displacement goes through zero. This is illustrated in Figure 4.21. The introduction of a length scale (the layer of finite thickness over the halfspace) allows for only a discrete number of Love-wave solutions. Since the medium has no horizontal length-scale, any possible value of k is allowed. But since there is a vertical length scale, in a sense, the Love wave has to “fit” inside of the layer. In Figure 4.21, the maximum displacement (over one period) is plotted as a function of depth, for the fundamental modes and three overtones of three distinct periods. The overtone number is equal to the number of zero crossings with depth. This is akin to the overtone number n used to describe the normal modes of the Earth. In addition, for the Earth, the number of surface wavelengths is discrete as well since the medium is bounded. We will see this in more detail when we discuss the normal modes of the Earth.

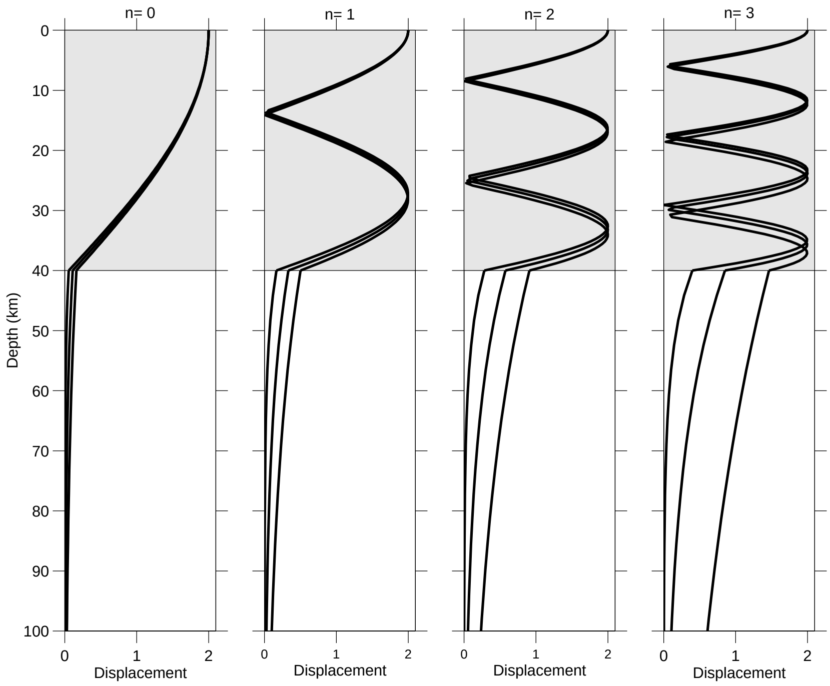


Fig. 4.21. Love-wave displacement magnitude for the layered system as in Figure 4.20, for the fundamental mode $n = 0$ and two overtones $n = 1 \rightarrow 3$, for three periods T=1, 2, and 3 s.

Love waves are observed only on the transverse component (parallel to x_2) of the displacement field.

4.15 Propagation speed of Rayleigh and Love waves

Make a table, perhaps? Must have mentioned anisotropy around Hooke.

From looking at data we can make an important observation: Love waves arrive before Rayleigh waves. Love waves propagate intrinsically faster than Rayleigh waves, see below, but the difference is not large enough to explain the observed advance of the Love wave arrival. Since Love waves involve only horizontal displacement whereas Rayleigh waves are composed of *P*-waves and vertically polarized *SV*-waves, the observed advance of the Love waves suggests a form of seismic anisotropy with faster wave propagation in the horizontal plane than in the vertical direction (a situation known as *transverse isotropy*).

It can be shown that for horizontally propagating waves to be evanescent they must travel with a propagation velocity c that is always smaller than the compressional wave speed α, $c = 1/p < \alpha$, and also smaller than the shear wave speed $\beta, c = 1/p < \beta$. If $1/p \rightarrow \beta$ the amplitude of the surface waves no longer decays with depth and conservation of energy is then achieved by the leaking of energy into the half space in the form of body waves (*SV* in the case of Rayleigh waves and *SH* in the case of Love waves). If this happens one speaks of **leaky modes**.

Airy.

So Rayleigh waves always propagate with a speed that is *lower* than the shear wave speed. For a half space with shear wave speed β_1, the propagation speed of the Rayleigh wave is about $0.9\beta_1$. (In the Earth the situation is more complicated because of the radial variation of both *P* and *S*-wave speed: if the wave speed gradually increases with depth from $c = \beta_1$ at the surface to $c = \beta_2$ in the half space: $0.9\beta_1 < c_{\text{Rayleigh}} < 0.9\beta_2$). We will see below that the surface-wave propagation speed depends on the wave length, and thus on frequency, of the wave (dispersion). For Love waves it is slightly different. Here it's the head wave that is evanescent; for high-frequency waves (short wavelengths) the evanescent head wave hardly penetrates into the half space (suppose a shear wave speed of β_2) so that the propagation speed is dominated by *SH*-propagation in the layer over the half space (propagation speed $c = \beta_1$). For longer period Love waves, the head wave is sensitive to as much larger depth range and the propagation speed gets closer to the shear wave speed in the half space (β_2). Thus: $\beta_1 < c_{\text{Love}} < \beta_2$.

4.16 Dispersion, phase and group speed

Treatment by Udias as taught in class before?

Naturally dispersive Love waves lead us to need to think about dispersion in general. Phase speed makes way to group speed.

Physical dispersion... anelasticity. Mention or ignore?

The dependence of the depth of penetration on the period is described by the sensitivity kernels. If the wave speed is constant in the half space the waves associated with different kernels travel with the same wave speed and thus arrive at the same time at a receiver at some distance from the source. But if, as is the case in Earth, the *P* and *S*-wave speed changes with depth, the longer period waves arrive at a different time than the shorter period waves. In Earth, the propagation speed of Rayleigh waves is thus frequency-dependent, and the waveform changes with increasing or decreasing distance from the source. This frequency dependence of propagation speed is called *dispersion*. Love waves are always dispersive since they cannot exist unless there is a layer over a half space, with the shear wave speed in the half space larger than in the overlying layer.

As a result of dispersion the surface waveform changes with varying distance from the source, and it is clear that one can no longer describe the wave propagation with a single wave speed. We describe the propagation velocity of the part of the waveform that remains constant, such as the onset of the phase arrrival, a peak, or a trough (see discussion of plane waves) with the **phase velocity** $c = \omega/k$. Wave packages with different frequencies travel at different velocities and their interference results in a phenomenon known as **beating**: the propagation velocity of the envelope, which is related to the energy, of the resulting wave train is called the **group velocity U**.

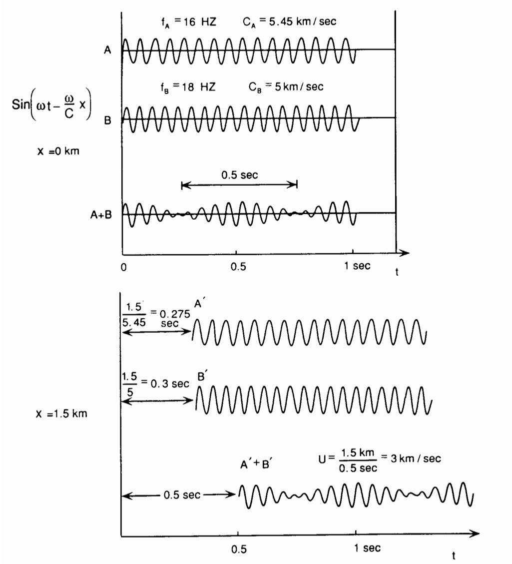


FIGURE 4.12 Example of the interference of two waves of the form (4.36) at two positions $x = 0$ and $x = 1.5$. The envelope of the interference pattern moves with group velocity $U = 3$ km/s. (Courtesy of H. Kanamori.)

Fig. 4.22. Two harmonic waves with the same amplitude but slightly different frequencies.

Consider two harmonic waves with the same amplitude but slightly different frequencies (ω_1 and ω_2), wave numbers k_1 and k_2, and phase velocities $k_1 = \omega_1/c_1$ and $k_2 = \omega_2/c_2$ (see Fig. 4.22). These waves combine to give the total displacement

$$u(x,t) = \cos(k_1 x - \omega_1 t) + \cos(k_2 x - \omega_2 t). \tag{4.184}$$

If we define ω as the average between ω_1 and ω_2 so that $\omega_1 + \delta\omega = \omega = \omega_2 - \delta\omega$, and $k_1 + \delta k = k = k_2 - \delta k$, with $\delta\omega \ll \omega$ and $\delta k \ll k$, insert it into (4.184) and apply the cosine rule $2\cos x \cos y = \cos(x+y) + \cos(x-y)$, we obtain

$$u(x,t) = 2\cos(kx - \omega t)\cos(\delta kx - \delta\omega t) \tag{4.185}$$

This is the product of two cosines, the second of which varies much more slowly than the first. The second cosine "modulates" the amplitude of the first. The propagation speed of this 'envelope' is given by $U(\omega) = \delta\omega/\delta k$. In the limit as $\delta\omega \to 0$ and $\delta k \to 0$,

$$U(\omega) = \frac{d\omega}{dk} = c + k\frac{dc}{dk} = c - \lambda\frac{dc}{d\lambda} \tag{4.186}$$

The group velocity is related to interference of waves with slightly different phase velocities; in other words U depends on c and on how c varies with frequency (or wavelength or wave number). In the earth $dc/d\lambda > 0$ so that the group velocity is typically smaller than the phase velocity.

Peaks or troughs in the wave form, or the onset of a particular phase arrival in the seismogram, all propagate with the **phase velocity**. In fact, we have seen this before when we discussed travel time curves of the body waves, which depend on the phase velocity. The phase velocity can thus be measured directly from travel time curves (recall that the horizontal slowness p can be determined from the slope of the travel time curve at a certain distance).

In Fig. 4.23 the dashed lines through A, B, etc. are travel time curves for those phases. But note that the frequency of those phases change with distance, so that the waveform changes. For instance, with increasing distance, the first arriving phase (A) is composed of waves with larger frequencies (because they sample deeper).

The **group velocity** is constant for a given frequency ($d\omega = 0$). Thus the group velocity of surface waves of a particular frequency defines a straight line through the origin and through the signal of that particular frequency on records of ground motion at different distances. The group velocity decreases as the frequency increases. As a result, high frequency phases become less and less pronounced with increasing distance from the source (or time in the seismogram).

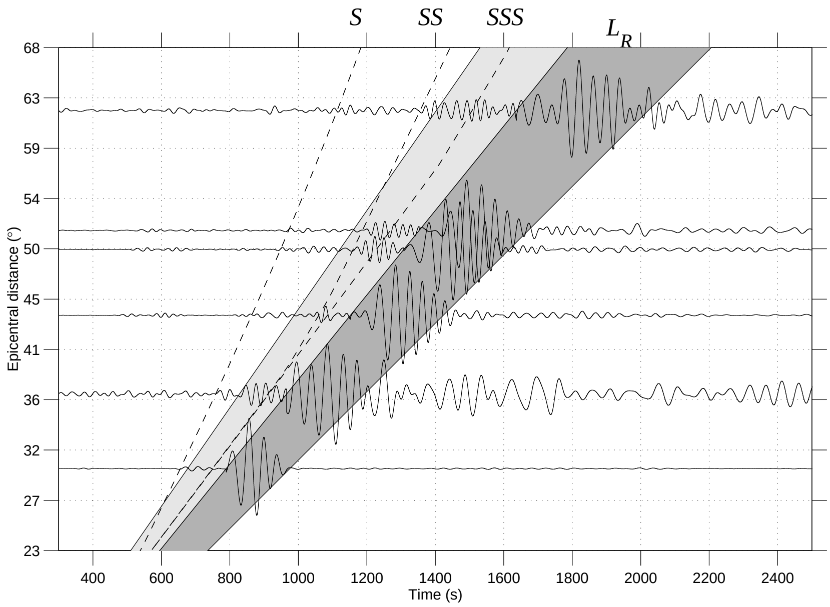


Fig. 4.23. Group velocity windows and phase velocity curves.

The group velocity is very important: the energy in surface waves propagates mainly in the constructively interfering wave packets, which move with the group velocity.

Narrow-band filtering can isolate the wave packets with specific central frequencies (see Fig. 4.24), and the group velocity for that frequency can then be determined by simply dividing the path length along the surface by the observed travel time. This technique can be used for the construction of dispersion curves (see Sec. **??**).

4.17 Free oscillations

We had neglected gravity. Now we won't.

See how much I would need to amend the wave equation to impart a simple understanding of how this would work. Coriolis? Gravity?

4.18 The earthquake source

We had neglected the forcing. Now we don't.

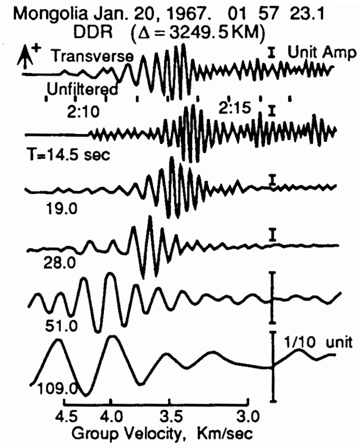


FIGURE 4.B6.2 Wave packets in narrow frequency bands obtained by filtering a Love-wave recording. The unfiltered record is shown at the top. Narrowband records with central periods shown on the left are plotted below the original seismogram, with varying amplitude scale. Note that each narrowband-filtered trace has the appearance of a wave packet. (From Kanamori and Abe, 1968.)

Fig. 4.24. Frequency-band filtering of seismograms.

Does this have a simple treatment for a beginner's course? Maybe via Udias or Herrmann?

4.19 Earthquake location

Talk again about particle motion. Introduce inverse methodology.

4.20 Earth structure from seismology in one dimension

This now should make sense.... via the various kinds of waves.

4.21 Earth structure from seismology in three dimensions

Seismic tomography. Talk about inverse theory. Should have talked a tiny bit about it in the context of density determination from gravity, or magnetization in the context of geomagnetism. Some simple examples and some nice results.

4.22 Time-variable seismology

A field with not much history and perhaps not much future, yet very interesting. In global seismology. Perhaps tie in with 4D industry monitoring.

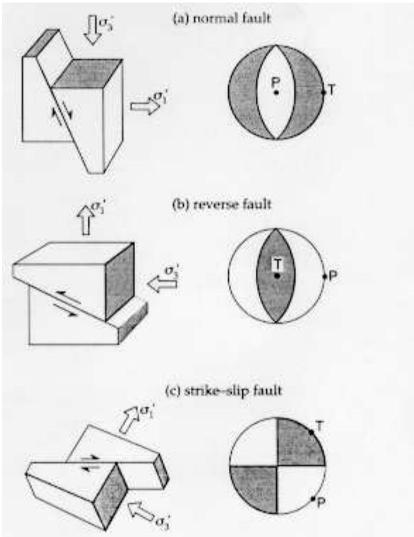


Fig. 3.36 The three main types of fault and their focal mechanisms. *Left*: the orientations of each fault-plane and the principal deviatoric stresses, σ'_1 and σ'_2. *Right*: focal mechanisms and orientations of P- and T- axes.

Fig. 4.25.

Need to say something about Green's functions, Helmholtz equation. The connection to gravity and magnetic Green's functions. Freeden stuff. Yomogida92 easy paper. Nolet's book.

The word Fresnel zone must appear here.

Here be pictures of some reference Earth models, maybe like S40RTS?

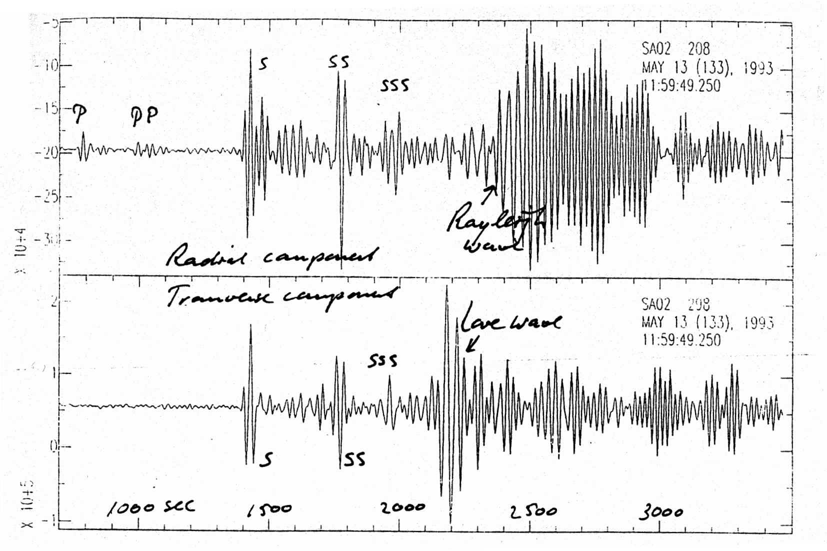

SA02 208
MAY 13 (133), 1993
11:59:49.250
P
PP
S
SS
SSS
Rayleigh wave
Radial component
Transverse component
Love wave
SSS
S
SS
1000 sec
1500
2000
2500
3000

Fig. 4.26.

5
Heat

The Earth is **hot**. Its *primordial heat* is the result of its formation from humble origins by the accretion of interstellar dust, chemical differentiation, core formation, and gravitational settling of the mantle, all of which liberated heat. Earth also continues to actively *generate* heat, by the decay of *radioactive elements*, which are concentrated mostly in the crust, and via some exothermic phase transformations. The Earth is also **cooling**. Whatever its early temperature, and whatever its subsequent evolutionary path, as a geologically active planet it is constantly in the business of ultimately, slowly, losing its heat. The primary mechanisms are *conductive cooling* (in the crust), *convective redistribution* (especially in mantle and core), and, to a minor extent *radiation*.

What is the Earth's internal temperature? Dig a mine, drill a hole, and it gets hot rather quickly. But what is that behavior beyond the deepest mine (some 4 km) or the longest drill hole (some 12 km)? We will study the mechanisms of heat transfer in steady-state, both with and without the presence of heat-producing elements, before moving on to time-dependent problems. As with the magnetic induction equation, which described the time-dependence of the field, we require the fundamental equations that describe how Earth's temperature evolves over time. As in previous chapters, our story starts with the luminaries of yesteryear. In particular Fourier, whose 1822 *Théorie Analytique de la Chaleur* paved the way for the study of heat conduction.

The discovery of radioactivity took another hundred years, and that the mantle should convect was a similarly foreign concept. Lord Kelvin's famous calculation of the age of the Earth—some 60 million years, considering conductive cooling alone—was so demonstrably at odds with the growing evidence of geology and paleontology that it now literally is a textbook example of how *physics* stands to gain from respecting evidence from other fields, even those that are less quantitative. With the wrong assumptions, even a correct calculation is likely to turn up... er, garbage [17]. *Caveat emptor!*.

5.1 Making heat

5.2 Losing heat

Pekeris 65 writes what we need to know in about a page. Good start.

We'll be shortcutting this somewhat.

Maybe just write the equation ones, then talk about v being difficult, like for geomagnetism, Bunge & Kennett maybe, then just focus on the conductive part. So we ignore the convection itself (and yet wait with the adiabat until after since it's lower down in the Earth?

Some underbraces!

5.3 Principles/Basic Theory/Conservation Equations

According to **Fourier's law of conduction**, for a system characterized by a temperature T, at a given location $\mathbf{r}$, the *rate of heat energy transfer per unit of surface* (thus measured in Wm^{-2}), namely the **heat flow** or **heat flux**, $\mathbf{q}(\mathbf{r})$, is proportional to the temperature gradient at that point: T $\mathbf{q}$

$$\mathbf{q}(\mathbf{r}) = -\mathbf{k} \cdot \boldsymbol{\nabla} T(\mathbf{r}), \tag{5.1}$$

where $\mathbf{k}$ is the thermal conductivity tensor. For simplicity, we quote eq. (5.1) with a scalar proportionality constant, the *thermal conductivity* k, as in $\mathbf{k}$ k

$$\mathbf{q}(\mathbf{r}) = -k\boldsymbol{\nabla} T(\mathbf{r}). \tag{5.2}$$

From the units of $\mathbf{q}$ follows that the thermal conductivity has units of $\mathrm{WK}^{-1}\mathrm{m}^{-1}$. See Table 5.1 for more details. The outward flux of heat through a surface per unit of volume, which, via the divergence theorem (2.46),

$$\int_V \boldsymbol{\nabla} \cdot \mathbf{q}\; dV = \int_{\partial V} \hat{\mathbf{n}} \cdot \mathbf{q}\; \mathrm{d}\Sigma. \tag{5.3}$$

is equal to the divergence of the heat flow inside the volume bounded by this surface, must be balanced by what is being generated internally, on a per-volume basis—and ultimately by the rate at which the object reduces or increases its *internal thermal energy* Q, barring any work done. The former is the *volumetric heat production rate*, H, the latter involves the *specific heat capacity* at constant pressure P, the intensive property which, for a certain mass m is given by Q H c_P

$$c_P = \frac{1}{m}\left(\frac{\partial Q}{\partial T}\right)_P. \tag{5.4}$$

Together, this balance gives the **thermal diffusion equation** as

$$\rho c_P \frac{\partial T}{\partial t} = -\boldsymbol{\nabla} \cdot \mathbf{q} + H, \tag{5.5}$$

ρ which involves the familiar mass density, ρ. Here, the volumetric heat production rate H has units of Wm^{-3}, and the specific heat capacity c_P is in J kg^{-1}K^{-1}: the energy required to raise the temperature of a unit mass by one degree Kelvin.

Combined with Fourier's laws eqs (5.1)–(5.2) the diffusion equation is written as the **parabolic** differential equation

$$\frac{\partial T}{\partial t} = \frac{k}{\rho c_P}\nabla^2 T + \frac{H}{\rho c_P}, \tag{5.6}$$

κ, h or indeed, in the scaled variables $\kappa = k/(\rho c_P)$ and $h = H/(\rho c_P) = H\kappa/k$,

$$\boxed{\frac{\partial T}{\partial t} = \kappa \nabla^2 T + h,} \tag{5.7}$$

noting that h is in Ks^{-1} and the *thermal diffusivity* κ is in m^{2}s^{-1}, and of course $h/\kappa = H/k$. Again, see Table 5.1 for a table of quantities and their units.

Special cases of the thermal diffusion equation, eq. (5.7) that are immediately of interest are the **steady-state** (time-independent) regime satisfying

$$\nabla^2 T = -\frac{h}{\kappa}, \tag{5.8}$$

which expresses the balance of heat production with a spatial gradient of temperature.

In the absence of heat production, we recover the plain-vanilla **transient** (time-dependent) heat diffusion equation

$$\frac{\partial T}{\partial t} = \kappa \nabla^2 T. \tag{5.9}$$

A *non-homogeneous* differential equation (with a right-hand side in standard form) is solved by adding a *particular* solution to the *general* solution to the homogeneous equation (with a zero-right hand side). This gives the *general* solution to the overall equation, after which one proceeds to determine the integration constants so as to obtain the particular solution to the full equation.

5.4 Steady-state geotherms

Oceanic vs continental. Crust vs mantle. Heat flow vs heat production.

Heat (energy)	Q	J	The original "stuff"
Heat (flow, flux)	$\mathbf{q}$	Wm^{-2}	The "flux of stuff"
Heat (specific)	c_P	$\mathrm{Jkg}^{-1}\mathrm{K}^{-1}$	An intensive property
Heat (production)	H	Wm^{-3}	"Volumetric rate of creation of stuff"
Heat (production)	h	Ks^{-1}	The scaled version of H
Temperature	T	K	The property of being "hot" or "cold"
Time	t	s	The passage of time
Conductivity	k	$\mathrm{Wm}^{-1}\mathrm{K}^{-1}$	Heat flow scales with thermal gradient
Diffusivity	κ	$\mathrm{m}^2\mathrm{s}^{-1}$	The scaled version of k

Table 5.1. *Quantities, symbols, and units in the study of heat flow, and some colloquialisms by which to make sense of them.*

5.4.1 Conductive, and with heat-producing elements

Let us specialize eq. (5.8) by restricting the Laplacian to the special and most important *vertical* dimension. Hence, in *steady-state*, and for a *constant heat production* term h, we obtain the differential equation

$$\boxed{\frac{d^2T}{dz^2} = -\frac{h}{\kappa}.} \tag{5.10}$$

We also of course still know from Fourier's law, eq. (5.2), in its scalar and one-dimensional form and choosing z to be positive downward, that the *magnitude* of the heat flux, $q = |\mathbf{q}|$, relates to the scaled first spatial derivative of the temperature, according to q

$$\frac{dT}{dz} = \frac{q}{k}. \tag{5.11}$$

Eq. (5.11) supplies our first measurable *boundary condition* for eq. (5.10). At the surface, the observable **temperature** is given by

$$T|_{z=0} = T_S, \tag{5.12}$$

and the surface **geothermal gradient** dT/dz

$$\left.\frac{dT}{dz}\right|_{z=0} = \frac{q_S}{k}. \tag{5.13}$$

To solve for the steady-state geotherm that satisfies (5.10) in an unbounded medium we postulate for the general form of the solution the *Ansatz*

$$T(z) = az^2 + bz + c, \tag{5.14}$$

with the constants a, b and c yet to be determined. Using eq. (5.14) with eq. (5.10) returns $a = -h/(2\kappa) = -H/(2k)$. Using the boundary conditions, combining eq. (5.14) with eq. (5.13) we obtain $b = q_S/k$, and finally, the combination with eq. (5.12) yields $c = T_S$.

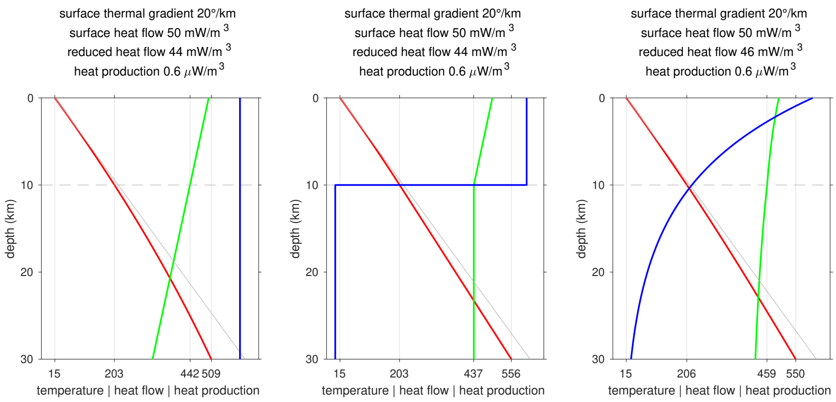


Fig. 5.1. Steady-state geotherms. (*Left*) Constant heat producing elements throughout, eqs. (5.15)–(5.16). (*Middle*) Constant heat producing elements in the top layer eqs (5.15)–(5.16) and none eqs (5.21)–(5.22) in the lower layer (*Right*) an exponential distribution of heat producing elements throughout,x eqs (5.26)–(5.27)

Thus, the *steady-state conductive geotherm in a medium with constant heat production* is given by the parabolic form

$$\boxed{T(z) = T_S + \frac{q_S}{k} z - \frac{H}{2k} z^2,} \tag{5.15}$$

with a *heat flux* that is linear in depth, given via eq. (5.11) by

$$\boxed{q(z) = q_S - Hz.} \tag{5.16}$$

The above derivations apply in a halfspace. Nevertheless, at a certain depth $z = D$, we obtain two specific values,

$$T|_{z=D} = T_D = T_S + \frac{q_S}{k} D - \frac{H}{2k} D^2. \tag{5.17}$$

and the geothermal gradient

$$\left.\frac{dT}{dz}\right|_{z=D} = \frac{q_D}{k} = \frac{q_S}{k} - \frac{HD}{k}, \tag{5.18}$$

which we can use as boundary conditions for the region below $z = D$, if we should wish to consider the solutions for a layer-over-a-halfspace regime, next. Note the linear relation between surface heat flow and the heat production term:

$$q_D = q_S - HD. \tag{5.19}$$

5.4.2 Conductive, and without heat-producing elements

Let us now specialize eq. (5.8) to the homogeneous version of eq. (5.10), suitable at depths deeper than $z \geq D$,

$$\boxed{\frac{d^2T}{dz^2} = 0.} \tag{5.20}$$

Again, we start from the general form (5.14), and use eq. (5.20) to find that, in this case, $a = 0$. We then use the new boundary condition eq. (5.18) to find that $b = q_D/k$ and then, using the boundary condition eq. (5.17), that $c = T_D - q_D D/k$. We conclude that the steady-state temperature profile is given by

$$\boxed{T(z) = T_D + \frac{q_D}{k}(z - D).} \tag{5.21}$$

In the absence of heat-producing elements, the quadratic temperature behavior of eq. (5.15) gives way to the linear behavior of eq. (5.21). As to the heat flow in the half-space, it is of course constant,

$$\boxed{q(z) = q_D.} \tag{5.22}$$

In a layer-over-a-halfspace regime where the difference between the layer and the halfspace beyond the presence/absence of a heat-production term should *also* involve a change in thermal conductivity, eqs (5.15)–(5.16) and (5.21) would have to be trivially adjusted to reflect the different thermal conductivities and diffusivities. The full patched solution, which we think of as applicable to the Earth's ocean crust, in that case, would be

$$T(z) = T_S + \left(\frac{q_S - HD/2}{k_1}\right) D + \frac{q_D}{k_2}(z - D) \quad \text{for} \quad z \geq D. \tag{5.23}$$

5.4.3 A patched solution for the continents

Now give it a little more realism with an exponential decay of heat-producing elements in the "above" portion. The problem to solve is

$$\boxed{\frac{d^2T}{dz^2} = -\frac{h_o}{\kappa}e^{-z/\tau},} \tag{5.24}$$

for some characteristic spatial length scale τ. Our temperature *Ansatz* is

$$T(z) = ae^{-z/\tau} + bz + c. \tag{5.25}$$

The form of the particular solution is a constant times the functional form present in the non-homogeneous term. We can do this because this term itself is indeed a solution of the equation, again we only have to determine the constant—hence the name "Method of undetermined coefficients" (see, for example, Bender and Orszag, 1999, p 19).

To *be* a solution is to satisfy eq. (5.24), hence $a = -\tau^2 h_0/\kappa = -\tau^2 H_0/k$. The boundary conditions eqs (5.12)–(5.13) return $b = q_S/k - \tau h_0/\kappa = q_S/k - \tau H_0/k$ and $c = T_S + \tau^2 h_0/\kappa = T_S + \tau^2 H_0/k$. Thus, the *steady-state conductive geotherm in a medium with exponential heat production* is given by

$$\boxed{T(z) = T_S + \left(\frac{q_S - \tau H_0}{k}\right) z + \tau^2 \frac{H_0}{k}\left(1 - e^{-z/\tau}\right),} \tag{5.26}$$

with a corresponding heat flux

$$\boxed{q(z) = q_S - \tau H_0 \left(1 - e^{-z/\tau}\right).} \tag{5.27}$$

At some depth $z = D$ we would have, in this particular case,

$$T|_{z=D} = T_D = T_S + \left(\frac{q_S - \tau H_0}{k}\right) D + \tau^2 \frac{H_0}{k}\left(1 - e^{-D/\tau}\right), \tag{5.28}$$

and a geothermal gradient

$$\left.\frac{dT}{dz}\right|_{z=D} = \frac{q_D}{k} = \frac{q_S}{k} - \tau \frac{H_0}{k}\left(1 - e^{-D/\tau}\right). \tag{5.29}$$

If the evaluation depth D is large compared to the characteristic depth τ, the exponential term drops out of eq (5.27), and we again obtain a quasi-linear relation, as we did in eq. (5.19),

$$q_D = q_S - \tau H_0 \left(1 - e^{-D/\tau}\right) \approx q_S - \tau H_0 \quad \text{for} \quad D \gg \tau. \tag{5.30}$$

We will call the right-hand side the **reduced heat flow**

$$\boxed{q_m = q_S - \tau H_0.} \tag{5.31}$$

Observations made on many continents lend credence to the fact that surface heat flow q_S *does* seem to scale linearly with the measured heat-production h, as in eq. (5.31). There thus is *some* basis for the interpretation of the intercept of the linear relation between h and q_S as the "reduced" or "mantle" heat flow, and the slope as a characteristic length scale for the presence of heat producing elements in the crust. But also eq. (5.19) was in this exact same linear form, so this is no validation of the particular *exponential* shape function of the distribution of heat-producing elements in the crust. Except for the fact that assuming the exponential preserves the linear relation between heat flux and heat-producing elements under erosion, which is used as an argument in support of it.

Whatever the interpretation, the reduced heat flow remains the "basal" heat flow, at the point where the geotherm becomes linear and the concentration of radio-isotopes is effectively zero.

As we did in Section 5.4.2 we now move to the regime *below* the heat-producing layer in the continents. Again we solve the homogeneous version of eq. (5.24), which means that if we stick to the form of eq. (5.25) we have $a = 0$, and from the boundary conditions in eqs (5.28)–(5.29) we are getting expressions identical to eqs. (5.21)–(5.22), but of course T_D and q_D now derive from eqs (5.28)–(5.29).

Finally, return to a two-conductivity regime, to conclude that, at $z \geq D$,

$$T(z) = T_S + \left(\frac{q_S - \tau H_0}{k_1}\right) D + \frac{q_D}{k_2}(z - D) + \tau^2 \frac{H_0}{k_1}\left(1 - e^{-D/\tau}\right). \quad (5.32)$$

Next up is another steady-state geotherm, but of a very different nature.

5.4.4 Convective, under adiabaticity

Wasn't going to tell you much about the velocities themselves (back of the env? from red book?) but let's see what geotherm we be getting, just about.

Return to eq. (5.4) for the *specific* heat capacity at constant pressure and now get rid of the "per unit of mass", an *intensive* characterization, to turn it into an *extensive* property. Then, make the thermodynamic identification, the definition of entropy, that

$$dS = \frac{\delta Q}{T}. \quad (5.33)$$

Funny that the word "entropy" does not appear in Jaupart and Mareschal, at least not in the index. With this, rewrite

$$c_P = \frac{T}{m}\left(\frac{\partial S}{\partial T}\right)\bigg|_P. \quad (5.34)$$

Now consider a no-heat added, reversible, isolated, equilibrium, isentropic process for which entropy is conserved,

$$dS = 0, \quad (5.35)$$

and then write the differential change in entropy as the total differential

$$dS = \left(\frac{\partial S}{\partial T}\right)\bigg|_P dT + \left(\frac{\partial S}{\partial P}\right)\bigg|_T dP \tag{5.36}$$

We have

$$\left(\frac{\partial S}{\partial T}\right)\bigg|_P = \frac{mc_P}{T} \tag{5.37}$$

$$\left(\frac{\partial S}{\partial P}\right)\bigg|_T = \left(\frac{\partial V}{\partial T}\right)\bigg|_P = V\alpha \tag{5.38}$$

We thus have, with the mass density $\rho = m/V$,

$$T\frac{V\alpha}{mc_P} = \left(\frac{\partial S}{\partial P}\right)\bigg|_T \bigg/ \left(\frac{\partial S}{\partial T}\right)\bigg|_P = \left(\frac{\partial T}{\partial P}\right)\bigg|_S \tag{5.39}$$

And then in function of the radius:

$$\left(\frac{\partial T}{\partial r}\right)\bigg|_S = \left(\frac{\partial T}{\partial P}\right)\bigg|_S \left(\frac{\partial P}{\partial r}\right)\bigg|_S \tag{5.40}$$

$$= \frac{\alpha g T}{c_P}. \tag{5.41}$$

since $\partial P/\partial r = \rho g$.

And so on. Then also define **potential temperature** which means integrate up the adiabat. Entropy is in J per K. We have used a MAXWELL relation.

Super/subadiabatic profiles according to notes, to make the argument that the mantle really does convect.

5.5 Time-variable geotherms

We done the continents. They ain't being created so much to today. But the oceans, they is! We bring time back!

But we take out the heat production term, so we have a thermal diffusion equation with κ the diffusivity in dimensions L^2T^-1. So immediately we notice that $sqrt(kt)$ has dimensions of length. Remember it is $k/\rho/c_P$.

$$\boxed{\frac{\partial T}{\partial t} = \kappa\nabla^2 T.} \tag{5.42}$$

We specify

$$T(0,t) = 0 \tag{5.43}$$

$$T(\infty,t) = T_m \tag{5.44}$$

$$T(z,0) = T_m \tag{5.45}$$

Temperature on the sea bottom is just one or two degrees above 0. Ambient temperature in the mantle is much hotter, on the order of 1300∘C. Temperature at zero age is T_m at all depths.

Equations (5.43)–(5.45) are just the right boundary conditions and initial conditions for the heat flow equation, a parabolic partial differential equation, to work.

5.5.1 Cooling of a half-space

The one-dimensional solution to eq. (5.42), after some analysis [18], is:

$$\boxed{T(z,t) = \frac{2}{\sqrt{\pi}} T_m \int_0^{z/\sqrt{4\kappa t}} e^{-u^2}\, du.} \tag{5.46}$$

Let us verify that the solution (5.56) indeed works. Indeed,

$$T(0,t) = \int_0^0 du = 0 \tag{5.47}$$

satisfies eq. (5.43) the boundary condition at the top. Furthermore,

$$T(\infty,t) = T(z,0) = \frac{2}{\sqrt{\pi}} T_m \int_0^\infty e^{-u^2}\, du = T_m, \tag{5.48}$$

which satisfies the boundary condition eq (5.44) at the bottom, and the initial condition (5.45) throughout the mantle. Thus the temperature at zero age and at infinite depth are those of the ambient mantle. The integral in eq (5.56) is the Gaussian integral, its value is $\sqrt{\pi}/2$. Remember this fact separately, it's an important integral that crops up in different contexts!! Also talk about when $t = \infty$ which wasn't yet part of our problem.

Now verify that eq (5.56) *solves* eq. (5.42), restated in one dimension as

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2}. \tag{5.49}$$

Calculate the first temporal and the second spatial derivative and don't forget to use eq. (1.43).

On left of eq. (5.49), differentiating the upper boundary term, we have

$$\frac{\partial T}{\partial t} = -\frac{2}{\sqrt{\pi}} T_m e^{-z^2/4\kappa t} \frac{\partial}{\partial t}\left(\frac{z}{\sqrt{4\kappa t}}\right) \tag{5.50}$$

$$= \frac{T_m}{2\sqrt{\pi}} e^{-z^2/4\kappa t}\left(\frac{z}{\sqrt{\kappa t^3}}\right). \tag{5.51}$$

On the right-hand side of eq. (5.49), we have

$$\frac{\partial^2 T}{\partial z^2} = -\frac{2}{\sqrt{\pi}} T_m \frac{\partial}{\partial z}\left[e^{-z^2/4\kappa t}\frac{\partial}{\partial z}\left(\frac{z}{\sqrt{4\kappa t}}\right)\right] \tag{5.52}$$

$$= -\frac{2}{\sqrt{\pi}} T_m \frac{\partial}{\partial z}\left[e^{-z^2/4\kappa t}\frac{1}{\sqrt{4\kappa t}}\right] \tag{5.53}$$

$$= -\frac{2}{\sqrt{\pi}} T_m \frac{1}{\sqrt{4\kappa t}}\left(e^{-z^2/4\kappa t}\right)\frac{\partial}{\partial z}\left(\frac{-z^2}{4\kappa t}\right) \tag{5.54}$$

$$= \frac{T_m}{2\sqrt{\pi}} e^{-z^2/4\kappa t}\left(\frac{z}{\sqrt{\kappa t^3}}\right)\frac{1}{\kappa}, \tag{5.55}$$

which validates our solution, since eqs. (5.51) and (5.55) together imply eq. (5.49).

Amending eq. (5.56) to nonzero surface temperature

$$\boxed{T(z,t) = T_s + (T_m - T_s)\frac{2}{\sqrt{\pi}}\int_0^{z/\sqrt{4\kappa t}} e^{-u^2}\,du.} \tag{5.56}$$

Bring back the heat flow (using Leibniz?) and write

$$\boxed{q = k\frac{T_m - T_s}{\sqrt{\pi\kappa t}}} \tag{5.57}$$

Data show $T_m - T_s$ is 1350°±275°. So the fractional temperature excess is $(T(z,t) - T_s)/(T_m - Ts)$ and since $erf(1/2)$ is 0.5 we have that $z = \sqrt{\kappa t}$ is the depth down to which 50% of the temperature excess has been felt. Boundary layer, define z at which, say, only 90%, leads to about $z = 2.3\sqrt{\kappa t}$.

Put a figure here and then run to the end of the page.

5.5.2 Cooling of a plate with limit-thickness

This being a slight variation on plate cooling, motivated by observations that the plate appears to reach some kind of a limit thickness due to, ultimately, gravitational foundering [19]. *Fixed* L, not some growing “crust!”. FIGURE in 4 panels, $t < 0$, $t = 0$, $t > 0$ and $t = \infty$ (really? or very large?).

Initial conditions:

$$T(0,0) = T_S \tag{5.58}$$
$$T(z,0) = T_m, \qquad z > 0 \tag{5.59}$$

Boundary conditions:

$$T(0,t) = T_s \tag{5.60}$$
$$T(L,t) = T_m. \tag{5.61}$$

We partition the behavior. We will look for a solution that satisfies both the initial and the boundary conditions in the form

$$T(z,t) = T_1(z) + T_2(z,t) \tag{5.62}$$

where $T_2(z,t)$ is a solution to it with *homogeneous* boundary conditions, i.e.

$$T_2(0,t) = 0 \tag{5.63}$$
$$T_2(L,t) = 0. \tag{5.64}$$

Here $T_1(z)$ is some solution that satisfies the boundary conditions but without any initial conditions or time dependence.

Choosing the equilibrium linear gradient as

$$T_1(z) = T_s + (T_m - T_s)\frac{z}{L}, \tag{5.65}$$

which satisifes the BC but has no time evolution, we may find a general solution for $T_2(z,t)$ by separation of variables.

The initial conditions for $T_2(z,t)$ are then given by

$$T_2(0,0) = T(0,0) - T_1(0) = 0 \tag{5.66}$$
$$T_2(z,0) = T(z,0) - T_1(z) = (T_m - T_s)\left(1 - \frac{z}{L}\right), \qquad z > 0, \tag{5.67}$$

Separation of variables to solve eq. (5.42), or rather, eq. (5.49),

$$T_2(z,t) = F(z)G(t), \tag{5.68}$$

leads to, using the dash for the spatial and the overdot for the time derivative,

$$F\dot{G} = \kappa G F''. \tag{5.69}$$

Consequently,

$$\frac{F''}{F} = \frac{\dot{G}}{\kappa G} = -p^2 \tag{5.70}$$

must hold separately, which in turns implies

$$F'' + p^2 F = 0 \tag{5.71}$$

$$\dot{G} + \kappa p^2 G = 0 \tag{5.72}$$

As to the first of those, let us realize that

$$F = A\cos pz + B\sin pz \tag{5.73}$$

with the homogeneous boundary conditions, though, that implies $A = 0$ and $B\sin pL = 0$ when $pL = n\pi$ with n integer for any B, therefore

$$F(z) = \sum_{n=1}^{\infty} \sin\left(\frac{n\pi z}{L}\right) \tag{5.74}$$

and then back to the G with this new information, now that's

$$\dot{G} + \kappa\left(\frac{n\pi}{L}\right)^2 G = 0 \tag{5.75}$$

which implies, still choosing constants,

$$G = \sum_{n=1}^{\infty} a_n \exp\left[-\kappa\left(\frac{n\pi}{L}\right)^2 t\right], \tag{5.76}$$

which combines the whole thing down to the form

$$T(z,t) = T_s + (T_m - T_s)\frac{z}{L} + \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi z}{L}\right) exp\left[-\kappa\left(\frac{n\pi}{L}\right)^2 t\right]. \tag{5.77}$$

The T_2 equation now has to satisfy eq. (5.67), which implies

$$\sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi z}{L}\right) = (T_m - T_s)\left(1 - \frac{z}{L}\right), \tag{5.78}$$

and thus we need to solve for the coefficients by integration!! It should be relatively obvious that $\sin(n\pi x/L)$ forms an orthogonal set on the interval zero to L! Orthogonal on $-L$ to L. Orthonormal on -1 to 1.

$$\int_0^L \sin\left(n\pi\frac{z}{L}\right)\sin\left(m\pi\frac{z}{L}\right)\,dz = \frac{L}{2}\delta_{nm}. \tag{5.79}$$

Try at home, find

$$a_n = \frac{2}{L}\int_0^L (T_m - T_s)\left(1 - \frac{z}{L}\right)\sin\left(\frac{n\pi z}{L}\right)\, dz \tag{5.80}$$

$$. \tag{5.81}$$

Note that $(-1)^{(n-1)}$ comes from $\cos(n\pi)$. This last bit needs an update to make it right. Also, should have mentioned integration by parts right in the first chapter, since we be using it here.

Reflect on the similarity of the solution. What when $L \to \infty$, we lose the discrete sum and go to the integral and we end back up with the erf?

5.6 The Adams-Williamson equation

I have good notes on that. Linking seismology to geodynamics.

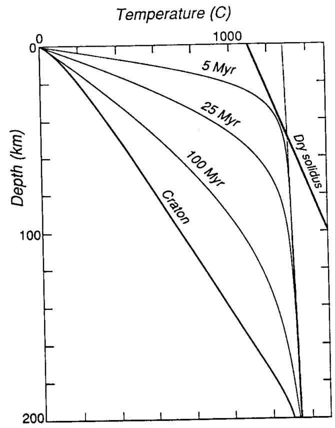


Figure 2. Representative geotherms, illustrating the cool thermal boundary layer at the earth's surface. Three oceanic geotherms are included, corresponding to sea-floor ages of 5, 25, and 100 Ma. The "craton" geotherm represents relatively thick continental lithosphere. A dry mantle solidus (McKenzie and Bickle 1988) is included, along with the extrapolation of the adiabat (gradient 0.3°K/km) to the surface.

Fig. 5.2.

Write down the advective term $\mathbf{v} \cdot \boldsymbol{\nabla} T$ at some point.

Write down as in Korenaga with Claude Herzberg some parameterized loss form.

5.7 What we've ignored

Super/supra adiabaticity?

Phase changes. Write something about that.

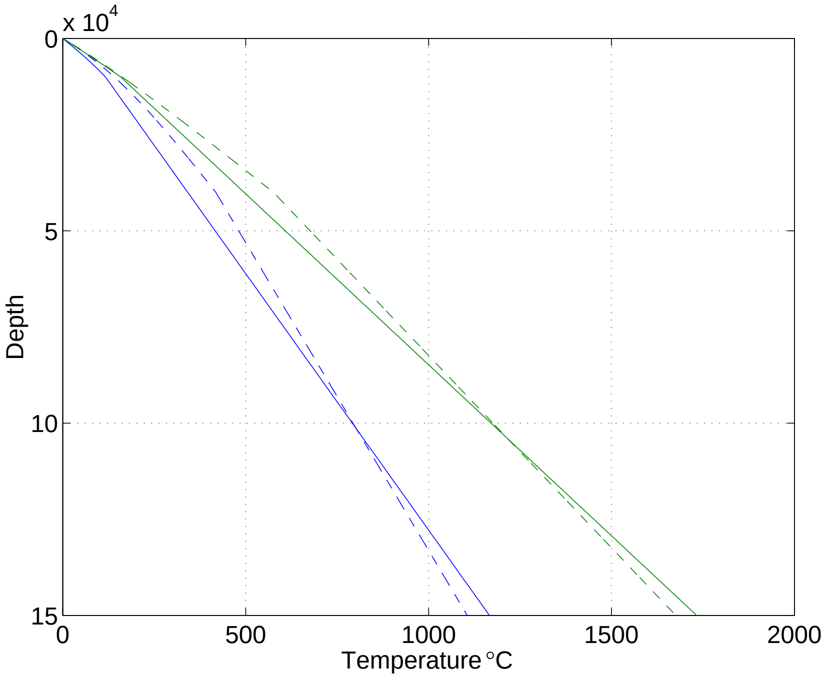


Fig. 5.3. Sclater

6
Topography

We've talked about it in so many words but not in so many words. Proper definition, choice words? Some first order questions: how deep the oceans, what happens when you load them with seamounts? What happens when you pull them down from a side?

What happens on continental lithosphere? Rember isostasy.

Topography *defined* by gravity. At the end we bring it all together and link topography *to* gravity in our special way.x

Got to begin with the big picture like we had for geoid, gravity, magnetism, (not heat flow yet!) and now topography. That could be a unifying theme, those awesome Mollweides? And knowing their spherical harmonic expansions? Refer to Mark W. Treatise.

6.1 How deep are the oceans?

Just cooling and isotasy.

Not waves. But still a relation that links time to space: the fact that the oceans spread, cool, thicken, and subside while they're at it.

Let us take the cooling of the halfspace one step further and see what happens, introduce **isostasy**. Probably should have mentioned it before, when we did gravity, but never mind. Archimedes lives on.

Return to eq. (5.56)

$$\boxed{T(z,t) = T_s + (T_m - T_s)\,\mathrm{erf}\left(\frac{z}{\sqrt{4\kappa t}}\right).} \tag{6.1}$$

Introduce coefficient of thermal volume expansion! then approximate it linearly

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P = -\frac{1}{\rho}\left(\frac{\partial \rho}{\partial T}\right)_P \tag{6.2}$$

$$= -\frac{1}{\rho_m}\frac{\rho - \rho_m}{T - T_m} \tag{6.3}$$

Little diagram. The concept of isostasy. The water depth is w, the growing thickness of the lithosphere is z_L. At the ridge, no water. At distance, have both. Something like $\alpha = 3.2 \times 10^{-5}$ per degree C.

$$\rho_m(w + z_L) = \rho_w w + \int_0^{z_L} \rho(z,t)\,dz \tag{6.4}$$

Work towards eq. (6.2), then use it, then substitute in eq. (6.1), hence end up with

$$w(\rho_w - \rho_m) = -\int_0^{z_L} [\rho(z,t) - \rho_m]\;dz \tag{6.5}$$

$$= \alpha\rho_m \int_0^{z_L} [T(z,t) - T_m]\;dz \tag{6.6}$$

$$= \alpha\rho_m(T_s - T_m)\int_0^{z_L}\left[1 - \operatorname{erf}\left(\frac{z}{\sqrt{4\kappa t}}\right)\right]\;dz \tag{6.7}$$

And now, with the complementary error function, we rewrite the before with (maybe introduce ΔT and $\Delta\rho$?

$$w = \frac{\alpha\rho_m(T_s - T_m)}{(\rho_w - \rho_m)}\int_0^{z_L} \operatorname{erfc}\left(\frac{z}{\sqrt{4\kappa t}}\right)\;dz \tag{6.8}$$

Remember, at z_L we have $T = T_m$ and $\rho = \rho_m$. Also, remember that $\int_0^\infty \operatorname{erfc}(q)\,dq = \pi^{-1/2}$, so how about we make a $z_L \to \infty$ approximation and we end up with

$$w(t) = \frac{2\sqrt{\kappa t}}{\sqrt{\pi}}\frac{\alpha\rho_m(T_s - T_m)}{(\rho_w - \rho_m)} \tag{6.9}$$

which is something like $350\sqrt{t}$ for $t < 70$ Ma. Of course, plus 2500 m, which is the depth to the ridge in the first place.

6.2 Flexure: It's Not Just Istostasy

Yes, but. Strength, dudes.

We have seen that upon rifting away from the MOR the lithosphere thickens (the base of the thermal lithosphere is defined by an isotherm, usually

$T_m \approx 1300°C$) and subsides, and that the cooled lithosphere is more dense than the underlying mantle. In other words, it forms a gravitationally unstable layer. Why does it stay atop the asthenosphere instead of sinking down to produce a more stable density stratification? That is because upon cooling the lithosphere also acquires **strength**. Its weight is supported by its strength; the lithosphere can sustain large stresses before it breaks. The initiation of subduction is therefore less trivial than one might think and our understanding of this process is still far from complete.

The strength of the lithosphere has important implications:

(i) it means that the lithosphere can support loads, for instance by seamounts
(ii) the lithosphere, at least the top half of it, is seismogenic
(iii) lithosphere does not simply sink into the mantle at trenches, but it *bends* or *flexes*, so that it influences the style of deformation along convergent plate boundaries.

Investigation of the bending or **flexure** of the plate provides important information about the mechanical properties of the lithospheric plate. We will see that the nature of the bending is largely dependent on the **flexural rigidity**, which in turn depends on the elastic parameters of the lithosphere and on the **elastic thickness** of the plate.

An important aspect of the derivations given below is that the thickness of the elastic lithosphere can often be determined from surprisingly simple observations and without knowledge of the actual load. In addition, we will see that if the bending of the lithosphere is relatively small the entire mechanical lithosphere behaves as an elastic plate; if the bending is large some of the deformation takes place by means of ductile creep and the part of the lithosphere that behaves elastically is thinner than the mechanical lithosphere proper.

6.2.0.1 Basic theory

To derive the equations for the bending of a thin elastic plate we need to

(i) apply laws for equilibrium: sum of the forces is zero and the sum of all moments is zero: $\sum F = 0$ and $\sum M = 0$
(ii) define the constitutive relations between applied stress σ and resultant strain ϵ
(iii) assume that the deflection $w \ll L$, the typical length scale of the system, and h, the thickness of the elastic plate $\ll L$. The latter criterion (#3) is to justify the use of **linear elasticity**.

In a 2D situation, i.e., there is no change in the direction of y, the bending of a homogeneous, elastic plate due to a load $V(x)$ can be described by the fourth-order differential equation that is well known in elastic beam theory in engineering:

$$D\frac{d^4w}{dx^4} + P\frac{d^2w}{dx^2} = V(x) \tag{6.10}$$

with $w = w(x)$ the deflection, i.e., the vertical displacement of the plate, which is, in fact, the ocean depth(!), D the **flexural rigidity**, and P a horizontal force.

The flexural rigidity depends on elastic parameters of the plate as well as on the thickness of the plate:

$$D = \frac{Eh^3}{12(1-\nu^2)} \tag{6.11}$$

with E the Young's modulus and ν the Poisson's ratio, which depend on the elastic moduli μ and λ (See Fowler, Appendix 2).

The bending of the plate results in **bending** (or *fiber*) **stresses** within the plate, σ_{xx}; depending on how the plate is bent, one half of the plate will be in compression while the other half is in extension. In the center of the plate the stress goes to zero; this defines the *neutral line or plane*. If the bending is not too large, the stress will increase linearly with increasing distance z' away from the neutral line and reaches a maximum at $z' = \pm h/2$. The bending stress is also dependent on the elastic properties of the plate and on how much the plate is bent; $\sigma_{xx} \sim$ elastic moduli $\times z' \times$ curvature, with the curvature defined as the (negative of the) change in the slope $d/dx(dw/dx)$:

$$\sigma_{xx} = -\frac{Eh^3}{1-\nu^2}z'\frac{d^2w}{dx^2} \tag{6.12}$$

This stress is important to understand where the plate may break (seismicity!) with normal faulting above and reverse faulting beneath the neutral line.

The *integrated effect* of the bending stress is the **bending moment** M, which results in the rotation of the plate, or a plate segment, in the $x - z$ plane.

$$M = \int_{\frac{h}{2}}^{-\frac{h}{2}} \sigma_{xx} z'\, dz' \tag{6.13}$$

Equation (6.10) is generally applicable to problems involving the bending of a thin elastic plate. It plays a fundamental role in the study of such problems as the folding of geologic strata, the development of sedimentary basins, the post-glacial rebound, the proper modeling of isostasy, and in the understanding of seismicity. In class we will look at two important cases: (1) loading by sea mounts, and (2) bending at the trench.

Before we can do this we have to look a bit more carefully at the dynamics of the system. If we apply bending theory to study lithospheric flexure we have to realize that if some load V or moment M causes a deflection of the plate there will be a hydrostatic restoring force owing to the replacement of heavy mantle material by lighter water or crustal rock. The magnitude of the restoring force can easily be found by applying the isostasy principle and the effective load is thus the applied load minus the restoring force (all per unit length in the y direction): $V = V_{\text{applied}} - \Delta\rho w g$ with w the deflection and g the gravitational acceleration. This formulation also makes clear that lithospheric flexure is in fact a compensation mechanism for isostasy! For oceanic lithosphere $\Delta\rho = \rho_m - \rho_w$ and for continental flexure $\Delta\rho = \rho_m - \rho_c$. The bending equation that we will consider is thus:

$$D\frac{d^4w}{dx^4} + P\frac{d^2w}{dx^2} + \Delta\rho w g = V(x) \tag{6.14}$$

6.3 Loading by sea mounts

How deep are the oceans at the smaller scales.

Let's assume a line load in the form of a chain of sea mounts, for example Hawaii.

Let V_0 be the load applied at $x = 0$ and $V(x) = 0$ for $x \neq 0$. With this approximation we can solve the homogeneous form of (6.14) for $x > 0$ and take the mirror image to get the deflection $w(x)$ for $x < 0$. If we also ignore the horizontal applied force P we have to solve

$$D\frac{d^4w}{dx^4} + \Delta\rho w g = 0 \tag{6.15}$$

The general solution of (6.15) is

$$w(x) = e^{\frac{x}{\alpha}}\left\{A\cos\frac{x}{\alpha} + B\sin\frac{x}{\alpha}\right\} + e^{-\frac{x}{\alpha}}\left\{C\cos\frac{x}{\alpha} + D\sin\frac{x}{\alpha}\right\} \tag{6.16}$$

with α the flexural parameter, which plays a central role in the extraction of structural information from the observed data:

$$\alpha = \left(\frac{4D}{\Delta\rho g}\right)^{\frac{1}{4}} \tag{6.17}$$

The constants $A - D$ can be determined from the boundary conditions. In this case we can apply the general requirement that $w(x) \to 0$ for $x \to \infty$ so that $A = B = 0$, and we also require that the plate be horizontal directly beneath $x = 0$: $dw/dx = 0$ for $x = 0$ so that $C = D$: the solution becomes

$$w(x) = Ce^{-\frac{x}{\alpha}} \left\{\cos\frac{x}{\alpha} + \sin\frac{x}{\alpha}\right\} \tag{6.18}$$

From this we can now begin to see the power of this method. The deflection w as a function of distance is an oscillation with period x/α and with an exponentially decaying amplitude. This indicates that we can determine α directly from observed bathymetry profiles $w(x)$, and from equations (6.18) and (6.11) we can determine the elastic thickness h under the assumption of values for the elastic parameters (Young's modulus and Poisson's ratio). The flexural parameter α has a dimension of distance, and defines, in fact, a typical length scale of the deflection (as a function of the "strength" of the plate).

The constant C can be determined from the deflection at $x = 0$ and it can be shown (Turcotte & Schubert) that $C = (V_0\alpha^3)/(8D) \equiv w_0$, the deflection beneath the center of the load. The final expression for the deflection due to a line load is then

$$w(x) = \frac{V_0\alpha^3}{8D} e^{-\frac{x}{\alpha}} \left\{\cos\frac{x}{\alpha} + \sin\frac{x}{\alpha}\right\} \qquad x \geq 0 \tag{6.19}$$

Let's now look at a few properties of the solution:

- The half-width of the depression can be found by solving for $w = 0$. From (6.19) it follows that $\cos(x_0/\alpha) = -\sin(x_0/\alpha)$ or $x_0/\alpha = \tan^{-1}(-1) \Rightarrow x_0 = \alpha(3\pi/4 + n\pi)$, $n = 0, 1, 2, 3\ldots$ For $n = 0$ the half-width of the depression is found to be $\alpha 3\pi/4$.
- The height, w_b, and location, x_b , of forebulge $\Rightarrow$ find the optima of the solution (6.19). By solving $dw/dx = 0$ we find that $\sin(x/\alpha)$ must be zero $\Rightarrow x = n\pi\alpha$, and for those optima $w = w_0 e^{-n\pi}$, $n = 0, 1, 2, 3\ldots$ For the location of the forebulge: $n = 1$, $x_b = \pi\alpha$ and the height of the forebulge $w_b = -w_0 e^{-\pi}$ or $w_b = -0.04w_0$ (very small!).

Important implications: The flexural parameter can be determined from the location of either the zero crossing or the location of the forebulge. No need to know the magnitude of the load! The depression is narrow for small α, which means either a weak plate or a small elastic thickness (or both); for a plate with large elastic thickness, or with a large rigidity the depression is very wide. In the limit of very large D the depression is infinitely wide but the amplitude w_0, is zero $\Rightarrow$ no depression at all! Once α is known, information about the central load can be obtained from Eq. (6.19)

Note: the actual situation can be complicated by lateral variations in thickness h, fracturing of the lithosphere (which influences D), compositional layering within the elastic lithosphere, and by the fact that loads have a finite dimensions.

6.4 Flexure at a deep sea trench

With increasing distance from the MOR, or with increasing time since formation at the MOR, the oceanic lithosphere becomes increasingly more dense and if the conditions are right† this gravitational instability results in the subduction of the old oceanic plate. The gravitational instability is significant for lithospheric ages of about 70 Ma and more. We will consider here the situation after subduction itself has been established; in general the plate will not just sink vertically into the mantlebut it will bend into the trench region.

This bending is largely due to the gravitational force due to the negative buoyancy of the part of the slab that is already subducted M_0. For our modeling we assume that the bending is due to an end load V_0 and a bending moment M_0 applied at the tip of the plate. As a result of the bending moment the slope $dw/dx \neq 0$ at $x = 0$ (note the difference with the seamount example where this slope was set to zero!). The important outcome is, again, that the parameter of our interest, the elastic thickness h, can be determined from the shape of the plate, in vertical cross section, i.e. from the bathymetry profile $w(x)$!, in the subduction zone region, without having to know the magnitudes of V_0 and M_0.

We can use the same basic equation (6.15) and the general solution (6.16) (with $A = B = 0$ for the reason given above)

$$w(x) = e^{-\frac{x}{\alpha}} \left\{ C \cos \frac{x}{\alpha} + D \sin \frac{x}{\alpha} \right\} \tag{6.20}$$

† Even for old oceanic lithosphere the stresses caused by the increasing negative buoyancy of the plate are not large enough to break the plate and initiate subduction. The actual cause of subduction initiation is still not well understood, but the presence of pre-existing zones of weakness (e.g. a fracture zone, thinned lithosphere due to magmatic activity — e.g. an island arc) or the initiation of bending by means of sediment loading have all been proposed (and investigated) as explanation for the triggering of subduction.

but the boundary conditions are different and so are the constants C and D. At $x = 0$ the bending moment‡ is $-M_0$ and the end load $-V_0$. It can be shown (Turcotte & Schubert) that the expressions for C and D are given by

$$C = (V_0\alpha + M_0)\frac{\alpha^2}{2D} \quad \text{and} \quad D = -\frac{M_0\alpha^2}{2D} \tag{6.21}$$

so that the solution for bending due to an end load and an applied bending moment can be written as

$$w(x) = \frac{\alpha^2 e^{-x/\alpha}}{2D}\left\{(V_0\alpha + M_0)\cos\frac{x}{\alpha} - M_0 \sin\frac{x}{\alpha}\right\} \tag{6.22}$$

We proceed as above to find the locations of the first zero crossing and the fore bulge, or **outer rise**.

$$w(x) = 0 \quad \Rightarrow \quad \tan(x_0/\alpha) = 1 + \alpha V_0/M_0 \tag{6.23}$$

$$dw/dx = 0 \quad \Rightarrow \quad \tan(x_b/\alpha) = -1 - 2M_0/\alpha V_0 \tag{6.24}$$

In contrast to similar solutions for the sea mount loading case, these expressions for x_0 and x_b still depend on V_0 and M_0. In general V_0 and M_0 are unknown. They can, however, be eliminated, and we can show the dependence of $w(x)$ on x_0 and x_b, which can both be estimated from the nathymetry profile. A perhaps less obvious but elegant way of doing this is to work out $\tan(1/\alpha(x_0 - x_b))$. Using sine and cosine rules (see Turcotte & Schubert, 3.17) one finds that

$$\tan\left(\frac{x_b - x_0}{\alpha}\right) = 1 \tag{6.25}$$

so that $x_0 - x_b = (\pi/4 + n\pi)\alpha$, $n = 0, 1, 2, 3, \ldots$ For $n = 0$ one finds that $\alpha = 4(x_0 - x_b)/\pi$, so that the elastic thickness h can be determined if one can measure the horizontal distance between x_0 and x_b.

After a bit of algebra one can also eliminate α to find the deflection $w(x)$ as a function of w_b, x_0, and x_b. The normalized deflection w/w_b as a function of normalized distance $(x - x_0)/(x_b - x_0)$ is known as the **Universal Flexure Profile**.

‡ At this moment, it is important that you go back to the original derivation of the plate equation in Turcotte & Schubert and realize they obtained their results with definite choices as to the signs of applied loads and moments — hence the negative signs.

$$\frac{w(x)}{w_b} = \sqrt{2}e^{\frac{\pi}{4}} \exp\left\{-\frac{\pi}{4}\left(\frac{x - x_0}{x_b - x_0}\right)\right\} \sin\left\{-\frac{\pi}{4}\left(\frac{x - x_0}{x_b - x_0}\right)\right\} \qquad (6.26)$$

In other words, there is a unique way to bend a laterally homogeneous elastic plate so that it goes through the two points $(x_0, 0)$ and (x_b, w_b) with the condition that the slope is zero at $x = x_b$. The example of the Mariana trench shown in Figure **??** demonstrates the excellent fit between the observed bathymetry and the prediction after Eq. (6.26) (for a best fitting elastic thickness h as determined from the flexural parameter calculated from equation (6.25).

6.5 How deep are mountains?

After two simple ocean models, something for the continents.

Bring in the cartoon of the Himalayas but now with three end members: infinitely stiff, weak, in-between.

The bending of a homogeneous, elastic plate due to a load $V(x)$ can be describedby a fourth-order differential equation that is well known in elastic beam theory. We assume a 2D situation, i.e., there is no change of any physical properties in the direction of y.

$$D\frac{d^4w}{dx^4} + H\frac{d^2w}{dx^2} + \Delta\rho g w = V(x) \qquad (6.27)$$

with $w = w(x)$ the deflection, i.e., the vertical displacement of the plate, D the **flexural rigidity**, and H a horizontal force (see Figure **??**). The restoring force is given by $\Delta\rho g w$, where the density difference amounts to $\rho_m - \rho_c$ for a crust-mantle (continental interface) and $\rho_m - \rho_w$ in the oceanic case. The flexural rigidity depends on elastic parameters of the plate as well as on its thickness:

$$D = \frac{Eh^3}{12(1-\nu^2)}, \qquad (6.28)$$

where E is Young's modulus and ν Poisson's ratio. Both depend on the elastic moduli μ and λ.

Let's consider a periodic load due to topography h with maximum amplitude h_0 and wavelength λ: $h = h_0 \sin(2\pi x/\lambda)$. The corresponding load is then given by

$$V(x) = \rho_c g h_0 \sin\left(\frac{2\pi x}{\lambda}\right). \qquad (6.29)$$

so that the flexure equation, in the absence of any horizontal (tectonic) forces, becomes, in standard form,

$$\frac{d^4w}{dx^4} + \frac{\rho_m - \rho_c}{D} gw = \frac{\rho_c g h_0}{D} \sin\left(\frac{2\pi x}{\lambda}\right). \tag{6.30}$$

This fourth-order differential equation is easy to solve. Here's a walk-through. Replacing $k = 2\pi/\lambda$, representing the differential operator as $\mathcal{D}$, and introducing the constants A and B we can represent eq. 6.30 as:

$$(\mathcal{D}^4 + A)w = B\sin(kx). \tag{6.31}$$

Exponentials of the general form e^{kx} are eigenfunctions of the $\mathcal{D}^4$ with eigenvalue k^4, because $\mathcal{D}^4 e^{kx} = k^4 e^{kx}$. This leads to the condition on k that $k^4 = -A$, which has four complex roots $k_{i=1\to4}$. The function

$$w_H = \sum_{i=1}^{4} C_i e^{k_i x} \tag{6.32}$$

solves the homogeneous equation $(\mathcal{D}^4 + A)w_H = 0$ for four undetermined coefficients $C_{1\to4}$ and the general solution $w_H + w_P = 0$ will solve the entire equation 6.31 if $(\mathcal{D}^4 + A)w_P = B\sin(kx)$. With w_P of the same form as the right-hand side of the original equation, i.e., $w_P = C_5\sin(kx)$,

$$C_5(k^4\sin(kx) + A\sin(kx) = B\sin(kx)) \tag{6.33}$$

must hold†. Therefore, C_5 is given by

$$C_5 = \frac{B}{k^4 + A}. \tag{6.34}$$

Hence, the general solution is obtained as

$$w = w_H + w_P = \sum_{i=1}^{4} C_i e^{k_i x} + \frac{B}{k^4 + A}\sin(kx). \tag{6.35}$$

† We've just illustrated the general fact that for non-homogeneous ordinary differential equations, the general solution is given by a combination of the solution to the homogeneous equation and one particular solution

Note that the determination of C_5 has nothing to do with the boundary conditions of the equation, but eq. 6.35 needs to be adjusted so the boundary conditions are valid‡. Since the loading is periodic in x, its is clear that the response pr deflection of the lithosphere will also vary sinusoidally in x with the same wavelength as the topography. Therefore, $C_{1\to 4} = 0$, and the solution is given by:

$$\boxed{w(x) = \frac{\rho_c g h_0}{Dk^4 + \Delta\rho g}\sin(kx).} \tag{6.36}$$

The amplitude of the deflection can be rewritten as follows:

$$w_0 = \frac{h_0}{\frac{\rho_m}{\rho_c} - 1 + \frac{D}{\rho_c g}\left(\frac{2\pi}{\lambda}\right)^4}. \tag{6.37}$$

In the short wavelength range

$$\lambda \ll 2\pi\left(\frac{D}{\rho_c g}\right)^{1/4}, \tag{6.38}$$

the denominator of eq. 6.37 becomes dominant, and the deflecton $w_0 \ll h_0$. The same happens for very a large flexural rigidity D (or very large elastic thickness h of the plate). In other words, short-wavelength loads cause virtually no deformation of the lithosphere. In contrast, for very long wavelengths ($\lambda \gg 2\pi(D/\rho_c g)^{1/4}$) or for a very weak (or thin) plate the maximum depression becomes

$$w = w_{0\infty} \approx \frac{\rho_c h_0}{\rho_m - \rho_c}, \tag{6.39}$$

which is the same as for a completely compensated mass (see Eq. **??**). In other words, the plate "has no strength" for long wavelength loads: the topography is fully compensated. The degree of compensation can be defined as the ratio of the deflection to the maximum hydrostatic deflection in the Airy-case. In function of the wavelength of the topography, this is plotted for varying elastic thicknesses in Fig. 6.1.

The importance of this formulation is evident if you realize that topography can be described by a Fourier series of periodic functions with different

‡ We are in fact, developing the *method of undetermined coefficients*. It works when the right-hand side of the differential equation itself (in our case, $\sin(kx)$) is a solution to a homogeneous linear equation with constant coefficients, that is, of the general form $f(x) = x^l e^{rx}$.

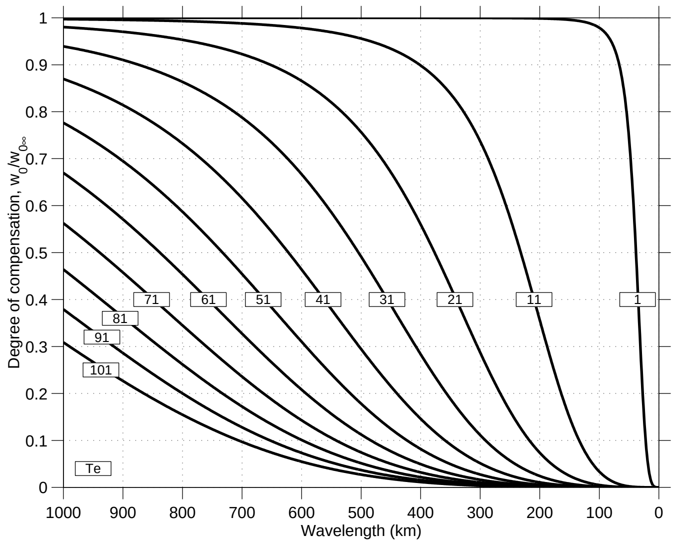


Fig. 6.1. Degree of compensation for varying elastic thicknesses.

wavelengths. One can thus use Fourier analysis to investigate the depression or compensation of any shape of load.

6.6 How strong is the crust?

Now we need to relate the last chapter to the first! And we've already used the middle ones!

7
Afterlude

This has been a story of differential geometry and what it's used for. Div, grad, curl, and all that! And of differential equations. Heat (parabolic, remember the quadratic geotherm!), $u_t = u + xx$. Wave (hyperbolic, remember the MOVEOUT curve!), $u_{tt} = u_{xx}$ and Laplaces' (elliptic, remember the ellipsoid!), $u_{xx} + u_{yy}$. We talked about boundary conditions more than about initial conditions, and about steady state more than about time dependence.

In that case we really should solve, at least once, 1/ Laplace's equation, perhaps like Kaula, short and to the poit. 2/ The wave equation (did that). 3/ The heat equation (we did that).

Every chapter to conclude with what the "Fundamental thing" was. For gravity, the fundamental theorem of geodesy, and how whole books are written about that. (Hoffman etc). For magnetism, the magnetic induction equation, and whole books! Hollerbach paper? For heat, the heat diffusion equation. (Carlslaw?)

Take a good look at Davies' bookS. Relook at Anderson Theory of the Earth.

Alternative table of contents: The vector calculs you apply learn. The differential equations that you learn.

Bibliography

R. Snieder, *A Guided Tour of Mathematical Methods for the Physical Sciences*, Cambridge Univ. Press, Cambridge, UK, 2 ed., 2004.

W. Lowrie, *A Students' Guide to Geophysical Equations*, Cambridge Univ. Press, Cambridge, UK, 1 ed., 2011.

H. M. Schey, *Div, grad, curl, and all that*, W. W. Norton, New York, 4 ed., 2005.

W. Appel, *Mathematics for Physics and Physicists*, Princeton Univ. Press, Princeton, N.J., 2007.

W. I. Newman, *Mathematical Methods for Geophysics and Space Physics*, Princeton Univ. Press, Princeton, N.J., 2016.

G. Strang, *Introduction to Applied Mathematics*, Wellesley-Cambridge, Wellesley, Mass., 1986.

W. Lowrie, *Fundamentals of Geophysics*, Cambridge Univ. Press, Cambridge, UK, 2 ed., 2007.

F. A. Dahlen and J. Tromp, *Theoretical Global Seismology*, Princeton Univ. Press, Princeton, N.J., 1998.

R. J. Blakely, *Potential Theory in Gravity and Magnetic Applications*, Cambridge Univ. Press, New York, 1995.

M. A. van der Sluijs, "A common miscitation of william gilbert," *Eos Trans. AGU* **95**(16), pp. 137, doi: 10.1002/2014EO160005, 2014.

M. Kono, "Geomagnetism in perspective," in *Treatise on Geophysics*, M. Kono, ed., **5**, pp. 1–31, doi: 10.1016/B978–044452748–6.00086–9, Elsevier, Amsterdam, Neth., 2007.

R. Hollerbach, "On the theory of the geodynamo," *Phys. Earth Planet. Inter.* **98**(3–4), pp. 163–185, doi: 10.1016/S0031–9201(96)03185–8, 1996.

R. A. Langel and W. J. Hinze, *The Magnetic Field of the Earth's Lithosphere: The Satellite Perspective*, Cambridge Univ. Press, Cambridge, UK, 1998.

E. Thébault, C. C. Finlay, C. D. Beggan, P. Alken, J. Aubert, O. Barrois, F. Bertrand, T. Bondar, A. Boness, L. Brocco, E. Canet, A. Chambodut, A. Chulliat, P. Coïsson, F. Civet, A. Du, A. Fournier, I. Fratter, N. Gillet, M. H. Brian Hamilton, G. Hulot, T. Jager, M. Korte, W. Kuang, X. Lalanne, B. Langlais, J.-M. Léger, V. Lesur, F. J. Lowes, S. Macmillan, M. Mandea, C. Manoj, S. Maus, N. Olsen, V. Petrov, V. Ridley, M. Rother, T. J. Sabaka, D. Saturnino, R. Schachtschneider, O. Sirol, A. Tangborn, A. Thomson, L. Tøffner-Clausen, P. Vigneron, I. Wardinski, and T. Zvereva, "International Geomagnetic Reference Field: the 12th generation," *Earth Planets Space* **67**, pp. 79, doi: 10.1186/s40623–015–0228–9, 2015.

S. Maus, C. Manoj, J. Rauber, I. Michaelis, and H. Lühr, "NOAA/NGDC candidate models for the 11th generation International Geomagnetic Reference Field and the concurrent release of the 6th generation Pomme magnetic model," *Earth Planets Space* **62**, pp. 729–735, doi: 10.5047/eps.2010.07.006, 2010.

F. E. Richart, J. R. Hall, and R. D. Woods, *Vibrations of Soils and Foundations*, Theoretical and Applied Mechanics, Prentice-Hall, Englewood Cliffs, New Jersey, 1970.

P. England, P. Molnar, and F. Richter, "John Perry's neglected critique of kelvin's age for the Earth: A missed opportunity in geodynamics," *GSA Today* **17**(1), pp. 4–9, doi: 10.1130/GSAT01701A.1, 2007.

D. L. Turcotte and E. R. Oxburgh, "Finite amplitude convective cells and continental drift," **28**(1), pp. 29–42, doi: 10.1017/S0022112067001880, 1967.

D. P. McKenzie, "Some remarks on heat flow and gravity anomalies," *J. Geophys. Res.* **72**(24), pp. 6261–6273, doi: 10.1029/JZ072i024p06261, 1967.