Potential Field Methods of Geophysical Exploration

Potential Field Methods of Geophysical Exploration

OPEN EDUCATIONAL RESOURCES FOR GEOL 7330 AT UH

JIAJIA SUN

XINYAN LI, FELICIA NURINDRAWATI, XIAOLONG WEI, AND KENNETH LI

Potential Field Methods of Geophysical Exploration by Jiajia Sun is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, except where otherwise noted.

Contents

CHAPTER 1

Chapter 1: Introduction

JIAJIA SUN, XINYAN LI, AND FELICIA NURINDRAWATI

1.1 LEARNING OBJECTIVES

Hello, everyone. Welcome to the potential field class! In this class, we will focus on the fundamental theory and commonly used data processing and interpretation techniques in potential field methods. This class consists of both lectures and lab exercises. After completion of the class, students can expect to

- Understand the fundamental theory behind potential field methods;
- Understand gravity and magnetic data acquisition practices, instrumentation and processing procedures;
- Understand various data processing techniques in Fourier domain;
- Be familiar with various interpretation methods for depth estimates;
- Be able to use Geosoft software to perform 2D basin modeling;

- Understand the 3D inversion theory and methods for potential field data; and
- Be familiar with various applications of potential field methods.

The target audience of this course is anyone who is interested in learning more about

- fundamental potential field theory;
- Processing and interpretation methods; and
- Application examples of potential field methods to
- Petroleum;
- Mineral;
- Geothermal;
- Geotechnical/Engineering (e.g., caves and tunnels);
- Regional to continental-scale geological studies; and
- Environmental (e.g., unexploded ordnance UXO)

1.2 POTENTIAL

Both gravity and magnetic fields can be described using a quantity called potential. Potential is a scalar and obeys Laplace's equation in source-free regions. Therefore, these two methods are collectively referred to as potential field methods in the geophysics community.

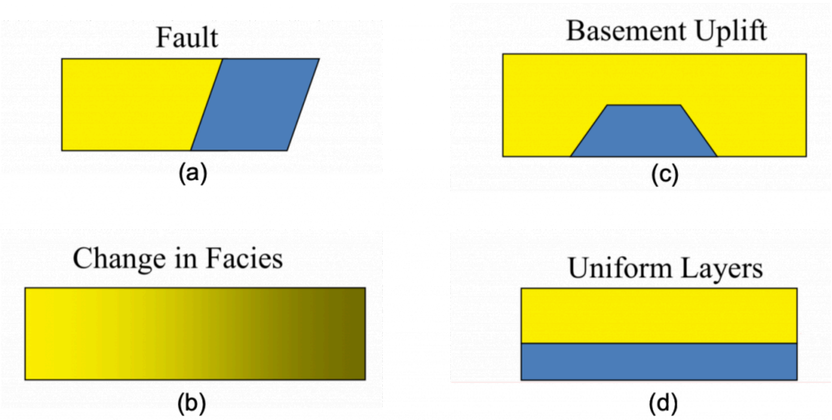
The fact that potentials are scalar quantities makes it easier to deal with the mathematics involved. After all, working with scalars is always easier than working with vectors. Note that, both gravity and magnetic fields are vector fields because gravity not only has a magnitude (such as 9.8 m/s^2) but also has a direction; the same is true for the magnetic field.

1.2.1 GEOPHYSICAL SIGNALS

Geophysical signals are mostly produced by **contrasts** in some physical properties of Earth materials. For example:

- Variations in electrical conductivities of the subsurface rocks produce secondary potentials that can be measured in a DC survey and used for subsurface imaging;
- Differences in the reflectivity/absorption of near-surface Earth materials result in useful remote sensing images;
- Changes in acoustic impedance between Earth materials result in seismic reflections.

Specifically, in gravity & magnetic methods, the material properties to be considered are density and magnetization (magnetic susceptibility), respectively. Signals result from ***horizontal*** changes in the density and magnetization of the Earth materials can be used for detecting *lateral changes* produced by either *vertical displacement* of homogeneous layers or by *inherent lateral inhomogeneities* in the Earth materials themselves, as model examples shown in figure (a, b, c) below. However, horizontal layers with constant properties (figure (d)) provide no signal and consequently are “invisible” to the gravity and magnetic methods.



Changes in density ad magnetization of the earth materials. Image courtesy of Stuart Hall at UH.

In fact, we should be aware that, the above statements are only true for **surface** gravity and magnetic measurement. If there are borehole gravity and magnetic measurements, potential field methods can be used for detecting physical property changes in the vertical direction. More details on borehole gravity can be found in https://wiki.aapg.org/Borehole_gravity, and some borehole gravity applications can be found here.

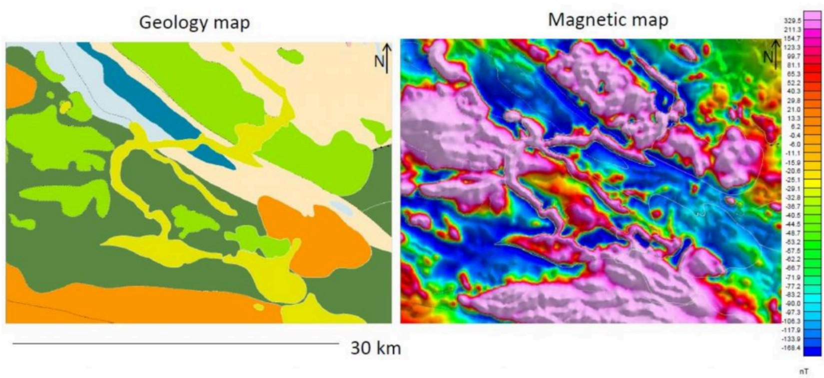
1.3 APPLICATION EXAMPLES

There are many applications of potential field methods. In the following, we will discuss some major geoscience applications.

1.3.1 APPLICATION TO OIL, GAS AND GROUNDWATER STUDIES

Example 1: Geologic mapping

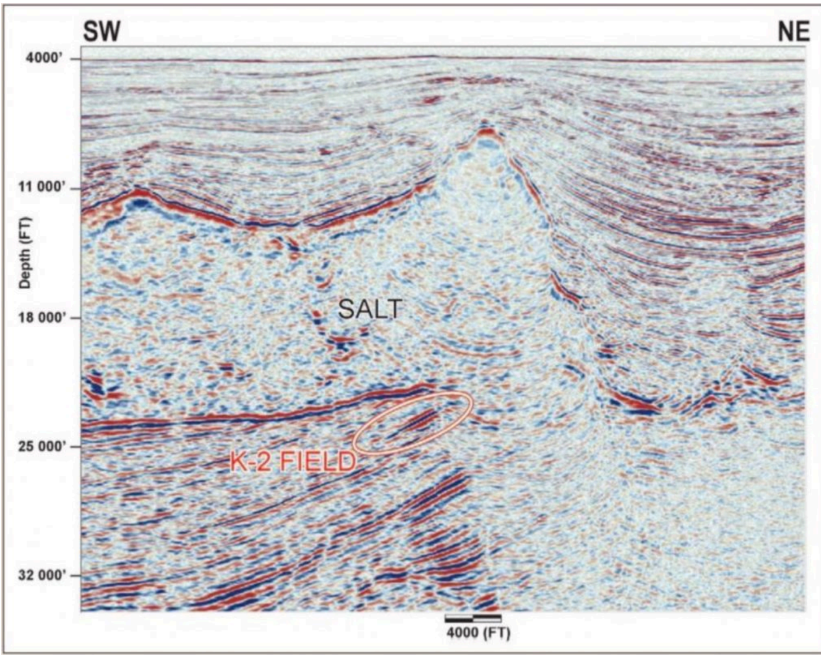
Geologic contacts can be quickly inferred from the magnetic maps, as illustrated in the figure below (image credit from Douglas Oldenburg at UBC).



Inferring from magnetic map to geologic contacts. Image credit: Douglas W. Oldenburg at UBC.

Example 2: Salt dome imaging

In many Gulf of Mexico prospects, salt plays a key role in acting as a structural trap. Overhanging salt often forms seals, and sediments on salt flanks can have structural and stratigraphic pinch-outs against the salt. The exact shape of the salt is critical in understanding these traps. Unfortunately, seismic imaging often tends to be poor in these prospects. The seismic image from O'Brien et al., 2005, TLE, illustrates the limitation of interpreting the salt dome using seismic imaging alone.

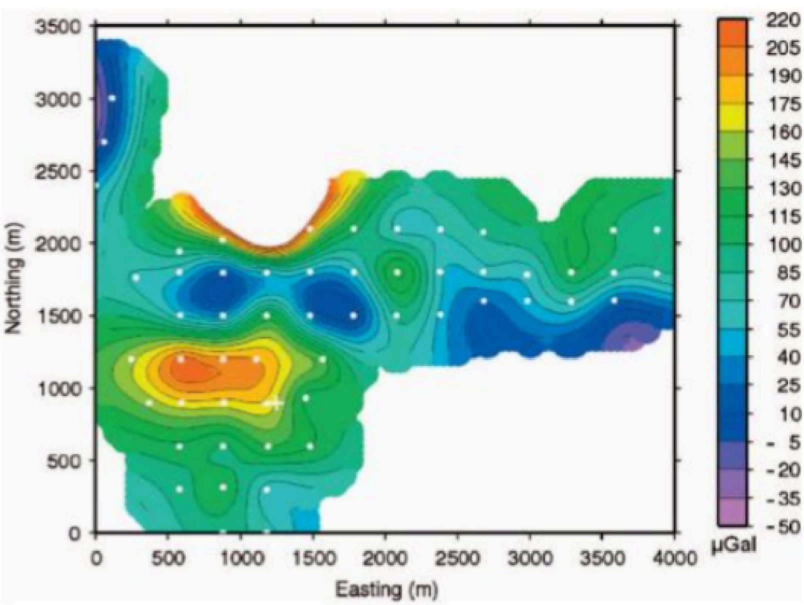


Prestack depth-migration profile through the K-2 salt body. Note the excellent image definition underlying the tabular portion of the salt body. However, the seismic image underlying the peak of the salt is very poor and interpretation of the base of salt and subsalt section is uncertain in this area, (O'Brien et al., 2005, TLE)

Example 3: Groundwater studies

Time-lapse gravity data, also known as 4D gravity data (with the 4th dimension being time) is often used to monitor the movement of water underground. Time-lapse data can be obtained by conducting multiple surveys in the same location but at different times. Besides gravity data, there are other methods that can be used to monitor groundwater movement. However, due to the delay in signal responses of the groundwater movement, some methods might be less effective than others. According to a study conducted by [Kenedy, 2016](), time-lapse gravity is “the most sensitive to movement of water through an unsaturated zone”.

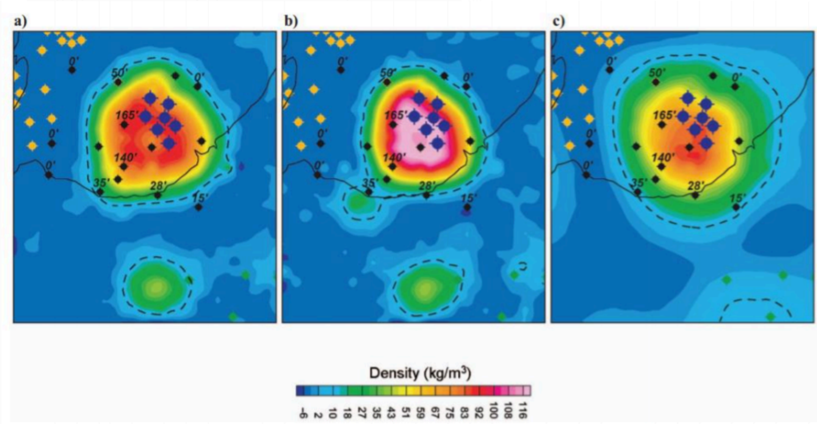
One example of groundwater monitoring with time-lapse gravity approach is shown in [Davis & Li, 2008](). In this paper, the movement of water injected into an underground artificial aquifer storage system is monitored using time-lapse microgravity data (meaning that the data obtained is in the orders of micro-Gal). The following figure shows the total gravity difference between April 2004 and February 2005, and the white (+) indicates the location of the injection well. We see that there is a positive gravity difference around the injection well, spanning towards the north-west side, which is the direction of the water injected into the reservoir.



Example 4: Reservoir waterflood surveillance

One of the most notable use of the 4D microgravity method is the time-lapse microgravity survey at Prudhoe Bay, Alaska ([Hare,]()

2008) In this case, the gas production in the field started to decline. It was suspected that there might be something wrong with the water injection which is used to maintain pressure, and therefore maintain production. The problem is that there were only a few wells available in the area that can be used to monitor the water movement in the reservoir. With that, they considered the use of gravity survey to monitor the waterflood movement in the reservoir. The figure below shows density models obtained through inversion (with each figure showing different inversion results using different inversion constraints). The method was successful in determining the behavior of the waterflood movement in the reservoir.



*If a denser feature replaces the position of a feature that is less dense, it will create a **positive** density contrast. If we inject water to a reservoir, the water would move and replaces the position of less denser features (gas/air). Thus, when water replaces gas, it will create a positive anomaly in our data.*

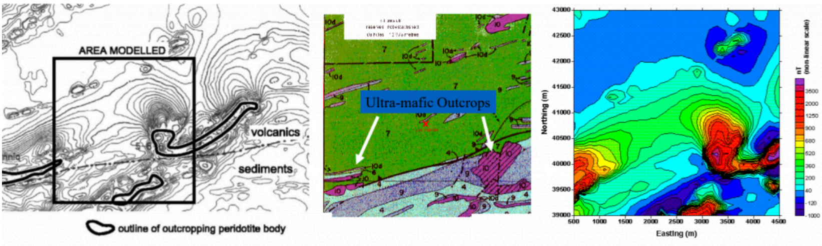
1.3.2 APPLICATION TO MINERAL EXPLORATION

Gravity and magnetic survey are especially useful in mineral explorations, as they have useful features that are not offered by seismic surveys, such as how gravity and magnetic surveys can cover a massive geographical region in just a couple of days with surveys conducted using airplanes. Examples for the use of gravity and magnetic surveys in mineral exploration are:

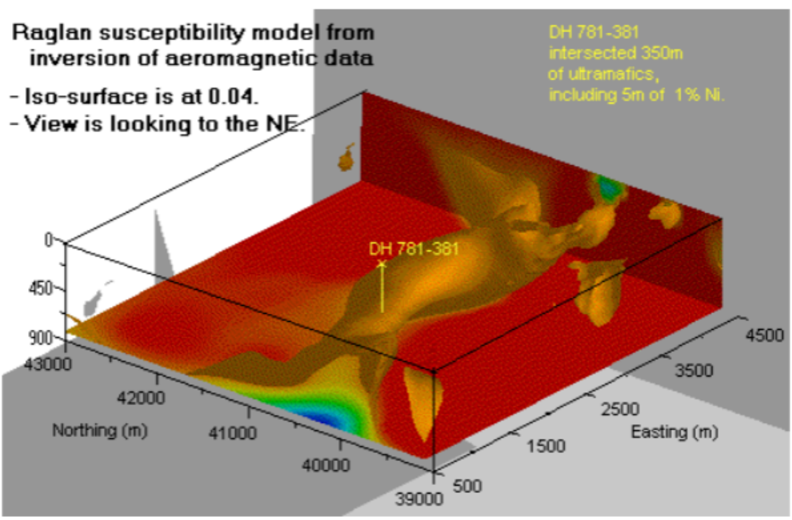
- **Sulfide exploration**

Massive sulfide deposits that came from volcanic sources and are rich in copper, zinc, lead and can contain other precious metals such as gold and silver (mining.com). The deposit is very high in density in comparison to the background, and thus a gravity survey is generally done for these types of target. The following figure is the gravity data obtained from a massive sulfide as well as the cross-section of the interpreted geology (credit: Yaoguo Li)

Due to the magnetic properties of the massive sulfide, magnetic surveys can also be done for this type of target. The following figure shows the aeromagnetic data of the Raglan deposit, a well-known nickel copper deposit in Canada (Watts, 1997). One of the questions posed for this region is whether the peridotite outcrops (appear in purple on the surface geologic map) are connected at depth or not. This is an important information to know the optimum location for drilling.

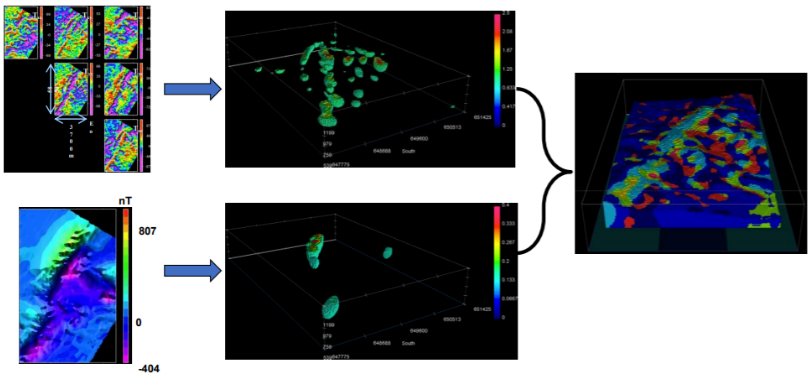


After inversion of the magnetic data, it is found that the outcrops are connected at depth and the drill hole can then be placed according to the interpretation from the inversion results.



- **Iron exploration**

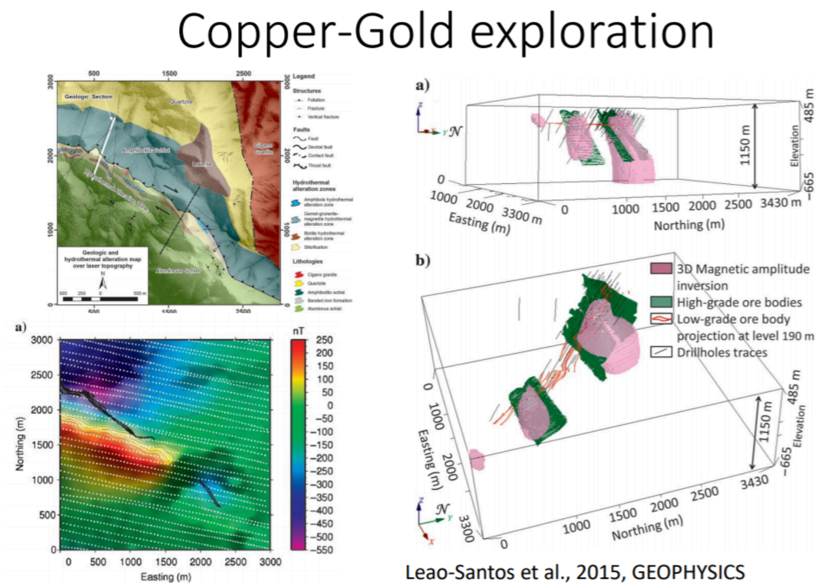
Other examples of gravity and magnetic data used in mineral exploration would be Iron exploration (Martinez and Li, 2015). The following Figure is the gravity gradient tensor data and the magnetic data for a geographical region as well as the inversion results.



Martinez and Li, 2015, Interpretation

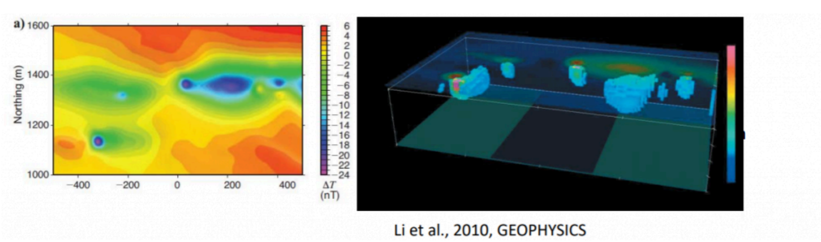
- **Copper-Gold exploration**

Another example is from Copper-Gold exploration (Leao-Santos, 2015). The following figure shows the geologic map as well as the magnetic data of the region. We see that the overall trend in the geologic map is also apparent in the magnetic data map. Figure (a) and (b) showcases the 3D inversion results of the magnetic data and the interpretation based on the inversion result.



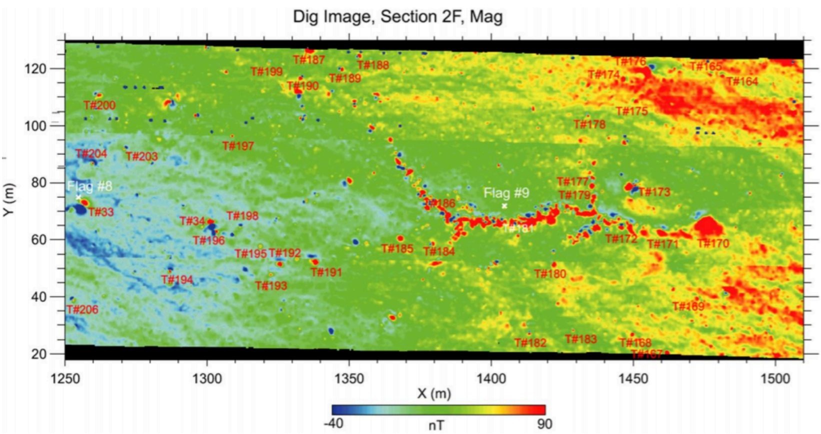
- **Diamond Exploration**

Diamonds are typically found in kimberlite pipes, which are igneous rocks with magnetic properties. From the following figure, we can see that the Total Magnetic Intensity (TMI) data has outlined an elongated body indicated with the negative anomaly values in Figure (a). After inverting the magnetic data, we can see the two broad elongated sources, interpreted as the kimberlite dikes. The vertically oriented bodies represents the kimberlite pipes in the region.



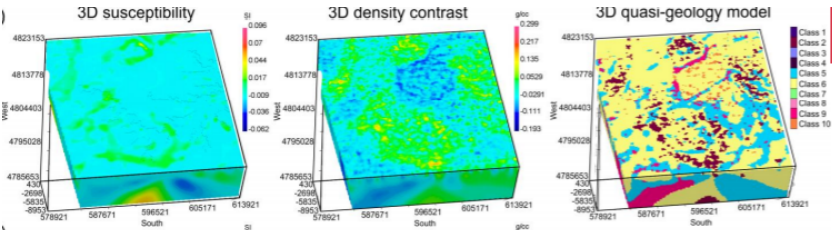
- **UXO Detection**

UXO (Unexploded ordnance) are explosive weapons, such as bombs, missiles, grenades, that did not explode during deployment, but poses high risk of detonation even decades after it was buried underground. Fortunately, UXOs are highly magnetic and thus we can detect them by conducting magnetic surveys over the region of interest. The following are examples of UXO data (credit: Yaoguo Li)



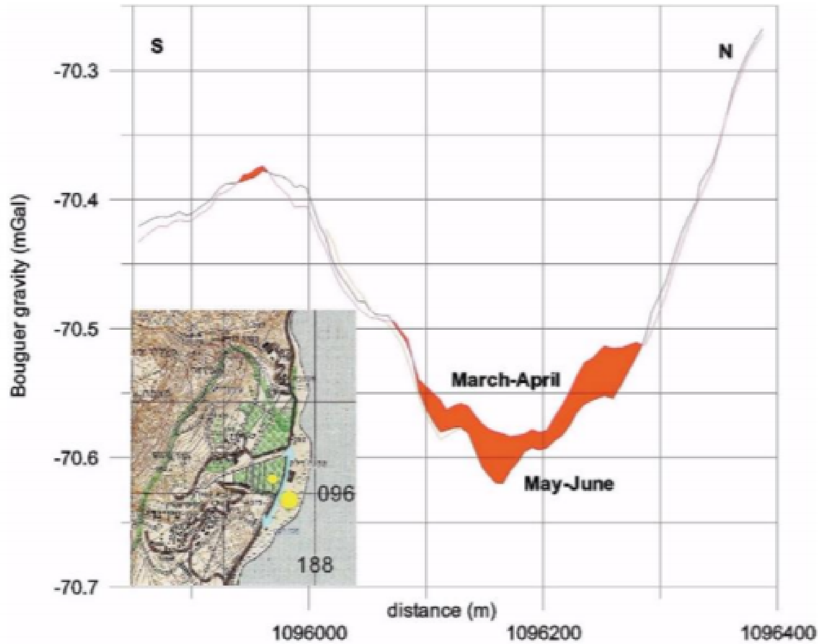
1.3.3 BASEMENT CHARACTERIZATION

Basement characterizations are done by interpreting the basement geology of a certain region based on the data obtained in the survey. Gravity and magnetic surveys can cover extensive geographical regions in a matter of days (using airplane or helicopter). Using inversion, the distribution of density and the magnetic susceptibility of the region can be found. Potential resources can be interpreted based on the recovered physical parameters, as certain geological units have a known range of density and magnetic susceptibility values. The following figure shows a 3D magnetic susceptibility model and 3D density contrast from inversion results and the interpreted geological units based on the the inversion results.



1.3.4 CAVE DETECTION EXAMPLE

Sinkholes are a serious environmental problem and hazard, especially around the Dead Sea area. The largest sinkhole found in the region was as large as 30m in diameter and 15 m deep. The recent collapse on both sides of the main highway at E'n Gedi, a popular destination for recreation and tourism, poses a risk for many people in the region. The sinkholes initially opened at a campsite east of the road, and subsequently, several sinkholes opened west of the road, which indicates that there might be a possible collapse of the highway in the near future. Gravity method is especially useful for this case. Having a hole underground surely will create a large density contrast (essentially replacing the rocks underground with air, a substance with a very low density). The following figure shows a time-lapse gravity data. The gap between the gravity data between March-April and May-June indicates the development of sinkholes happening in the region of study (Rybakov et al., 2001).

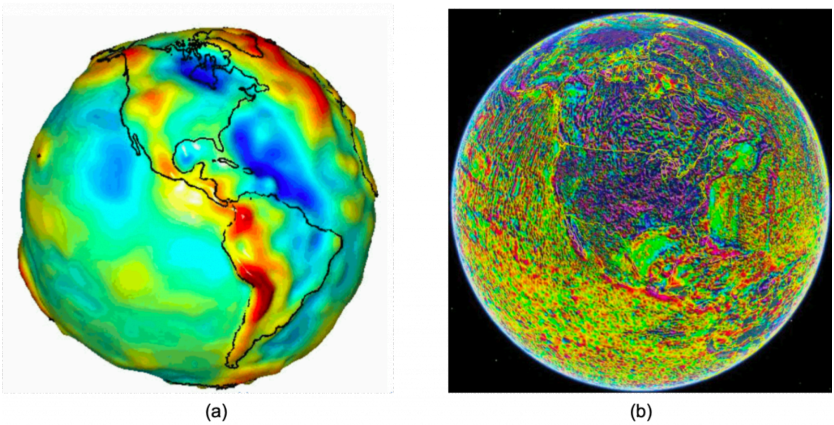


1.4 ADVANTAGES OF POTENTIAL FIELD METHODS

In comparison with other geophysical methods, such as seismic exploration, there are a lot of advantages of potential field methods that make it attracting, for example:

- it is inexpensive;
- measurements can be covered a large area in a short timeframe;
- signal sources are passive and non-destructive, we do not need to create any man-made sources;
- quick interpretation can be made;

- global coverage data, etc.



(a) Global gravity field from GRACE (https://en.wikipedia.org/wiki/Gravity_of_Earth). (b) Global magnetic field from EMAG2 (https://www.ngdc.noaa.gov/geomag/emag2.html).

CHAPTER 2

Chapter 2: Potential Field Theory

JIAJIA SUN, XINYAN LI, FELICIA NURINDRAWATI, XIAOLONG WEI, AND KENNETH LI

2.1 FUNDAMENTAL CONCEPTS

In this section, we will look at a few fundamental concepts that are highly relevant to the gravity and magnetics. These concepts include field, work, conservation field, and potential.

2.1.1 FIELD

A **field** is a set of functions of space and time. Mathematically, a field can be summarized as follows:

$$y = f(x, y, z; t)$$

There are two types of fields that we are concerned about within this course, **material fields** and **force fields**.

A few examples of material fields include

- A density field that describes the density value at each point of a material (e.g., the Earth) at a given time;
- A porosity field that describes the porosity value at each

point of a material (e.g., a reservoir) at a given time;

- A temperature field that describes the temperature value at each point of a material (e.g., a turkey) at a given time.

As you can tell, a material field describes some physical property of a material at each point of the material and at a given time.

The second type of field that is relevant to this course is the force field. A force field describes some kind of forces that act at each point of space at a given time. Some examples of force fields include

- Gravitational field that describes the gravitational force that acts at each point of space at a given time;
- Magnetic field that describes the magnetic force that acts at each point of space at a given time.

Moreover, the field can also be classified as either a **scalar field** or a **vector field**.

A scalar field is a single function of space and time, mathematically denoted by

$$y = f(x, y, z; t)$$

For example, a scalar field could be:

- displacement of a stretched string;
- temperature of a volume of gas;
- density within a volume of rock.

A vector field is a vector function of space and time, which can be written as

$$\vec{y} = \vec{f}(x, y, z; t) \quad \text{or} \quad \boldsymbol{y} = \boldsymbol{f}(x, y, z; t)$$

Since it is a vector function, the vector field is characterized by three functions of space and time. Mathematically, these functions can be written as:

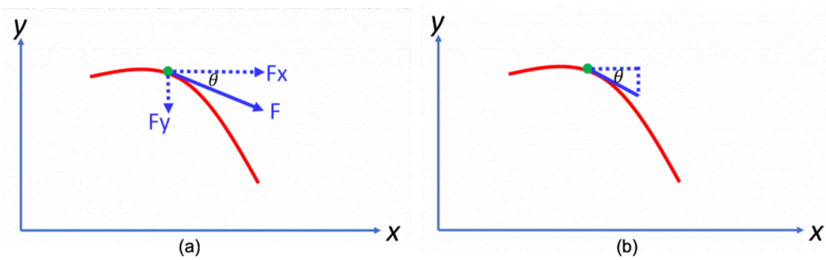
$$y_1 = f_1(x, y, z; t)$$

$$y_2 = f_2(x, y, z; t)$$
$$y_3 = f_3(x, y, z; t)$$

Examples of vector fields including:

- flow of heat;
- velocity of a fluid;
- gravitational attraction of a mass.

To visualize vector fields, we can use **field lines**, aka *lines of flow*, or *lines of force*. These are lines that are tangential to the vector field at every point. Therefore, scalar fields do not have field lines. Since field lines are tangential to the vector field at every point, it follows that small displacement along a field line must have x, y, and z components proportional to the corresponding x, y, and z components of the field at the point of its displacement. This can be proved by the illustration and derivations in the following.



The red curve shows a field line. At a location (i.e., green dot), the blue arrow is tangential to the red curve. The length of the blue arrow corresponds to the magnitude of the force at the green dot. The dashed blue arrows are two decomposed components of the force F. The angle is between Fx and F.

From the above figure (a), it is straightforward to derive the equation

$$\tan(\theta) = \frac{F_y}{F_x} \quad (1)$$

Now, let's forget about the force in figure (a). Just consider this red curve in figure (b) as a function $y = g(x)$. Then at the same green dot location, what is the derivative of the function $y = g(x)$ at this point? According to the definition of derivative evaluated at a point, we can write the following equation:

(2) $$g'(x_0) = \frac{dy}{dx} = \tan(\theta)$$

Since $\tan(\theta)$ is common for both equation (1) and (2), we can then link these two equations by

(3) $$\tan(\theta) = \frac{F_y}{F_x} = \frac{dy}{dx}, i.e., \frac{F_y}{F_x} = \frac{dy}{dx}$$

Rearranging equation (3), we will have:

(4) $$\frac{dx}{F_x} = \frac{dy}{F_y}$$

Further extending this equation to 3D space, it can be mathematically expressed as:

(5) $$\frac{dx}{F_x} = \frac{dy}{F_y} = \frac{dz}{F_z}$$

Therefore, if a vector field $\boldsymbol{F}$ is continuous, its field lines can be mathematically described by the differential equation (5).

Exercises

Find the gravitational attraction of a uniform sphere of mass M, centered at point Q, and observed outside the sphere at point P, through the given equation of

$$g = -\gamma M \frac{\hat{\boldsymbol{r}}}{r^2}$$

where γ is a constant, r is the distance from Q to P, and $\hat{\boldsymbol{r}}$ is the unit vector directed from Q to P.

Let Q be at the origin. Use the differential equation to describe the gravitational field lines at each point outside the sphere.

2.1.2 WORK

Let us recall what we have learned in college Physics courses, the **work** is defined as the product of force and distance. For example, as the cartoon image shown below, assuming a constant force $\boldsymbol{F}$ acting on the airplane from time equals zero to some distance $\boldsymbol{s}$ after some time interval, the work done on the airplane, denoted by $\boldsymbol{W}$, can be mathematically calculated by:

$$W = \boldsymbol{F} * \boldsymbol{s}$$

For this simple example, the force is assumed to be along with the same direction of the displacement. The unit of work is Joules, which is equal to Newton – meter is the **SI** unit system.

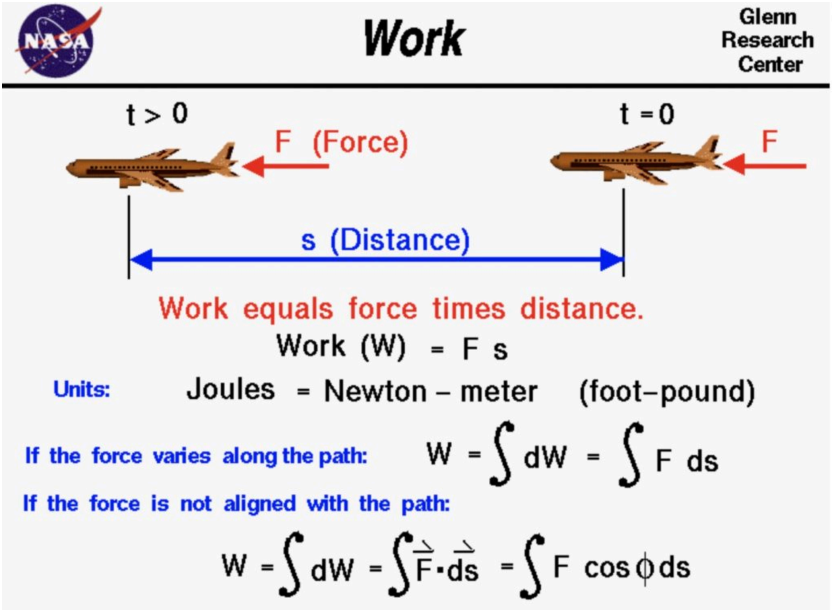


Image credit: https://www.grc.nasa.gov/WWW/K-12/airplane/work.html

However, if the force is no longer a constant value, but varies along the displacement, then the work is the integrated value of the force along the distance. Its math expression is as follows:

$$W = \int dW = \int \boldsymbol{F} d\boldsymbol{s}$$

Moreover, for the vector quantity force, if the force direction is not parallel with the moving path, then work is the integrated value of the force component along the direction of the path. Let the angle between the force and the displacement be ϕ, the work can be mathematically calculated by the following:

$$W = \int dW = \int \vec{\boldsymbol{F}} \cdot d\vec{\boldsymbol{s}} = \int \boldsymbol{F} \cos(\phi) d\boldsymbol{s}$$

Let us consider a more realistic example, illustrated by the figure below. A particle of mass m moves from position P_0 to P under the influence of force field $\boldsymbol{F}$. What is the work done by the force field in this example?

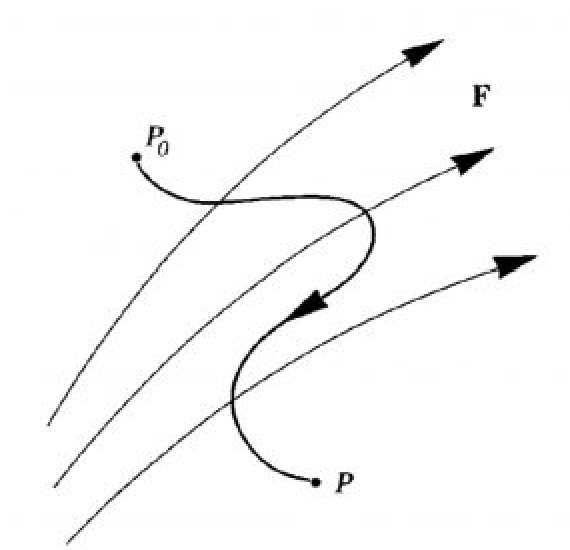


Image credit: Blakely, 1996, p4

In this given example, the force direction is not aligned with the particle displacement path. Therefore, the work required to move the particle from position P_0 to P is the integrated value of the force component along the path direction, which can be mathematically represented as the following:

$$W(P_0, P) = \int_{P_0}^{P} \boldsymbol{F} \cdot \boldsymbol{ds}$$

2.1.3 CONSERVATIVE FIELD

In general, the work depends upon the path taken by the particle. However, for some fields, the work is ***independent*** of the path of the particle. These fields are said to be **conservative**. Please keep in mind that, since we are talking about work and force, the fields we are talking about are vector fields, which are vector functions of space and time.

Now let us consider another example of work done in the conservative field. The scenario is illustrated in the figure below. Assuming a particle of mass moves through a conservative field, first from position P_0 to P in an irregular path, then parallel to the x axis with an additional small distance Δx. What is the work?

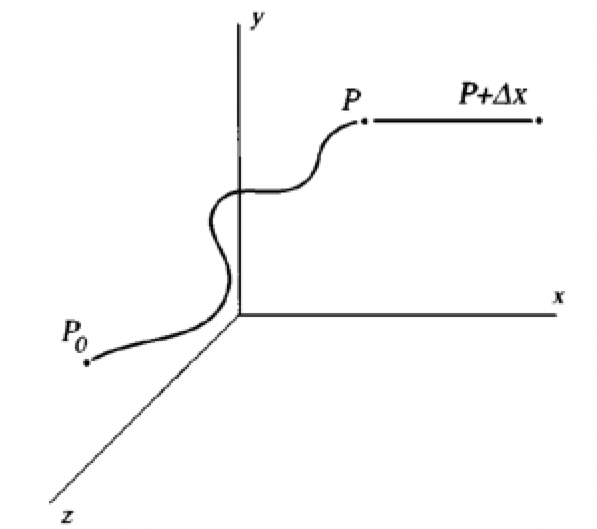


Image credit: Blakely, 1996, p5-6

We can deal with work done from position P_0 to $P + \delta x$ by summing the work from P_0 to P with work from P to $P + \Delta x$. Its math expression is as follows:

(1) $$W(P_0, P + \Delta x) = W(P_0, P) + W(P, P + \Delta x)$$

Rearranging the terms in equation (1), we will have equation (2) below:

(2) $$W(P_0, P + \Delta x) - W(P_0, P) = W(P, P + \Delta x)$$

Since the path from position P to $P + \Delta x$ is parallel with x axis, we can calculate the work done along this path segment by the integration as follows:

(3) $$W(P, P + \Delta x) = \int_{P}^{P+\Delta x} F_x(x, y, z)dx$$

Therefore, combining equation (2) and (3), we will have the following equation:

(4)

$$W(P_0, P + \Delta x) - W(P_0, P) = \int_{P}^{P+\Delta x} F_x(x, y, z)dx$$

By applying the *Mean value theorem* to the integral calculation, equation (4) can then be written as follows:

(5)
$$W(P_0, P + \Delta x) - W(P_0, P) = F_x(x + \varepsilon\Delta x, y, z)\Delta x$$

Dividing both sides of equation (5) by the small distance Δx, we will have the new equation as below:

(6) $$\frac{W(P_0, P + \Delta x) - W(P_0, P)}{\Delta x} = F_x(x + \varepsilon\Delta x, y, z)$$

Now if we make Δx approach 0, that is,

(7)

$$\lim_{\Delta x \to 0} \frac{W(P_0, P + \Delta x) - W(P_0, P)}{\Delta x} = \lim_{\Delta x \to 0} F_x(x + \varepsilon\Delta x, y, z)$$

For the left-hand side of the equation (7), the limit is defined to be the derivative of the function work W evaluated at point P ; the right-hand side of the equation is defined to be the force component along x axis. Therefore, equation (7) can then be expressed as follows:

$$\frac{\partial W}{\partial x} = F_x \tag{8}$$

Similarly, we can repeat the same derivation for the y and z directinos, and obtain the following:

$$\frac{\partial W}{\partial x} = F_x$$
$$\frac{\partial W}{\partial y} = F_y$$
$$\frac{\partial W}{\partial z} = F_z \tag{9}$$

In a more compact form, equation (9) can then be summarized as follows:

$$\nabla W = \boldsymbol{F}(x, y, z) \tag{10}$$

How to interpret this equation? It can be explained in the following aspects:

- gradient of work (or, work functino) (i.e., the left-hand side of equation (10)) is equal to force (i.e., the right-hand side of equation (10));
- derivative of work in any directino is equal to the component of force in that direction (e.g., the x component in this example);
- the vector force field $\boldsymbol{F}$ is completely specified by the scalar field W.

In brief summary so far, a **conservative field**, $\boldsymbol{F}$, is given by the **gradient of its work function**, W. Or vice versa, any vector field that has a work function satisfying the relation $\nabla W = \boldsymbol{F}(x, y, z)$ is conservative.

2.1.4 POTENTIAL

Potential ϕ of a vector field $\boldsymbol{F}$ is defined as the **work** function (or its negative). Its math representation is as follows:

$$\boldsymbol{F} = \nabla\phi$$

Usually, potential is defined at the infinity to be 0. And the potential at point P is defined as

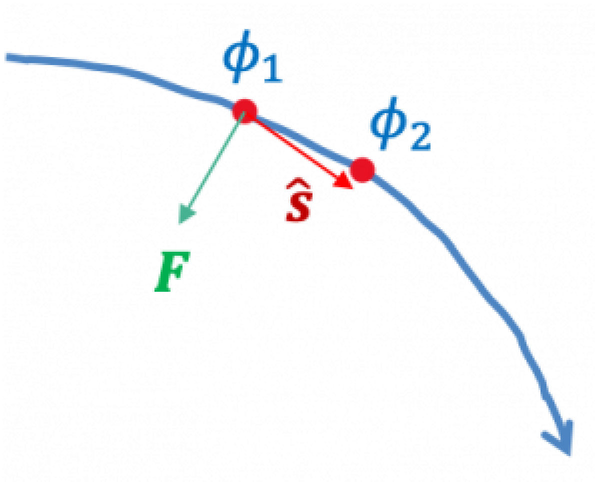
$$\phi(P) = \int_{\infty}^{P} \boldsymbol{F} \cdot \boldsymbol{ds}$$

2.1.5 EQUIPOTENTIAL SURFACE

As the name implies, an **equipotential surface** is a surface on which the **potential remains** *constant*. That is

$$\phi(x, y, z) = constant$$

Let us suppose $\hat{\boldsymbol{s}}$ is a unit vector that is tangential to an equipotential surfacee of $\boldsymbol{F}$, which is illustrated in the figure below.



The blue arrow curve represents an equipotential surface of F, the red arrow is a unit vector tangential to the surface, the green arrow is the field line at that point.

Since the potential is constant at any point along the equipotential surface, then we will have

$$\hat{\boldsymbol{s}} \cdot \boldsymbol{F} = \frac{\partial \phi}{\partial \hat{\boldsymbol{s}}} = 0$$

That is, the dot product of a unit vector $\hat{\boldsymbol{s}}$ with force field $\boldsymbol{F}$ is equal to the derivative of the potential with respect to the unit vector, $\frac{\partial \phi}{\partial \hat{\boldsymbol{s}}}$. We can think the derivative definition as the finite difference, i.e., the potential difference between ϕ_1 and ϕ_2, and if the distance between them is approaching to 0, then the derivative equals 0. Therefore, the dot product of the unit vector with force field equals 0.

According to the math definition of dot product between two vectors, since the dot product of the unit vector with force field equals 0, then the unit vector is perpendicular to the force field. That is,

$$\boldsymbol{F} \perp \hat{\boldsymbol{s}}$$

Therefore, the **field lines** at any point must be **perpendicular** to their **equipotential surface**. Conversely, any surface that is everywhere perpendicular to all field lines must be an equipotential surface. And no work is done when moving a test particle along an equipotential surface.

2.2 HELMHOLTZ DECOMPOSITION

Helmholtz theorem can be expressed as the following equation:

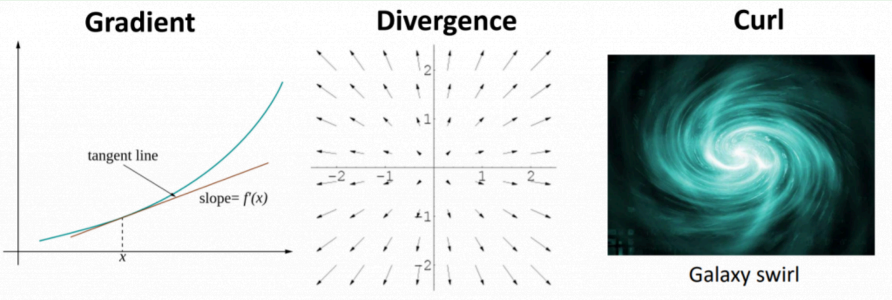
$$F = \nabla \phi + \nabla \times A$$

This basically says that *"any **vector field** (F) can be represented as the **gradient of a scalar** (ϕ) and the **curl of a vector** (A)".* The details and proof behind the above equation will be explained in this section.

Recap on important concepts

Before moving on to the next part, make sure that you understand the following points in order to make it easier to understand the derivation of the Helmholtz theorem.

- The difference between **gradient** **(∇)**, **divergence** **(∇.)**, and **curl** **($\nabla \times$)**



- The <u>curl</u> of a conservative field (such as gravity field) is always zero $\nabla \times \nabla f = 0$

- The divergence of a curl is always zero $\nabla.(\nabla \times U) = 0$

Please review the previous section if the points above does not make sense to you.

2.2.1. LAPLACE OPERATOR

Laplace operator is defined as the **divergence** of the gradient of a function. It can be expressed using the following expression:

$\nabla.\nabla f$

where *f* is the function. To mathematically understand what the Laplace operator means, let us consider a Cartesian coordinate system with 3 directions *(x,y,z)*. Thus the above expression can also be expressed as such:

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

The **Laplacian** (Laplacian operator) notation can also be further simplified as $\nabla^2 f$ or Δf.

Thus the Laplacian operator itself in a Cartesian coordinate can be written as:

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

An Example of a Laplacian

One example of a Laplacian is the Laplacian of an inverse distance, which is 0 everywhere except at $r = r'$. The Laplacian of the inverse distance can be expressed as such:

$$\nabla^2 (\frac{1}{|r - r'|}) = -4\pi\delta(r - r')$$

The delta function ($\delta(r - r')$) represents the following:

$$\delta(r - r') = \begin{cases} \infty, r = r' \\ 0, otherwise \end{cases}$$

The derivation of this function will be covered in the next section.

2.2.2. POISSON'S EQUATION

Poisson's equation is expressed as the following:

$$\nabla^2 \phi = f$$

Usually, f is given, while we want to find what ϕ is. The solution to the above equation can be expressed as an integral over all of space:

$$\phi = -\frac{1}{4\pi} \int \frac{f}{r} dv$$

2.2.2.1. Example of Poisson's equation in gravity application

The following is an example of Poisson's equation:

$$\nabla^2 \phi = -4\pi\gamma\rho$$

Usually, we know what ρ is and we want to know ϕ. The solution to the above equation can be expressed as such *(just replace* f *with* $-4\pi\gamma\rho$*):*

$$\phi = -\frac{1}{4\pi}\int \frac{f}{r}dv$$

$$\phi = -\frac{1}{4\pi}\int \frac{-4\pi\gamma\rho}{r}dv$$

$$\phi = \gamma\int \frac{\rho}{r}dv$$

Given a density distribution, this is how you can calculate gravitational potential everywhere in space.
In the above equation, ϕ represents the gravitational potential, while ρrepresents density. Note the Laplacian operator in the left hand side of the equation.
In this example, we can define a forward problem if we know ρ and want to know ϕ. If we know ϕ and want to know ρ, we would call this an inverse problem.

2.2.2.2. Relation to Laplace's equation

In a region of space that is not occupied by sources (i.e. mass), the following are true:

$$\nabla^2\phi = 0$$

$$\nabla^2\phi = -4\pi\gamma\rho = 0$$

Note that having the *Laplacian* of the gravitational potential as zero **does not** mean that the potential itself is zero.

2.2.3. HELMHOLTZ THEOREM

2.2.3.1. Definition

Previously, we have learned that a **conservative field (F)** can be represented as the gradient of a scalar (ϕ)

$$F = \nabla\phi$$

The above equation is actually a subset of the Helmholtz theorem, which states that:

Any vector field F that is continuous and zero at infinity can be expressed as the sum of the gradient of a scalar and the curl of a vector:

$$F = \nabla\phi + \nabla \times A$$

ϕ: scalar potential of **F**

A: vector potential of **F**

In the following sections, we will be discussing on the proof of the Helmholtz theorem.

2.2.3.2. Proof of Helmholtz Theorem

First, let's construct the following integral:

$$W(P) = \frac{1}{4\pi}\int \frac{F(Q)}{r}dv$$

- Q: point of integration
- r: distance between point P and point Q
- W: vector that we want to find (unknown)
- F: vector that we already have (known)

The above equation can be further split into 3 components in three-dimensional space as such:

$$W_x(P) = \frac{1}{4\pi}\int \frac{F_x(Q)}{r}dv$$

$$W_y(P) = \frac{1}{4\pi}\int \frac{F_y(Q)}{r}dv$$

$$W_z(P) = \frac{1}{4\pi}\int \frac{F_z(Q)}{r}dv$$

Recall that in **section 2.2.2: Poisson's equation**, we found the solution to Poisson's equation($\nabla^2\phi = f$):

$$\nabla^2\phi = f \Longleftrightarrow \phi = -\frac{1}{4\pi}\int \frac{f}{r}dv$$

Notice the similar form of equation that we have in the integral that we formed and the solution to the Poisson's equation.

Thus, we can do the same with our integral and reformat it as such:

$$W(P) = \frac{1}{4\pi}\int \frac{F(Q)}{r}dv = \Longleftrightarrow \phi = \nabla^2 W = -F$$

Note that W and F in the above equations are still vectors.

Now, using the following vector identity:

$$\nabla \times \nabla \times W = \nabla(\nabla . W) - \nabla^2 W$$

and therefore:

$$\nabla^2 W = \nabla(\nabla . W) - \nabla \times \nabla \times W$$

we can form the following equation from $\nabla^2 W = -F$:

$$\nabla^2 W = \nabla(\nabla . W) - \nabla \times \nabla \times W$$

$$-F = \nabla(\nabla . W) - \nabla \times \nabla \times W$$

$$F = -\nabla(\nabla . W) + \nabla \times \nabla \times W$$

In the above equation, we know that $\nabla . W$ is a **scalar** and $\nabla \times W$ is a **vector**

In other words, we have proved that the vector field **F** can be expressed as the gradient of a scalar ($\nabla . W$) plus the curl of a vector ($\nabla \times W$).

Lastly, if we define:

- $\phi = -\nabla . W$
- $A = \nabla \times W$

We can form the equation as:

$$F = \nabla \phi + \nabla \times A$$

2.2.3.2. Scalar and Vector Potential

Recall that at the beginning of our proof, we established the integral:

$$W(P) = \frac{1}{4\pi}\int \frac{F(Q)}{r}dv$$

That means, we can express **scalar potential** as:

$$\phi = -\nabla.W = -\frac{1}{4\pi}\int \frac{\nabla.F(Q)}{r}dv$$

While **vector potential** can also be expressed as:

$$A = -\nabla \times W = -\frac{1}{4\pi}\int \frac{\nabla \times F(Q)}{r}dv$$

2.2.3.3. Consequences of Helmholtz theorem

A vector field is **irrotational** (conservative) in a region if its **curl vanishes** everywhere

$$\nabla \times F = 0$$

According to Helmholtz theorem, the **vector potential** of this vector field becomes zero, and therefore:

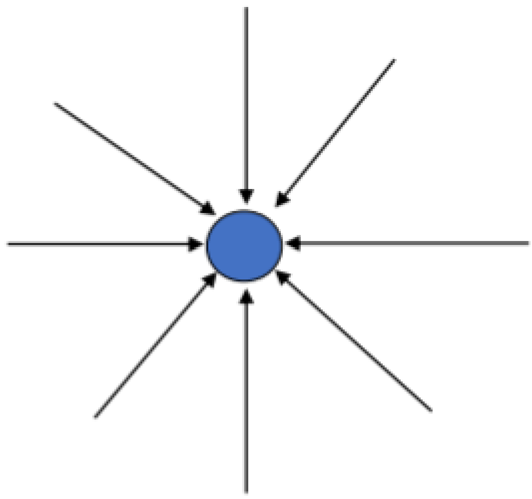
$$F = \nabla\phi$$

Conversely, it is easy to prove that $F = \nabla\phi$, then $\nabla \times F = 0$ by using the vector identity (shaded box in the previous section).

Therefore,

$$\nabla \times F = 0 \Longleftrightarrow F = \nabla\phi$$

One example of an irrotational field would be the **gravity field** which has the following properties. The lines of a gravity field do not form loops, thus the curl of the gravity field is zero ($\nabla \times g = 0$) and therefore $g = \nabla\phi$ where ϕ is the gravitational potential.



Similarly, a vector field is called **solenoidal** in a region if its **divergence vanishes** everywhere.

$$\nabla.F = 0$$

Since the scalar potential becomes zero according to Helmholtz theorem, we can conclude that:

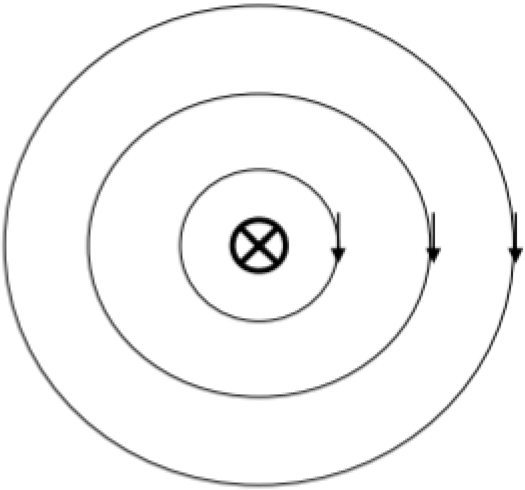
$$F = \nabla \times A$$

for a solenoidal field. The above can be easily proven by using the vector identity introduced in the previous section.

Therefore,

$$\nabla.F = 0 \Longleftrightarrow F = \nabla \times A$$

One example of a solenoidal field is a **static magnetic field.** The field lines do not emanate from or converge to any point, and thus the divergence is zero ($\nabla.B = 0$), and thus $B = \nabla \times A$ where A is a vector potential.



2.3 GREEN'S IDENTITIES

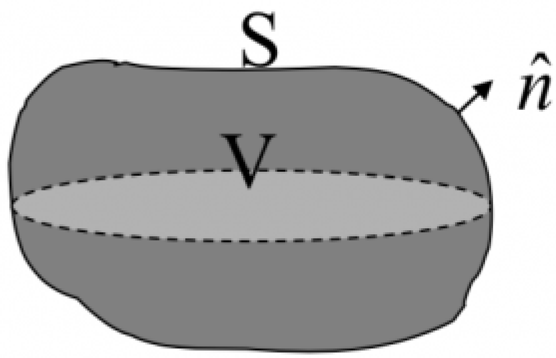
In this section, we will discuss about the Divergence theorem and the Green's first, second, and third identities, which were first published in 1828 by an English mathematician George Green. More information about him can be found in the following links: https://uh.edu/engines/epi1924.htm, https://sites.math.washington.edu/~morrow/334_19/green.pdf, and https://cosmosmagazine.com/mathematics/this- week-in-science-history-england-s-enigmatic- mathematician-is-born

2.3.1 DIVERGENCE THEOREM

Mathematically speaking, the Divergence theorem can be written as the following,

$$\iiint_V (\nabla \cdot \mathbf{F}) dV = \oiint_S (\mathbf{F} \cdot \mathbf{n})\, dS$$

where the left-hand side of this equation represents the volume integral of the divergence over the region inside the surface, while the right-hand side of the equation represents the outward flux of a vector field $\boldsymbol{F}$ through a closed surface. The physical meaning can be illustrated through the cartoon image below.



Physical meaning of the divergence theorem.

Intuitively, it states that the sum of all sources (with sinks regarded as negative sources) gives the net flux out of a region.

2.3.1.1 Application to gravity

Now, if we consider the vector field $\boldsymbol{F}$ as the gravity field $\boldsymbol{g}$, so that $\boldsymbol{F} = \boldsymbol{g}$, then let's substitute it into the above divergence theorem equation, we'll get the following:

$$\iiint\limits_V (\nabla \cdot \mathbf{g}) dV = \iint\limits_S (\mathbf{g} \cdot \hat{\mathbf{n}}) dS$$

From the previous section, we know that $\mathbf{g} = \nabla\phi$, therefore, replacing $\boldsymbol{g}$, we will get the equation as follows:

$$\iiint_V (\nabla \cdot \nabla\phi) dV = \iint_S (\mathbf{g} \cdot \hat{\mathbf{n}}) dS$$

$$i.e., \quad \iiint_V \left(\nabla^2\phi\right) dV = \iint_S (\mathbf{g} \cdot \hat{\mathbf{n}}) dS$$

Remembering that

$$\nabla^2\phi = -4\pi\gamma\rho$$

therefore, the above equation can be re-arranged into the format below:

$$-4\pi\gamma \iiint_V \rho dV = \iint_S (\mathbf{g} \cdot \hat{\mathbf{n}}) dS$$

It can be noticed that the left-hand side integration gives the quantity of mass, M, thus, the above equation can be further simplified as follows:

$$-4\pi\gamma M = \iint_S (\mathbf{g} \cdot \hat{\mathbf{n}}) dS$$

$$\because \quad M = -\frac{1}{4\pi\gamma} \iint_S (\mathbf{g} \cdot \hat{\mathbf{n}}) dS$$

where the right-hand side of this equation contains the surface integral of the flux of the gravity field through a closed surface.

This equation implies that, if we do an integration of the gravity map over a study area, the integration gives an estimate of the mass underneath the gravity map, that is responsible for the measured gravity data.

Before moving to the Green's identities, let's review two equations that will be used later.

$$\nabla^2 \left(\frac{1}{r}\right) = -4\pi\delta\left(r\right) = 0,\ \text{if } r \neq 0$$

and

$$\nabla^2 V = -4\pi\gamma\rho$$

2.3.2 GREEN'S IDENTITIES

The three identities can be derived from the vector calculus and the Laplace's equation. Each of the three identities has different utility and implications for potential field study, and their common starting point is the **divergence theorem** (aka., *Gauss's theorem*) discussed above.

2.3.2.1 Green's first identity

Let's assume that there are

- two continuous functions U, V with continuous first-order partial derivative;
- and U also second-order derivative that's also continuous;

- Then defining an arbitrary vector $\boldsymbol{A} = V\nabla U$

Applying the divergence theorem, i.e., replacing vector field $\boldsymbol{F}$ with $\boldsymbol{A}$, we will have the following equation:

$$\iiint\limits_{Vol} \nabla \cdot (V\nabla U) dV = \iint\limits_{Surf} V\nabla U \cdot \hat{\mathbf{n}} dS$$

Then, let's use the vector identity

$$\nabla \cdot (\alpha \mathbf{V}) = \alpha \nabla \cdot \mathbf{V} + \nabla \alpha \cdot \mathbf{V}$$

to expand the above divergence theorem equation, we will then get the Green's first identity as follows:

Green's first identity

$$\iiint\limits_{Vol} \left(V\nabla^2 U + \nabla V \cdot \nabla U\right) dV = \iint\limits_{Surf} V \frac{\partial U}{\partial \hat{\mathbf{n}}} dS$$

Implication for gravity – 1

If making some simplifications by setting $V = 1$, and choose U such that $\nabla^2 U = 0$ that satisfies the Laplacian equation (i.e., U is harmonic and its second-order derivative is continuous), then, based on Green's first identity written above, we will have the following expression

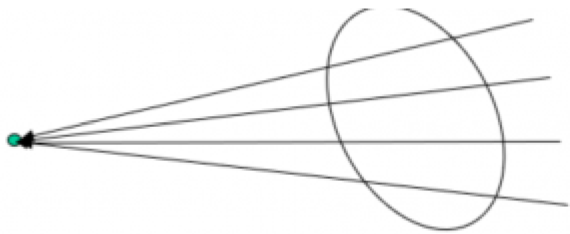
$$\iint\limits_{S} \frac{\partial U}{\partial \hat{\mathbf{n}}} dS = 0$$

which can be interpreted as, the normal derivative of a harmonic function averages to 0 on a closed surface.

Specifically for the case of gravity, considering a gravity field $\boldsymbol{g}$ in regions of space not occupied by mass, it is associated with a potential U which is harmonic. Thus, the above equation can be written as follows:

$$\iint_S \frac{\partial U}{\partial \hat{\mathbf{n}}} dS = \iint_S \nabla U \cdot \hat{\mathbf{n}} dS = \iint_S \mathbf{g} \cdot \hat{\mathbf{n}} dS = 0$$

That is, the gravity field has a net zero flux over any closed surface in any source-free regions (i.e., regions not occupied by sources). This interpretation can be illustrated by the image below.



Green's first identity implication for gravity.

In other words, the normal component of gravity field (or, in general, any conservative field) averages to zero over any closed surface in source-free regions.

Implication for gravity – 2

Another interesting consequence of applying Green's first identity to potential field is that, if letting U be harmonic and $U = V$, then we will have

$$\iiint_{Vol} (\nabla U)^2 dV = \iint_{Surf} U \frac{\partial U}{\partial \hat{\mathbf{n}}} dS$$

Consider the above equation when $U = 0$ on the surface S

. Then the R.H.S vanishes, and because $(\nabla U)^2$ is positive and continuous in the region, then $(\nabla U)^2 = 0$. Therefore, U is constant. Moreover, because $U = 0$ on the surface and U is continuous, the constatn must be 0. Therefore, if U is harmonic and continuously differentiable in $\mathbf{R}$, and if U vanishes everywhere on the surface S, U must also vanish everywhere within the volume.

Furthermore, let U_1 and U_2 be harmonic in R and have identical boundary conditions, that is, $U_1(S) = U_2(S)$. The function $U_1(S) - U_2(S)$ must be harmonic. But $U_1 - U_2$ vanishes on S. Based on previous consequence, $U_1 - U_2$ must vanish at every point in R. Therefore, U_1 and U_2 are identical. Thus, a function that is harmonic and continuously differentiable in R is uniquely determined by its values on S.

2.3.2.2 Green's second identity

Based on the assumption that functions U and V are continuous functions with continuous second-order derivatives, we can start from the Green's first identity

$$\iiint\limits_{Vol} \left(V\nabla^2 U + \nabla V \cdot \nabla U\right) dV = \iint\limits_{Surf} V \frac{\partial U}{\partial \hat{\mathbf{n}}} dS$$

then exchange function U with V to get the function as follows:

$$\iiint\limits_{Vol} \left(U\nabla^2 V + \nabla U \cdot \nabla V\right) dV = \iint\limits_{Surf} U \frac{\partial V}{\partial \hat{\mathbf{n}}} dS$$

If subtracting the above equation from the original first identity, we will get the Green's second identity defined as follows:

$$\iiint\limits_{Vol} \left(V\nabla^2 U - U\nabla^2 V\right) dV = \iint\limits_{Surf} \left(V\frac{\partial U}{\partial \hat{\mathbf{n}}} - U\frac{\partial V}{\partial \hat{\mathbf{n}}}\right) dS$$

That is,

Green's second identity

$$\iiint\limits_{Vol} \left(V\nabla^2 U - U\nabla^2 V\right) dV = \iint\limits_{Surf} \left(V\nabla U - U\nabla V\right) \cdot \hat{\mathbf{n}} dS$$

Implication for gravity – 1

So how can this identity be related with potential field methods? If assuming $V = 1$, then the Green's second identity can be re-written as follows:

$$\iiint\limits_{Vol} \nabla^2 U dV = \iint\limits_{Surf} \nabla U \cdot \hat{\mathbf{n}} dS$$

and assuming U is the potential of some vector field $\boldsymbol{F}$, i.e., $\nabla U = \boldsymbol{F}$, then the above equation can be further written as:

$$\iiint\limits_{Vol} \nabla^2 U dV = \iint\limits_{Surf} \mathbf{F} \cdot \hat{\mathbf{n}} dS$$

In source-free regions, $\nabla^2 U = 0$, therefore, we have

$$0 = \iint\limits_{Surf} \mathbf{F} \cdot \hat{\mathbf{n}} dS$$

which implies that, the net flux of gravitational field over a surface with source-free equals zero, the same conclusion we derived before.

Implication for gravity – 2

In Green's second identity, the surface is a closed surface bounding the volume. For gravity application, it is a surface bounding the mass. Now let's consider a special surface: an ***equipotential surface***. Assuming function V as gravity potential, and consider a point P outside the surface, and r is the distance from P. Let's consider what happens when $U = \frac{1}{r}$.

We can substitute $U = \frac{1}{r}$ into the second identity, then we will have the following:

$$\iiint\limits_{Vol} \left(V\nabla^2 \left(\frac{1}{r}\right) - \frac{1}{r}\nabla^2 V \right) dV = \iint\limits_{Surf} \left(V\nabla \left(\frac{1}{r}\right) - \frac{1}{r}\nabla V \right) \cdot \hat{\mathbf{n}} dS$$

Recall the two equations defined above, which are $\nabla^2 \left(\frac{1}{r}\right) = 0, \text{ if } r \neq 0$ and $\nabla^2 V = -4\pi\gamma\rho$, then the above equation's left-hand side can be simplied as follows:

$$\iiint\limits_{Vol} \left(V\nabla^2 \left(\frac{1}{r}\right) - \frac{1}{r}\nabla^2 V \right) dV = 4\pi\gamma \iiint\limits_{Vol} \frac{\rho}{r} dV$$

For the right-hand side, since gravity potential V is constant on the equipotential surface, so V can be moved out of the integral;

then we can apply the divergence theorem, so that the surface integral can be changed to colume integral. Their mathematical expressions are as follows:

$$\iint\limits_{Surf} V\nabla\left(\frac{1}{r}\right)\cdot\hat{\mathbf{n}}dS = V\iint\limits_{Surf} \nabla\left(\frac{1}{r}\right)\cdot\hat{\mathbf{n}}dS$$

$$\iint\limits_{Surf} \nabla\left(\frac{1}{r}\right)\cdot\hat{\mathbf{n}}dS = \iiint\limits_{Vol} \nabla^2\left(\frac{1}{r}\right)dV$$

Now, recall the Laplacian of inverse distance

$$\iint\limits_{Surf} V\nabla\left(\frac{1}{r}\right)\cdot\hat{\mathbf{n}}dS = 0$$

Thus, the second identity can be written as follows:

$$4\pi\gamma\iiint\limits_{Vol}\frac{\rho}{r}dV = -\iint\limits_{Surf}\frac{1}{r}\nabla V\cdot\hat{\mathbf{n}}dS$$

If we look carefully on the left-hand side of the above equation, it contains the calculation of the gravitational potenital given a density distributino, which equals $\gamma\iiint\limits_{Vol}\frac{\rho}{r}dV$, therefore, we have,

$$4\pi V_p = 4\pi\gamma\iiint\limits_{Vol}\frac{\rho}{r}dV = -\iint\limits_{Surf}\frac{1}{r}\nabla V\cdot\hat{\mathbf{n}}dS$$

dividing 4π on both sides, we will have

$$V_p = \gamma \iiint\limits_{Vol} \frac{\rho}{r} dV = -\frac{1}{4\pi} \iint\limits_{Surf} \frac{1}{r} \nabla V \cdot \hat{\mathbf{n}} dS$$

This equation implies that for a set of gravitational field data, it can be interpreted by two methods: at any point outside S, the potential caused by a 3d source inside S is the same as the potential caused by a material that is spread over the equipotential surface S with a surface density of $-\frac{1}{4\pi}\frac{\partial V}{\partial \hat{\mathbf{n}}}$.

2.3.2.3 Green's third identity

Its derivation begins with second identity

$$\iiint\limits_{Vol} \left(V\nabla^2 U - U\nabla^2 V\right) dV = \iint\limits_{Surf} (V\nabla U - U\nabla V) \cdot \hat{\mathbf{n}} dS$$

Again, let's make simplifications by letting $U = \frac{1}{r}$ where the second identity will then be written as follows

$$\iiint\limits_{Vol} \left(V\nabla^2 \frac{1}{r} - \frac{1}{r}\nabla^2 V\right) dV = \iint\limits_{Surf} \left(V\nabla \frac{1}{r} - \frac{1}{r}\nabla V\right) \cdot \hat{\mathbf{n}} dS$$

Remembering that $\nabla^2 \frac{1}{r} = -4\pi\delta\left(r\right)$ in general case (i.e., r can be zero within the volume). Then the above equation can be re-written as follows:

$$\iiint\limits_{Vol} \left(-4\pi V\delta\left(r\right) - \frac{1}{r}\nabla^2 V\right) dV = \iint\limits_{Surf} \left(V\nabla\frac{1}{r} - \frac{1}{r}\nabla V\right) \cdot \hat{\mathbf{n}} dS$$

Moving the second term on the left to the R.H.S., we have

$$\iiint\limits_{Vol} -4\pi V\delta\left(r\right) dV = \iiint\limits_{Vol} \frac{1}{r}\nabla^2 V dV + \iint\limits_{Surf} \left(V\nabla\frac{1}{r} - \frac{1}{r}\nabla V\right) \cdot \hat{\mathbf{n}} dS$$

Diving by -4π on both sides, and expand further, the new equation will be

$$\iiint\limits_{Vol} V\delta\left(r\right) dV = -\frac{1}{4\pi}\iiint\limits_{Vol} \frac{1}{r}\nabla^2 V dV - \frac{1}{4\pi}\iint\limits_{Surf} V\nabla\frac{1}{r}\cdot\hat{\mathbf{n}} dS + \frac{1}{4\pi}\iint\limits_{Surf} \frac{1}{r}\nabla V\cdot\hat{\mathbf{n}} dS$$

Based on the definition of derivatives, the R.H.S. of the previous equation can be re-written as

$$\iiint\limits_{Vol} V\delta\left(r\right) dV = -\frac{1}{4\pi}\iiint\limits_{Vol} \frac{1}{r}\nabla^2 V dV - \frac{1}{4\pi}\iint\limits_{Surf} V\frac{\partial}{\partial\hat{\mathbf{n}}}\frac{1}{r} dS + \frac{1}{4\pi}\iint\limits_{Surf} \frac{1}{r}\frac{\partial V}{\partial\hat{\mathbf{n}}} dS$$

According to the definition of the Dirac delta function,

$$\delta(x) = \begin{cases} +\infty, & x = 0 \\ 0, & x \neq 0 \end{cases}$$
$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1$$
$$\int_{-\infty}^{\infty} f(x)\, \delta(x)\, dx = f(0)$$

Through applying the Dirac delta function property, the above equation will be the Green's third identity. For simplicity, we assumed the origin to be the point of observation, therefore, we will have:

$$V(0) = -\frac{1}{4\pi}\iiint\limits_{Vol} \frac{1}{r}\nabla^2 V dV - \frac{1}{4\pi}\iint\limits_{Surf} V\frac{\partial}{\partial \hat{\mathbf{n}}}\frac{1}{r} dS + \frac{1}{4\pi}\iint\limits_{Surf} \frac{1}{r}\frac{\partial V}{\partial \hat{\mathbf{n}}} dS$$

In general, for a continuous function V, the Green's third identity is

Green's third identity

$$V\left(P\right)=-\frac{1}{4\pi}\iiint\limits_{Vol}\frac{1}{r}\nabla^{2}V\,dV-\frac{1}{4\pi}\iint\limits_{Surf}V\frac{\partial}{\partial\hat{\mathbf{n}}}\frac{1}{r}dS+\frac{1}{4\pi}\iint\limits_{Surf}\frac{1}{r}\frac{\partial V}{\partial\hat{\mathbf{n}}}dS$$

Understanding Green's third identity

The first integral on the R.H.S. of the Green's third identity can be understood as the potential due to a volume distribution with density $\rho\left(Q\right)=-\frac{1}{4\pi\gamma}\nabla^{2}V$, that is, the gravitational potential due to a volume density distribution $\rho(Q)$ is

$$V\left(P\right)=\gamma\iiint\limits_{Vol}\frac{\rho\left(Q\right)}{r}dV$$

The third integral has the same form as the potential due to a surface density distribution σ where $\sigma=\frac{1}{4\pi\gamma}\frac{\partial V}{\partial\hat{\mathbf{n}}}$. That is, the gravitational potential due to a surface density distribution $\sigma(Q)$ is

$$V\left(P\right)=\gamma\iint\limits_{Surf}\frac{\sigma\left(Q\right)}{r}dS$$

The second integral can be understood as the magmatic potential due to a surface distribution of magnetization $\mathbf{M}\left(Q\right)=-\frac{V}{\mu_{0}}$, which is spread over surface S, and directed normal to S. The magnetic potential is

$$V\left(P\right)=\frac{\mu_0}{4\pi}\iint\limits_{Surf}\mathbf{M}\left(Q\right)\hat{\mathbf{n}}\cdot\nabla_Q\frac{1}{r}dS=\frac{\mu_0}{4\pi}\iint\limits_{Surf}\mathbf{M}\left(Q\right)\frac{\partial}{\partial\hat{\mathbf{n}}}\frac{1}{r}dS$$

In summary, *any function* with sufficient differentiability can be expressed as the sum of three potentials:

- The potential due to a volume distribution of density $\rho\left(Q\right)=-\frac{1}{4\pi\gamma}\nabla^2V$
- The potential due to a surface distribution of density $\sigma(Q)=\frac{1}{4\pi\gamma}\frac{\partial V}{\partial\hat{\mathbf{n}}}$
- The potential due to a surface distribution of magnetization $\mathbf{M}\left(Q\right)=-\frac{V}{\mu_0}$

In other words, any function with sufficient differentiability is a potential!

Furthermore, if considering the situation when V is harmonic, i.e., $\nabla^2U=0$, then Green's third identity will become the following

$$V\left(P\right)=-\frac{1}{4\pi}\iint\limits_{Surf}V\frac{\partial}{\partial\hat{\mathbf{n}}}\frac{1}{r}dS+\frac{1}{4\pi}\iint\limits_{Surf}\frac{1}{r}\frac{\partial V}{\partial\hat{\mathbf{n}}}dS$$

$$i.e.,\ V\left(P\right)=\frac{1}{4\pi}\iint\limits_{Surf}\left(\frac{1}{r}\frac{\partial V}{\partial\hat{\mathbf{n}}}-V\frac{\partial}{\partial\hat{\mathbf{n}}}\frac{1}{r}\right)dS$$

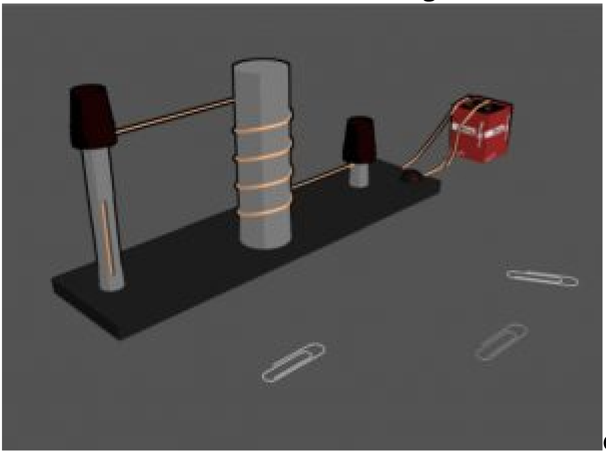
This is a representation formula, where a harmonic function can be calculated at any point simply from its values and normal derivatives on the boundary. This is the theoretical basis for the **upward continuation** and the **equivalent source technique**, which will be introduced later in this course.

2.4 GRAVITATIONAL POTENTIAL

2.4.1 NEWTONIAN POTENTIAL

There are *four basic forces* known presently to physics, which are strong force, electromagnetic force, weak force, and gravitational force.

- **Strong** force could hold protons and neutrons together in the atomic nucleus and have extremely short range, but a hundred times more powerful than electrical forces.
- **Electromagnetic** force, for instance, the electrostatic force following the Coulomb's law, could produce the everyday phenomenon of static electricity, or lightning strike due to sudden electrostatic discharge. Another



examplе
is the magnetic force, generated by a magnet. In a simple

experiment setting shown in the figure on the right, if the coil wired around the metal is powered with electricity, that will generate a magnetic field, which will then make this metal becomes serving as a magnet.

- **Weak** force accounts for certain kinds of radioactive decay.
- **Gravitational** force, which is the force that we will focus on in this course. Unfortunately, we do not fully understand gravity, despite Newton's law of gravitational attraction and Einstein's general relativity.

In this course, we will focus on Newton's law of gravitational attraction to discuss the gravitational force.

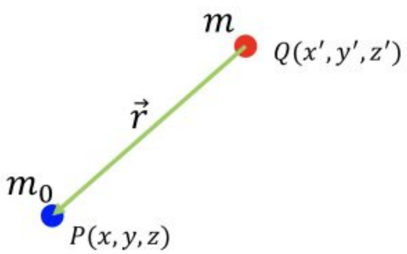
2.4.1.1 Gravity attraction

Issac Newton was an English mathematician, physicist, astronomer, theologian and author who lived in the time period of 1642-1726. He had made fundamental contributions to classical mechanics, optics and calculus, and published the famous Mathematical Principle of Natural Philosophy in 1687. In this book, Newton formulated the laws of motion and the law of gravitational attraction.

Newton's law of gravitational attraction

Newton's law of gravitational attraction states that two masses would attract each other, and this attraction is called ***gravitational force***. The magnitude of the force is proportional to each mass and inversely proportional to the square of their distance.

The simple illustration of Newton's law of gravitational attraction is shown in the figure below.



The mutual force between the two masses m and m_0 is mathematically represented as:

$$F = \gamma \frac{m m_0}{r^2}$$

$$r = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}$$

where r is the distance between two masses, $\gamma = 6.674 \times 10^{-11}\ m^3 kg^{-1} s^{-2}$ is the universal gravitational constant.

Let us consider m_0 to be a *test particle* with unit mass (i.e., the existence of it doesn't affect the force caused by mass m). Then the gravitational attraction produced by mass m at the location of the test particle is

$$\vec{g}\,(P) = -\gamma \frac{m}{r^2} \hat{\mathbf{r}}$$

Note m_0 is gone since we have assumed it is a test particle with unit mass, and the unit vector $\hat{\boldsymbol{r}}$ is pointing from m to m_0. The minus sign in this equation is necessary because $\vec{r}$, following convention, is directly from the source m to the observation point m_0.

The unit of this gravitational attraction $\boldsymbol{g}(P)$ can be derived from the following:

$$\mathbf{g}(P) = \frac{F}{m_0} = \frac{\gamma \frac{m m_0}{r^2}}{m_0} = \gamma \frac{m}{r^2}$$

i.e., $\boldsymbol{g}(P)$ is force divided by mass, therefore it has the unit of acceleration, m/s^2.

Therefore, the gravitational attraction is also called gravitational acceleration, and it is a vector since the gravitational force has a direction. Its values vary depending on the measured locations on Earth, its conventional standard value is about 9.8 m/s^2.

After-class Exercises

Given

$vecg\left(P\right)=-\gamma\frac{m}{r^2}\hat{\mathbf{r}}$

Prove

$\nabla\times\mathbf{g}=0$

*Hint: it is easier to do the proof in spherical coordinate.

Once we have proved ourselves after doing the exercise listed on the right box, that

$$\nabla\times\mathbf{g}=0$$

we can interpret and make sense of it since, for Earth, its gravitational field lines are all pointing to its center from 360^o degrees, thus there is no rotation, so the curl must be zero.

That brings us to the concept of the **irrotational field** when we talked about the Helmholtz theorem. An irrotational field is a vector field in a region if its curl vanishes everywhere, i.e.,

$$\nabla\times\mathbf{F}=0$$

that is, when the curl of a vector field is zero, the field is irrotational. Since the gravitational field satisfies this condition, therefore, the gravitational field is irrotational.

Moreover, according to the Helmholtz theorem, the vector potential becomes zero. Therefore,

$$\mathbf{F}=\nabla\phi$$

which is a scalar potential, or **gravitational potential**.

In brief summary, because $\nabla\times\mathbf{g}=0$, the Gravitational/ Newtonian potential is irrotational and is conservative. It can be fully described by a scalar potential $\mathbf{g}(P)=\nabla U(P)$, where U is called gravitational potential, and it can be mathematically represented as follows:

$$U=\gamma\frac{m}{r}$$

which decays as a function of distance r.

2.4.1.2 Gravitational potential of continuous matter

For the sake of convenience and robustness, we want to be able to calculate the gravitational potential due to any distribution of density with any geometries. Thus, we need to apply the **Principle of Superposition**, so that the gravitational potential of a collection of masses can be calculated as a sum of the gravitational potentials due to each individual mass.

Similarly, the gravitational field/acceleration, which is a vector, of a collection of masses is the vector sum of the gravitational acceleration due to each individual mass.

If we want to calculate the gravitational attraction due to a continuous distribution of matter, that is density $\rho(x, y, z)$ varies spatially, we can use the principle of superposition, so that the continuous distribution of mass m is simply a collection of many very small masses. For each small mass, its individual mass can be calculated by the product of constant density $\rho(x, y, z)$ of this small mass with its volume occupied by the tiny mass dv, that is,

$$dm = \rho(x, y, z)dv$$

The gravitational potential due to this single small mass is

$$U(P) = \gamma \frac{dm}{r}$$

Then the potential observed at location P due to all small masses can be then calculated by summation of all individual potentials, i.e., the integration over the whole volume occupied by mass at each point of Q within the volume, which is mathematically written as follows:

$$U(P) = \gamma \int_v \frac{dm}{r} = \gamma \int_v \frac{\rho(Q)dv}{r}$$

According to Helmholtz theorem, we know that

$\boldsymbol{g}(P) = \nabla U(P)$. Thus, in order to calculate gravity, we need to calculate

$$\frac{\partial U(P)}{\partial x} = \gamma \frac{\partial}{\partial x} \int\limits_v \frac{\rho(Q)dv}{r} = \gamma \int\limits_v \frac{\partial}{\partial x} \frac{1}{r} \rho(Q)dv = -\gamma \int\limits_v \frac{x - x'}{r^3} \rho(Q)dv$$

Similarly, following the same procedure, we can calculate $\frac{\partial U(P)}{\partial y}$, $\frac{\partial U(P)}{\partial z}$, i.e.,

$$\frac{\partial U(P)}{\partial x} = -\gamma \int\limits_v \frac{x - x'}{r^3} \rho(Q)dv$$

$$\frac{\partial U(P)}{\partial y} = -\gamma \int\limits_v \frac{y - y'}{r^3} \rho(Q)dv$$

$$\frac{\partial U(P)}{\partial z} = -\gamma \int\limits_v \frac{z - z'}{r^3} \rho(Q)dv$$

If we collect the three terms on the left hand side in the above three equations into a vector, we get $\nabla U(P)$ which is equal to $\boldsymbol{g}(P)$. The three integrals on the right hand side in the above three equations can also be collapsed into a more compact form. Consequently, we arrive at the following equation:

$$\mathbf{g}(P) = -\gamma \int\limits_V \rho(Q)\frac{\vec{r}}{r^3}dv$$

$$i.e., \mathbf{g}(P) = -\gamma \int\limits_V \rho(Q)\frac{\hat{r}}{r^2}dv$$

where $\hat{r}$ is the unit vector in the same direction as the vector $\vec{r}$.

Please be noted that the previous derivation of the gravity field based on gravitational potential assumes the observation point is outside the distribution of mass. What about the potential inside the mass?

If the observation point P is inside the mass, the integrand in equation becomes singular and the integral is improper (see P47 in Blakely's book). However, Kellogg (1953) shows that the integral

$$I(P) = \int\limits_V \frac{\rho}{r^n}dv$$

is convergent for P inside a volume V and is continuous throughout V if $n < 3$, V is bounded and ρ is piecewise continuous. Therefore, both $U(P)$ and $\boldsymbol{g}(P)$ exist and are continuous everywhere, both inside and outside the mass if the density in the volume is well behaved (Blakely, 1996, p48). In addition, Kellogg (1957) also shows that $\boldsymbol{g}(P) = \nabla U(P)$ for P inside the mass.

In summary, the following two equations hold true for any bounded distribution of piecewise-continuous density:

$$\mathbf{g}(P) = -\gamma \int\limits_V \rho(Q)\frac{\hat{r}}{r^2}dv$$

and

$$\boldsymbol{g}(P) = \nabla U(P).$$

2.4.1.3 Poisson's equation

So far, we can represent the gravitational potential in two math equations, one is based on Helmholtz theorem, we have mathematically derived at

$$U(P) = -\frac{1}{4\pi}\int \frac{\nabla \cdot \mathbf{g}(Q)}{r} dv$$

another equation is due to Newton's law of gravitational attraction, which is written as

$$U(P) = \gamma \int_V \frac{\rho(Q)dv}{r}$$

By equalling those two equations, we will arrive at the Poisson's equation, which is expressed as follows:

$$\nabla \cdot \mathbf{g} = -4\pi\gamma\rho$$

This is valid for observation point both inside and outside the mass distribution.

A special case of Poisson's equation is Laplace's equation, which is denoted as below:

$$\nabla \cdot \mathbf{g} = 0$$

which is valid in regions of space not occupied by mass. For most of the geophysical potential field data acquisition, we are dealing with this equation, since the data measurement region is at the source-free region, for example, airborne data acquisition, satellite data collection, etc.

Brief Summary

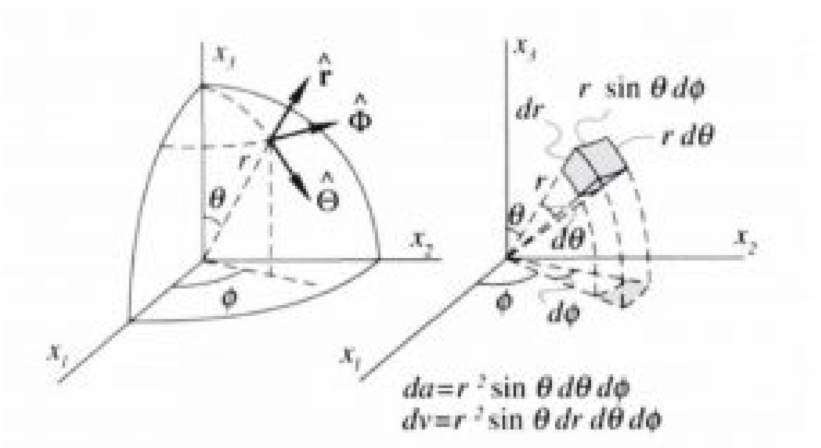
- Gravitational potential: $U(P) = \gamma \int\limits_{v} \frac{dm}{r} = \gamma \int\limits_{v} \frac{\rho(Q)dv}{r}$
- Gravitational field/acceleration: $\mathbf{g}(P) = -\gamma \int\limits_{V} \rho(Q) \frac{\hat{\mathbf{r}}}{r^2} dv$
- Poisson's equation $\nabla \cdot \mathbf{g} = -4\pi\gamma\rho$

2.4.2 EXAMPLES OF GRAVITY DUE TO SIMPLE OBJECTS

Before delving into the details of deriving the gravitational response due to simple geometric objects, it is important to note that the following derivations are much simplified due to having to deal with the problem in spherical coordinates. Thus, firstly let us recap on the concept of the spherical coordinate system.

2.4.2.1. Spherical Coordinate System

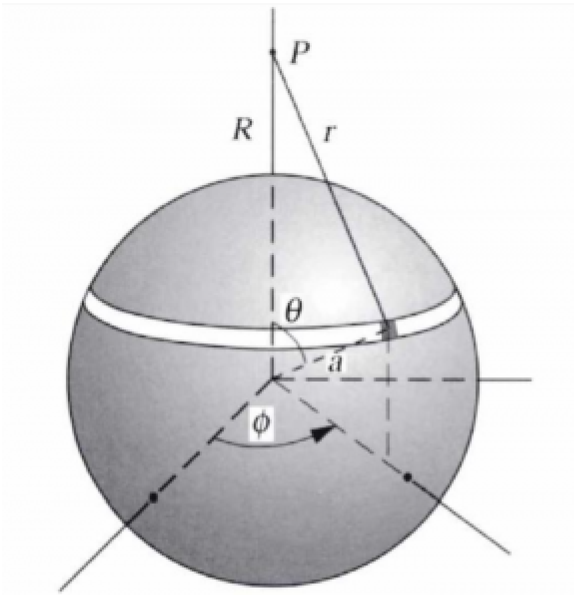
Normally, most problems are done in the Cartesian coordinate system (x, y, z). However, to simplify the integration in our derivation, we use the spherical coordinate system instead. The following figure gives a visualization of the coordinate system. A point in space can be expressed as the point (r, θ, ϕ), where **r** is the distance between the point and the origin, ϕ is the declination of the line connecting the point to the origin, and θ is the inclination (angle made by the line that crosses the origin with the horizontal). With simple trigonometrical calculations, we can find that the small differential area da can be expressed as $da = r^2 sin\theta d\theta d\phi$ and the differential volume as $dv = r^2 sin\theta dr d\theta d\phi$



Spherical coordinate system

2.4.2.2. Gravity due to Spherical Shell

First, let us consider a thin-walled, spherical shell with radius α and uniform surface density σ (think about those hollow colorful plastic balls in a kids ball pit, but with much much thinner skin). The ball is perfectly symmetric, and thus we can use this property to form the problem in a spherical coordinate system. We will derive the gravity in point **P** which is outside of the shell with a distance of **R** from the shell (r in the figure is the distance of any point in the shell to P)**.**



A spherical shell

Recall that the gravitational potential equation from the previous section can be expressed as such:

$$U(P) = \gamma \int_S \frac{\sigma(Q)}{r} dS$$

This is true for a mass distribution that spreads over a vanishingly thin surface, where σ has a unit of mass per unit area (kg/m^2 in SI).

Transforming the above equation to our spherical coordinate system (that is, substituting $dS = a^2 sin\theta d\theta d\phi$ and assuming that $sigma$ is constant at the differential area), the above equation can also be expressed as:

$$U(P) = \gamma \int_0^{2\pi} \int_0^{\pi} \frac{\sigma(Q)}{r} a^2 sin\theta d\theta d\phi = \gamma\sigma a^2 \int_0^{2\pi} \int_0^{\pi} \frac{sin\theta}{r} d\theta d\phi$$

With simple trigonometrical calculations, we can find that $r = (R^2 + a^2 - 2aRcos\theta)^{\frac{1}{2}}$

Since the following relation is true:

$$\frac{dr}{d\theta} = \frac{aRsin\theta}{r}$$

$$sin\theta d\theta / r = dr/aR$$

That means we can further simplify the equation as:

$$U(P) = \gamma\sigma a^2 \int_0^{2\pi} \int_0^{\pi} \frac{1}{aR} dr d\phi$$

With that in mind, we can substitute the integral bounds in terms of the radius, with $R - a$ and $R + a$ being the minimum and the maximum value of r respectively, thus turning our gravity potential equation to:

$$U(P) = \frac{2\pi\gamma\sigma a}{R} \int_{R-a}^{R+a} dr$$

Thus if we evaluate the integral, we can get the **gravity potential at point P due to a spherical shell, if <u>P is outside the shell</u>:**

$$U(P) = \gamma \frac{4\pi a^2 \sigma}{R}$$

Since $M = 4\pi a^2 \sigma$ where M is mass, then the above equation can also be expressed as:

$$U(P) = \gamma \frac{M}{R}$$

Notice that this is the same gravity potential from a point source. In other words:

"Gravity potential at <u>any point outside a uniform shell</u> is the same as the potential of a <u>point source located at the center of the shell</u> with mass equal to the total mass of the shell"

We can conclude the same with the **gravitational attraction** ($g(P) = \nabla U(P) = -\gamma \frac{M}{R^2}\hat{r}$), thus the **gravitational attraction at any point outside a uniform shell is the same as the attraction of a point mass.** This can be easily verified with $\nabla^2 U(P) = 0$

This emphasize the non-uniqueness problem in gravity, since this implies that the same gravity observation can be form by either point source or spherical shell equally well.

Meanwhile, if point P is inside the shell, then the above derivation still holds true but with a small difference in the bounds in the integral, which is:

$$U(P) = \frac{2\pi\gamma\sigma a}{R} \int_{a-R}^{a+R} dr = \gamma\sigma 4\pi a = \gamma \frac{M}{a}$$

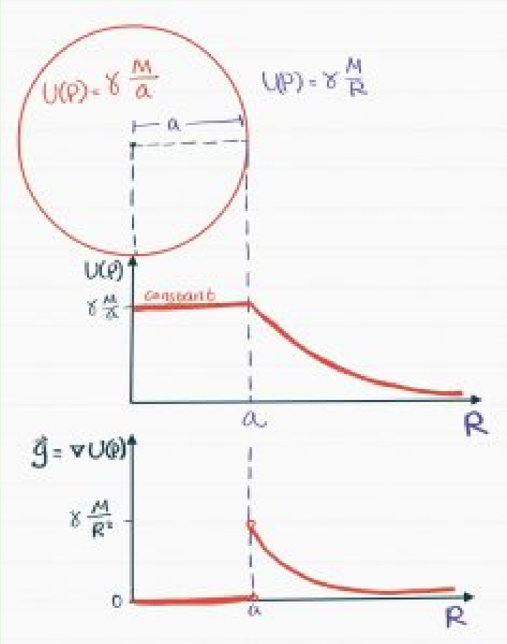
In this case, the **gravitational potential is constant everywhere inside a uniform shell.** Thus the gravitational attraction is:

$$g(P) = \nabla U(P) = \nabla(\gamma \frac{M}{a}) = 0$$

Thus, the above derivations can be simplified in the following summary:

Gravity due to a uniform spherical shell

- If **P** is **outside** the shell, then the following holds true ($M = 4\pi a^2 \sigma$):
  - Gravity potential: $U(P) = \gamma \frac{M}{R}$
  - Gravitational attraction: $g(P) = \nabla(U) = -\gamma \frac{M}{R^2} \hat{r}$
- If **P** is **inside** the shell, then the following holds true:
  - Gravity potential: $U(P) = \gamma \frac{M}{a}$
  - Gravitational attraction: $g(P) = 0$



A visual summary of the gravity potential (top) and gravity acceleration (bottom) due to a spherical shell in terms of R (distance)

2.4.2.2. Gravity due to a Uniform Solid Sphere

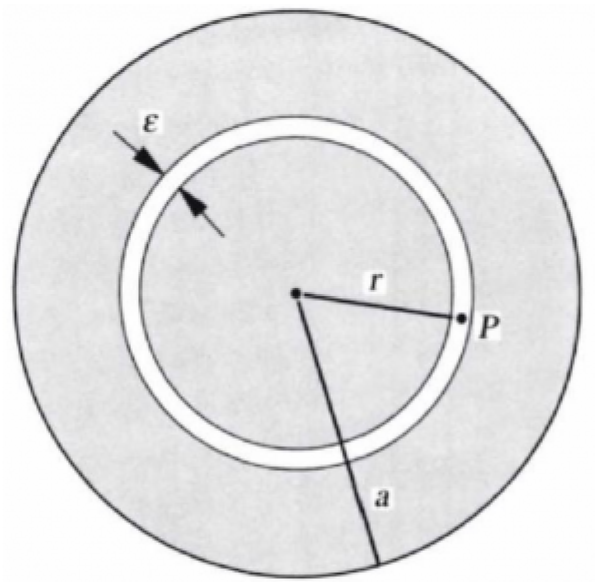
For **P** outside the sphere, we can consider the sphere as a collection of concentric, thin-walled shells with radii ranging from 0 to a (think of an infinite matryoshka, Russian nested doll, but instead it's spherical and the skin are much more thin). Thus, we can apply the **superposition** principle in this case. In other words:

"The potential of a solid sphere at any location outside the sphere is the same as a point mass at the center of the sphere with mass equal to the total mass of the sphere."

$$U(P) = \gamma \frac{M}{R} = \gamma \frac{\frac{4}{3}\pi a^3 \rho}{R}$$

The derivation in the case of **P** being inside the sphere requires a bit more understanding.

Suppose that **P** is in a narrow, spherical cavity of radius *r* and thickness ϵ indicated in the following figure:



Solid spherical sphere

The potential at **P** is due to two parts:

1. The **inner** part of the sphere with radius less that $r - \frac{\epsilon}{2}$
2. The **outer** part of the sphere with radius greater than $r + \frac{\epsilon}{2}$

Let's first looks at the **inner** part (first part). Since in the figure above, P is outside of the first part, we can define the gravitational potential as:

$$U_1(P) = \gamma \frac{\frac{4}{3}\pi(r - \frac{\epsilon}{2})^3 \rho}{R}$$

Now, for the **outer** part (second part), the derivation gets more complex, as we can express it as such:

$$U_2(P) = \gamma\pi\gamma\rho(a^2 - (r + \frac{\epsilon}{2})^2)$$

Adding the two parts together, while setting the thickness $\epsilon \to 0$:

$$U(P) = U_1(P) + U_2(P) = \frac{2}{3}\pi\gamma\rho(3a^2 - r^2)$$

Thus, the gravitational attraction can be expressed as:

$$g(P) = \nabla U(P) = -\frac{4}{3}\pi\gamma\rho r\hat{r}$$

Thus, the above derivation can be summarized as:

Gravity due to a uniform solid sphere

- If **P** is **outside** the sphere, then the following holds true ($M = \frac{\frac{4}{3}\pi a^3 \rho}{R}$):
  - Gravity potential: $U(P) = \gamma \frac{M}{R}$
  - Gravitational attraction: $g(P) = \nabla(U) = -\gamma \frac{M}{R^2}\hat{r}$
  - Laplacian: $\nabla^2 U(P) = 0$
- If **P** is **inside** the sphere, then the following holds true:
  - Gravity potential: $U(P) = \frac{2}{3}\pi\gamma\rho(3a^2 - r^2)$

- Gravitational attraction: $g(P) = \nabla U(P) = -\frac{4}{3}\pi\gamma\rho r\hat{r}$
- Laplacian: $\nabla^2 U(P) = -4\pi\gamma\rho$

2.4.3 Green's Equivalent Layer

In practice, we don't normally consider our objects as uniform spheres. Thus, the previous section where we derived the gravitational attraction of the simple objects aren't normally used in geophysics application. However, it does help us understand the concept of **Green's equivalent layer.**

Recall Green's second identity in section 2.3, which is:

$$\int\int\int_{vol}(V\nabla^2 U - U\nabla^2 V)dv) = \int\int_{surf}(V\frac{\partial U}{\partial \hat{n}} - U\frac{\partial V}{\partial \hat{n}})ds$$

This can in turn be further simplified as:

$$V_p = \gamma\int\int\int_{vol}\frac{\rho}{r}dv = -\frac{1}{4\pi}\int\int_{surf}\frac{1}{r}\nabla V.\hat{n}ds$$

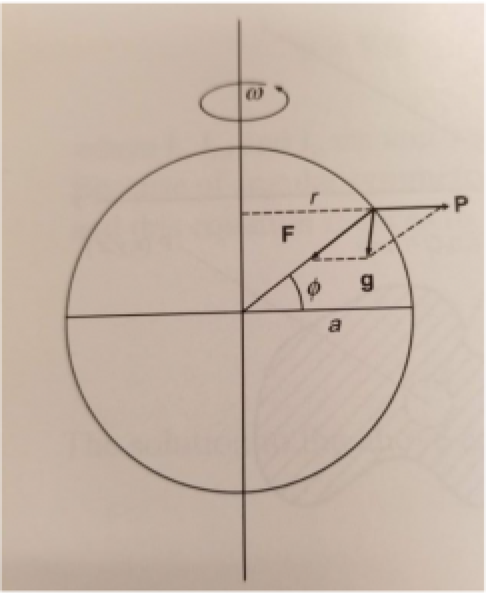
The above equation implies that at any point outside of surface S, the potential caused by a source inside S is the same as the potential caused by a material that is spread over the equipotential surface S with a surface density of $-\frac{1}{4\pi}\frac{\partial V}{\partial \hat{n}}$. Thus, the reverse is also true:

"At any point outside S, the potential caused by a 3D density distribution is indistinguishable from the potential caused by a thin layer of mass spread over any of its equipotential surface S with a surface density of $-\frac{1}{4\pi}\frac{\partial V}{\partial \hat{n}}$*."*

2.4.4 The Earth's gravitational field

2.4.4.1 Centrifugal force and gravitational force

If Earth is a stationary non-rotating spherical body with uniform density distribution, the strength of gravitation acceleration would be constant over the surface. However, the Earth is rotating. The centrifugal force P is created by the rotation of Earth and the gravitational force F is created from non-rotating Earth. Thus, the gravitational force F and the centrifugal force P combine to yield the observed gravitational force g. Consequently, the Earth's gravity field decreases from poles to equator.



LaFehr and Nabighian, 2012, P10. P *is centrifugal force.* F *is gravitational force.* g *is observed gravitational force.*

2.4.4.2 Total gravity and theoretical gravity

The Earth's gravity is due to both the mass of the Earth and the centrifugal force caused by Earth's rotation, so the total potential is the sum of its self-gravitational potential U_g and its rotational potential U_r. The equation of total gravity potential is below:

$$U = U_g + U_r$$

where,

$$U_r = \frac{1}{2}\omega^2 r^2 \cos^2 \lambda$$

ω is angular velocity ($\omega = 7.292 * 10^{-5} rad/s$), r is the axial radius and λ is latitude.

Theoretical gravity takes into account the effect of Earth's spin on gravity. After gravity survey, the first thing we need to do is data processing. We need to remove the effects of general background (including the effect of Earth's rotation) by subtracting theoretical gravity from measurements.

$$g_0 = g_e \left(1 + \alpha \sin^2 \lambda + \beta \sin^2 (2\lambda)\right)$$

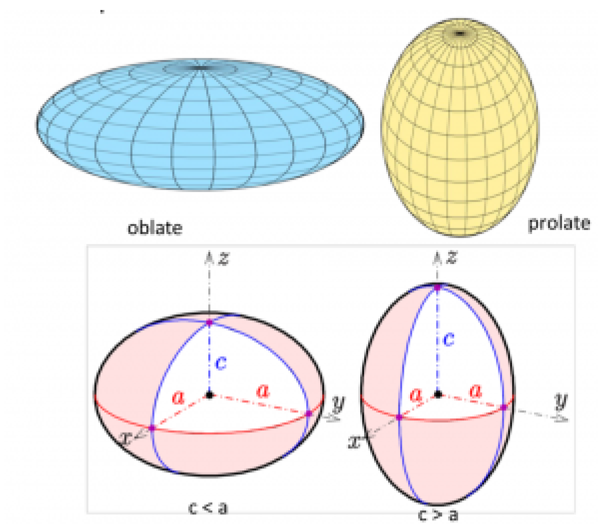
$$g_0 = g_e \left(\frac{1 + k \sin^2 \lambda}{\sqrt{1 - e^2 \sin^2 \lambda}} \right)$$

where, g_e is gravitational attraction at equator.

Note that the theoretical gravity is the gravity due to a rotating and uniformly dense spheroid (i.e., an idealized and overly simplified Earth).

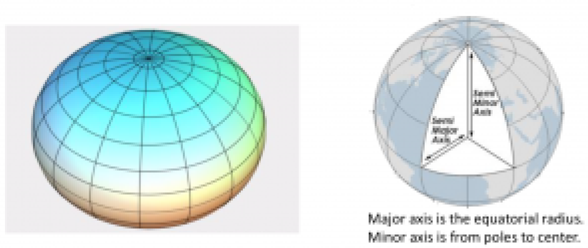
2.4.4.3 The shape of the Earth

As the Earth spins, the centrifugal force causes the Earth to bulge at the equator. So the Earth has a spheroidal shape. Spheroid is the surface obtained by rotating an ellipse about one of its principal axes. Spheroid is therefore an ellipsoid with two equal semi-axes. It has circular symmetry.



Spheroids with vertical rotational axes

Therefore, the Earth is an oblate spheroid shape (seeing figure below) and the difference between the major axis and the minor axis around 20 km.



Oblate spheroid planet Earth

Shape and axes length of the Earth

2.4.4.4 Reference ellipsoids and geodetic datums

Geodesists have adopted an ellipsoid model to determine latitude and longitude coordinates. There are a few such ellipsoid models, also called reference ellipsoids. Different ellipsoid models result in different geodetic datums.

A geodetic datum uniquely defines all locations on Earth with coordinates. Datums precisely specify each location on Earth's surface in latitude and longitude. NAD27, NAD83, and WGS84 are geodetic systems. NAD27 uses the Clarke Ellipsoid of 1866. NAD83 is the most current datum being used in North America. It uses reference ellipsoid GRS80. It forms the basis of coordinates of all horizontal positions for Canada and the US. WGS84 is the reference coordinate system used by GPS. It uses the WGS84 ellipsoid.

Rotation plays a significant role in the Earth. Here, we briefly summarize the rotation:

(a) Rotation causes centrifugal forces, which causes gravity to decrease from poles to the equator.

(b) Rotation causes a difference between polar and equatorial radii, which yields a larger gravitational attraction at the poles compared with the equator (because of the smaller distance to the center).

(c) The combined centrifugal force and flattening effect results in a difference of approximately 5.3 mgal between gravity measurements at poles and equator.

2.4.5 Geoid

Before talking about geoid, let's recall the relevant details about the equipotential surface.

An equipotential surface is a surface on which the potential remains constant. That is:

$$\phi(x, y, z) = cons\tan t$$

Suppose $\widehat{\mathbf{S}}$ is a unit vector that is tangential to an equipotential surface of F, then

$$\widehat{\mathbf{S}} \cdot \mathbf{F} = \frac{\partial \phi}{\partial \widehat{\mathbf{S}}} = 0$$

2.4.5.1 Basic concepts of geoid

(a) Geoid is an equipotential surface.

(b) The gravitational field is the norm to the geoid and defines the vertical direction at any location.

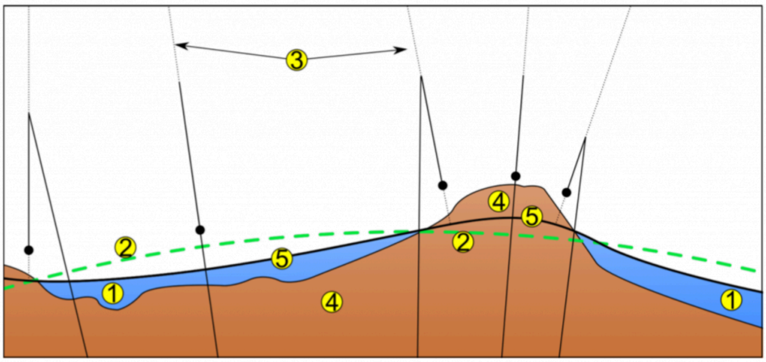
(c) Geoid is the equipotential surface that coincides with the mean ocean surface (or mean sea level) (assuming no tides, ocean currents, winds, etc.) and extends through the continents.

(d) The geoid at any point on land can be thought of as the level of water in an imaginary canal connected at each end with the ocean.

Here is a question. Why the mean ocean surface is an equipotential surface for the Earth's gravitational field?

The figure above is the ocean. If the ocean is not an equipotential surface, there will be a horizontal component of the Earth's gravitational field acting on the ocean water, which means gravity will move the ocean water. But in reality, ocean water is stationary (assuming no tides, winds). This phenomenon explains why the ocean is the equipotential surface.

Geoid is closely related to a spheroid. If the spheroid is rotating earth with uniform density, then the geoid and spheroid will coincide. However, we all know that the density within the Earth changes spatially. So, depending on the density distribution, the geoid will be higher or lower than the reference ellipsoid (shown in the figure below). In the figure below, we noticed that the ocean surface (number 1) is higher or lower than the reference ellipsoid (green dash line, the number 2). For a local excess mass, geoid will be higher than the spheroid (i.e., warp outward) in order to keep the potential constant.

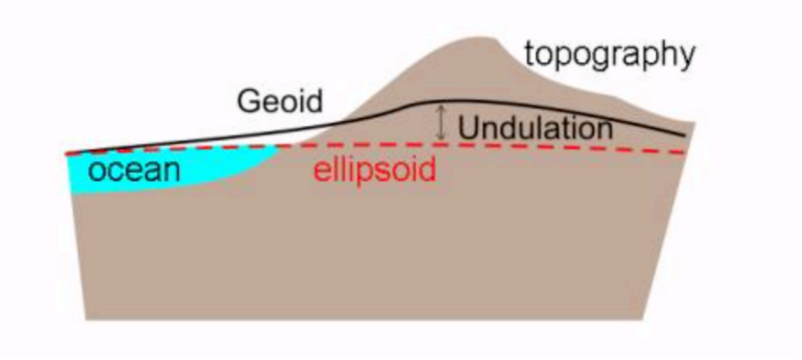


1. Ocean 2. Reference ellipsoid 3. Local plumb line 4. Continent 5. Geoid

Number 1 *is the ocean surface. Number* 2*, green dash line, is the reference ellipsoid. Number* 3 *is local plumb line. Number* 4 *Continent, we can regard this as rock. Number* 5 *is geoid.*

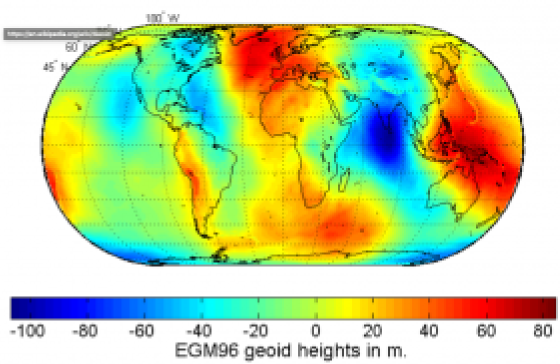
2.4.5.2 Geoid undulations

Geoid undulations are the differences between geoid and ellipsoid. Seeing figure below:



The relationship among geoid, ellipsoid, topography, and undulation.

Here are some examples of geoid undulations. All of the figures have the same trend. The dark blue color is the area where geoid is lower than ellipsoid, which means something has a low density in this area. In comparison, red color is the area where geoid is higher than ellipsoid, which means the high density underneath this area.



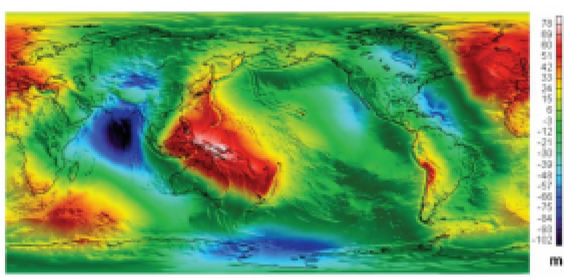
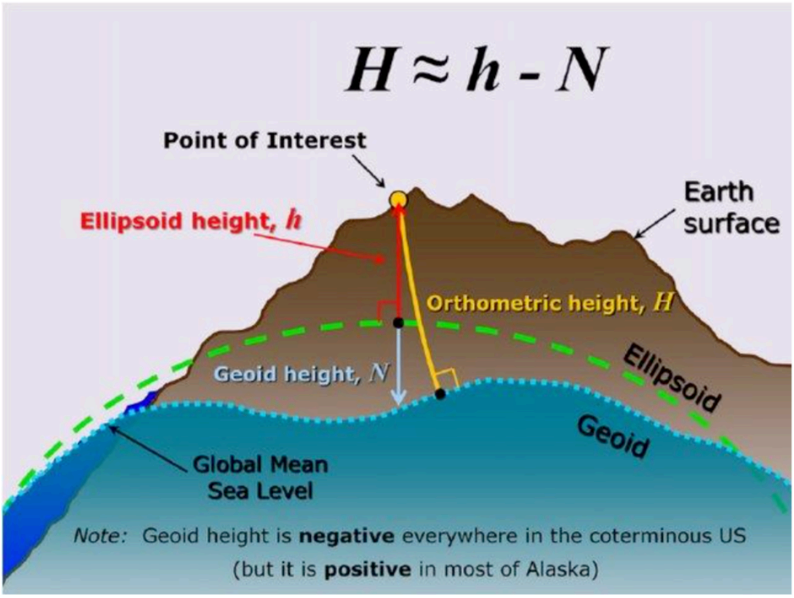
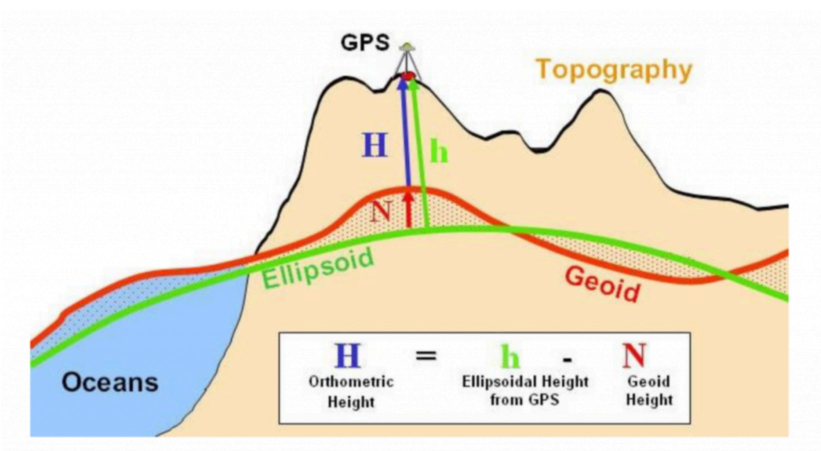


FIG. 3. The 15' × 15' global geoid undulations produced by EGM96 (Lemoine et al., 1998). The undulations range from −107 m to 85 m. Black lines indicate coast lines.

The heights obtained from GPS are typically ellipsoid height, which means the distance between point of interest and ellipsoid. The height displayed on most consumer handheld GPS receiver is, however, the orthometric height, height above mean sea level (LaFehr and Nabighian, 2012, P12s).





Here is an equation to compute orthometric height:

$$H = h - N$$

H is orthometric height means the distance between the observation point to the geoid surface. h is ellipsoidal height means the local plumb distance from the observation point to ellipsoid. The ellipsoidal height is currently obtained from a handheld GPS receiver. N is the geoid height. The geoid height is negative when the geoid surface is lower than ellipsoid, and geoid height is positive when the geoid surface is higher than ellipsoid.

2.5 GRAVITATIONAL POTENTIAL & GRAVITY GRADIOMETRY

2.5.1 A FEW GRAVITY EXAMPLES

2.5.1.1 Unit of gravity attraction/acceleration

(a). In the International System (SI), mass has a unit of kilograms (kg), distance is meters (m).

(b). Gravitational acceleration $\mathbf{g}$ has a unit of $m/^2$. This is a very large unit to use. So, we use cm/sec^2 instead. The unit cm/sec^2 is also refereed to as Gal (short for Galileo).

(c). Unit of gravitational constant γ is $m^3 \cdot kg^{-1} \cdot s^{-2}$.

Derivation

$$\mathbf{g}(p) = -\gamma \frac{m}{r^2} \widehat{\mathbf{r}}$$

If $\mathbf{g}$ has a unit of $\frac{m}{sec^2}$, and m has a unit of kg, and r has a unit of m. The unit of two sides is:

$$\frac{m}{sec^2} =? \frac{kg}{m^2}$$

So, we obtain γ is $6.67 \times 10^{-11} m^3 \cdot kg^{-1} \cdot s^{-2}$. The unit of γ is $m^3 \cdot kg^{-1} \cdot s^{-2}$.

(d). Gal is a more appropriate unit to use, but even Gal is too large for geoscience applications, so we currently use $mGal$.

where, $1Gal = 1000mGal$.

(e). Sometimes, even $mGal$ is too large, in such cases, we use μGal.

where, $1mGal = 1000\mu Gal$.

(g). Microgravity measures minute change in the earth's gravitational filed, which can be used for time-lapse hydrogeophysical studies, aquifer recharge and depletion, fluid monitoring in a petroleum reservoir, fluid flow in geothermal reservoirs, underground cavities.

Microgravity

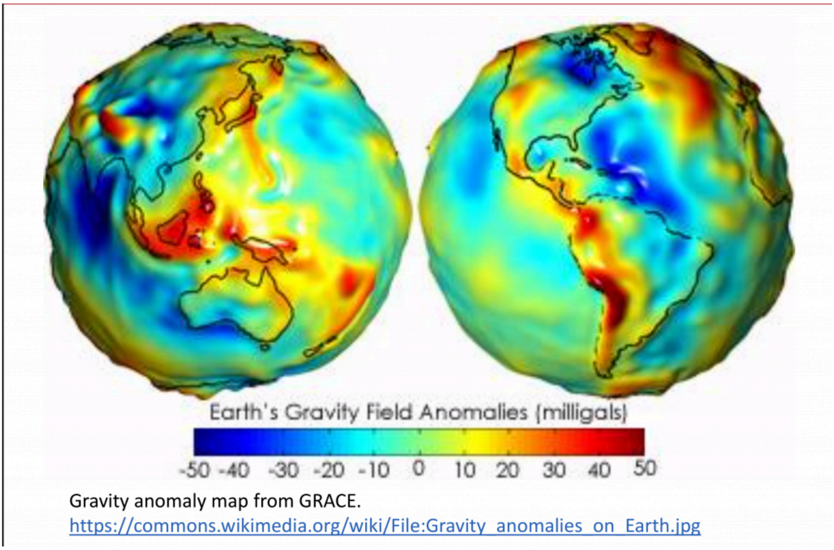
Not to be confused with microgravity experienced by astronauts in space, where microgravity is the condition in which people or objects appear to be weightless.

https://www.nasa.gov/centers/glenn/shuttlestation/station/microgex.html

2.5.1.2 Examples

Example 1

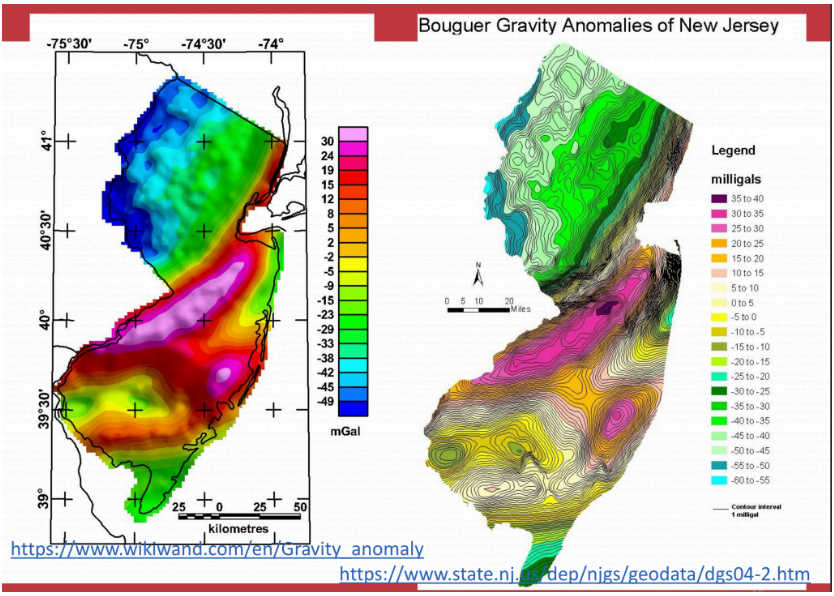
$mGal$ is most common used unit in gravity measures. For global scale, the gravity anomalies is hundreds of $mGal$. Seeing the figure below, the total range of color bar is 100.



Earth's gravity field anomalies (mGal).

Example 2

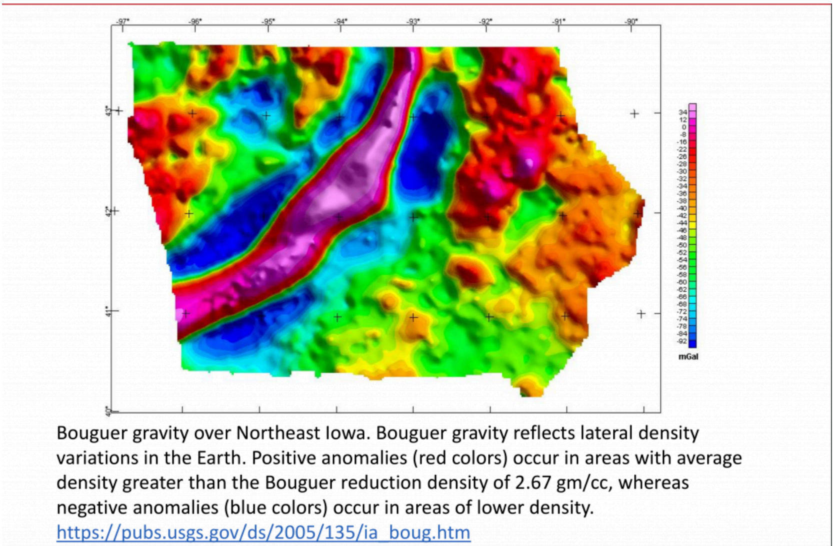
The entire range of bouguer gravity anomalies is hundreds of $mGal$, this range is widely used in geophysics, and we rarely use the Gal. In gravity data processing, we firstly have to correct bouguer anomalies that will remove all effects of mountain and terrain. The warm color reflects high density, the cold color reflects low density. The 0 value here roughly corresponds to average density. In the New Jersey area, the highest feature is characterized as a Northeast to Southwest anomalies caused by basalt intrusion.



Bouguer gravity anomalies of New Jersey

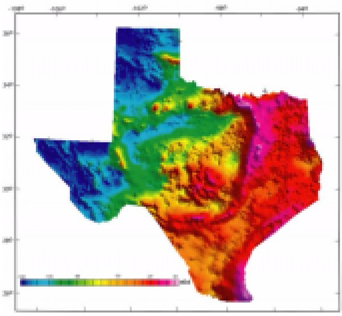
Example 3

The highest anomalies are closely related to mid continent rift, which means something subsurface has really high density.

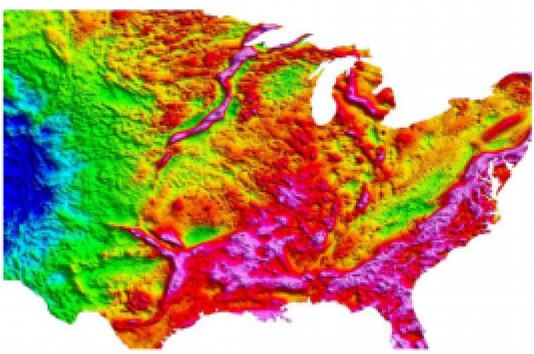


Bouguer gravity over Northeast Iowa.

Example 4



Bouguer anomaly over Texas. https://pubs.usgs.gov/ds/2006/232/texas_boug.htm

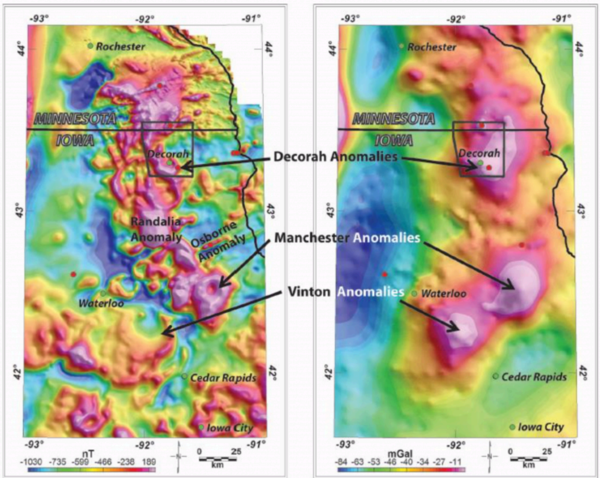


https://source.wustl.edu/2016/09/a-terrible-rift/

Bouguer anomaly over Texas (left) and bouguer anomaly over America.

Example 5

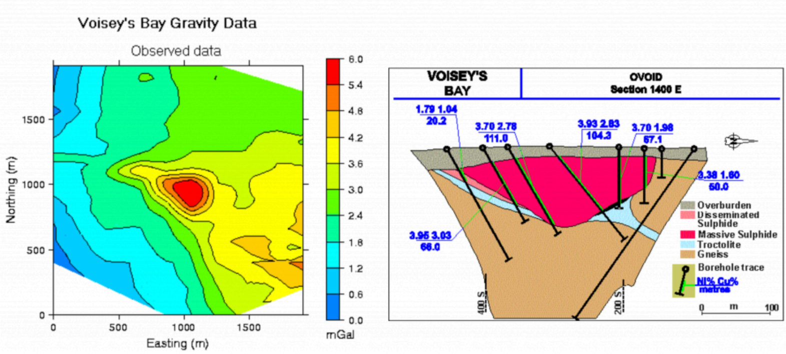
The figure below (left) is reduced-to-pole total field aeromagnetic anomalies. The figure below (right) is bouguer ground gravity data. The range of color bar for bouguer gravity is around $1()()$, which is currently used in mineral exploration.



(Left) Reduced-to-pole total-field aeromagnetic anomalies. (Right) Complete Bouguer ground gravity data. (Drenth et al., 2015)

Example 6

The figure below (left) is observed gravity data, the figure below (right) is geological model built based on drill hole data.



Credit: Yaoguo Li @ CSM

Example 7

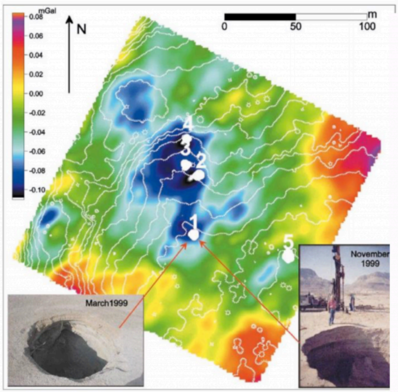


Figure 9. Residual gravity map of the Hever site superimposed by detailed topography (white contours - 0.25-m interval). White circles denote open sinkholes. Photos show growth of sinkhole 1 between March and November 1999.

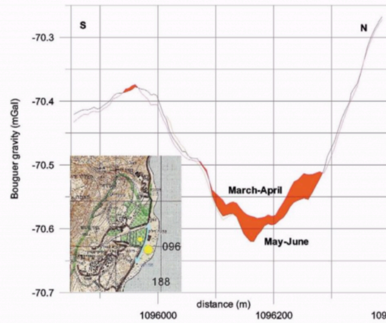


Figure 11. E'n Gedi site. Microgravity monitoring along the road No. 90 during March-June 1999. Insert shows the location of the gravity survey (blue line) and open sinkhole areas (yellow circles).

Rybakov et al., 2001, TLE

Example 8

Time-lapsed gravity measure can be used to monitor fluid flow subsurface.

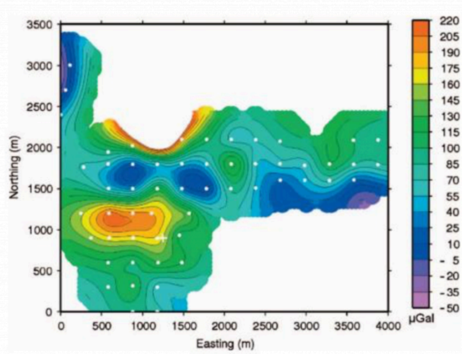


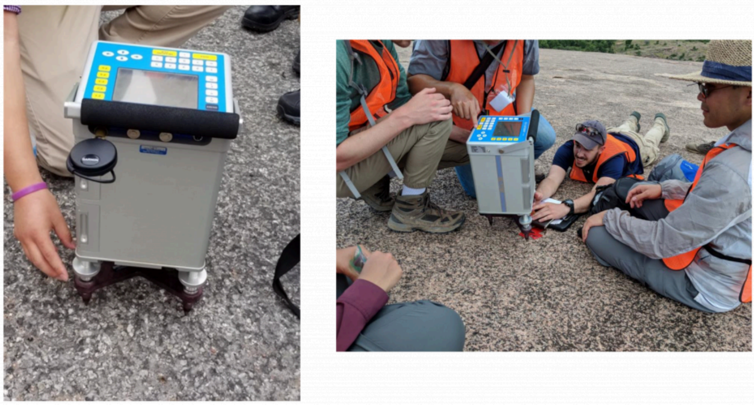
Figure 7. The total gravity difference (µGal) between April 2004 and February 2005. The white + indicates the approximate location of the injection well, and white dots show where data were acquired.

ing, and quantitative interpretation.

Davis et al., (2008)

Example 9

CG-5 is used to measure gravity anomaly in field camp.



Photos from 2018 Geophysical Field Camp

2.5.2 GRAVITY GRADIOMETRY

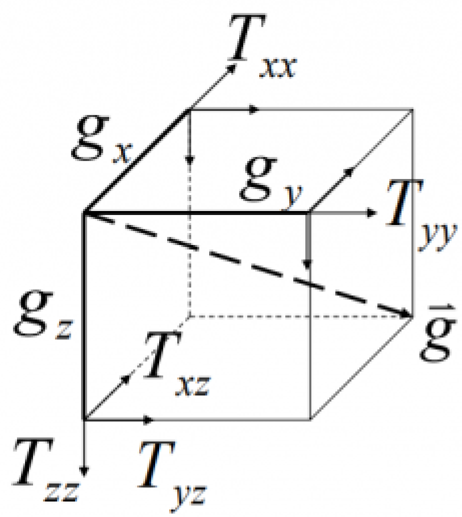
2.5.2.1 Basic Concepts of Gravity Gradiometry

Gravity gradiometry is defined as a measure of the **spatial changes** in gravitational acceleration **g**. In order to better understand gravity gradiometry, we must return to the concept of gravitational acceleration being a vector field. As a vector, **g** is composed of 3 components: $\mathbf{g_x}$, $\mathbf{g_y}$, and $\mathbf{g_z}$. Normally, the influence of the $\mathbf{g_x}$ and $\mathbf{g_y}$ components of gravity are minuscule and commonly disregarded in gravity surveys as most gravimeters measure gravity along the $\mathbf{g_z}$ component. A **spatial change** of $\mathbf{g_z}$ in the **x**-direction is defined as $\frac{dg_z}{dx}$ which measures how fast $\mathbf{g_z}$ changes in the **x**-direction. Similarly, $\frac{dg_z}{dy}$ and $\frac{dg_z}{dz}$ describe the **spatial change** of $\mathbf{g_z}$ in the **y**-direction and **z**-direction.

2.5.2.2 Gravity Gradiometry Tensor

The **spatial change** relationships listed above for $\mathbf{g_z}$ also apply to $\mathbf{g_x}$ and $\mathbf{g_y}$, resulting in 9 components that can be organized into the following tensor:

$$T = \begin{Bmatrix} \frac{dg_x}{dx} & \frac{dg_x}{dy} & \frac{dg_x}{dz} \\ \frac{dg_y}{dx} & \frac{dg_y}{dy} & \frac{dg_y}{dz} \\ \frac{dg_z}{dx} & \frac{dg_z}{dy} & \frac{dg_z}{dz} \end{Bmatrix}$$



This can also be expressed in the simpler form of

$$T = \nabla^T g$$

Recall that gravity is the gradient of gravitational potential: $g = \nabla\phi$. Therefore, the gravity gradient is the **second order** derivative of the gravitational potential and can be expressed as the following:

$$T = \nabla\nabla^T\phi = \begin{Bmatrix} \frac{d^2\phi}{dx^2} & \frac{d^2\phi}{dydx} & \frac{d^2\phi}{dzdx} \\ \frac{d^2\phi}{dxdy} & \frac{d^2\phi}{dy^2} & \frac{d^2\phi}{dzdy} \\ \frac{d^2\phi}{dxdz} & \frac{d^2\phi}{dydz} & \frac{d^2\phi}{dz^2} \end{Bmatrix}$$

While the gravity gradient tensor may look complex with 9 components, there are a few important properties of gravitational potential that simplify the tensor down to only 5 independent components.

- Gravitational potential is well-behaved outside the source region and the tensor is therefore symmetric:
  $$\frac{d^2}{dxdy} = \frac{d^2}{dydx}, etc.$$
- The potential is harmonic outside the source region, meaning that the tensor has zero trace $Tr(T) = 0$:
  $$\nabla^2\phi = 0$$

As a reminder, the trace of a matrix is the sum of its elements along the main diagonal.

One example of the five independent components of the gravity gradient tensor is as follows: $T_{xx}, T_{xy}, T_{xz}, T_{yy}$ and T_{yz}

2.5.2.3 Units of Gravity Gradients

In order to find the units of gravity gradient measurements, we must first start with the units of gravity which is m/s^2. The gravity gradient measures the **spatial change** in gravity **over a short distance**, meaning that it has a unit of $1/s^2$ in SI units, or also $mGal/m$ and $mGal/km$. These units are normally too large to use, so units of Eotvos or Eö are used instead, with 1 eotvos

= $10^{-9}/s^2$ = 0.1 mGal/km = 0.1 μ Gal/m. This is an extremely sensitive unit of measurement!

2.5.2.4 Calculating Gravity Gradient

In order to find the gravity gradient, we must first start with the equation of gravity due to a point source:

$$g(P) = \gamma m \nabla \frac{1}{r}$$

Recall that $T = \nabla^T g$

The gravity gradient of a point source can therefore be represented as:

$$T = \gamma m \nabla \nabla^T \frac{1}{r}$$

In the general case, gravity gradient can be expressed as:

$$T = \gamma m \nabla \nabla^T \frac{\vec{1}}{|\vec{r} - \vec{r'}|}$$

As an example, we will now take a look at the equation for gravity gradient at T_{xx}:

$$T_{xx} = \gamma m \frac{2(x-x')^2 - (y-y')^2 - (z-z')^2}{[(x-x')^2 + (y-y')^2 + (z-z')^2]^{5/2}}$$

Additionally, T_{xz} is represented below:

$$T_{xz} = \gamma m \frac{3(x-x')(z-z')}{[(x-x')^2 + (y-y')^2 + (z-z')^2]^{5/2}}$$

Notes:

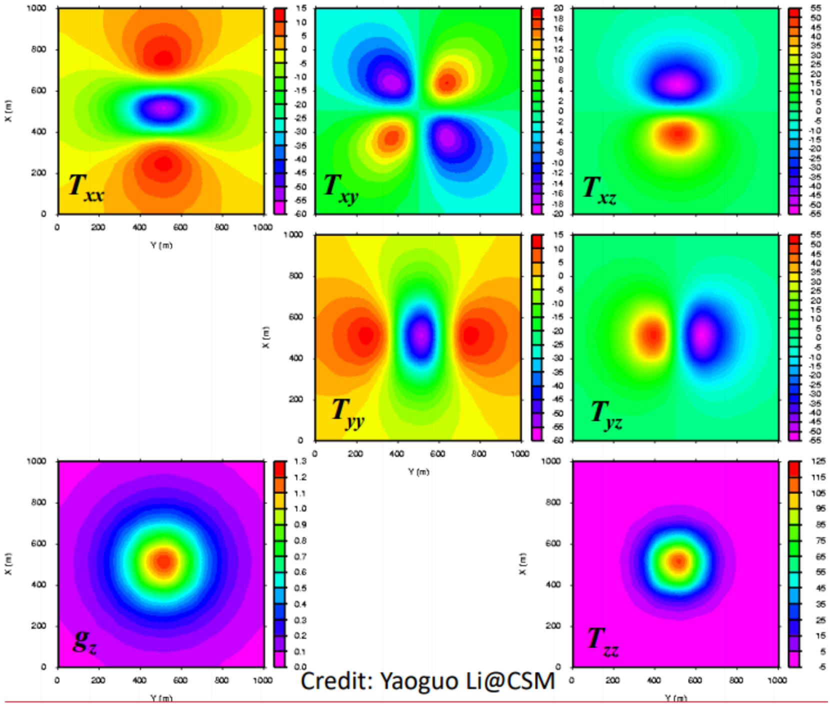
We will see $\nabla\nabla^T \frac{\vec{1}}{|\vec{r}-\vec{r'}|}$ again later in regards to the magnetic field due to a small current loop!

Additionally, **T** and **G** are interchangeable representations of gravity gradient.

2.5.2.4 Examples

Example 1

The image below is of the gravity and gravity gradient responses to an anomaly in the subsurface. g_z is the gravity response of the body, while all components of T represent the gravity gradient response. Notice that T_{zz} is easier to interpret compared to the other components of the gravity gradient.

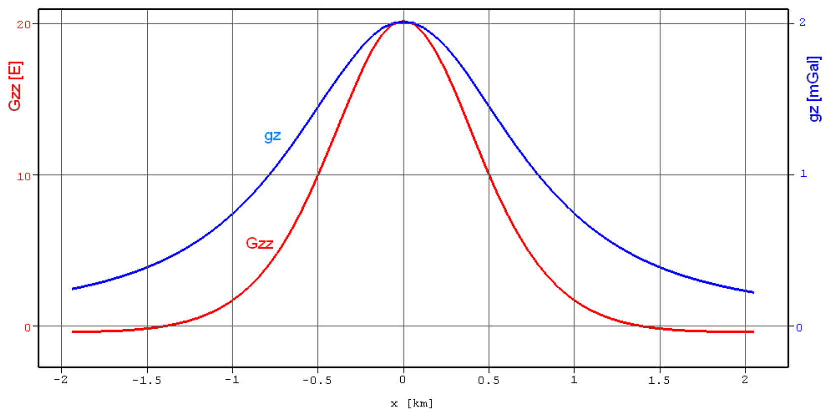


Credit: Yaoguo Li@CSM

Synthetic Gravity Gradient Data. The gravity gradient data is in units of Eotvos and the gravity data is in units of $mGal$

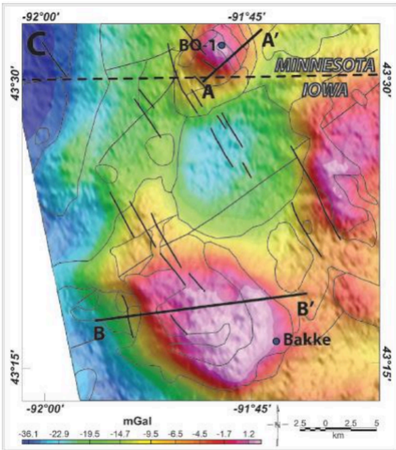
Example 2

Vertical gravity and gravity gradient signals from a point source buried at 1km depth. While the two signals do not share the same units, notice that the responses of g_z and G_{zz} are fairly similar, although the signal of G_{zz} is narrower and better constrained over the anomaly.

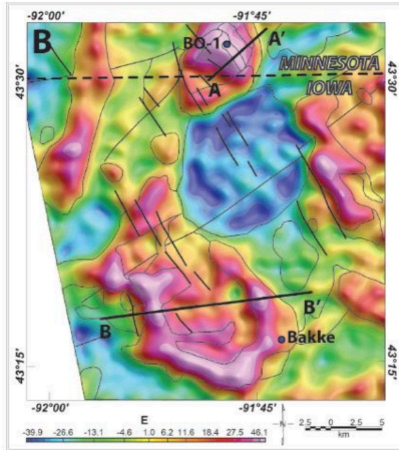


Example 3

Notice the difference between the collected Bouguer gravity and G_{zz} component gravity gradient data collected over the same region. The boundaries of different anomalies in G_{zz} are better defined than in the Bouguer gravity data. AGG stands for Airborne Gravity Gradiometry.



Complete Bouguer gravity computed from AGG and ground gravity datasets.

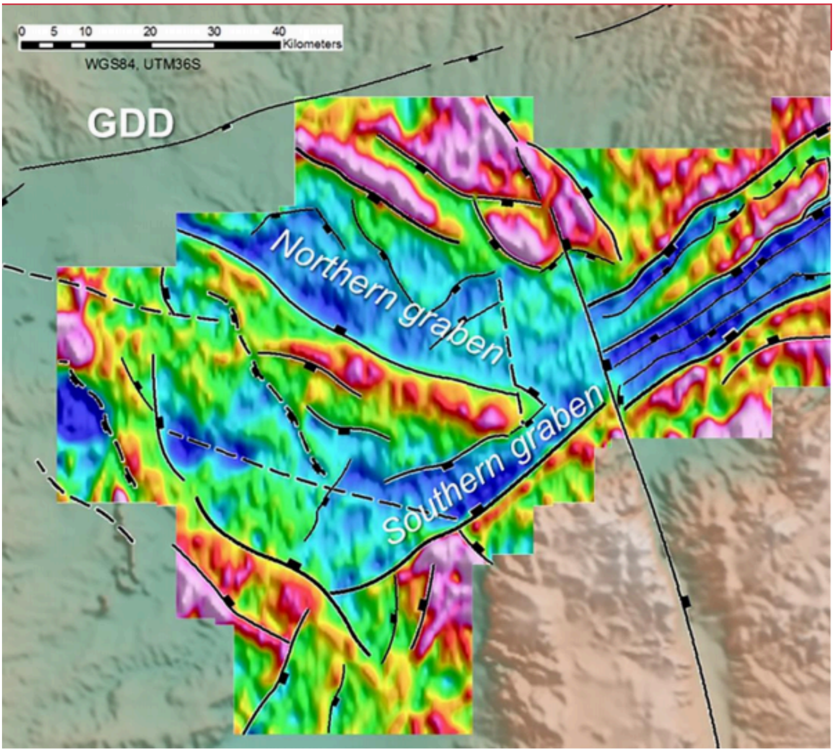


Gzz data math filtered to remove effects of noise and overlying sedimentary rocks

Drenth et al., 2015

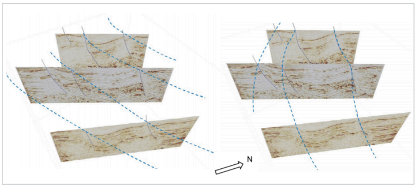
Example 4

Falcon gravity gradient map ($G_{DD} = G_{zz}$) of the western Eyasi rift basin. A 3D model of the area was produced by combining the responses from the Falcon AGG and airborne magnetic data. Basement structural features and sediment depo-centers are easily identified in the model.

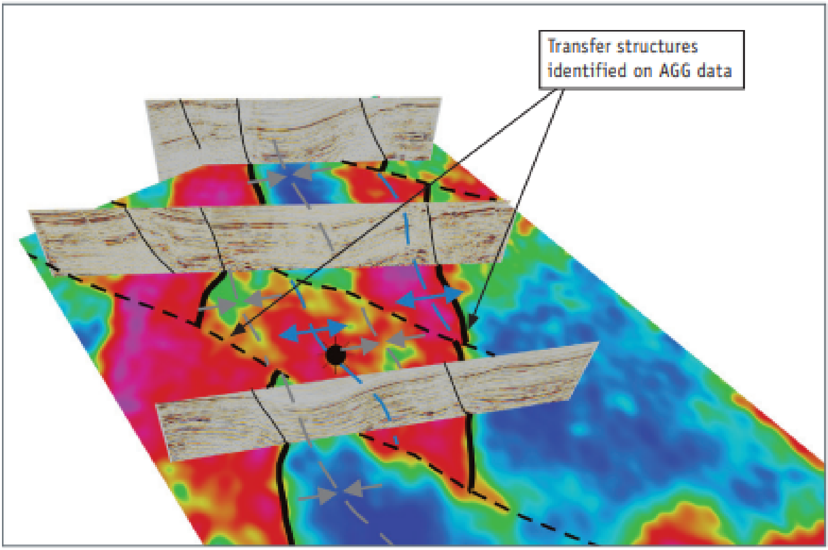


Example 5

The figure below shows two alternative and equally valid correlations of structures on adjacent seismic lines. It is not possible to determine which of the NW/SE or N/S orientation of structures is correct using the seismic alone.



Ambiguity in the interpretation of wide-line-spacing 2D seismic data.



FALCON data removes ambiguity in seismic interpretation.

Example 6

Both figures below are gravity gradiometry surveys conducted over the Vinton Salt dome.

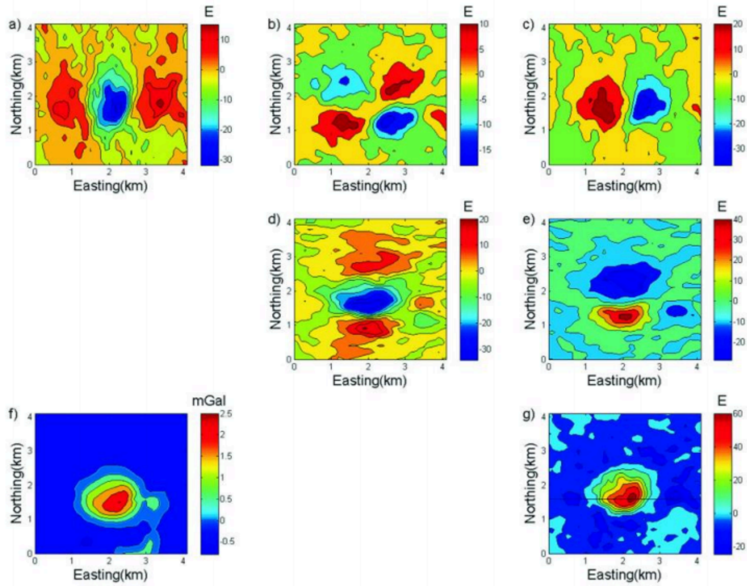


Fig. 18. Observed data over the Vinton dome, USA, (a) g_{xx}, (b) g_{xy}, (c) g_{xz}, (d) g_{yy}, (e) g_{yz}, (g) g_{zz} and (f) g_z. The solid black line in (g) shows the location of the profile in Fig. 20.

Qin et al. (2016, JAG)

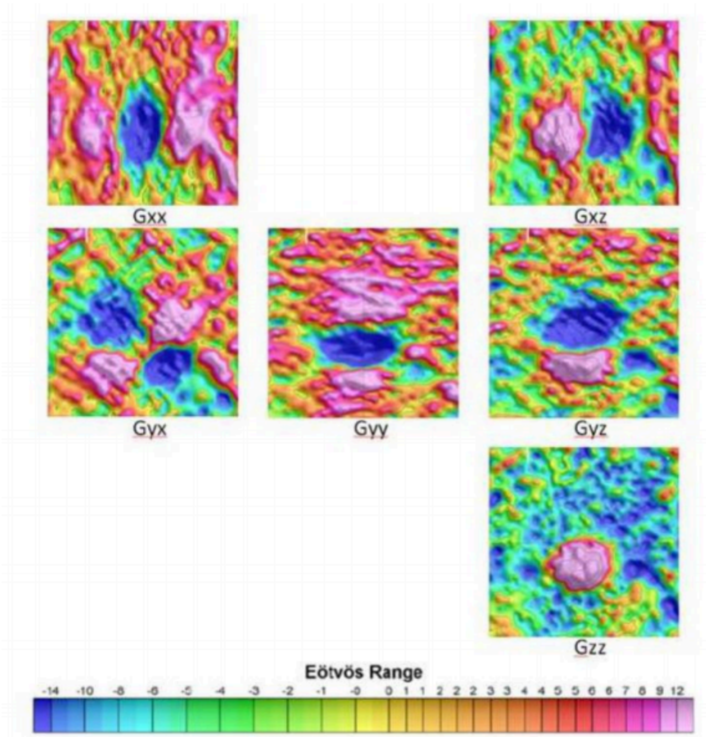
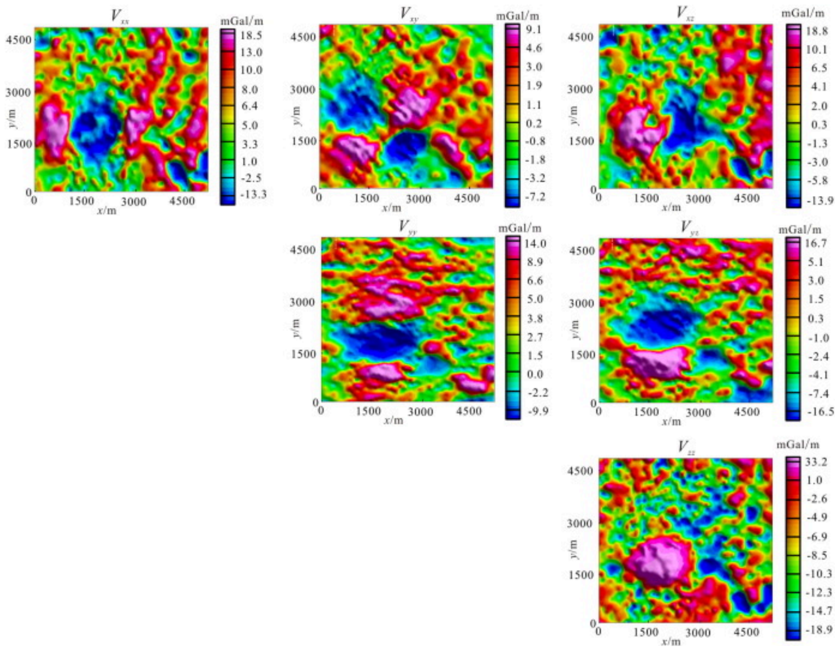


Figure 1. Measured gravity gradients over the Vinton salt dome (Bell Geospace).

Ennen and Hall (2011, SEG)

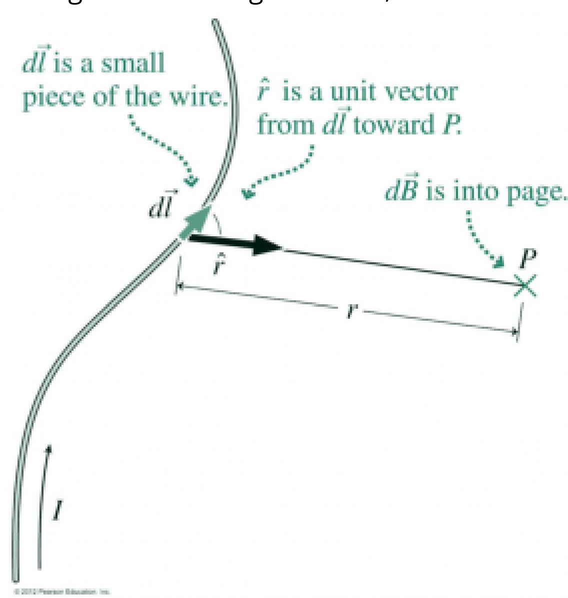
Example 7

Another gravity gradiometry figure of the Vinton salt dome, but note the error made with the units.



2.6. MAGNETIC POTENTIAL

2.6.1. BIOT-SAVART LAW

From classical electromagnetism, we know that electricity and magnetism are inter-related. For example, if we have a wire of current, it will generate a magnetic field, as illustrated in the figure below. 

In order to calculate the current in the wire which induced the magnetism, we can use Biot-Savart law, which is mathematically written as follows:

$$d = \frac{\mu_0}{4\pi} \frac{I dl \times \hat{\mathbf{r}}}{r^2}$$

Integrating around a loop b yields the following expression:

$$\mathbf{B} = \frac{\mu_0}{4\pi} I_b \oint \frac{dl_b \times \hat{\mathbf{r}}}{r^2}$$

Where μ_0 is the magnetic constant permeability of the vacuum (free space), dl is the “current element” directed along the current in the wire, the unit vector $\hat{\mathbf{r}}$ is pointing from the current element towards the observation point where the magnetic field $\mathbf{B}$ is to be calculated. Magnetic field vector $\mathbf{B}$ induced at observation location P due to the steady current in the wire is at right angles to both dl

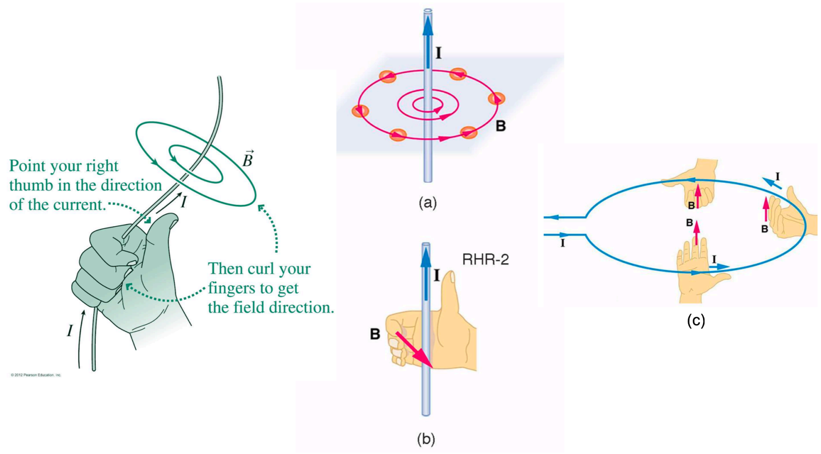
and $\hat{\mathbf{r}}$, and its direction can be determined through the Right-Hand Rule.

The magnetic field of a loop of current can be called as

- Magnetic induction
- Magnetic flux density
- Magnetic field

2.6.1.1.1 Right-hand rule

The directions of the current and magnetism can be represented by our right hand. As shown in the figure, we can point the right thumb in the direction of the current, then curl your figures to get the magnetic field directions.

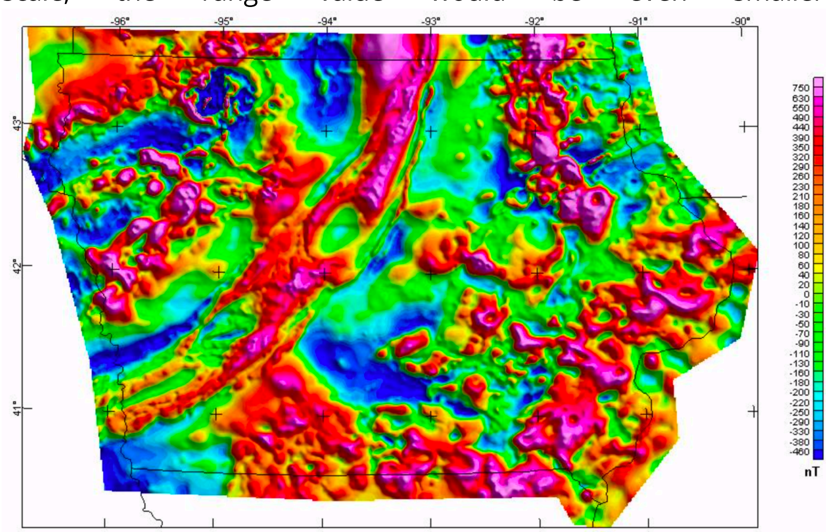


2.6.1.1.2 Magnetic field unit

The SI unit for the magnetic field $\mathbf{B}$ is Tesla denoted as T. However, this unit is too large to represent the actual field measurements. Therefore, we often use nanotesla (nT) instead, where

$$1nT = 10^{-9}T]$$

The Iowa magnetic anomaly map is shown here to demonstrate the value range of the measured magnetic field data, which is about 1300 nT for a regional-scale study area. If in the explorational scale, the range value would be even smaller.



2.6.1.1.3 Implications from Helmholtz Theorem

Before we move on to the magnetic potentials, let us be reminded about the consequences of the Helmholtz theorem. We have learned that a vector field is a solenoidal field in a region if its divergence vanishes everywhere, i.e.,

$$\nabla \cdot \mathbf{F} = 0$$

According to the Helmholtz theorem, the scalar potential becomes zero. Therefore,

$$\mathbf{F} = \nabla \times \mathbf{A}$$

An example of the solenoidal field is the static magnetic field, i.e., a magnetic field that does not change with time. As illustrated in the (figure), magnetic field lines do not emanate from or converge to any point, and since there is no source, these field lines are closed loops. Therefore, its divergence is zero, and magnetic field $\mathbf{B}$ is solenoidal

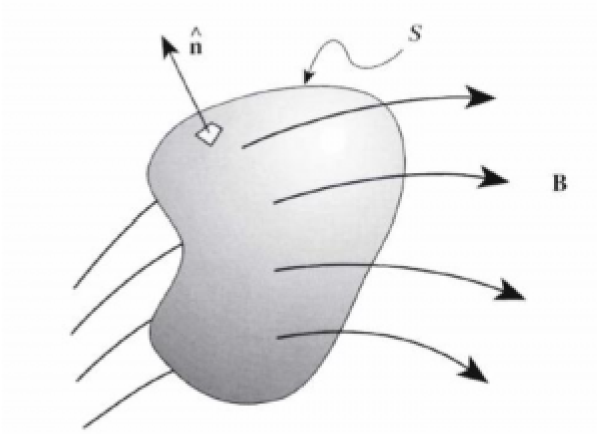
$$\nabla \cdot \mathbf{B} = 0$$

This statement is true everywhere, even within magnetic media.

If applying divergence theorem, we will have an equation as follows:

$$\oint \nabla \cdot \mathbf{B} dv = \oint_S \mathbf{B} \cdot \hat{\mathbf{n}} ds = 0$$

It implies that the net magnetic flux over any closed surface is always zero. And there are no net sources (or sinks) anywhere in the space. Therefore, magnetic monopole does not exist. These implications are demonstrated in the figure below, where all magnetic field lines entering and exiting the shaded area no matter how many field lines there are. Thus, there is no magnetic volume.



Furthermore, according to Helmholtz theorem, we will then have

$$\mathbf{B} = \nabla \times \mathbf{A}$$

Where $\mathbf{A}$ is a vector potential and always exists.

However, under certain circumstances, the scalar potential for the magnetic field also exists at the same time (note that vector potential still always exists). In order to explain this phenomenon, we need to review the Ampere's Law.

2.6.1.2. Ampere's Law

In the previous section, you see **Biot-Savart law** that gives you the equation of the magnetic field (**B**), which is the following:

$$\mathbf{B} = \int \frac{\mu_0}{4\pi} \frac{dl \times \hat{\mathbf{r}}}{r^2}$$

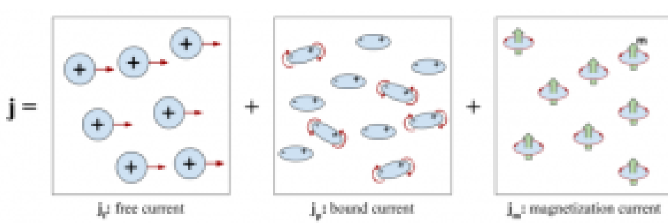
Ampere's law gives the following relation between magnetic field **B** and current density **j**:

$$\nabla \times B = \mu_0 j$$

Both the Biot-Savart and Ampere's law essentially expresses the same concept but in different forms. Both equations tell us that *if you have current, then you will have a magnetic field (and vice versa).*

To learn more about how to derive Ampere's law from Biot-Savart law, please refer to Page 229-233 in David J. Griffith's book (Introduction to Electrodynamics, Fourth Edition).

Note that **j** in Ampere's law refers to the **total current density,** which consists of three parts: free current(j_f), bound current (j_p), and magnetization current (j_m). The division between the current densities are not so relevant for this chapter, but if you wish to learn more about this topic, you can do so through this link.



Different types of current density

However, a question arises from Ampere's law, **what if there is no**

current? Putting it on the perspective of Ampere's law, if there is no current in the region of investigation, then this holds true:
$\nabla \times B = 0$

This implies that the divergence of the magnetic field, and thus **B is irrotational.** Recall Helmholtz theorem. If **B** is irrotational, then that means there exists a **scalar potential V** that satisfies $B = -\nabla V$.

The above equation is only true *if* the region of study does not have currents. In many geophysical situations, electrical currents are negligible in regions where the magnetic field is measured. Therefore, in general, the scalar potential exists outside of magnetic materials. This is good news for geophysics applications because we do not have to worry about dealing with vector potentials (which is more complicated). We just need to derive the scalar potential to get **B!** This can be done by deriving the scalar potential V and then taking the negative gradient ($B = -\nabla V$).

2.6.2. MAGNETIC FIELD DUE TO THE MAGNETIC DIPOLE

In this section, we will talk more about magnetic dipoles and how it can help us find the scalar potential of the magnetic field.

2.6.2.1 Magnetic Field due to a Small Loop

When we talk about the magnetic field due to a small loop, we are talking about the magnetic field **outside of the loop,** and therefore the **scalar potential exists.** The basic strategy, in this case, is to derive the scalar potential first.

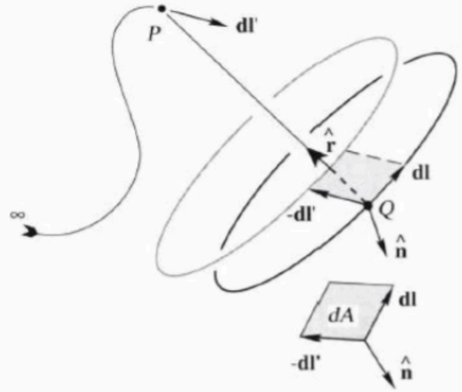


Fig. 4.5. Point P in the vicinity of a current loop. Moving P along $\mathbf{dl}'$ is equivalent to moving the loop along $-\mathbf{dl}'$.

The potential change caused by moving the test particle Q along with the line segment d**l’** is:

$$dV(P) = -B.dl'$$

where **B** is the magnetic field due to the loop. Recall Biot-Savart law:

$$B = \frac{\mu_0}{4\pi} I \oint \frac{dl \times \hat{r}}{r^2}$$

Substituting B from Biot-Savart can give us:

$$dV(P) = -\frac{\mu_0}{4\pi} I \oint \frac{dl \times \hat{r}}{r^2}.dl'$$

Where dl' is a constant and can be moved inside the integral. Also, use the vector identity $A.(B\ timesC) = (A \times B).C$. Thus, the equation becomes:

$$dV(P) = -\frac{\mu_0}{4\pi} I \oint \frac{dl \times (-dl').\hat{r}}{r^2}$$

We introduce the concept of **solid** angle(Ω), which is the solid angle subtended at P by the entire ribbon of wire. In geometry, a solid angle is a measure of the amount of the **field of view** from some particular point that a given object covers. That is, it is a measure of how large the object appears to an observer looking from that point. The point from which the object is viewed is called the *apex* of the solid angle, and the object is said to **subtend** its solid angle from that point.

Thus, the equation can then be rewritten into:

$$dV(P) = \frac{\mu_0}{4\pi} I d\Omega$$

If we assume that the loop is very small in diameter compared to r, we do not need to use the integral. Otherwise, we would need to use the integrals for the solid angle. Thus, with this assumption, the following equation for the magnetic scalar potential due to a small loop can be rewritten as:

$$V(P) = \frac{\mu_0}{4\pi} I \frac{\hat{n}.\hat{r}}{r^2} \Delta S$$

Where Δs is the area of the loop, and $\hat{n}$ is the normal vector of that area.

Now, let us define the **dipole moment** (a vector),mathematically expressed as such:

$$m = I \Delta s \hat{n}$$

The equation for the scalar potential in terms of **dipole moment (m)** is then:

$$V(P) = \frac{\mu_0}{4\pi}\frac{m.\hat{r}}{r^2} = -\frac{\mu_0}{4\pi}m.\nabla_p\frac{1}{r}$$

Thus the **magnetic scalar potential** is:

$$V(P) = \frac{\mu_0}{4\pi}\frac{m.\hat{r}}{r^2}$$

The above equation implies that the function decays as the square of distance. Recall that gravity field also decays as the function of square of distance ($1/r^2$), while the gravitational potential decays as a function of distance ($1/r$). This means, given a fixed dipole moment, the potential value depends on the **distance** and $\hat{r}$. The potential is positive when the angle between **m** and $\hat{r}$ is smaller than 90, and negative when the angle is larger than 90.

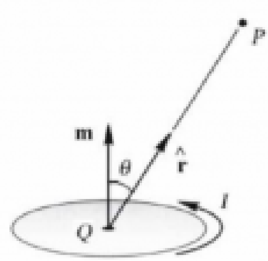
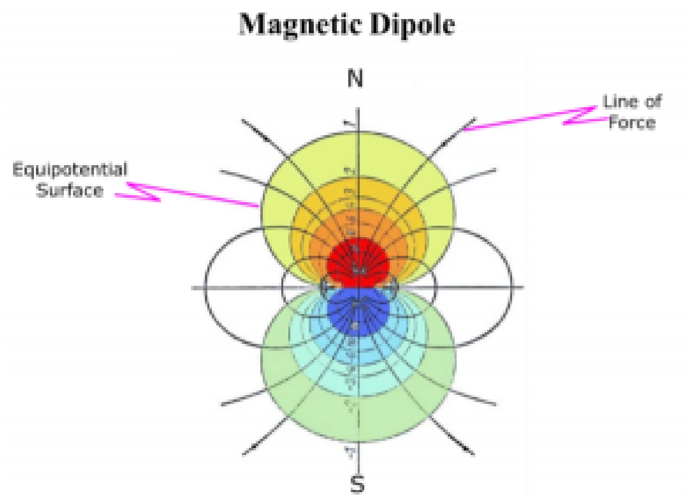


Fig. 4.6. Current loop observed at point P. Vector **m** has direction $\hat{\mathbf{n}}$ and magnitude equal to the current I times the area of the loop.

The following pictures will help you understand the signs of the potential field value better. Suppose the current loop is placed in the middle (origin) of the following picture and the dipole moment is pointing to the North. Then if you place your point **P** in any of the regions, the color of the region in the picture below tells you the sign of potential due to the magnetic dipole in point **P.** In the picture, red indicates more positive values and blue indicates more negative values.



Red is more positive and blue is more negative

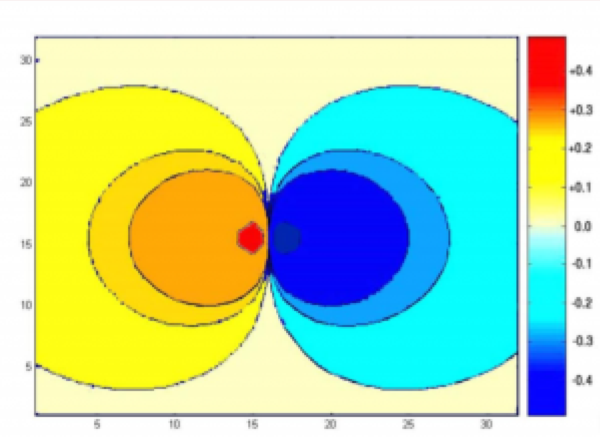
From the above picture, we can see that:

- V(P) is **positive(+)** if the angle between **m** and $\hat{r}$ is **<90 degrees.**
- V(P) is **negative(-)** if the angle between **m** and $\hat{r}$ is **>90 degrees.**
- V(P) is **zero(0)** if the angle between **m** and $\hat{r}$ is **90 degrees.**

For example, we can see that the smallest potential value is achieved when the angle is between **m** and $\hat{r}$ is 180 (the most negative you can get).

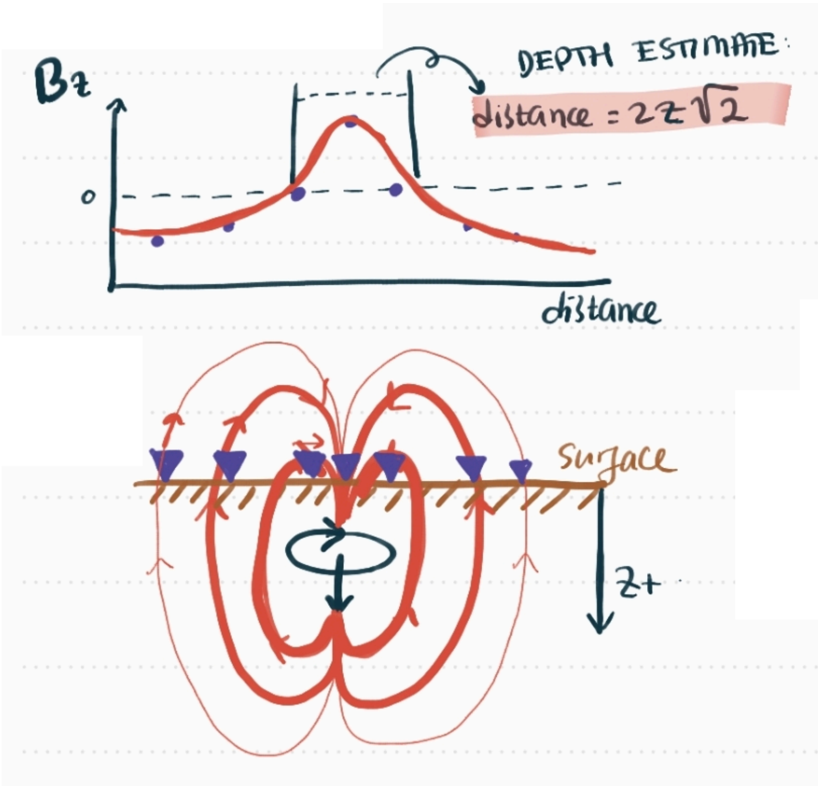
Test your understanding

From the picture below, where the color bar represents potential, can you tell where the direction of the **dipole moment (m)** is?



2.6.2.1 Depth Estimation with Magnetic Dipoles

Suppose a magnetic dipole is buried underground, and you are tasked to figure out the depth of it. We can do so by measuring the field response.

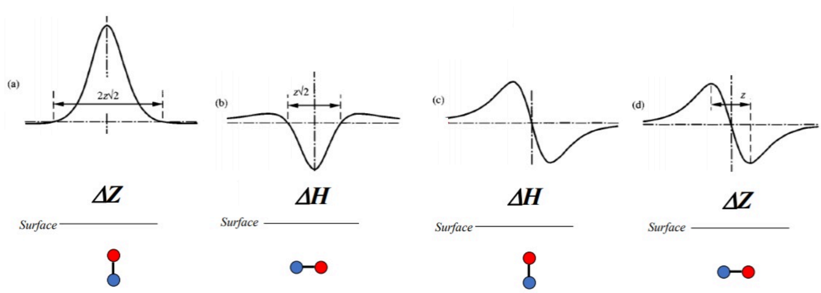


For example, in the above figure, we have a magnetic dipole buried underground and oriented downwards such that the magnetic field lines are oriented vertically. Suppose the blue triangles represent the observation location, which measurements are recorded in the Bz graph above it. The interaction between the field lines that hits the observation location (blue triangle) with the direction of the positive Z axis (pointing downwards) is what is recorded in the Bz graph. For example, the observation station on the left-most side has the field line pointing approximately upwards, and the positive Z axis is pointing downwards. Adding the two directions will lead to having a negative reading in Bz. There are two points in the above figure where the field lines are perfectly orthogonal (the two blue triangles next to the middle triangle) with the positive Z-axis, and thus the measurement of Bz is 0. Lastly, the observation right in the middle of the dipole experiences the most positive Bz because the field lines are in the same direction as the positive Z-axis. This leads to the profile we see in the graph.

It turns out the distance between the zero-crossing in the Bz graph is proportional to the depth of the dipole, specifically, the distance between the two zero-crossing is proportional to $2z\sqrt(2)$ where z is the depth of the dipole.

Similarly, if we have a dipole oriented in the horizontal direction (so the field lines are mostly pointing horizontally), we can find the same response between the positive x axis with the field lines to create the graph for Bx. In this case, the distance between the zero-crossings are exactly the same as $z\sqrt(2)$, where z is the depth of the horizontally-oriented dipole.

Below are other orientation of the dipoles and the relation between the zero-crossing in the graph with the depth (z) of the magnetic dipole.



The different responses of Bx (ΔH) and Bz (ΔZ) from different orientation of the magnetic dipole (two connected circles below the surface)

The above figure tells us that the broadness of the contours and profiles depends on the depth of the dipole ad provides a way of estimating the depth to the source. Please note that the above approximation is only true when we have a single magnetic dipole underground. There will be more complicated calculations involved for more complex cases. However, this approximation is still useful in certain cases. For example, even today we approximate the magnetic response from a seamount is approximately that of a single dipole moment. The shape of the anomaly depends on the shape of the seamount and the direction of the mean magnetization vector.

2.7. MAGNETIZATION

We have learned and derived the magnetic field due to a **magnetic dipole** (i.e., a vanishingly small loop of electrical current). However, this is not very realistic compared to real-life problems. Thus the question is asked: How about the magnetic field due to a volume of magnetic materials? You might have guessed it already, but the field due to a volume of magnetic material is the **sum** of the magnetic effects of all dipoles within that volume.

We define **magnetization** as the **amount of dipoles within a certain volume.** Mathematically expressed as:

$$M = \frac{1}{V}\sum_i m_i$$

Thus, **magnetization** is the **vector sum** of all individual **dipole moments** (m_i) divided by the **volume (V).**

From the above equation, we can tell that the unit of magnetization is **Ampere per meters (A/m)**.

2.7.1 MAGNETIC FIELD DUE TO A VOLUME

Before we sum the magnetic effects of all dipoles, we need to figure out the magnetic effects of only a tiny portion of the volume. So the question is: what is the magnetic field due to a small volume (dv)?

First, recall that the magnetic potential due to small dipole at a certain point (P) is the following:

$$V(P) = -\frac{\mu_0}{4\pi} m.\nabla_p \frac{1}{r}$$

Recall also that the equation for magnetization is:

$$M = \frac{1}{V} \sum_i m_i$$

Thus, rearranging the equation, we can see that the **amount of dipoles(m)** within a certain small volume (dV), is equivalent to **M.dV.**

$$m = M dv$$

That means, the magnetic potential due to a tiny volume is:

$$dV(P) = -\frac{\mu_0}{4\pi} M.\nabla_p \frac{1}{r} dv$$

Note the difference between **dV** and **dv!** Remember that **dV** is the magnetic potential due to a tiny volume and **dv** is the tiny volume of magnetic material.

What about magnetic potential due to big volume? Since a big volume consists of many tiny volumes, then we can **sum/integrate** them to obtain:

$$V(P) = -\frac{\mu_0}{4\pi} \int_{vol} M(Q).\nabla_p \frac{1}{r} dv$$

Where **Q** is the position of dv, and **P** is the position of observation.

Keep in mind the following identity:

$$\nabla_p \frac{1}{r} = -\nabla_Q \frac{1}{r}$$

where:

$$r = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2 + (z_p - z_q)^2}$$

With the above identity, we can then rearrange the equation as:

$$V(P) = \frac{\mu_0}{4\pi} \int_{vol} M(Q).\nabla_Q \frac{1}{r} dv$$

Thus, the magnetic field due to a big volume would be:

$$B(P) = -\nabla_P V(P) = -\frac{\mu_0}{4\pi} \nabla_P \int_{vol} M(Q).\nabla_Q \frac{1}{r} dv$$

2.7.2 MAGNETIC FIELD INTENSITY

2.7.2.1 Total Current Density

Recall Ampere's law which is:

$$\nabla \times B = \mu_0 j$$

where **j** is the **total current density** consisting of **free current** (

j_f), **bound current** (j_p), and **magnetization current** (j_m). Thus, with the different parts of total current density, Ampere's Law can also be expressed as:

$$\nabla \times B = \mu_0(j_m + \nabla \times M + \frac{\partial D}{\partial t})$$

Each part has the following meaning:

- j_m: macroscopic currents (e.g., **free current** caused by moving charges)
- $\nabla \times M$: **magnetization currents** due to the motion of electrons in atoms
- $\frac{\partial D}{\partial t}$: displacement currents (*can be ignored for most geophysical applications*)

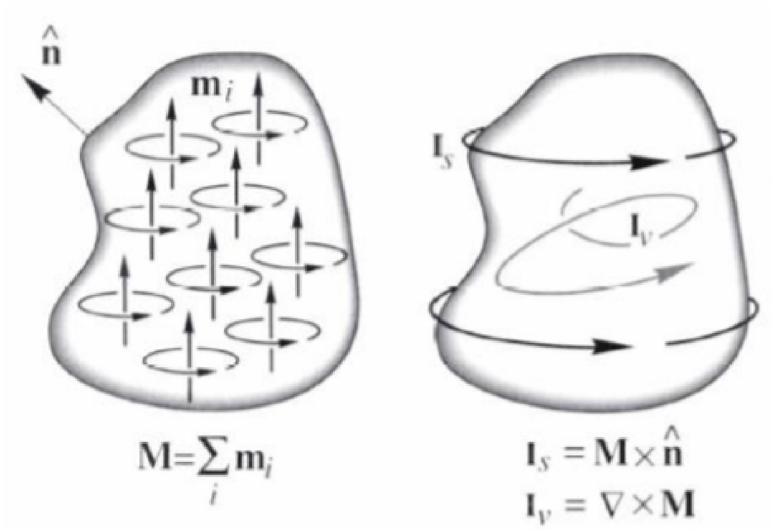
Knowing so, we can also write Ampere's law as:

$$\nabla \times B = \mu_0(j_m + \nabla \times M)$$

2.7.2.2 Magnetization Current

When exposed to an external magnetic field, the dipoles will align themselves accordingly. If the **dipoles are all parallel to each other and have identical magnitude**, circulating current of one dipole will **cancel the current of its neighboring dipoles,** resulting in a **net surface current.** If the magnetization is not uniform, a volume current will exist where circulating elemental currents fail to cancel.

To understand this concept, take a look at the Figure below. The dipoles are all oriented in the same direction and are uniform throughout the volume. The current *inside* the volume will cancel each other out, and thus leaving the *outside* volume to have a surface current (pictured ont he right), which in this case is counter-clockwise.



With simple mathematical arrangements:

$$\nabla \times B = \mu_0(j_m + \nabla \times M)$$

We move μ_0 to the left hand side:

$$\nabla \times \frac{B}{\mu_0} = j_m + \nabla \times M$$

Then, we move M to the left hand side:

$$\nabla \times (\frac{B}{\mu_0} - M) = j_m$$

We can then define the following:

$$H = \frac{B}{\mu_0} - M$$

To form the following:

$$\nabla \times H = j_m$$

The unit of **H** is the same as **M,** and therefore it's **A/m**. We call **H** as the **magnetic field intensity**.

2.7.2.3 Magnetization Field Intensity

As previously stated, the **magnetic field intensity (H)** is mathematically defined as:

$$H = \frac{B}{\mu_0} - M$$

Magnetic field intensity is a hybrid vector with two components with quite different physical meanings.

The major difference between **B** (magnetic induction) and **H** (magnetic field intensity) is that the former originates from both free currents (macroscopic) and magnetization currents (atomic), as summarized by the following equation

$$\nabla \times B = \mu_0(j_m + \nabla \times M)$$

whereas the latter arises *only* from **true currents,** as shown by the following equation.

$$\nabla \times H = \mathbf{j}_m$$

Also, be noted that magnetization **M** is subtracted from the definition of **H**. In other words, magnetization currents are excluded from the definition of magnetic field intensity, leaving behind only macroscopic currents.

Therefore, we can understand **H** as **magnetic induction** (except for a factor μ_0) **minus the effects of magnetization (M)**.

Outside of the magnetic materials, It holds true that $\mathbf{H} = \mathbf{B}/\mu_0$.

2.7.2.4 Magnetic scalar potential

According to Amphere's law,

$$\nabla \times H = \mathbf{j}_m$$

In the absence of such currents, we have:

$$\nabla \times H = 0$$

Thus a scalar potential exists such that:

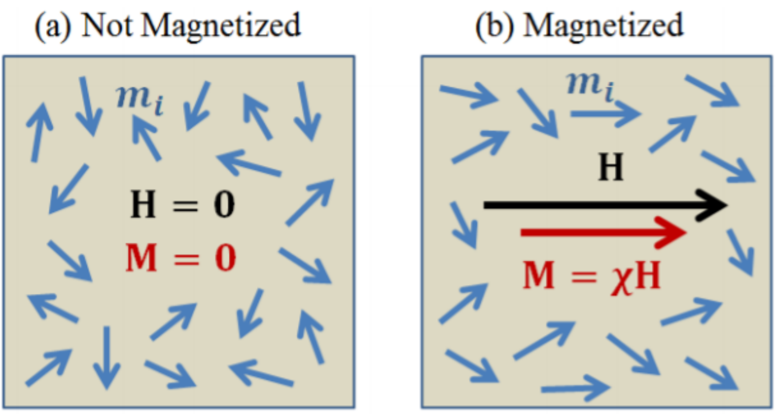
$$H = -\nabla V'$$

2.7.2.5 Magnetized Materials

Magnetic materials can be magnetized in the presence of an external magnetic field. What we mean by **magnetized** is that the **dipole moments are aligned** instead of randomly distributed. Mathematically this means:

$$M = \frac{1}{V}\sum_i m_i \neq 0$$

In the following picture, we see the difference between a non-magnetized (a) and magnetized (b) material. In the non-magnetized material, the arrows representing the dipoles are randomly oriented, and both **H** and **M** are zero. On the other hand, magnetized materials have the dipoles generally pointing on the same direction. In this case, the material is magnetized by an external field **H** pointing to the right, and **M** (which is a product of a material constant and **H**) is also pointing to the right. This leads to the magnetic dipoles (m_i) pointing generally also to the right.



2.7.3 INDUCED MAGNETIZATION

In the section below, we will cover induced magnetization and a couple of the properties associated with it. A couple key concepts that must be understood first include when a volume of magnetic materials acquires a non-zero net magnetization due to an external field, this is called **induced magnetization**. The external magnetic field applying the magnetization is called the **inducing field**.

2.7.3.1 Magnetic Susceptibility

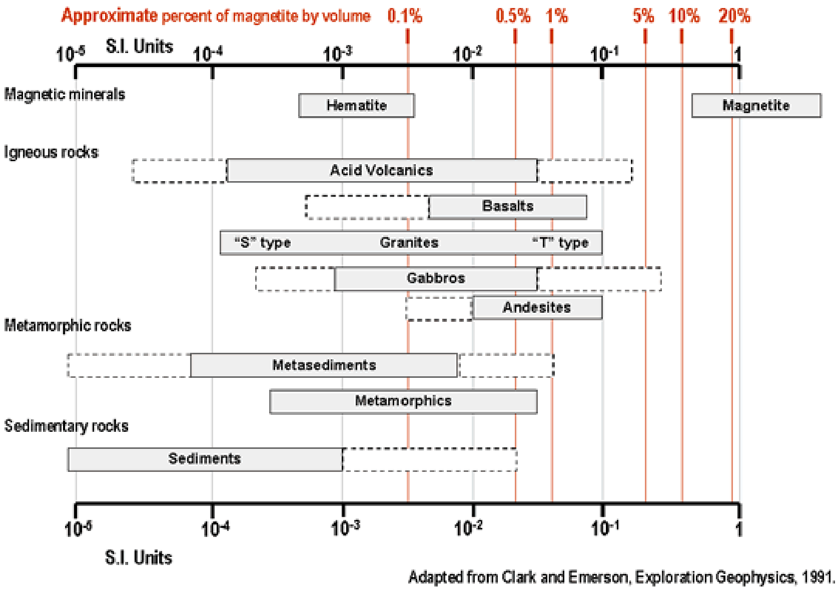
When an **inducing magnetic field** (**H**) is applied to a material, the resulting **induced magnetization** (**M**) an be determined based the material's **magnetic susceptibility** (χ or κ). Magnetic susceptibility defines how easily a material (e.g., a rock) becomes magnetized under an inducing field and is unitless. The following equation describes this relationship:

$$M = \kappa H$$

.

However, this relationship is only true for low-amplitude inducing fields such as the Earth's magnetic field.

From the relationship above, it's clear that materials with high magnetic susceptibilities will also generate strong magnetic signals. In a geological context, **magnetite-rich** rocks are usually the main source of magnetic signals. However, in engineering **buried metal objects** such as water pipes are usually the sources for magnetic signatures. It is important to note that magnetic susceptibility can be positive or negative, and this will be covered in a following section on the types of magnetization.



A figure showing the range of magnetic susceptibility values of a couple different materials. https://www.eoas.ubc.ca/ubcgif/iag/foundations/properties/magsuscept.htm

Earth's Magnetic Field
https://gpg.geosci.xyz/content/magnetics/magnetics_physical_property.html

- The Earth's magnetic field is an inducing field
- Earth's magnetic field is not uniform and will vary depending on location
- Objects will get magnetized differently depending upon where it is situated
- For example, the magnetic signals generated by a steel drum buried at the North pole will be very different compared to an equivalent drum buried at the equator.

Geologic Relationships of Magnetic Susceptibility
Blakely, 1996, p90

- In general, mafic rocks are more magnetic than silicic rocks
  - Basalts are usually more magnetic than rhyolites
  - Gabbros are usually more magnetic than granites
- Extrusive rocks generally have higher remanent magnetization and lower magnetic susceptibility than an intrusive rock of the same composition
- Sedimentary and metamorphic rocks often have low remanent magnetizations and magnetic susceptibilities

2.7.3.2 Magnetic Permeability

Magnetic permeability μ describes the ease at which magnetic flux can pass through a material, similar to how electric conductivity describes the ease at which electricity can pass through a material. The **permeability of free space** which describes the magnetic permeability of a vacuum μ_0 is a constant value of about $1.2566 * 10^{-6} H/m$. Its units are in Henry per meter or $kg * m * s^{-2} * A^{-2}$ in SI units. We can use magnetic permeability to derive the relation between an **inducing magnetic field** (**H**)and the resulting **magnetic field** (**B**) of an object.

We start off using the previously established equation for **magnetic field intensity** (**H**):

$$H = \frac{B}{\mu_0} - M$$

We can then move **magnetization** (**M**) to the same side as **H** before multiplying μ_0 through:

$$B = \mu_0(H + M)$$

We can use the previously established relation $M = \chi H$ to replace **M:**

$$B = \mu_0(H + \kappa H)$$

This can be rearranged into the following form:

$$B = \mu_0(1 + \kappa)H$$

$\mu_0(1 + \kappa)$ is a constant value and can be replaced with μ for the resulting equation:

$$B = \mu H$$

However, outside magnetic materials **magnetic susceptibility** $\kappa = 0$ so the equation becomes:

$$B = \mu_0 H$$

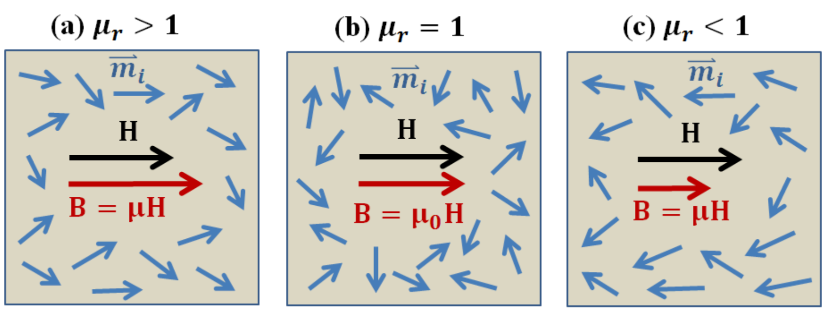
In the derivation above, we replaced $\mu_0(1 + \kappa)$ with μ, and this relationship can be defined as **relative permeability** μ_r**:**

$$\mu_r = \frac{\mu}{\mu_0} = 1 + \kappa$$

2.7.3.3 Kinds of Magnetization

Using the equation established above $\mu_r = \frac{\mu}{\mu_0} = 1 + \kappa$, we can define three relationships between **relative permeability** μ_r and **magnetic susceptibility** κ:

- If $\mu_r > 1$ 1" title="Rendered by QuickLaTeX.com" height="16" width="50" style="vertical-align: -4px;">, $\kappa > 0$ 0" title="Rendered by QuickLaTeX.com" height="12" width="43" style="vertical-align: 0px;">
- If $\mu_r = 1$, $\kappa = 0$
- If $\mu_r < 1$, $\kappa < 0$



https://em.geosci.xyz/content/physical_properties/magnetic_permeability/index.html

The figure above describes partial alignment of magnetic dipole moments under the influence of an inducing field for various cases.

- (a). **Paramagnetic** ($\kappa > 0$ 0" title="Rendered by QuickLaTeX.com" height="12" width="43" style="vertical-align: 0px;">): Dipole moment aligned parallel to the applied field and produce a net magnetization in the direction of the external field.

- (b). **Non-permeable or non-magnetic** ($\kappa = 0$).
- (c). **Diamagnetic** ($\kappa < 0$): Dipole moments aligned opposite to the external field, reducing the magnetic flux density.

Most rocks are **paramagnetic** ($\kappa > 0$ 0" title="Rendered by QuickLaTeX.com" height="12" width="43" style="vertical-align: 0px;">) but some are **diamagnetic** ($\kappa < 0$) with both forms of magnetization being fairly weak. This means that they are insignificant contributors to the geomagnetic field. This is in contrast to **ferromagnetic** rocks, which can produce magnetic signals that are many times greater than paramagnetic or diamagnetic rocks. The strength of ferromagnetic materials comes from neighboring dipole moments interacting strongly with each other which produces a quantum mechanical effect called **exchange energy.** These interactions allow ferromagnetic materials to **retain a magnetization** even with the absence of an inducing field. This type of magnetization is called **remanent magnetization** or **remanence**.

2.7.4 Remanent magnetization

2.7.4.1 Theory for remanent magnetization

(a). Even if you take it to outer space where there is no inducing field, it is still magnetized.

(b). Ferromagnetic materials acquire remanence when they cool through its Curie temperature.

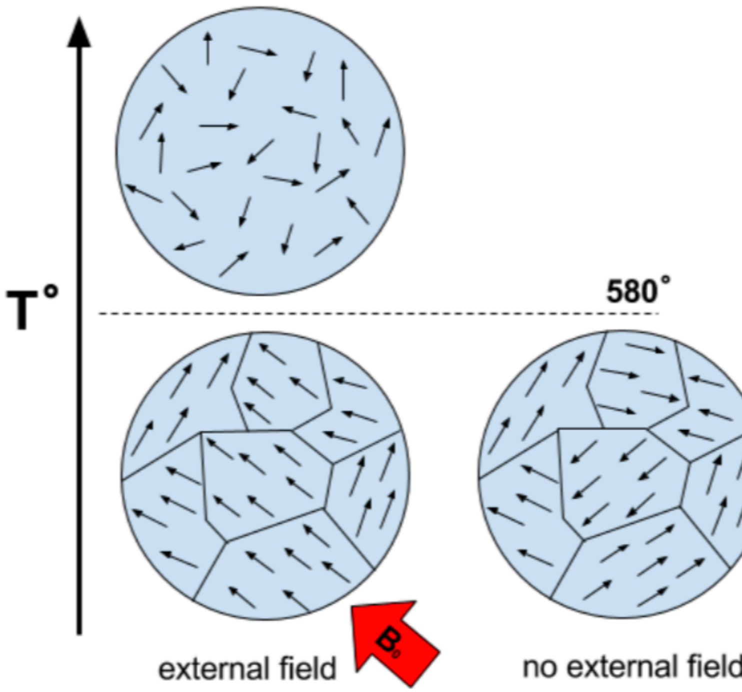
(c). Above Curie temperature, thermal energy prevents dipoles from aligning with the external field.

(d). As the materials cool down, and eventually below Curie temperature, the magnetic dipoles start to align and stay aligned.

Curie temperature
(https://en.wikipedia.org/wiki/Curie_temperature
https://www.britannica.com/science/Curie-point)

- Curie point
- Named after a French physicist, Pierre Curie, who showed that magnetism was lost at a critical temperature.
- For iron, the Curie temperature is 770 $^{\circ}C$ $(1,418^{\circ}F)$.
- Cobalt $(1,121^{\circ}C$ $(2,050^{\circ}F))$, one of the highest Curie point.
- Below Curie point, atoms (behave as tiny magnets) spontaneously align themselves in direction of external magnetic field (they are permanently magnetized).
- Above Curie point, materials lose their permanent magnetic properties.

As the material cools the magnetic particles can stay aligned and eventually lock into place in a domain structure. Each domain has all of its constituent dipoles locked into a single direction. This structure stays in place after the ambient field is removed and the object will have a net remanent magnetism.

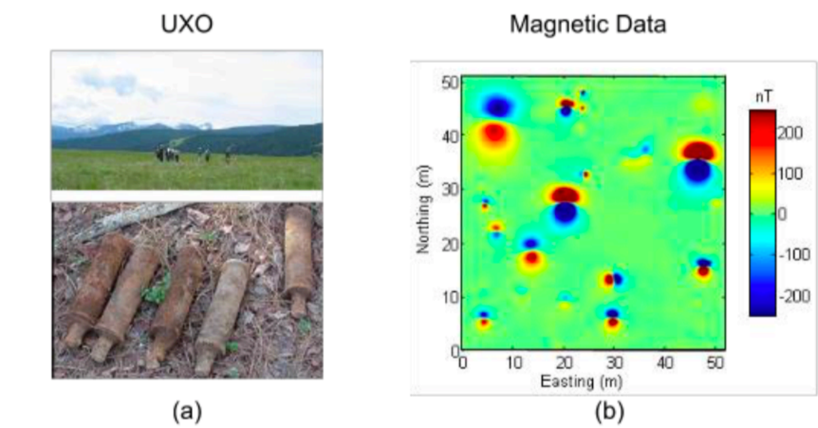


Remanent magnetization. Above 580 thermal energy prevents dipoles from aligning with external field. Below 580, the magnetic dipoles start to align and stay aligned. https://gpg.geosci.xyz/content/magnetics/magnetics_basic_principles.html#magnetization

The temperature within the Earth increases with depth. We know the surface temperature, and we also know roughly how temperature increases with depth. So, a simple calculation shows that at approximately $25km$ depth, the temperature in the Earth would be higher than the Curie temperature of almost all known ferromagnetic materials.

2.7.4.2 Application for remanent magnetization

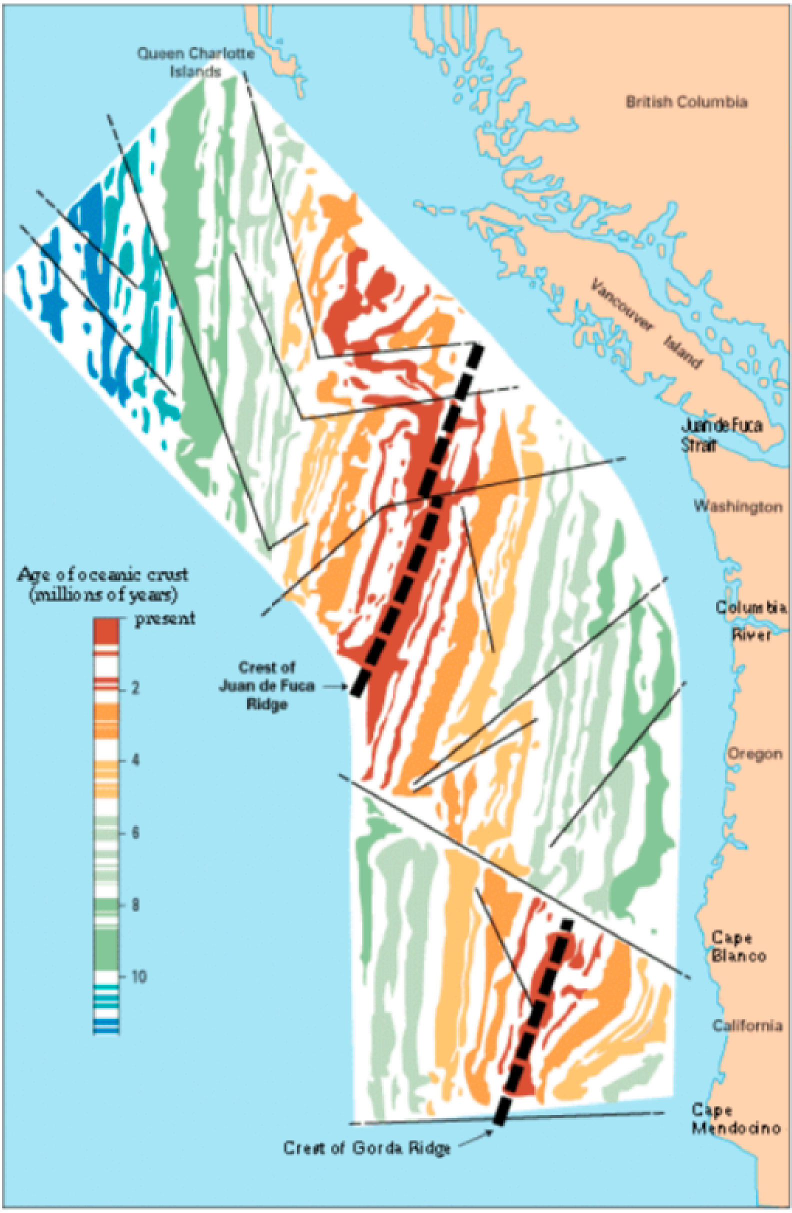
The ordnance items and other man-made objects can be detected through remanence. In the figure (b) below: these kinds of magnetic anomalies cannot be explained by induced magnetization.



(a) Upper: a typical UXO site. There are no surface indications of ordnance items. Lower: typical ordnance items. (b) Magnetic field data over a site contaminated with UXO. https://gpg.geosci.xyz/content/magnetics/magnetics_basic_principles.html#magnetization

From magnetization to plate tectonics, when magnetics materials are magnetized, the magnetization direction is consistent with the direction of the external magnetic field when the magnetic materials were magnetized. Ferromagnetic materials such as basalt lava flow able to lock in a record of the direction and intensity of the magnetic field when they form (i.e., permanently magnetized). These records can be transported, rotated and faulted by plate tectonics. Therefore, these records provide information on not only the past behavior of Earth's magnetic field and but also the past location of tectonic plates. Paleomagnetism: the study of such records in rocks, which played an instrumental role in establishing continental drift and plate tectonics as science.

On the following figure, magnetic stripping offshore Pacific Northwest, which centered around Juan de Fuca Ridge which is a mid-ocean spreading center and divergent plate boundary between Pacific plate and the Juan de Fuca plate. Magnetic striping on either side of the ridge helps data the rock and determine the spreading rate and age of the plate.


Queen Charlotte Islands
British Columbia
Vancouver Island
Juan de Fuca Strait
Washington
Age of oceanic crust (millions of years)
present
2
4
6
8
10
Columbia River
Crest of Juan de Fuca Ridge
Oregon
Cape Blanco
California
Cape Mendocino
Crest of Gorda Ridge

Magnetic stripping (https://pubs.usgs.gov/gip/dynamic/magnetic.html)

2.7.5 Total magnetization

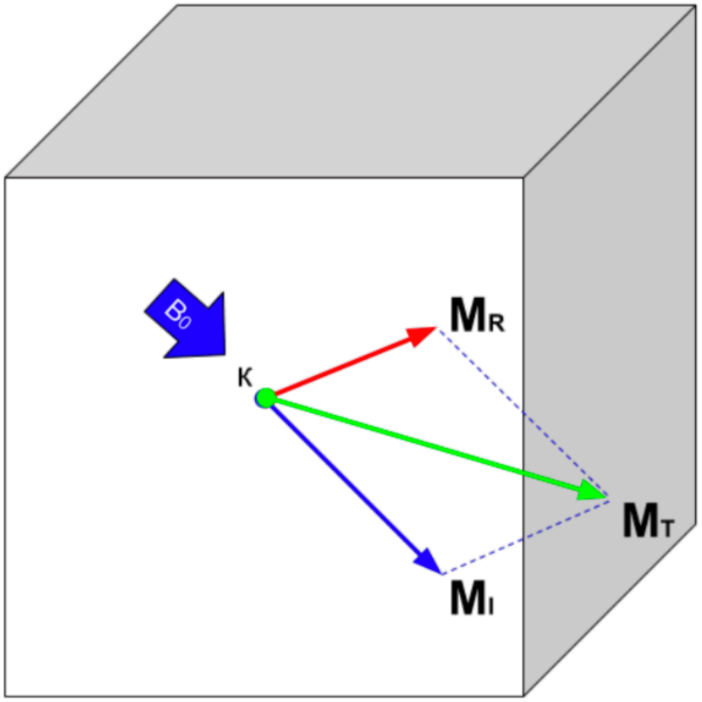
The proper understanding is that the magnetization is composed of two parts: (a) An induced portion ($\boldsymbol{M}_{\mathbf{i}}$) and (b) remanent portion ($\boldsymbol{M}_{\mathbf{r}}$). The net magnetization is:

$$\boldsymbol{M} = \boldsymbol{M}_{\mathbf{i}} + \boldsymbol{M}_{\mathbf{r}}$$

The relationship between magnetization $\boldsymbol{M}$ and the source $\boldsymbol{H}$ (earth's magnetic field) is given by:

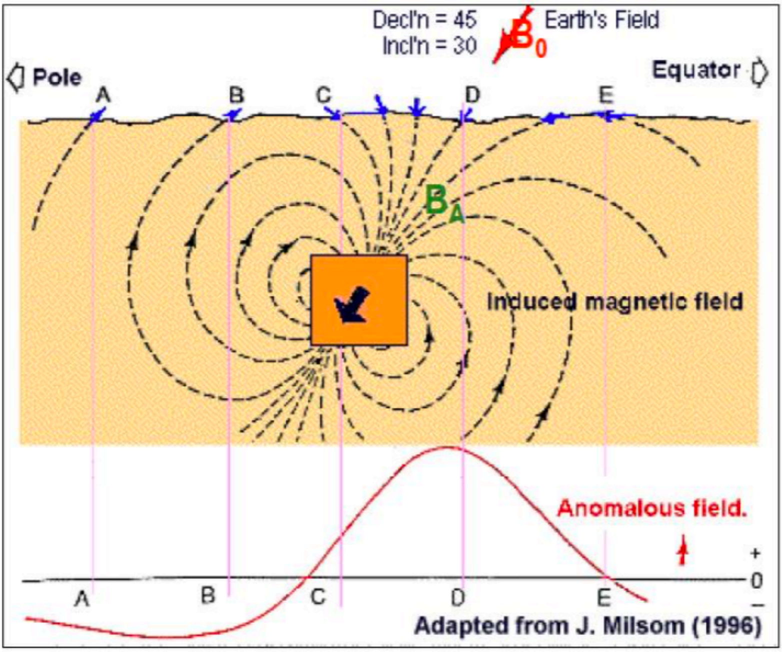
$$\boldsymbol{M} = \kappa \boldsymbol{H} + \boldsymbol{M}_{\mathbf{r}}$$

where κ is the magnetic susceptibility.



https://gpg.geosci.xyz/content/magnetics/magnetics_basic_principles.html#magnetization

The composite field is below:



Credit: Doug Oldenburg @ UBC

Composite field:

$$\boldsymbol{B} = \boldsymbol{B_0} + \boldsymbol{B_A}$$

where $\boldsymbol{B}$ is a vector, $\boldsymbol{B} = [\mathbf{B_x}, \mathbf{B_y}, \mathbf{B_z}]$

Total field:

$$|\boldsymbol{B}| = |\boldsymbol{B_0} + \boldsymbol{B_A}|$$

Majority of instrumentation measures $|\boldsymbol{B}|$, i.e., total magnetic intensity, or TMI. Newer platforms can acquire three-component magnetic field. Latest instrumentation measures magnetics tensor.

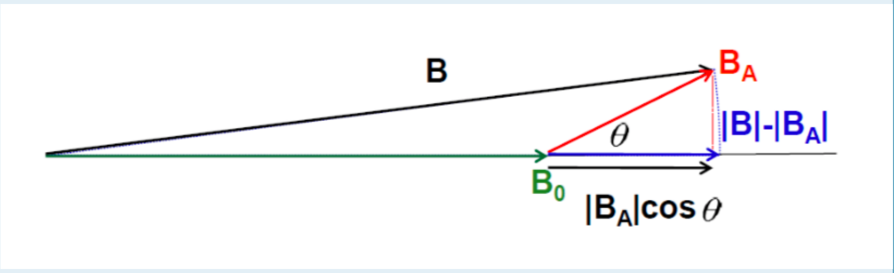
Measured field: $\boldsymbol{B} = \boldsymbol{B_0} + \boldsymbol{B_A}$

The total field anomaly: $\triangle \boldsymbol{B} = |\boldsymbol{B}| - |\boldsymbol{B_0}|$

if $|\boldsymbol{B_A}| << |\boldsymbol{B_0}|$, then that is , total field anomaly $\triangle \boldsymbol{B}$ is the projection of the anomalous field onto the direction of the inducing field.

$$\triangle\overrightarrow{B} \simeq \triangle\overrightarrow{B_A} \cdot \widehat{B}_0$$

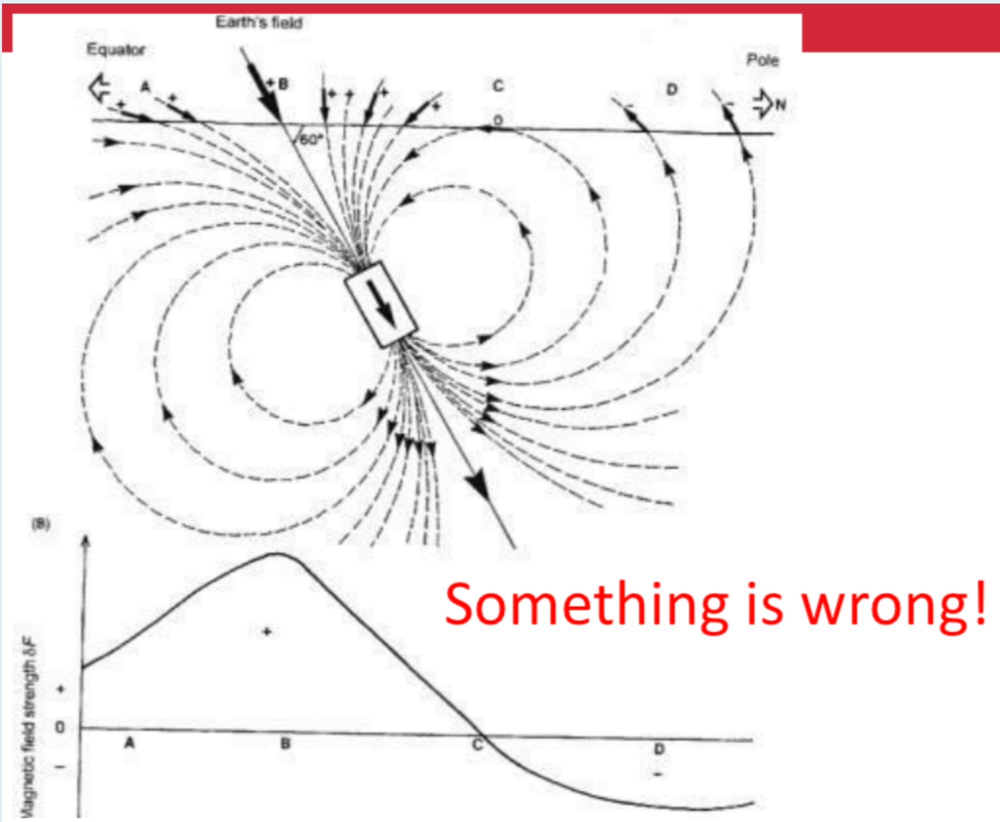
Tips for ploting magnetic anomalous field

- First, find the location with zero field (C and E location above the figure).
- Second, decide $\boldsymbol{B_A}$ direction and calculate the projection of $\boldsymbol{B_A}$ to $\boldsymbol{B_0}$.
- 

Credit: Doug Oldenburg @ UBC

A wrong example

Something is wrong!!!
The magnetic anomaly field is not zero at C location. It should be a small positive value.


Earth's field
Equator
Pole
A
B
C
D
N
60°
(B)
Magnetic field strength δF
+
0
−
Something is wrong!

CHAPTER 3

Chapter 3: Data Acquisition and Reduction

JIAJIA SUN, FELICIA NURINDRAWATI, KENNETH LI, XINYAN LI, AND XIAOLONG WEI

3.1 GRAVITY INSTRUMENT

By now, we have accumulated the basic understandings of the gravity theories, now let us put things in perspective. Assume earth's gravity is approximately 9.8 m/s^2, that is, 980 Gal. The anomalies that we typically see in geoscience applications are typically less than 100 mgal, which is less than 0.01% (one part in 10^4) of the Earth's average gravity field. Therefore, if our interested anomaly is 10 mgal, we are then looking for a tiny signal that is one part in 10^5! If our signal is 10 μgal, then the tiny signal we will be looking for is only about one part in 10^8!!

3.1.1 HOW TO MEASURE GRAVITY?

So how to measure gravity to catch the tiny signals caused by our interested anomaly targets? There are generally three kinds of measurements to get either the absolute gravity or the relative gravity data. These methods are

- Falling body measurements
    - Which is simply to drop an object, and measure the distance and time, then calculate the (absolute) gravitational acceleration.
- Pendulum measurements
    - Which is to measure the period of oscillation of a pendulum.
- Mass on spring measurements
    - Which is to suspend a mass on a spring and measure the amount of the stretch of the spring under the force of gravity.

Let us discuss these methods and the gravity instruments designed based on them in more detail.

3.1.1.1 Falling body measurements

The basic principle applied in the falling body measurements is the free fall, which is any motion of a body where gravity is the **only** force acting upon it, as illustrated in the cartoon images shown below.

From what we have learned from Physics lectures, we have known that the final velocity at a later time can be calculated from the velocity as an earlier time, along with gravity acceleration and the traveling time, which is

$$V(t_2) = V(t_1) + g(t_2 - t_1)$$

$$\therefore \; g = \frac{V(t_2) - V(t_1)}{t_2 - t_1}$$

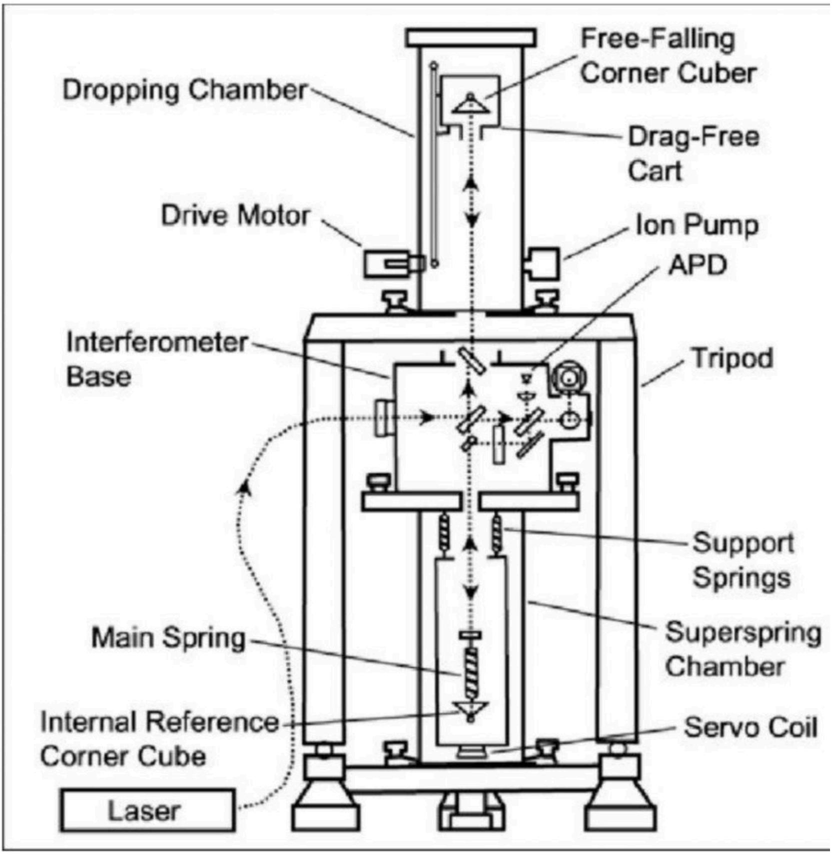
Where g is the time rate of change of speed (i.e., acceleration). Therefore, we can calculate the gravity acceleration through the time difference and the speeds.

However, in the actual practice, it is very hard to measure velocity accurately. But we can measure the traveled distance more precisely and easily. So, if we assume the initial velocity is zero, then the distance traveled can be expressed as follows:

$$d = \frac{1}{2}gt^2$$
$$\therefore g = \frac{2d}{t^2}$$

Thus, the absolute gravity g can be then calculated from the measurement of distance and time.

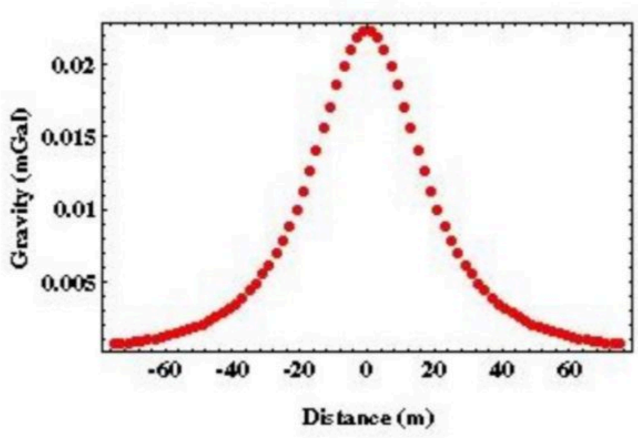
One of the example instruments designed based on the free fall principle is the Absolute gravimeter FG-5 by the Micro g LaCoste company. Its schematic diagram is shown below, which consists of three sections. The upper section is a vacuum dropping chamber, where a mass drops at roughly 100 cycles/minute. The middle section contains an interferometer to measure the position of the falling mass. The lower section has spring systems that prevent the free-fall system from being affected by Earth vibrations, therefore it functions as a cushion center to stabilize the whole gravimeter system. The reported absolute accuracy of this instrument is $\pm 2\mu gal$, with the measurement precision of $\pm 1\mu gal$.

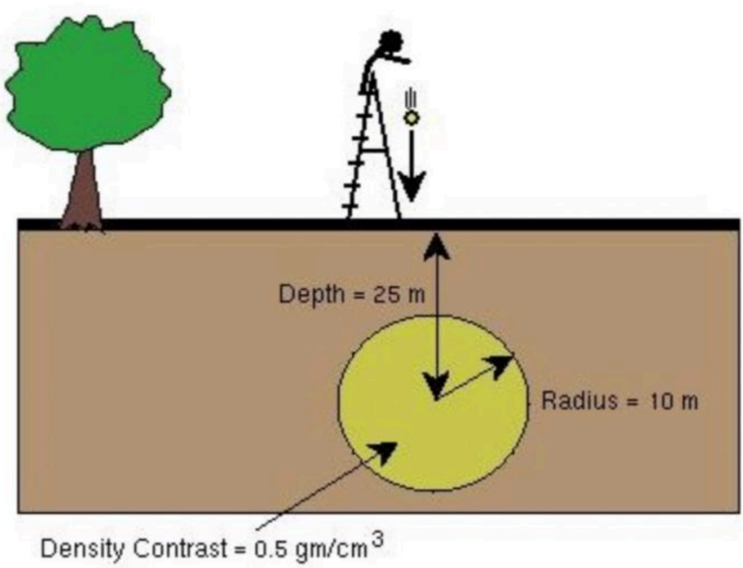


Schematic diagram of the FG-5 absolute gravity meter (Timmen et al., 2007)

Measuring absolute gravity is very very challenging! To measure gravity down to 1 part in 40 million (i.e., 25 μgal) using an instrument of reasonable size (e.g., one that allows an object to drop 1 meter), we need to be able to measure changes in distance down to 1 part in 10 millions and changes in time down to 1 part in 100 millions!

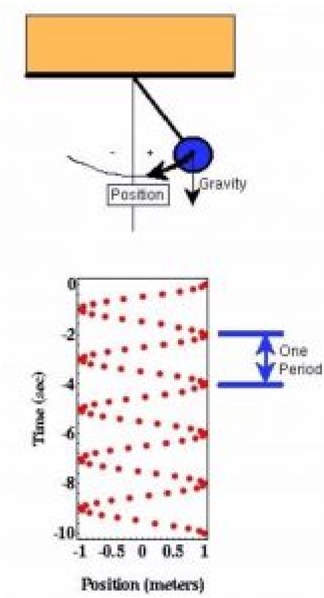
For example, let us assume a small ore body (indicated as the yellow circle in the graph below), in a spherical shape with a radius of 10 meters, centered at the depth of 25 meters below the ground surface. The density contrast of it with its surrounding sediments is 0.5 g/cc. The measured gravity data along a line profile above the ore body will be symmetrical about the center of the sphere. The maximum value is quite small, with about 25 μgal, and the gravity data approach 0 at a further distance away from the anomaly center (here beyond 60 m).





3.1.1.2 Pendulum measurements

Another method to measure absolute gravity is through the pendulum measurements. The principle is that we can build a simple pendulum by hanging a mass from a rod and displacing it from the vertical, like the setting shown in the picture on the



left. The mass will begin to oscillate back and forth. The gravity will determine the period of its oscillation; if the gravity is small, there is less force pulling the mass downward, therefore, the pendulum will move slowly toward the vertical, and resulting a longer oscillation period. The math expression for gravity and the period is as follows:

$$T = 2\pi\sqrt{\frac{k}{g}}$$

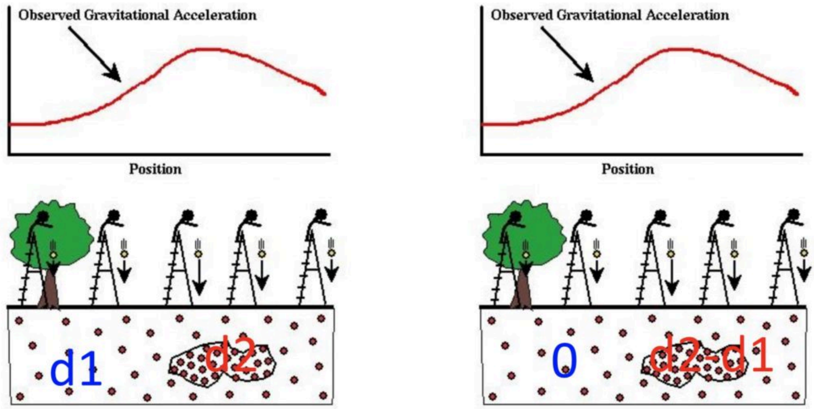
The equation is assumed to be true when there is no friction involved in the motion. k is a constant controlled by the physical characteristics of the pendulum such as its length and the distribution of the mass. This equation will give us **absolute** gravity measurements.

Historically, pendulum measurements were used extensively to measure the gravitational acceleration around the globe. To minimize the error, many periods of oscillations were observed then the average of them would be used.

Unfortunately, the measurement of constant k is not accurate enough to allow us to make gravity measurements down to 1 part in 40 million precision. Often the accuracy is limited to roughly 0.1 $mgal$ (Hinze et al., 2017, p103).

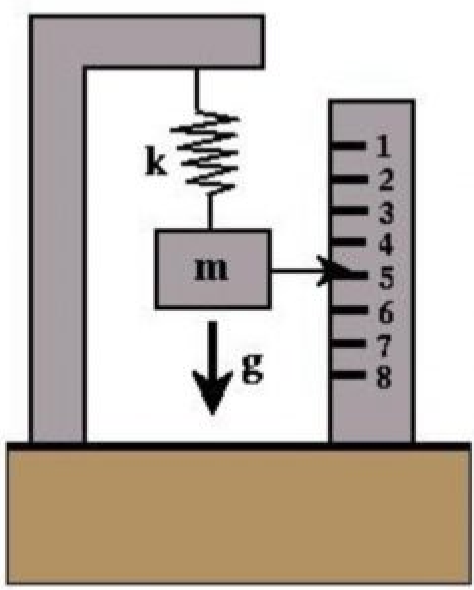
Although we cannot measure k accurately, if we use the same pendulum system, the k value is constant, then by using the same pendulum to measure the periods of oscillation at two different locations, we can estimate the change in gravitational acceleration at these two locations, without knowing k. This method gives us the **relative** gravity measurement.

For example, the figure below shows the comparison of absolute and relative gravity measurements for the same study region. Notice that, the shapes of the two gravity profiles are the same, the only difference is a constant offset; the anomaly absolute density value d_2 was replaced with the density contrast value $d_2 - d_1$ in the relative gravity measurement setting. Relative gravity measurements contain all the information we need to identify the location and shape of the ore body since what matters is the change in gravity, not the absolute gravity values. Thus the relevant parameter is the density contrast, not the density itself.



3.1.1.3 Mass on spring measurements

Gravity can also be measured with a mass-spring system. As shown

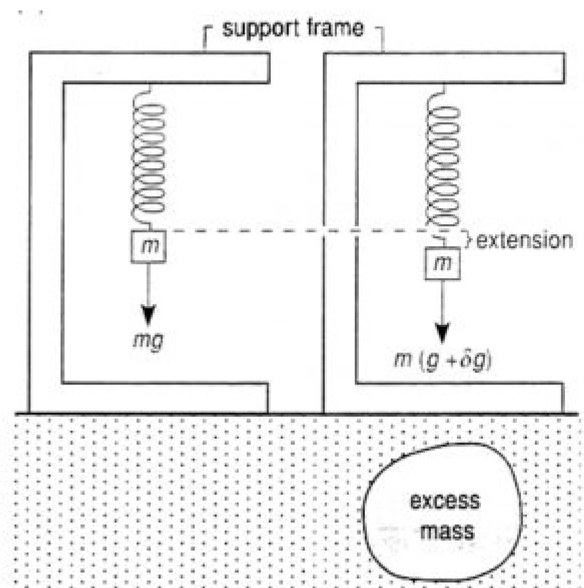


in the figure on the left, the principle is that, if a mass is hung on a spring, the spring will be stretched because of the gravity force acting upon the mass. The stretch is proportional to gravity, mathematically expressed as follows:

$$x = \frac{mg}{k}$$

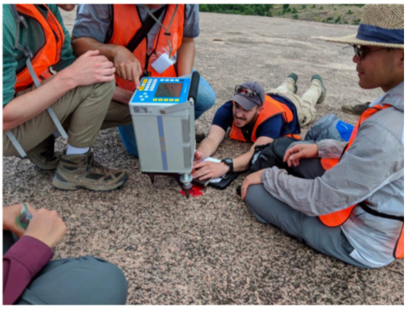
Where k is the stiffness of the spring, the larger the k is, the stiffer the spring is, and the less the spring would be stretched.

Again, we cannot measure k value accurately enough to make sure the measured gravity precision down to 1 part in 40 million. However, the extension is proportional to the change in gravity caused by the change in density. Therefore, we can measure relative gravity (i.e., gravity change) by measuring the extension, as illustrated in the experiment set in the graph below.



Principle and photograph of a gravimeter. (Mussett & Khan, 2000)

One of the gravimeter examples using this mass-spring measurement system is the CG5 gravimeter, used often in the UH Geophysical Field Camp, as shown in the picture below.



Photos from 2018 UH Geophysical Field Camp.

3.1.1.4 Absolute vs. Relative gravimeter

In terms of measurement accuracy, relative gravimeters usually have an accuracy of > $10\mu gal$, while absolute gravimeters can have accuracies on the order of 1 μgal. For the instrument itself, absolute gravimeters are usually larger and heavier than relative gravimeters. Therefore, using absolute gravimeters will take longer measurement time, and more expensive.

However, absolute gravity measurements are more accurate, and the measured data are free from instrument drift corrections. Therefore, their measured data can be used to establish reference points to tie together individual surveys, to both national and international datums, and can be used to establish gravity benchmarks. They are also useful for studies involving high-precision time variations in gravity, for instance, it is possible to observe a 3 mm crustal uplift by monitoring the change in gravity at a single station.

3.1.1.5 Satellite gravimeter

Gravity measurements can also be taken by designed satellites. One of the well-known examples is the Gravity Recovery and Climate Experiment (GRACE) mission, which was a joint mission of NASA and the German Aerospace Center launched in March 2002 and ended in October 2017. This mission was operated by two twin satellites which could take detailed measurements of Earth's gravity field and its changes over time. Their measured data were used for studying Earth's ocean, geology, climate, and hydrology, etc.

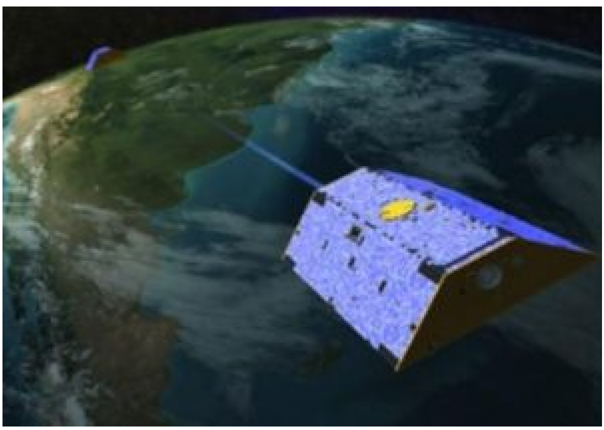


Image credit: NASA

Their working principle is that the two identical satellites, each about the size of a car, were separate 220 km (137 miles) apart, one following the other around the Earth. A microwave ranging system in the satellites could measure the distance between them to within a micrometer (0.001 mm), which is smaller than a red blood cell. By measuring the tiny changes in distance between them, we were able to measure the subtle changes in gravity, since their distance changes are caused by each of them speeds up and slows down in response to the gravitational force. A successful case study done by using the GRACE data is the measurement of the water storage in Amazon.

3.2 GRAVITY DATA PROCESSING:

3.2.1 CONTRIBUTIONS TO GRAVITY

The observed gravity measurement that we get from gravity instruments (i.e. CG-5) are due to many factors. The observed gravity measurements are a summation of the following contributors:

1. **Attraction of the reference ellipsoid:** Accounts for 99% of your gravity measurements. It is the gravity due to the earth, which has an ellipsoidal shape. Recall that the gravitational attraction in the equator is smaller than the poles because due to the ellipsoidal shape of the earth.
2. **Effect of earth's rotation:** The gravity due to the earth spinning on its axis. Recall that the centrifugal force is stronger in the equator, which makes the gravitational attraction in the equator to be smaller than in the poles.
3. **Instrument drift:** The gravity measurements from relative gravity instruments (such as CG-5) is not consistent with time due to the spring-mass system inside the instrument. The change of k (spring constant) in the spring affects the gravity readings from the instrument (not a real change in the gravity field)
4. **Effect of elevation:** The gravitational attraction in different observation elevation are different simply because the observation locations will have different distances towards the center of the earth

5. **Effect of mass above sea level:** The earth is not a simple flat plain, it also has mountains and other features above the sea level that will also have its own gravitational attraction to the gravity instrument and account for the observed gravity measurement
6. **Effect of topography:** The earth's surface is subject to different topography (i.e. mountains, troughs, valleys, etc.) and therefore our measurements are also subject to this effect as well (similar reasoning to mass above sea level).
7. **Effect of compensating masses at depth:** The earth's crust has different density and thickness throughout the earth. This also accounts for the gravity measurements.
8. **Effect of moving platform:** If we measure gravity in a moving platform, such as airplanes, helicopter, boat, etc, then even the movements and rumblings that is caused by the movement can contribute to the gravity measurements.
9. **Effect of local geology of interest:** Gravity due to density variations in the crust. This is caused by the density contrasts in the subsurface, which gives clues on what geological features exist in the subsurface. This effect is what we want to isolate and use for interpreting geology.

In geophysics, the gravity that we are interested in is the effect of **local geology of interest**, which will give us information on the density contrast in the subsurface, and interpret the geology based on the measurements. However, in order to get this information, *that means we need to correctly find the gravity contribution from the 8 other effects, and subtract the observed gravity with them*. Putting it in a more mathematical perspective:

Effect of local geology *= Observed Gravity – (Effect of ellipsoid + Effect of rotation + Instrument drift + Effect of Elevation + Effect of mass above sea level + Effect of Topography + Effect of masses at depth + Effect of moving platform)*

The process of subtracting these contributions from the observed gravity is called **gravity data correction**. This involves modelling/finding each of the gravity contributions so that we can subtract them from the observed gravity to find the gravity from the geology of interest. There are different types of corrections that account for each of these effects, which we will learn in the next section.

- **Theoretical Gravity:** Gravity attraction from the reference ellipsoid and effect of the earth's rotation. Generally, this correction has something to do with the latitude position of the measurements. This has been discussed in the previous section, as the theoretical gravity can be expressed as:

$$g_o = g_e\left(\frac{1 + ksin^2\lambda}{\sqrt{1 - e^2sin^2\lambda}}\right)$$

Where ge is the gravitational attraction in the equator, λ is the latitude.

- **Temporal Correction:** Like the name suggested, this correction has something to do with changes in gravity measurements with time. This includes the correction for effects of tides and instrument drift.

- **Free-air Correction:** This correction accounts for the effect of elevation in our gravity measurements.

- **Bouguer slab correction:** The correction simplifies the different topography of the earth as a uniform slab with a uniform density. This accounts for the effect of mass above sea level and effect of topography. We want to remove these effects so that our gravity data is from the same elevation/topography throughout.

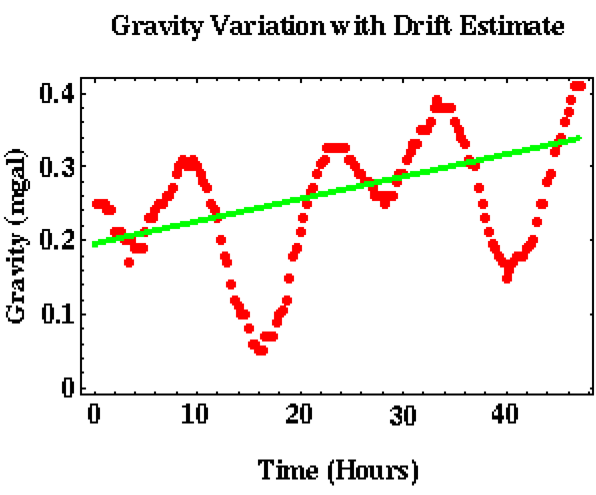
- **Isostatic Correction:** The correction accounts for the effect of compensating masses at depth.
- **Eotvos Correction:** The correction accounts for the effect of moving platform (is done for gravity measurements using airplanes, helicopters, boat, ship, etc.)

In the next section, we will go over each of these corrections in more detail.

3.2.2 TIDE AND DRIFT CORRECTION

3.2.2.1 Instrument drift effect

In Geophysics, the most common instruments used to measure gravity are relative gravimeters due to their lower cost and better portability to absolute gravimeters. Relative gravimeters use a spring-mass system to measure gravity and are prone to instrument drift. Instrument drift in relative gravimeters is a gradual and **unintentional change** in readings due to **material property change**. This change can be due to the **spring stretching** repeatedly over time and changes caused by **temperature**. Relative gravimeters are built to minimize the effect of temperature through temperature control or are built out of materials that are relatively insensitive to temperature changes. However, relative gravimeters can still drift as much as **0.1 mGal** per day.



https://pburnley.faculty.unlv.edu/GEOL442_642/gravity/notes/GravityNotes13InstrumentDrift.htm

The figure above shows real gravity measurements collected at the same site in Tulsa, Oklahoma over 48 hours. There is clear **oscillatory** behavior in the measurements due to the **tidal attraction** of the moon and the sun. It is also clear that underneath the oscillations there is a **linear increasing trend** due to instrument drift.

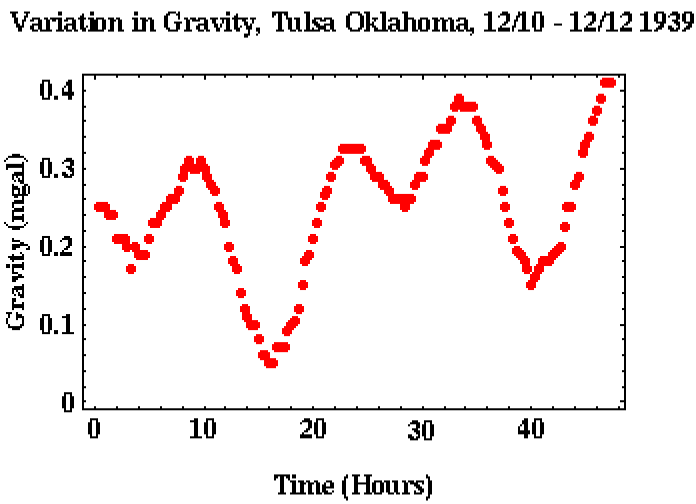
3.2.2.2 Tidal effect

The gravitational attraction of the Moon and Sun produces significant distortions on the Earth such as tides, and these changes can be measured as variations in gravity observations. This effect is referred to as **tidal effect**, and must be considered in gravity observations as they can overwhelm gravity anomalies. It is important to make a clear distinction between instrument drift and tidal effects. Instrument drift is simply due to temporally varying material properties and does not reflect real changes in gravity. However, tidal effects are real and measurable changes in gravitational acceleration. Unfortunately, changes from tidal effects are not related to local geology and are considered as noise.

In regards to tidal effect, there are two types of tides that must be considered.

1. **Ocean Tides**: These are distortions of the ocean due to the attraction of the Sun and Moon and can be measured in **meters.**
2. **Solid Earth Tides**: These are distortions of solid earth such as rocks due to the attraction of the Sun and Moon and can be measured in **centimeters.**

Regardless of the type of tides, the significance of tidal effect is both **time and latitude** dependent, with the greatest effects at low latitudes over a period of roughly 12 hours. Tidal effect never exceeds **0.3 mGal** but should be accounted for in high-resolution gravity surveys where gravity anomalies can be measured in **μGal.** A formula exists for computing the tidal effect at any point or time on the Earth's surface.



https://pburnley.faculty.unlv.edu/GEOL442_642/gravity/notes/GravityNotes14EarthTides.htm

Using the previously mentioned gravity measurements in Tulsa, a few observations can be made on tidal effect. The oscillation period roughly corresponds to 12 hours, and in this location the tidal effects ranged about 0.15 mGal. This is close to the influence of instrument drift, which had an influence of about 0.12 mGal on the recorded data.

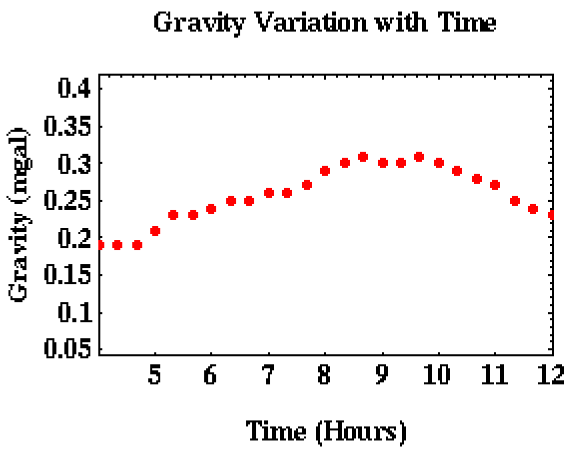
3.2.2.3 How to correct for tides and drift

An unfortunate consequence of tidal effect and instrument drift is that relative gravimeters will record different measurements at the same location. This type of noise is slowly varying with time. This has lead to the contemplation of two strategies in order to correct for tides and drift.

The **first strategy** employs two gravimeters with one permanently located at the base station of a survey with the second gravimeter used to take measurements for the survey. This strategy uses the first gravimeter to continuously monitor gravity changes over the period of the entire survey. However, this strategy has a few significant problems that have kept it from being used in most gravity surveys. The first is the expensive cost of performing a survey in such a manner, as it would require two gravimeters and two field crews to operate. Additionally, this strategy disregards that each gravimeter has a unique instrument drift so this strategy can only remove tidal effects in gravity measurements.

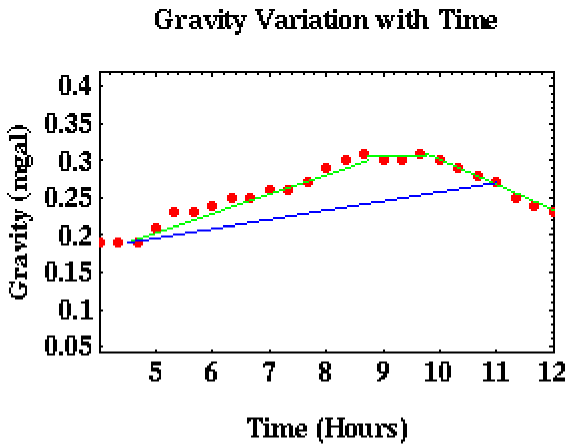
In comparison to the first strategy, a **second strategy** uses only one gravimeter where the measurements are periodically retaken at the base station of a survey rather than continuous monitoring. This means that the survey will **loop** back to the base station in between data measurements before continuing onward. The advantage of using such a method is that it is cheaper and easier to operate with only a single gravimeter. Additionally, using a common gravimeter allows for the correction of both tidal effect and instrument drift.

In the following example, we will show how to approximate tidal effects and drift. We first start off with recorded gravity measurements and observe the variations in regard to time.



https://pburnley.faculty.unlv.edu/GEOL442_642/gravity/notes/GravityNotes15TemporalVariationCorrectionStrategy.htm

In order to correct for the slow changes in time of tidal effects and instrument drift, we can start off by first approximating the changes with a few straight lines. This appears as the green lines in the figure below. Once these approximations are made, the endpoints of the green line segments can be taken as measurements. By assuming linear variation in between the measurements, linear interpolation along the blue line below can be used to predict the gravity due to tidal effects and drift at any time. It is important to note that the time interval between two consecutive measurements at the base station must be short enough for the linear approximation to be valid.

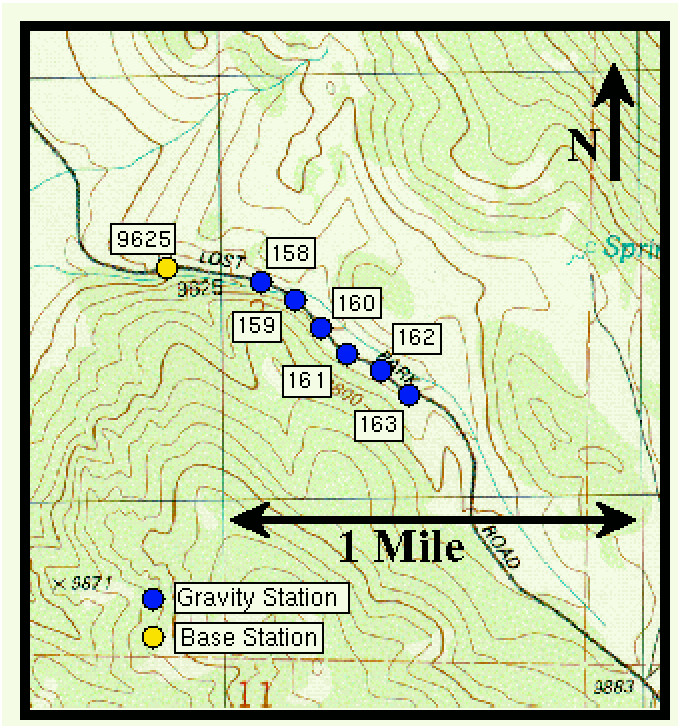


Approximated tidal and drift variation

A common practice in geophysics is to consider tidal effect as part of instrument drift due to the mixed effects that both have on gravity measurements, despite the different origins of the effects. This means that drift correction in practice will often refer to both tidal and drift correction.

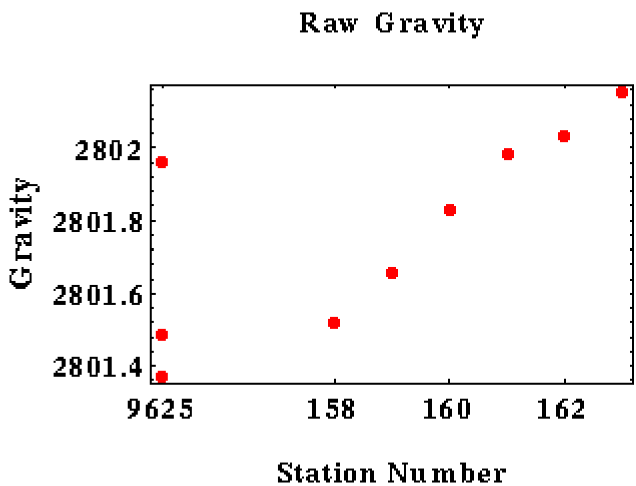
3.2.2.4 A Field Example: Looping Procedure

In order to better understand how to perform tidal effect and drift correction, we will use a field example of a gravity survey. The gravity survey uses the looping strategy we discussed previously as the correction strategy and the stations are shown in the following figure:



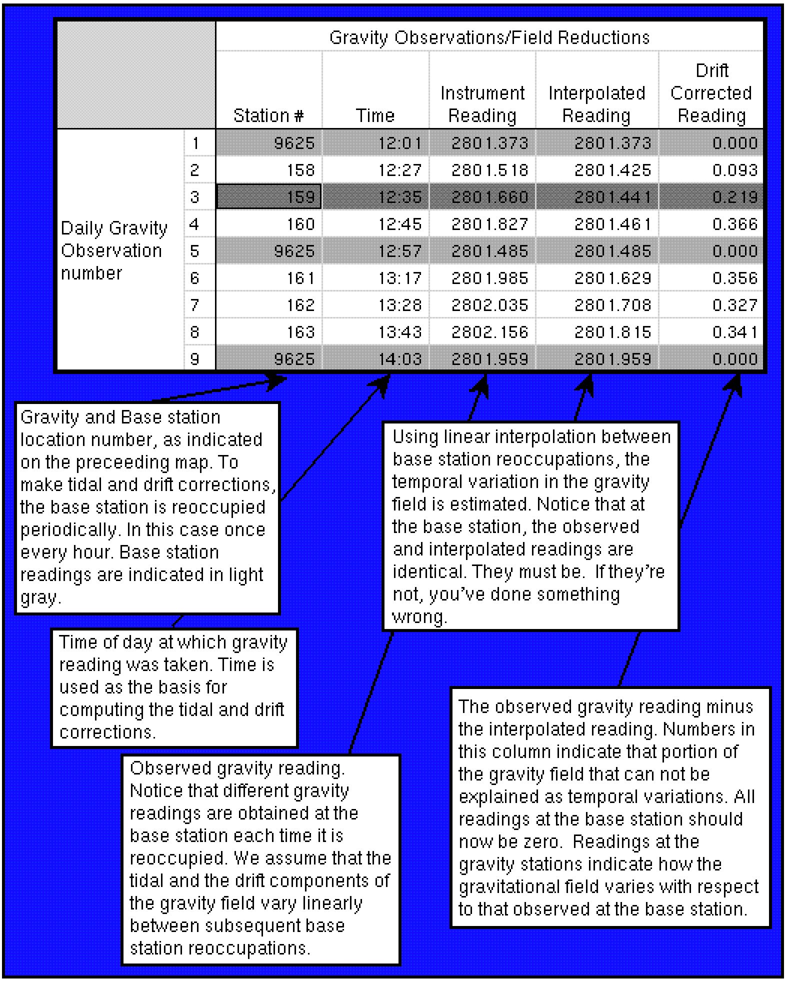
https://pburnley.faculty.unlv.edu/GEOL442_642/gravity/notes/GravityNotes16TemporalVariationFieldProcedure.htm

In the figure above, the yellow dot indicates the location of the base station. Establishing a base station is fundamental to a gravity survey as a point of reference for future tidal and drift corrections. The next few step involves establishing a set of gravity (survey) stations to collect data at and the points are indicated with the blue dots. After the survey line is designed, the gravity survey itself starts with gravity and time measurements at the base station. After taking the initial measurements at the base station, the procedure used for this gravity survey was to take measurements at stations 158-163, with gravity measurements and time measured at each station. However, continuing down the line of stations was often interrupted after about an hour to two hours in order to return to the base station. If there were additional stations more measurements could be made, but in this case reaching the end of the survey line meant returning to the base station for the final measurements.



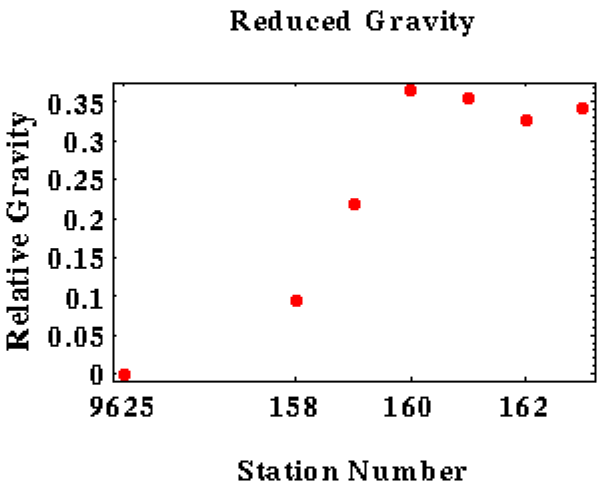
The raw gravity measurements of the example survey

In the raw data collected above, it is important to note the three gravity measurements taken at the base station. The influence of time-varying effects is quite clear in the gravity measurements of the base station, and highlights the importance of correcting the data. We won't cover the specific equations and methods used to construct the linear interpolation required to correct the data, but the figure below does show the difference between the raw and corrected data.



Daily Gravity Observation number	Gravity Observations/Field Reductions				
	Station #	Time	Instrument Reading	Interpolated Reading	Drift Corrected Reading
1	9625	12:01	2801.373	2801.373	0.000
2	158	12:27	2801.518	2801.425	0.093
3	159	12:35	2801.660	2801.441	0.219
4	160	12:45	2801.827	2801.461	0.366
5	9625	12:57	2801.485	2801.485	0.000
6	161	13:17	2801.985	2801.629	0.356
7	162	13:28	2802.035	2801.708	0.327
8	163	13:43	2802.156	2801.815	0.341
9	9625	14:03	2801.959	2801.959	0.000

Plotting the corrected data results in the following figure:



3.2.3 LATITUDE CORRECTION

The Earth's gravitational field varies with latitudes because of its ellipsoidal shape and its rotation. Even for a uniform Earth, the gravity measurements will change as a function of latitudes. Therefore, some of the spatial changes that we observe a gravity data map come purely from the differences in the latitudes where the measurement were taken. These spatial changes, if not corrected for, will be incorrectly interpreted as being due to subsurface geological features, leading to incorrect interpretations.

To correct for latitudes, we can simply subtract the theoretical gravity from the gravity measurements.

3.2.4 FREE AIR CORRECTION

The free air gravity anomaly takes into account the latitudinal changes in gravity. It measures the vertical change in gravity between that reference datum and the observation height assuming that the gravity station is located in free air, hence the name free air anomaly. In this anomaly, the intervening space between the observation and the height datum is assumed to have no mass and no gravitation effect, which is unlike other anomalies no assumptions are made about the Earth's masses in free air anomaly.

The free air correction is applied to remove the effects caused by the elevation. After the free air correction, the measurements of gravity would be adjusted at a reference level. For the Earth, the reference is commonly taken as mean sea level.

The first order approximation can be used to estimate and correct the free air anomaly. The gravity field varies by $-0.3086mGal/m$ at the surface of the Earth. Minus sign comes from the fact that gravity decreases when elevation increases. $-0.3086mGal/m$ means if we want our gravity measurement to have a precision of $0.01mGal$, the precision of elevation measurements must be around $3cm$. Obtaining accurate elevation measurements is one of the primary cost of high resolution gravity survey.

Why 0.3086

$$g = \gamma \frac{M}{R^2}$$

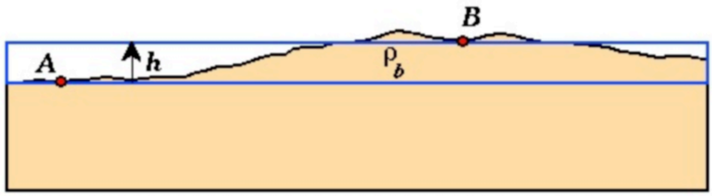
$$\frac{\partial g}{\partial R} = -2\gamma \frac{M}{R^3} = -2\frac{g}{R}$$

$$g = 9.807m/s^2$$
$$R = 6357km$$

3.2.4 BOUGUER SLAB CORRECTION

3.2.4.1 Gravity variations due to excess mass

We observe several basic information from the figure below:



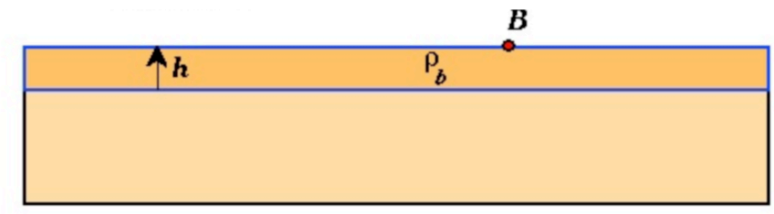
https://pburnley.faculty.unlv.edu/GEOL442_642/gravity/notes/GravityNotes22SlabEffects.htm

It is apparently shown in the figure above, gravity varies from A point to B point. There are two reason to explain the gravity variation. One reason is that A and B have the difference in topography. Another reason is the different amount of excess mass underneath cause the gravity variation from A to B. The excess masses would contribute more to the gravity anomaly.

The gravity variation caused by the excess mass underneath the B point can be approximated as a slab of material with thickness h and density ρ. Obviously, this description does not accurately describe the nature of mass below the point B, because the topography is not of uniform thickness and the density varies with location. But in this section, we assume the slab of material is regular (thickness h) and uniform (density ρ), the more detailed and complicated correction will be considered next.

3.2.4.2 Correct for excess mass: Bouguer slab correction

The method we used to correct (or, remove) the spatial variations in gravity that are due to the differences in the amount of the excess mass underneath each station (seeing figure above) is Bouguer slab correction, which is based on this simple slab approximation. We assume that the excess mass underneath B can be approximated by a slab of uniform density and thickness (seeing figure below).



https://pburnley.faculty.unlv.edu/GEOL442_642/gravity/notes/GravityNotes22SlabEffects.htm

We remember gravity decays as a function of distance squared, so, the mass directly under the gravimeter and vary near areas contributes most to the measurements. The slab approximation, thus, can adequately describe how much of the gravity anomalies associated with excess mass.

In the Bouguer slab correction, the vertical gravitational acceleration associated with a flat slab can be simply written as $-0.04193\rho h$. Where ρ is the density of the slab, h is the elevation difference. h is positive for observation point above the reference level and negative for observation points below the reference level. The negative sign of the Bouguer slab correction is make sense, because if an observation point is at a higher elevation, there is excess mass underneath. Our gravity reading, thus, is larger due to the excess mass, and we would subtract a gravity anomaly value to move the observation point back down the reference level.

In the Bouguer slab correction, we need to know the elevation of all observation points and the density of the slab used to approximate the excess mass. We currently use an average density for the rocks in the survey area. For a density of $2.67g/cc$, the slab correction is about $0.11mGal/m$.

Summary for the Bouguer slab correction

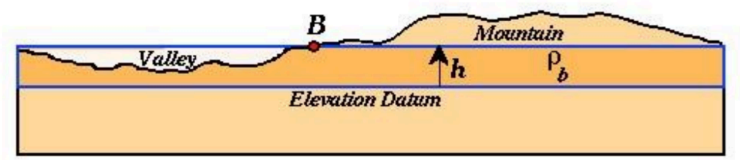
- Bouguer slab correction is rough approximation, but it is simple.
- Gravity due to a slab is simply $-0.04193\rho h$.
- We currently use average density for the rocks in the survey area. For a density of $2.67g/cc$, the slab correction is about $0.11mGal/m$.

(https://pburnley.faculty.unlv.edu/GEOL442_642/gravity/notes/GravityNotes22SlabEffects.htm)

3.2.5 TERRAIN CORRECTION

3.2.5.1. The problem with Bouguer slab correction

In the previous section, we talked about how excess mass in the surface of the earth can lead to gravity variations. Having our stations be in different levels of topography can lead to changes in elevation and different amount of excess mass underneath them, which in turn leads to these gravity variations. For example, gravity measured on top of a mountain will be larger simply because there are more masses underneath it. Recall that we can simplify the variations in topography by approximating it as a uniform slab with uniform thickness. This is only a rough approximation, and is also known as **simple Bouguer correction.** This rough approximation is adequate for simple gentle slopes, but there is a problem in Bouguer slab that arises if we have rougher topography. The problems can be classified into two:



https://pburnley.faculty.unlv.edu/GEOL442_642/gravity/notes/GravityNotes24TopographicEffects.htm

- **Valleys**

Consider the figure above and focus on the left-most part underneath the text that says "Valley". We see that the area below the blue line (the Bouguer slab) in the left-most side does not account for the lack of mass. However, with our gravity correction scheme, you still subtract it with with the gravity due to the uniform slab. Thus, you will end up with **overcorrection** of gravity measurements near the valleys. The gravity measurement near the valley is already low even without the slab correction, but because we still subtract it in this process, you will end up with a lower gravity value than what you should have.

- **Mountains**

As before, consider the figure above, but focus on the right-most part above the text that says "Mountain". We see that the area above the blue line (the Bouguer slab) in the right-most side does not account for the excess of mass. With the gravity correction introduced in the previous section, you subtract the gravity measured at point B with the gravity due to the uniform slab. What's missing is that it doesn't take into account that the mountain will have its own gravitational attraction acting upon point B, which has an upward direction. This means that the gravity measurement that we measure in point B is supposed to be lower than it's supposed to, due to the gravitational attraction of the excess mass above the slab. This means that we **undercorrected** our measurements.

The above factors poses a problem for our correction, which brings us to the need to do terrain correction.

3.2.5.2. Complete Bouguer Anomaly

Knowing the above factors, we must then find a way to account for the topography and terrain in our gravity correction. This can be done using **terrain correction. Terrain correction** is the small adjustment we make to our Bouguer slab correction to account for topography. This will produce the **complete Bouguer Anomaly.** In other words, terrain correction is the gravity correction due to the excess/deficit of mass in the Bouguer slab, accounting for contribution from the valleys and mountains. In mathematical term, this can be expressed as:

[Simple Bouguer Correction] **+** [Terrain Correction] **= Complete Bouguer Anomaly**

In order to perform a terrain correction, we need a high resolution Digital Elevation Model (DEM). This tells us the x,y,z locations of all observation locations. We also need to estimate the densities of the terrain in order to calculate the gravity of terrain at all observation locations. However, this an be computationally expensive and time consuming.

Fortunately, digital topography maps are available worldwide. However, they are typically not fine-sampled enough for the near-zone terrain correction in areas of extreme topographic relief, or where high-resolution gravity observations are required. What we mean by **near-zone terrain correction** is corrections for topography located very close (within 558 ft) to a gravity station. If the topography is rough, there needs to be a very accurate elevation model. This requires very time consuming and expensive process.

Thus we usually use **LIDAR**s to quantify the topography of the area. LIDARs can be mounted into an airborne survey (helicopter, airplane, etc.). Airborne surveys can cover vast regional areas in a short period of time. The method is accurate to about a few millimeters (mm).

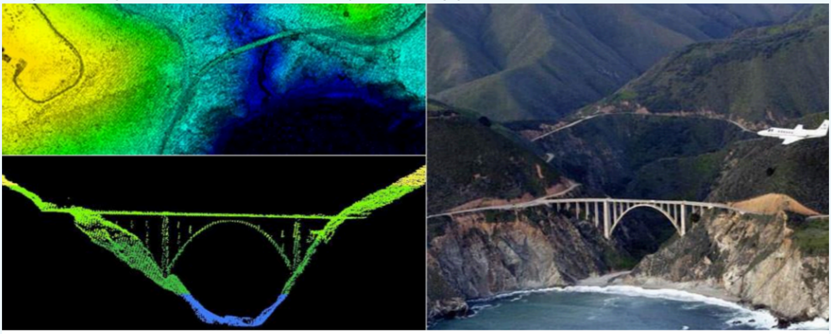
The excerpt below talks about the LIDAR in depth, which can be found from https://oceanservice.noaa.gov/facts/lidar.html.

LIDAR

"LIDAR stands for Light Detection and Randing. It is a remote sensing method to examine the surface of the earth and uses light in the form of a pulsed laser to measure ranges/variable distances to earth. These light pulses, combines with other data recorded by the airborne system, generate precise 3D information about the shape of the earth.

A LIDAR instrument principally consists of a laser, a scanner, and a specialized GPS receiver. Airplanes and helicopters are the most commonly used platforms for acquiring LIDAR data over broad areas. Two types of LIDAR are topographic and bathymetric. Topographic LIDAR typically uses a near-infrared laser to map the land, while bathymetric lidar uses water-penetrating green light to also measure seafloor and river bed elevations.

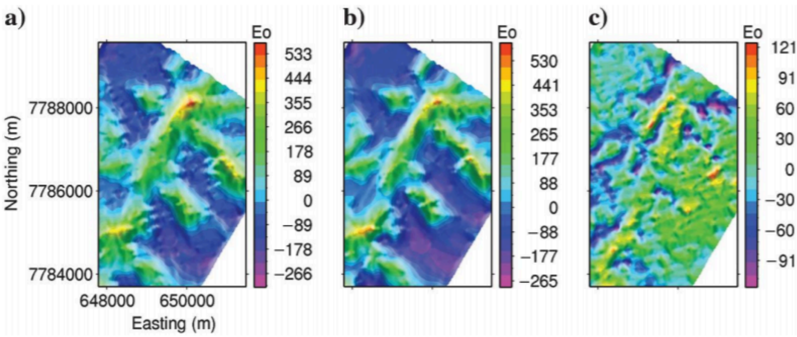
LIDAR can be used to make accurate shoreline maps, make digital elevation models for use ingeographic information systems, to assist in emergency *response operations, and in many other applications."*



LIDAR data collected over Bixby Bridge in California reveals a top-down (top left) and profile view of Bixby Bridge.

3.2.5.3. Example of Terrain Corrections

Consider the following figure:



(a) Observed Tzz; (b) Terrain effect calculated for Tzz component; and (c) Tzz after the terrain effect is removed.

Martinez et al. (2013, GEOPHYSICS)

Figure (a) is the observed Tzz component of the gravity data while (b) is the terrain effect of the survey area. Notice that (a) and (b) are almost identical. What this indicate is that **the contribution to the gravity data in (a) are mostly (or, completely) from terrain effects.** This does not give us any important information about the subsurface density distribution. After applying terrain correction, we get Figure (c), which is the better data to be used for further interpretation processes (like inversion).

3.2.5 HAMMER NET

3.2.5.1. Terrain Correction History

Before the technological advancement that we have today, (that is, before we have LIDAR), researchers would need to manually calculate the terrain correction themselves. Terrain correction was first considered by Hayford and Bowie around 1912. Cassinis, Bullard, and Lambert in 1930s tackled the problem of how to estimate terrain correction. Hammer developed a practical approach for performing terrain corrections out to about 22km from the station. Bullard (1936) broke the terrain correction into three parts:

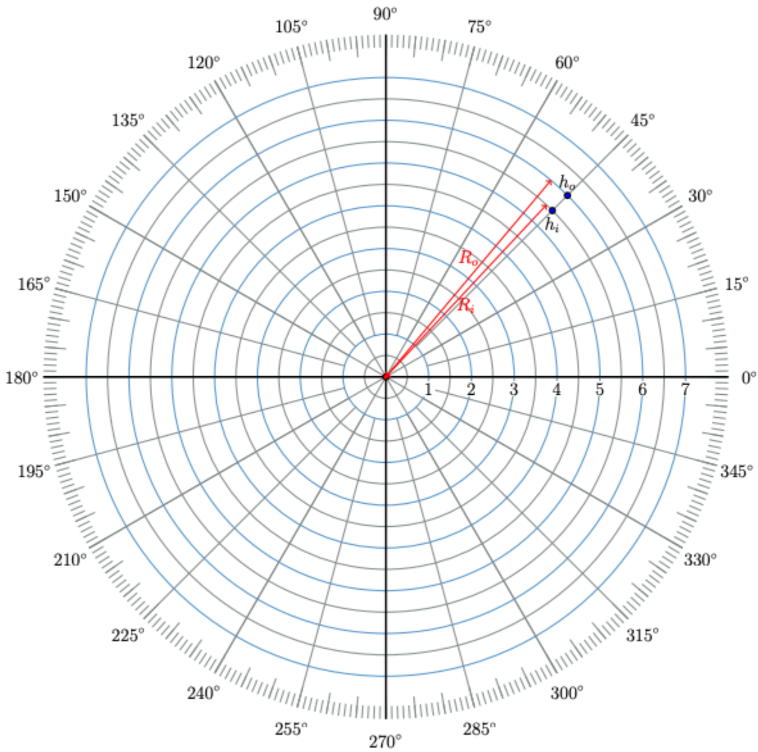
The first part is similar with simple Bouguer correction, which is also referred to Bullard A. This method approximates topography as an infinite horizontal slab of thickness equal to the height of the station above the reference ellipsoid or another datum plane.

The second part (Bullard B) is developed based on the curvature of the Earth, which reduces infinite Bouguer slab to a spherical cab of the same thickness with a surface radius of 166.735 km – 1.5 degree.

The third part (Bullard C) is terrain correction which takes undulations of topography into account. Topographic variations results in the upwards attraction of hills above the plane of the station and valleys below, which decrease the observed value of gravity, so both of these effects must be added to readings to correct for topography.

3.2.5.2 Hammer net

Hammer improved on the method of Hayford to simply terrain corrections. His "Hammer net" was widely used to make the terrain correction in past decades. This method involves compartmentalizing the area surrounding the measurement point using a template that is termed as Hammer net. Specifically speaking the Hammer net divides the area around a gravity station (shown in the figure below) using a template on the printed topographic maps.



A Hammer net (http://www.cas.usf.edu/~cconnor/pot_fields_lectures/Lecture7_graity_terrain.pdf)

During the terrain correction process, for each gravity station (central red point in the figure above), a Hammer net is centered around it. The elevation of the centered station is determined through a known terrain model. The elevation difference, therefore, is estimated in each compartment by the Hammer net. The mathematical expressions of elevation estimation is following:

$$h_i^{'} = h_s - h_i$$

$$h_o^{'} = h_s - h_0$$
$$h = \frac{h_i^{'} + h_o^{'}}{2}$$

where h_s is the station elevation, h_i and h_o are two known elevations of different compartment, h is the elevation difference of compartments.

Clearly note that topography affects gravity data the most and $h_i = h_s$ in the innermost zone, so different inner zone correction is normally used and Hammer net correction is currently referred to an outer zone terrain corrections.

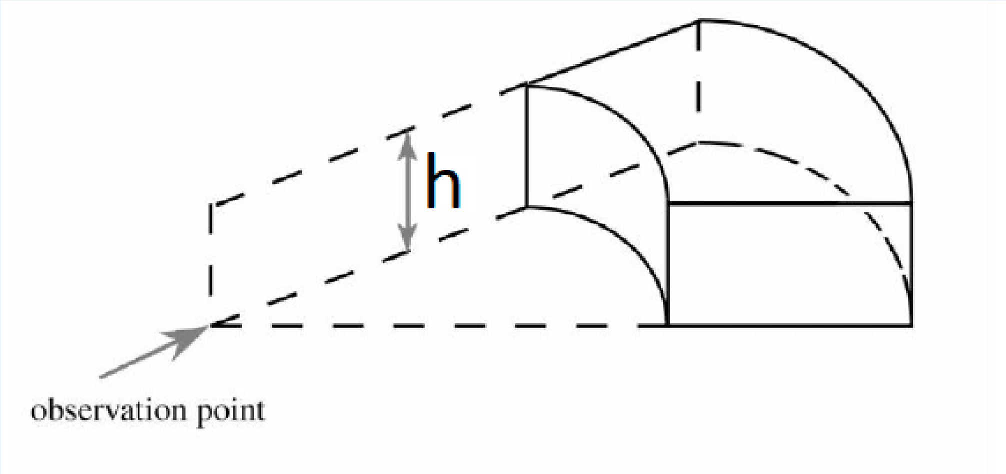
What is gravity effect of a radial component of a hollow vertical cylinder with a flat top?

$$g = \gamma\rho\theta\left[R_o - R_i + \sqrt{R_i^2 + h^2} - \sqrt{R_o^2 + h^2}\right]$$

where, ρ:bulk density for the compartment, θ: angle subtended by the two radial lines bounding the segment, R_i: inner radius, R_o: outer radius.

Total terrain effect:

$$g_{terrain} = \sum_{k=1}^{M} g$$



http://www.cas.usf.edu/~cconnor/pot_fields_lectures/Lecture7_graity_terrain.pdf

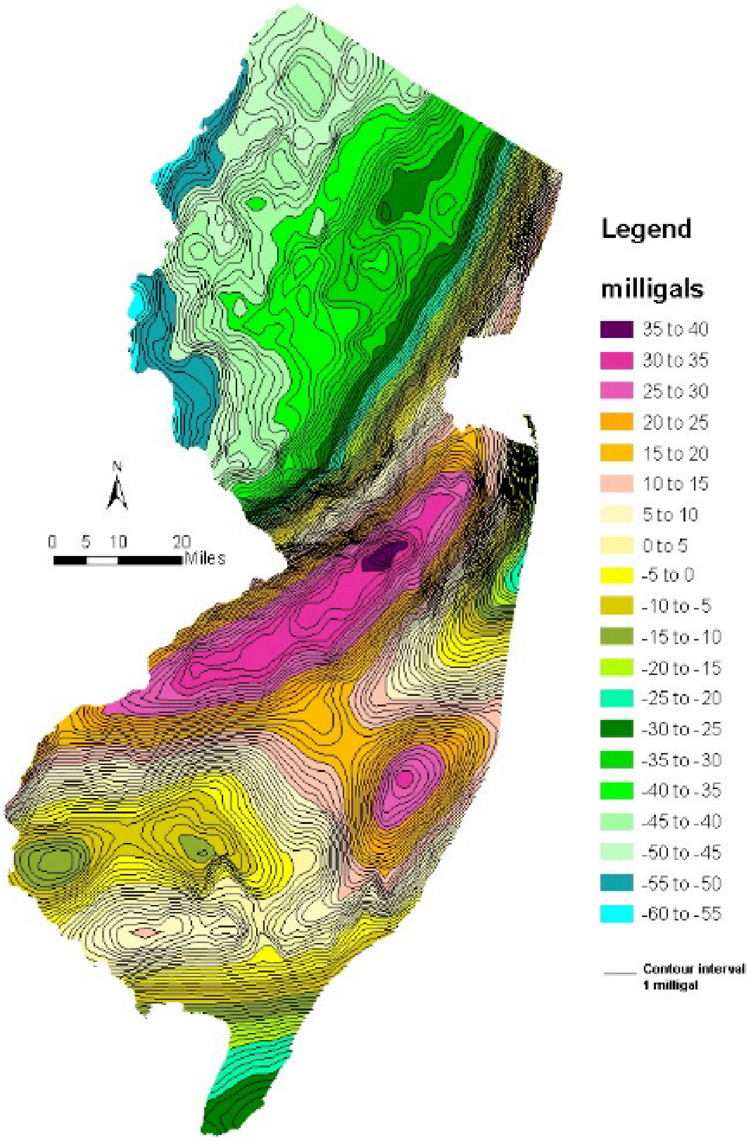
3.2.5.3. Complete Bouguer Anomaly

Complete Bouguer anomaly is associated with observed gravity, free air correction, slab correction, terrain correction and latitude correction, which reflects the subsurface density variations. The complete Bouguer anomaly can be expressed as:

$$\triangle g_{cb} = g_{obs} - g_{fa} - g_{sb} - g_t - g_0$$

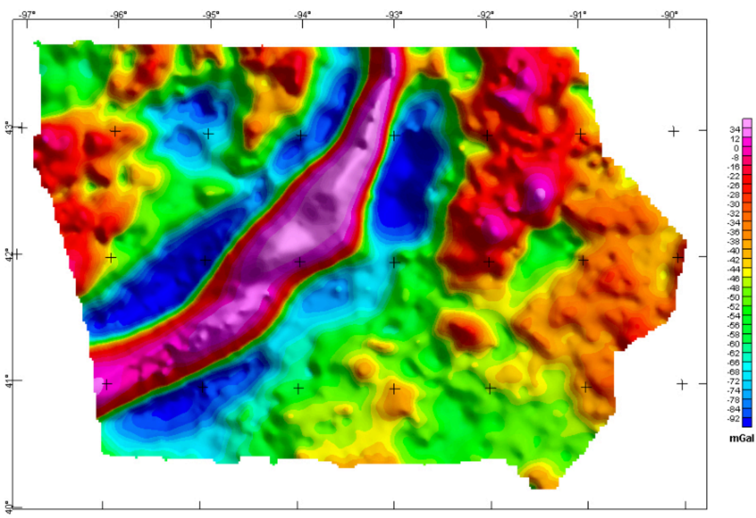
where, g_{obs} is measured gravity, g_{fa} is free air correction, g_{sb} is slab correction, g_t is terrain correction and g_0 is latitude correction.

Some Bouguer anomaly examples are shown below:



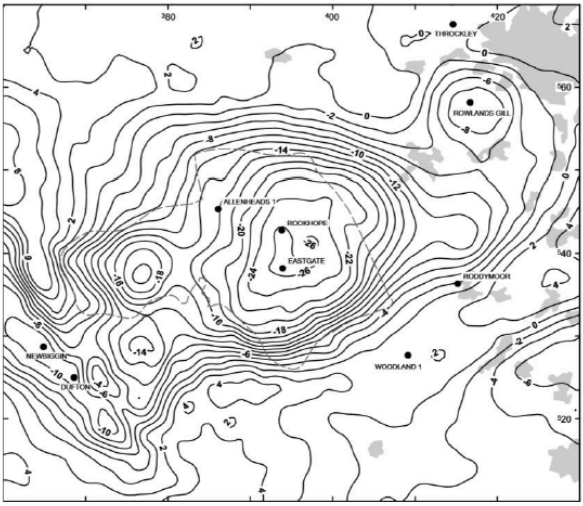
Bouguer Gravity Anomalies of New Jersey. (https://www.state.nj.us/dep/njgs/geodata/dgs04-2.htm)

The figure shown above contains two GIS shape files of bouguer gravity contours, lines and polygons, at 1 and 5 milligal intervals respectively. The contours are based on gravity data in New Jersey and vicinity. The bouguer anomalies at 1 milligal interval range from a low of -58 milligals to a high of +37milligals. At 5 milligal intervals they have lows ranging from -55 to -60 milligals and highs ranging from +35 to +40 milligals.



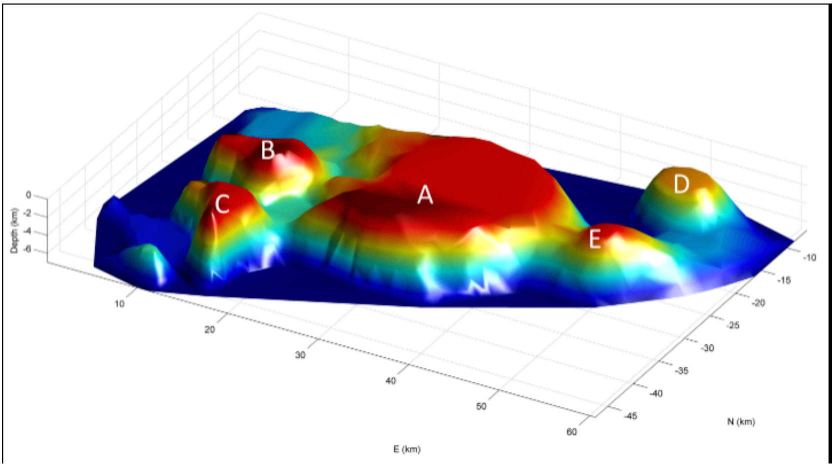
Bouguer gravity over Northeast Iowa. (https://pubs.usgs.gov/ds/2005/135/ia_boug.htm)

Bouguer gravity (figure shown above) reflects lateral density variations in the Earth. Positive anomalies (red color) occur in ares with average density greater than the Bouguer reduction density of 2.67 gm/cc, whereas negative anomalies (blue color) occur in areas of lower density. The highest anomalies are closely related to mid continent rift, which means something subsurface has really high density.



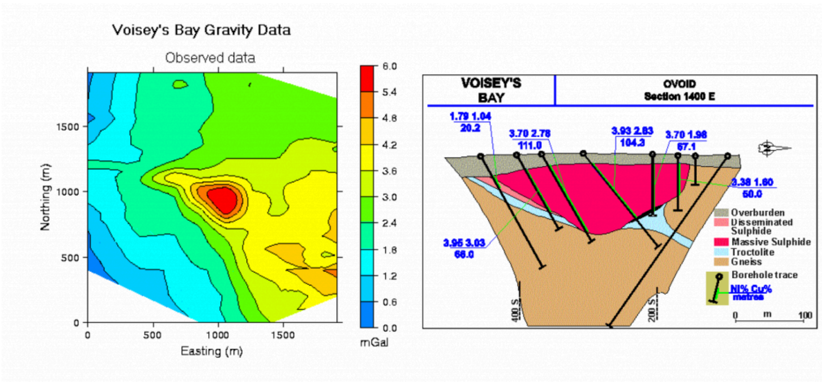
Observed Bouguer gravity anomalies in the North Pennine region.

Based on the stations from the British Geological Survey national gravity databank. Contour interval = 2 mGal. Dashed line is the edge of the fluorite zone (after Dunham 1990). Grey shading indiates urban areas; black dots are selected boreholes (Figure 2 in Kimbell et al, 2010, The north Pennie batholith (Weardale Granite) of northern England – new data on its age and form).



3D model of the north Pennine batholith. (https://en.wikipedia.org/wiki/North_Pennine_Batholith)

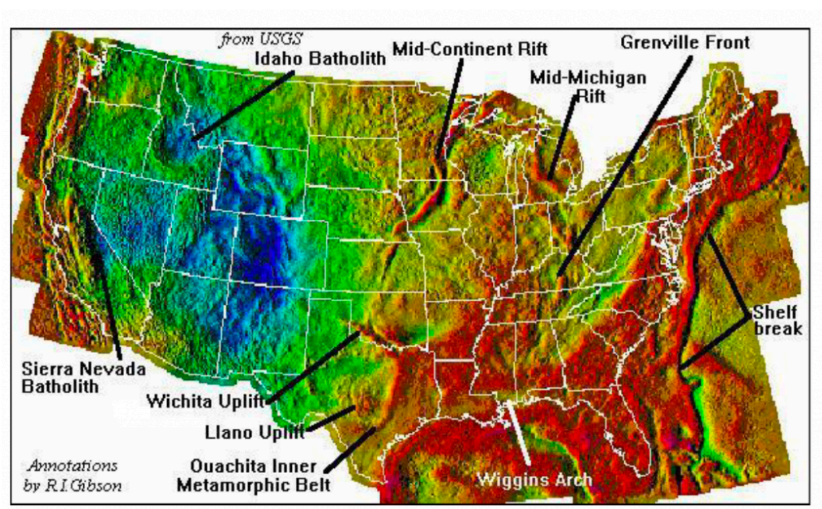
3D model of the North Pennine Batholith (aka the Weardale Granite), adapted from Kimbell et al. (2010). The North Pennine Batholith, also known as the Weardale Granite is a granitic batholith lying under northeast England, emplaced around 400 million years ago in the early Devonian. The batholith is composed of five plutons: A: Weardale Pluton B: Tynehead Pluton C: Scordale Pluton D: Rowlands Gill Pluton E: Cornsay Pluton.



Gravity anomaly of a massive sulphide body (credit: Yaoguo Li @ CSM)

The figure above (left) is observed gravity data, the figure below (right) is geological model built based on drill hole data.

3.2.5.4. A Few More Examples



Bouguer anomalies map of the conterminous US.

The negative (cold color) in the West US in basin areas, the positive (warm color) in the East US in the mountain areas. The anomaly's corresponding features are labeled in the figure above.

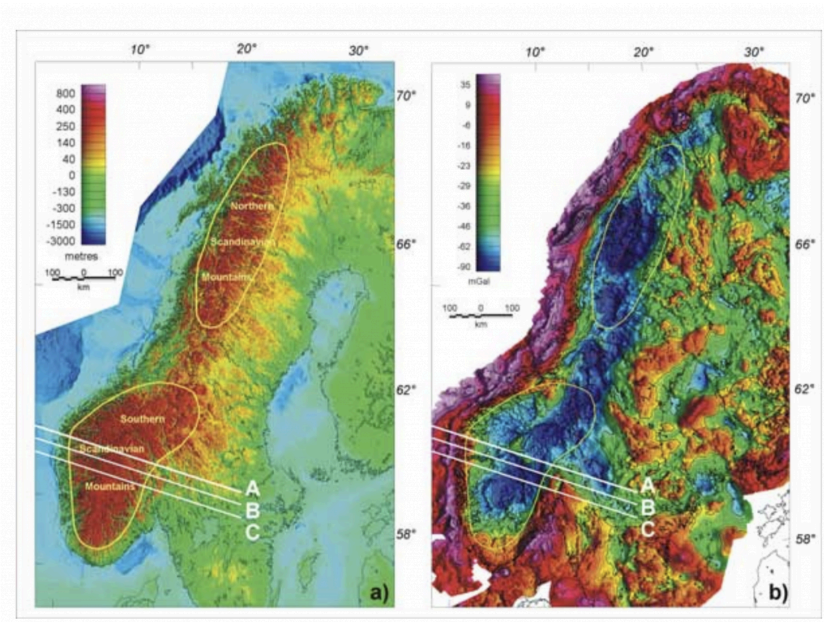
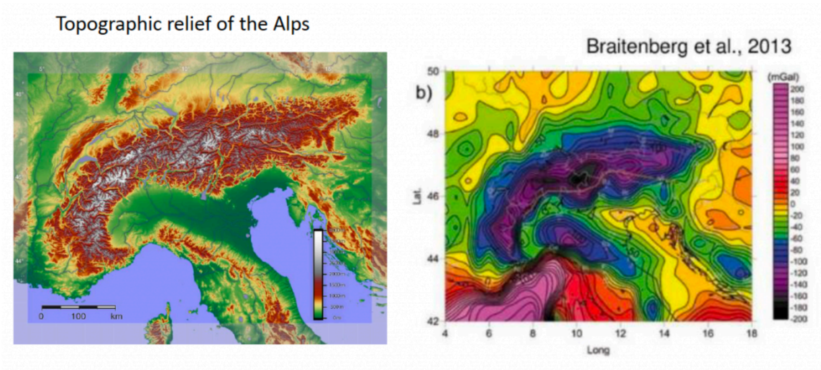
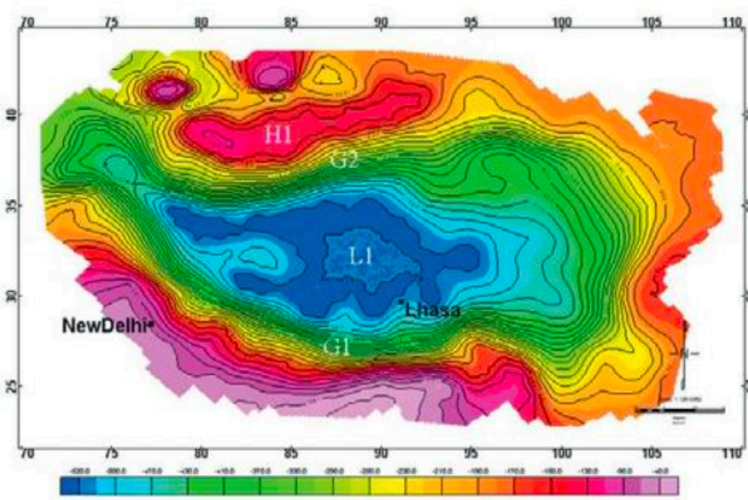


Figure 1: a) topography (Dehls et al., 2000) and b) Bouguer gravity anomalies of Fennoscandia (Skilbrei et al., 2000; Korhonen et al., 2002). A, B, C represent the three lines used in the gravity modelling. (http://www.mantleplumes.org/Scandes.html)

The observation from the figure above, the negative anomalies at location where the elevation is high.



Bouguer anomaly over Alps



Bouguer anomalies over Himalayas (Mishra & Kumar 2008)

Bouguer anomaly map of the Himalaya and Tibet obtained from the satellite free air anomaly map (Shin et al., 2007) showing a major gravity low (L1 area) over Tibet and gradients G1 and G2 coincide with Himalayan. thrust and suture zone (ITSZ) and Altyn Tagh fault, respectively. The gravity highs, H1 are related to Tarim basin. We noticed that the gravity is lower over the Himalaya mountain. The trend is the Bouguer gravity is generally lower over mountains.

Thinking

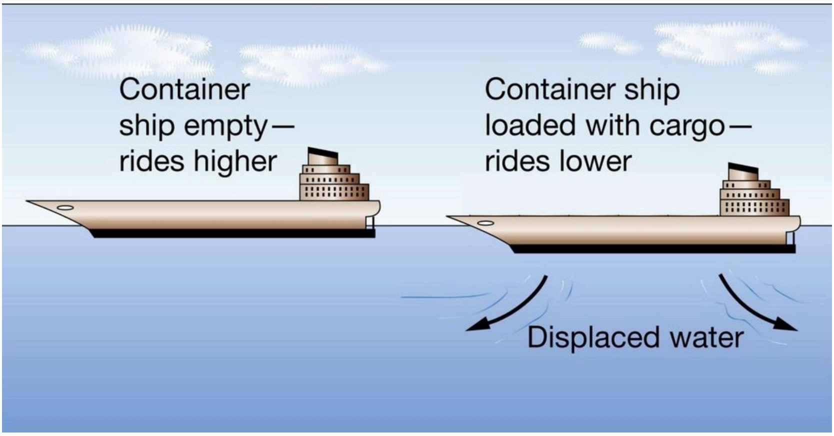
From the observation above, the Bouguer anomalies are usually negative over the mountains, why???

This question will be discussed in the following section.

3.2.6 ISOSTASY

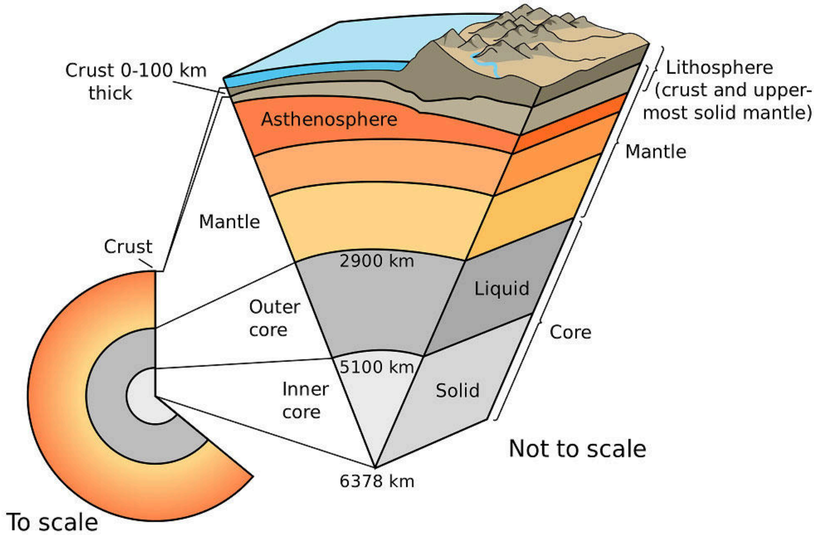
3.2.6.1. Concept of Isostasy

In order to understand the concept of isostasy, we will first start by discussing Archimedes' Principle. Archimedes' Principle states that "A floating body will displace a volume of fluid whose mass is equal to that of the body." Consider the following example of a cargo ship floating in the ocean:



Increase body density, body floats lower in fluid. Increase fluid density, body floats higher in fluid.

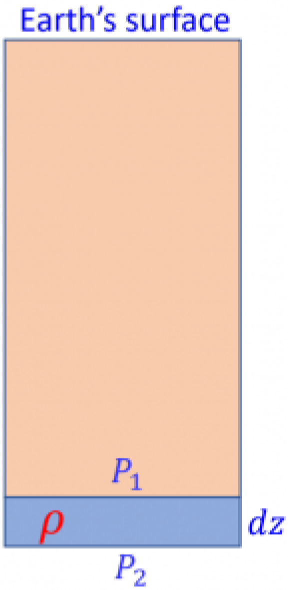
While the cargo ship is unburdened, it will float higher than after cargo is added. This is due to the increased density of the total body, which increases the water displaced and decreases the height that the ship floats at. In both situations, the ship is at **hydrostatic equilibrium** as stated by Archimedes' principle, and isostasy is a geophysical way of defining the same concept. The word **isostasy** is derived from Greek and means "equal standing." In geoscience, isostasy is the vertical positioning of the lithosphere so that the gravitational and buoyant forces balance one another. The body and fluid of Archimedes' principle is the low-density lithosphere that floats on denser underlying asthenosphere in isostasy.



Courtesy of the USGS

3.2.6.2. Hydrostatic Equilibrium

In the previous section, we briefly mentioned hydrostatic equilibrium, and we will further expand on the concept. Simply put, at hydrostatic equilibrium all mass involved in the system is in equilibrium and does not move. This means that the sum of all forces acting on a mass in this state must be zero according to Newton's second law. Consider the following example of a block of rock located deep in the Earth:



In the figure above, the block that we are considering is colored blue and has a density ρ and length dz. In this case, the force on the top of the rock is $P_1 a$. The force at the bottom of the rock is the sum of the force on top plus the gravity of the cube itself, and can be defined in the following equation:

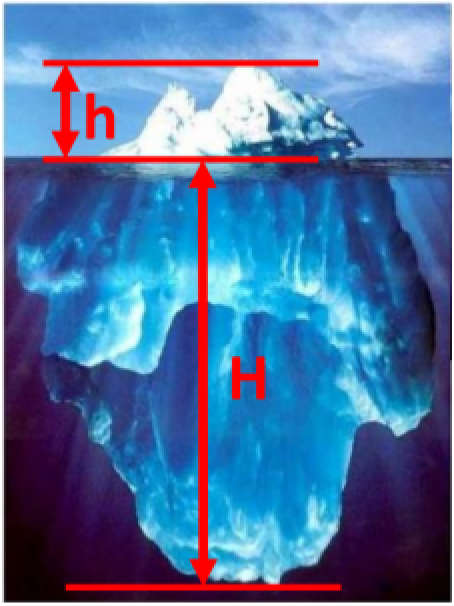
$$P_2 a = P_1 a + *dz * g$$
$$P_2 - P_1 = \rho * dz * g$$
$$dP = \rho * dz * g$$

$$\frac{dP}{dz} = \rho * g$$

If ρ and g are both functions of depth, then to solve for pressure at some depth R, we must integrate from the surface down to depth R.

$$P(z) = \int_0^R \rho(z)g(z)dz$$

Now, we will consider the following example of an iceberg at hydrostatic equilibrium:



Suppose that the iceberg above sticks out of the water by a height h and extends below the waterline at a length H. By using the known densities of $\rho_{ice} = 0.9$ and $\rho_{water} = 1.0$, we can find the extent of H based on h using the following equation:

$$\rho_{ice}(H + h) * A = \rho_{water} H * A$$

By substituting in the known densities of water and ice, we get:

$$0.9(H + h) * A = H * A$$

At this point, we can define h as a ratio of H as area A cancels out:

$$h = \frac{0.1}{0.9} H = 0.11H$$

$$H = 9h$$

Exercise:

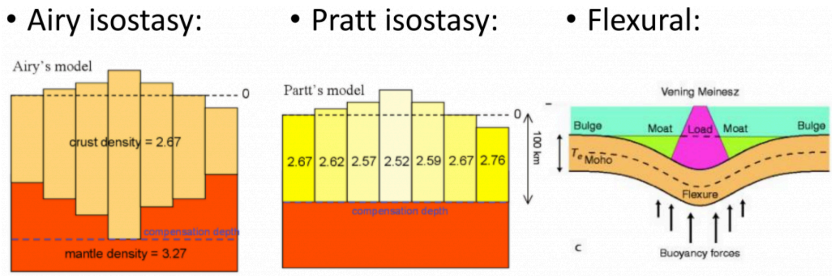
Suppose that an iceberg sticks out by H instead of h. How deep does it now extend below the waterline?

Now, imagine if we used a mountain instead of an iceberg in the example above. In this case, we would replace water with a sea of heavier mantle materials (e.g. olivine or pyroxene) but the same analysis applies if we assume that the mountain is also in hydrostatic equilibrium. However, we would apply a different term in geophysics as we would state that the mountain is in **isostatic equilibrium** or **isostasy** instead. Additionally, we must make the basic assumption that there is a depth at which the pressure due to the column of rocks above does not change laterally. This depth is called **compensation depth** or **level**.

3.2.6.3. Isostasy

There are a couple of important concepts and assumptions related to isostasy that we must discuss. The first comes from the section above, and this is that there is a compensation depth or level at which pressure from a column of rocks above does not change laterally. Regions of high topography on a surface represents an excess of mass and pressure, which must be compensated at depth by a deficit of mass with respect to the surrounding region. Additionally, mountain belts are often also regions of thickened crust. This means that conversely, topographic depressions are matched by mass excesses at depth.

The figure below depicts three types of isostasy:



3.3. MAGNETIC MEASUREMENTS

Measurements of magnetic fields can be done using **magnetic instruments.** In this section, we will explain physics behind the instruments that we used to measure magnetic field, commonly used in geophysics.

3.3.1. FLUXGATE MAGNETOMETER

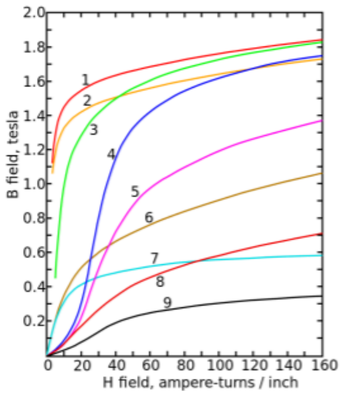
The fluxgate magnetometer was first invented in 1936 before WWII and was mostly used to detect submarines (which are metallic, and therefore would produce magnetic signals). The fluxgate magnetometer was also the type that was used to prove the theory of plate tectonics. As we have explained in the previous section, the evidence of plate tectonics and seafloor spreading were found from the magnetic measurements near mid-ocean ridges, that form what we call as “magnetic stripes”.

3.3.1.1. Concepts

Before going through the details on how Fluxgate Magnetometer works, let us review some concepts in magnetics that will be useful in this section. First concept is magnetic susceptibility (κ), a dimensionless property. If we assume an inducing field H (for example the earth’s magnetic field at the geographic region) that is applied to a magnetic material, then it will produce an induced magnetization M in the material. This is mathematically expressed as:

$$M = \kappa H$$

From the linear relationship above, theoretically we can say that the magnetization (M) increases if we increase the inducing field (H). However, some magnetic materials under the influence of an external magnetic field, will reach a state where an increase in the external magnetic field H does not increase the magnetization of the materials further. In other words, the magnetic flux levels off as you increase the external field, and continues to increase at a very slow rate due to vacuum permeability. This phenomenon is called magnetic saturation.



Magnetization curves of 9 ferromagnetic materials.

1. Sheet steel;
2. Silicon steel;
3. Case steel;
4. Tungsten steel;
5. Magnet steel;
6. Cast iron;
7. Nickel;
8. Cobalt;
9. Magnetite.

By Charles Proteus Steinmetz - Tracing of graph downloaded from Charles Steinmetz (1917) Theory and Calculation of Electric Circuits, McGraw-Hill, New York, USA, p.84, fig.42 on Google Books, Public Domain, https://commons.wikimedia.org/w/index.php?curid=4655202)

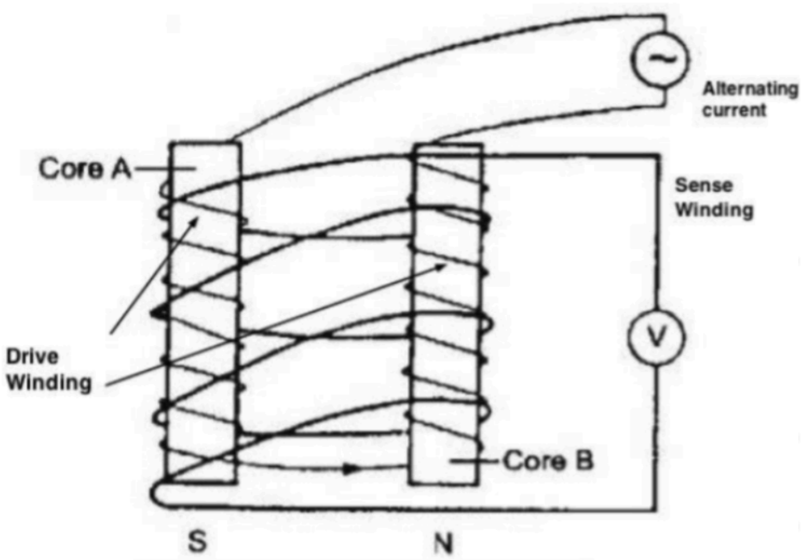
From the above picture, we see that different magnetic materials will have different curves associated with the relation between H field (inducing field) and B field (induced field, equivalent to M). Some curves saturate at a slower rate than others, and this depends on the materials.

Now that the above concepts are explained, we can move on to explaining more about the physics behind fluxgate magnetometers.

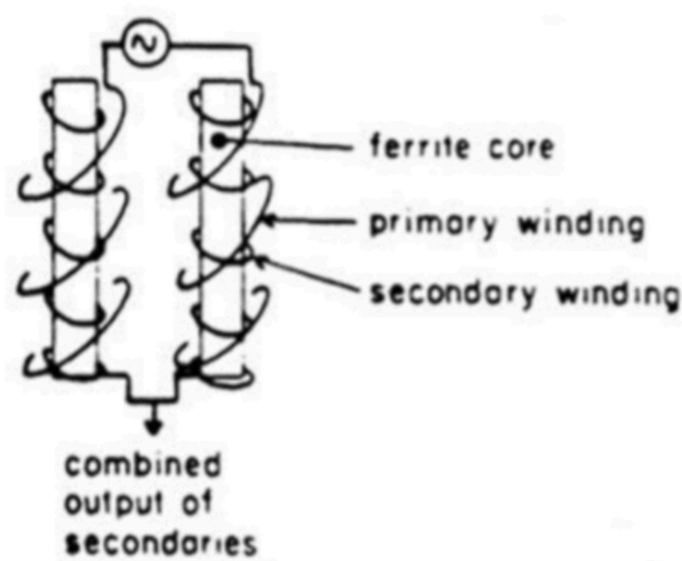
3.3.1.2. Instrument Parts

Primary Coil

A fluxgate magnetometer can be explained by the following diagram:



We can see that the magnetometer is made off two cores that are made off ferromagnetic materials, labelled as Core A and Core B. These cores are placed next to each other and have two loops surrounding it. One loop is called the primary coil/sense coil and has an AC (alternating current) going through it. This one coils around core A and core B as the following picture:



The AC current generates magnetic field H and gets the two cores magnetized. However, their magnetization varies with time. This is because AC current varies with time and thus the magnetic field H also varies in time. Notice that the primary coil is wound in the opposite senses around the two cores (ex: Core B is wound such that the magnetic field direction goes down, and core A is wound such that the magnetic field direction goes up). Because the inducing fields H are in the opposite direction in these two cores, therefore, these two cores get magnetized in opposite directions.

Secondary Coil

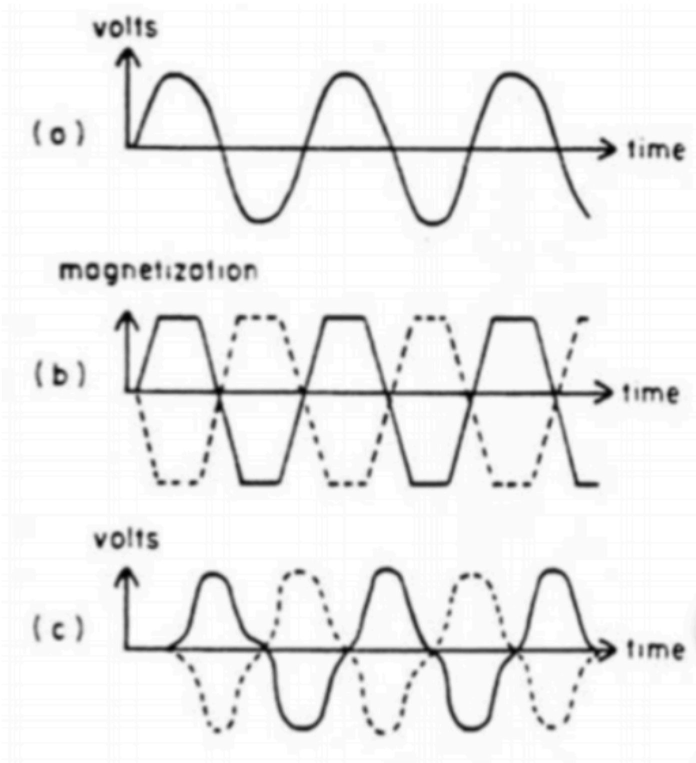
The other coil that wounds through both core A and core B is used to measure the voltage in the two cores (indicated by the letter 'v' in the schematic of the fluxgate magnetometer). This is the secondary coil and it measures the voltage using Faraday's Law. To understand this, we introduce the concept of electromotive force (emf). It is the force that drives currents to flow in a wire or in a conductive body, with unit of volts. It can be expressed mathematically as such:

$$\epsilon = -\frac{d\phi}{dt}$$

Where ϕ is the magnetic flux. Any change in magnetic flux ϕ induces an emf/voltage. This process is called electromagnetic induction. An emf, when applied to a conductor, generates current.

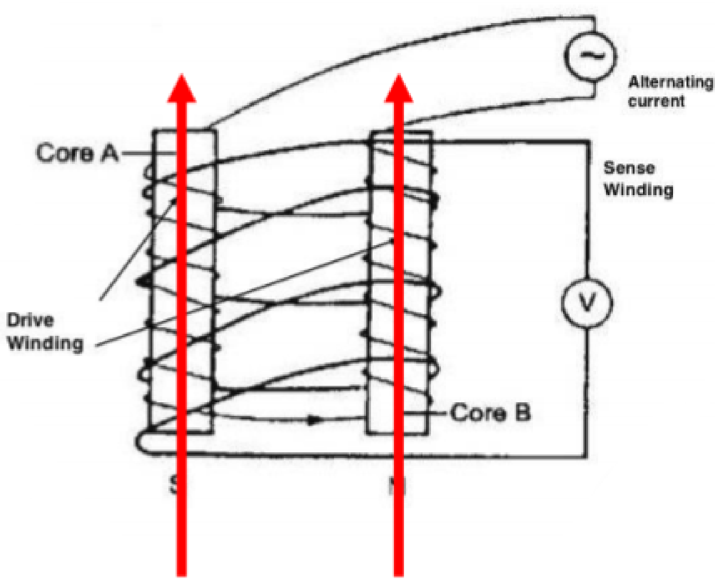
3.3.1.3. Measurements

Knowing that the secondary coil measures the emf, and that the changes in current throughout time also causes the magnetization to change throughout time, let us examine the following figure:

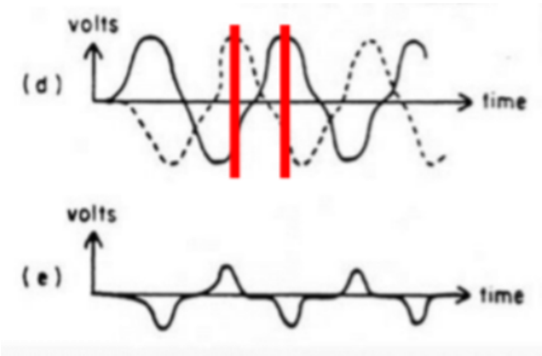


From the above picture, (a) is the emf from the AC current that changes in time. This in turn changes the magnetization of the cores as we see in (b), where the dash line is from core A and the solid line is from core B. Notice that the magnetization saturates to a flat line after a certain amount of time, and changes with the alternating current. In addition, (c) is the measured emf from the change in magnetization between core A and core B. Notice that if we add the dash line (emf from core A) and solid line (emf from core B), the summation will result in zero. Thus, **the sum of the voltages from the two cores is zero if there is no external magnetic field.**

In the presence of an external field, for example, if this two-core system is placed in the earth's magnetic field, then the sum of the voltages between the two cores is non-zero. Let's look at the figure below:



In this case, the earth's magnetic field is represented as the red line pointing upwards. The total magnetic field applied to core A increases, and thus the magnetization reaches saturation earlier. The total magnetic field applied to core B thus decreases (opposite directions) and the magnetization reaches saturation later than in core A. Consequently, the voltages from these cores will have a phase shift, producing a non-zero summation of the voltages. This phase shift due to the external magnetic field is captured in the voltage measurements as the following figure (d), and the summation of voltage can be seen in figure (e).



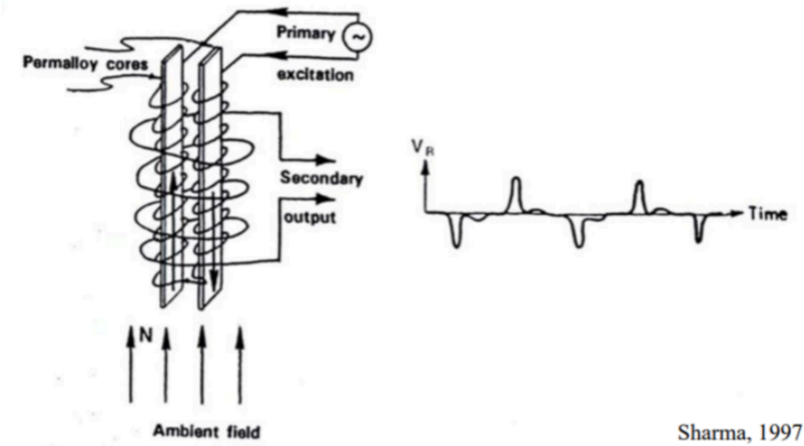
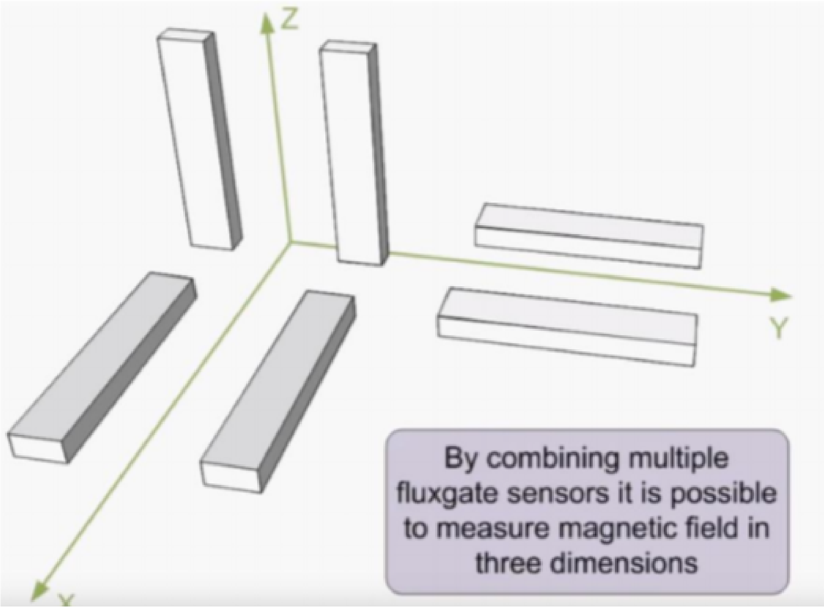


Fig. 3.6 Principle of the fluxgate magnetometer. The trace on the right shows the resultant voltage, V_R, induced in the secondary coil due to the net rate of the change of magnetic flux produced by the Permalloy cores. V_R is proportional to the strength of the ambient field.

In some systems, a **third coil** which passes a direct current is wrapped around the entire system. This produces a magnetic field that offsets the earth's field. When the output of the secondary coil is reduced to zero, then the DC field is exactly equal and opposite to that of the earth's magnetic field. Thus, by monitoring the DC required to maintain zero secondary output, variations in the earth's magnetic field can be measured.

Knowing how a fluxgate magnetometer works, the two-core system can be put in three directions in order to get the **three components** of the magnetic field (Bx, By, Bz).



Source: https://www.youtube.com/watch?v=_5d0qz_umuE

A fluxgate can measure the component of the earth's magnetic field in the direction of the axis of the system. By purposely pointing the axis into a direction, we can measure any component of the earth's magnetic field. Thus, to measure the total field, we can either:

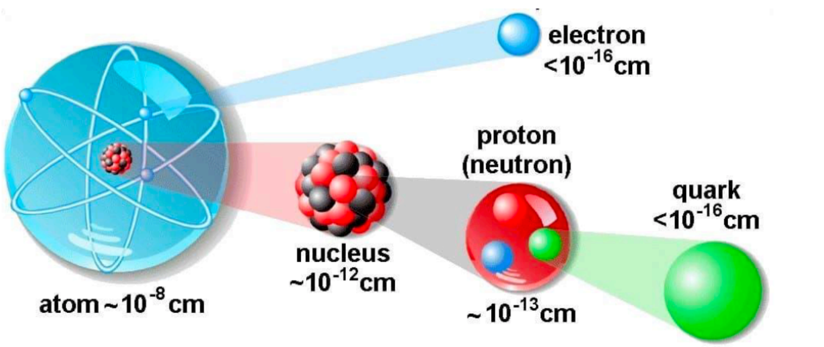
1. Measure the three component using three mutually perpendicular system
2. Orient a single system in the direction of the total field

3.3.2. PROTON PRECESSION

3.3.2.1 Atomic nucleus and proton

Proton is the fundamental part of the proton precession magnetometer. So, firstly, we briefly recall the atomic nucleus and proton.

The atomic nucleus is the small, dense region consisting of protons and neutrons at the center of an atom. An atom is composed of a positive-charged nucleus, with a cloud of negative-charged electrons surrounding it, the nucleus and electrons bound together by electrostatic force. Almost all of the mass of an atom is located in the nucleus, with a very small contribution from the electron cloud. Even though, the nucleus contains almost 99.9% of the mass of an atom but only occupies a volume whose radius is 1/100,000 the size of an atom. Protons and neutrons are bound together to form a nucleus by the nuclear force.

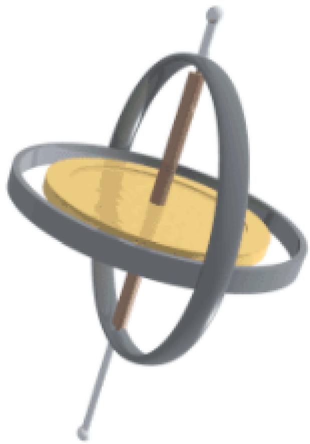


https://www.nuclear-power.net/wp-content/uploads/2014/12/Structure-of-Matter.jpg

Structure of an atom

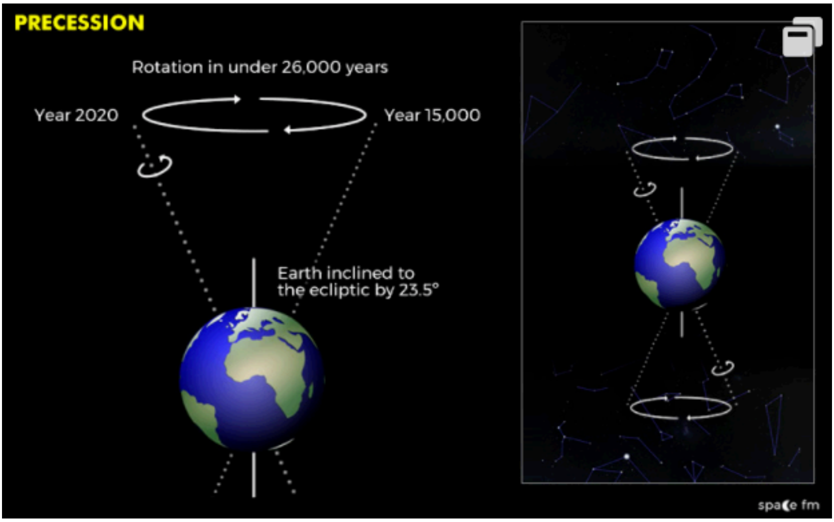
3.3.2.2 Precession

Precession is a change in the orientation the rotational axis of a rotating body. In other words, if the rotation axis of a body is itself rotating about a second axis, that body is said to be precessing about the second axis.



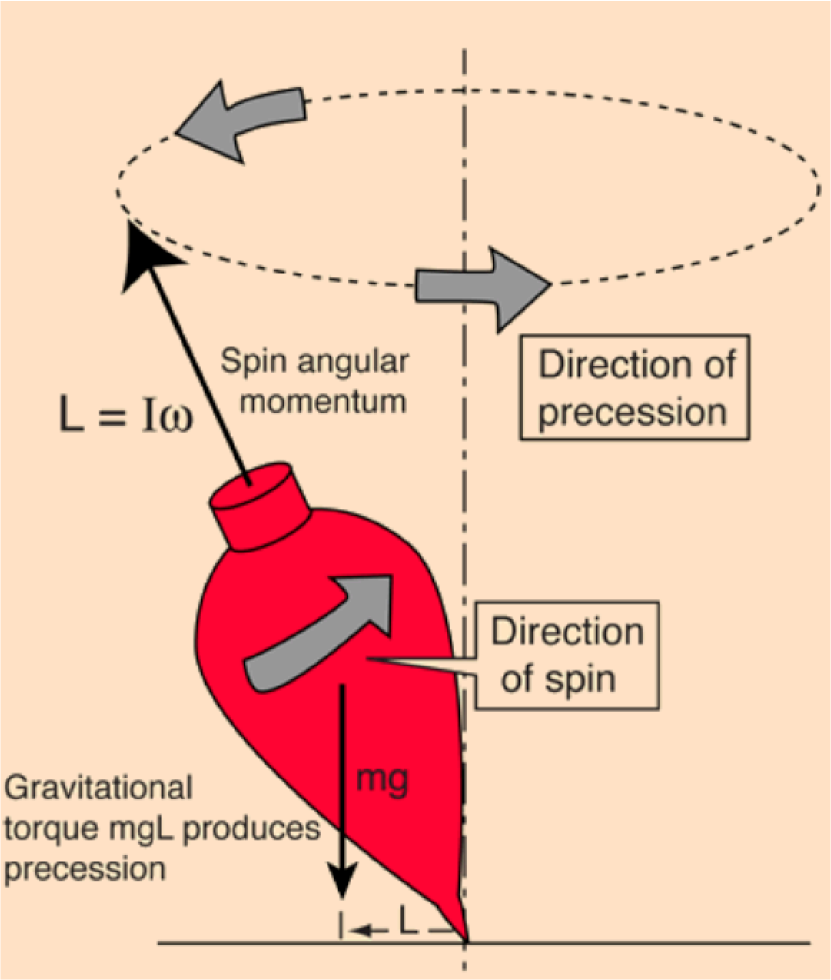
Precession of a gyroscope

The Earth rotates on its axis but has a slight ”wobble or ‘oscillation’ to be precise like a spinning top. This wobble takes approximately 26,000 years and has implications for how we view and measure the stars over a long period. The process is known as precession of the equinoxes or axial precession. Main reason for precession of Earth is that Earth is not a perfect sphere and is an oblate spheroid, it is slightly wider at the equator. The Sun and Moon have a gravitational influence on Earth and this combined with the Earth’s bulge causes a wobble in the Earth’s tilt.



Precession of Earth (https://www.space.fm/astronomy/planetarysystems/precession.html)

Here, we introduce the precession of a spinning top. A rapidly spinning top will precess in a direction determined by the torque exerted by its weight. The precession angular velocity is inversely proportional to the spin angular velocity, so that the precession is faster and more pronounced as the top slows down. Spin a top on a flat surface, and you will see it's top end slowly revolve about the vertical direction, a process called precession. As the spin of the top slows, you will see this precession get faster and faster. It then begins to bob up and down as it precesses, and finally falls over. In other words, any rotating objects (i.e., with an angular momentum) can undergo precession under the influence of a torque.



Precession of spinning top (http://hyperphysics.phy-astr.gsu.edu/hbase/top.html)

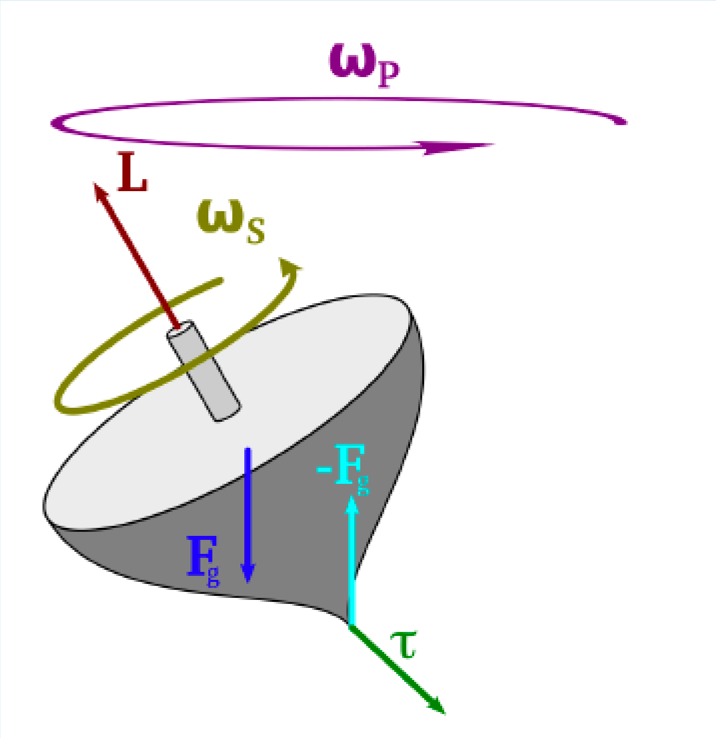
Precession of a top

The torque caused by the normal force $\mathbf{F}_g$ and the weight of the top causes a change in the angular momentum $\mathbf{L}$ in the direction of that torque. This cause the top to precess.

The mathematical equation for precession of a top is:

$$\tau = \frac{\mathrm{d}\,\mathbf{L}}{\mathrm{d}\,t}$$

where, τ: torque, $\mathbf{L}$: angular momentum.



https://en.wikipedia.org/wiki/Precession#/media/File:PrecessionOfATop.svg

It is important to note that if an object has an angular momentum (i.e., it is rotating), and if an external torque exists, the object will precess.

3.3.2.3 Precession of a proton

A proton has an angular momentum, resulting from spin of a proton. Just like a top has an angular momentum. The more specific information refer https://physicsworld.com/a/the-spin-of-a-proton/. Objects possessing momentum tend to maintain their motion unless acted up by an external force. In our cases, the gravity can create a torque. For magnetic data acquiring, a torque can be created by an external magnetic field. The equation is below:

$$\tau = \mathbf{m} \times \mathbf{B} = \gamma \mathbf{J} \times \mathbf{B}$$

where, τ: torque, $\mathbf{m}$: magnetic dipole moment, $\mathbf{B}$: external magnetic field, $\mathbf{J}$: angular momentum vector.

Therefore, a proton will precess. The external field with angular frequency knowns as Larmor frequency.

$$\omega = \gamma B$$

where, ω: angular frequency in radians / sec, γ: a particle-specific constant.

3.3.2.4 Proton precession magnetometer

The proton magnetometer, also known as the proton precession magnetometer uses the proton precession to measure the variation in the Earth's magnetic field. The cylinder at the top (white) contains hydrogen-rich fluid (e.g., kerosene, decane, water). There is also a solenoid. When direct current flows in the solenoid, a strong magnetic field is created. The protons align themselves with that field. Then the current is interrupted, and as protons realign themselves with the ambient magnetic field. The external Earth's field creates a torque, and causes the protons to precess and they precess at a frequency that is directly proportional the magnetic field.

G-856 proton precession magnetometer at base station. Picture taken at UH 2018 Geophysics Field Camp.

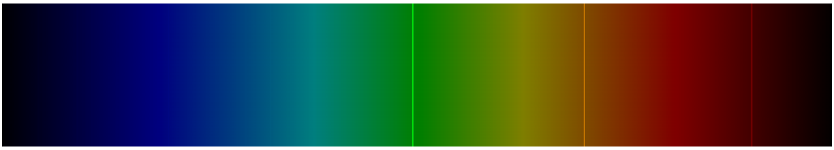
3.3.3. AKALI VAPOR

3.3.3.1. Basic Concepts of Alkali Vapor Magnetometers

The easiest way to understand **Alkali Vapor Magnetometers** is to start by deconstructing its name. **Alkali** is fairly simple, and refers to alkaline metals such as cesium, potassium, and rubidium that can be used in the construction of Akali Vapor Magnetometers. The magnetometer contains an alkali metal within a cell or chamber that will continuously heat the metal until it reaches a gaseous form. In the case of cesium, this **vapor** can be produced at temperatures ranging from approximately 45 to 55 degrees Celsius. The **vapor** is important as these magnetometers operate based on splitting electron energy levels in alkali metals by the Zeeman effect.

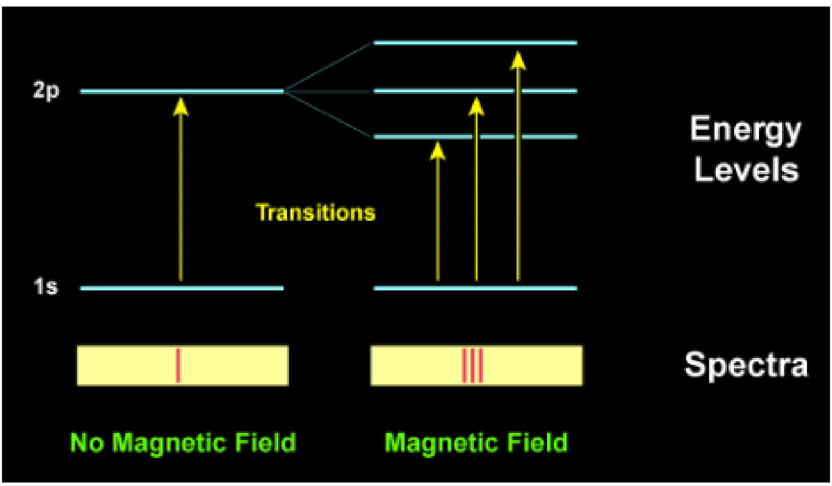
3.3.3.2. Zeeman Effect

In order to explain the **Zeeman Effect**, we will need to delve a bit into quantum mechanics. When an external magnetic field is applied to an atom, its atomic energy levels are split into a large number of levels. Under a continuous spectrum of light, spectral lines are distinct lines resulting from either light emission or absorption by atoms or molecules which form against the continuous background. These lines can be used as “fingerprints” to identify molecules or atoms, but are split into multiple components of slightly different frequencies when under the influence of a magnetic field. This is called the **Zeeman Effect** and was first observed by Dutch physicist Pieter Zeeman who shared the 1902 Nobel Prize in Physics with Hendrik Lorentz for this discovery.



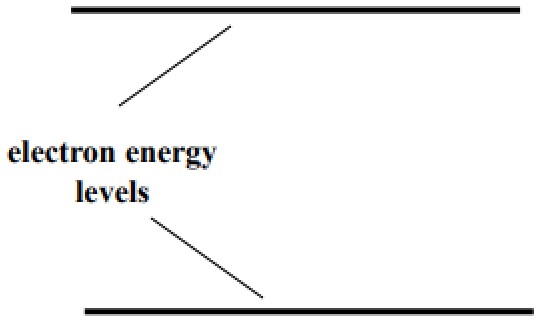
Spectral lines of Lithium, an akali metal. Produced by Neill Tucker, distributed under a CC-BY 0 license

The figure below shows the splitting of energy levels and the changes to the light spectra under a magnetic field:

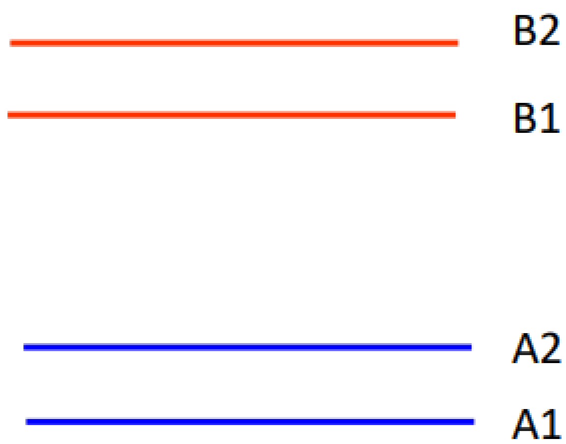


http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/zeeman.html

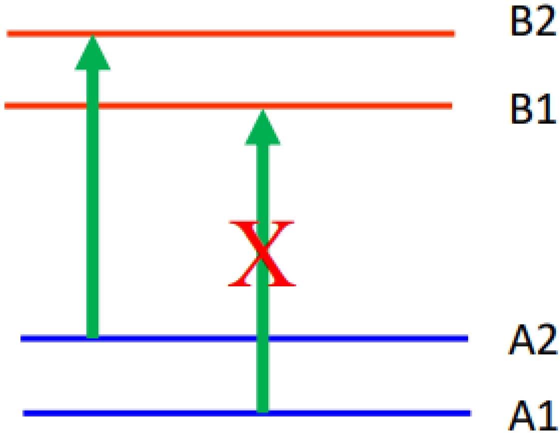
We will now do a deeper dive into the physics behind the Zeeman Effect. We will start with the atomic energy levels of an atom or molecule when there is no magnetic field.



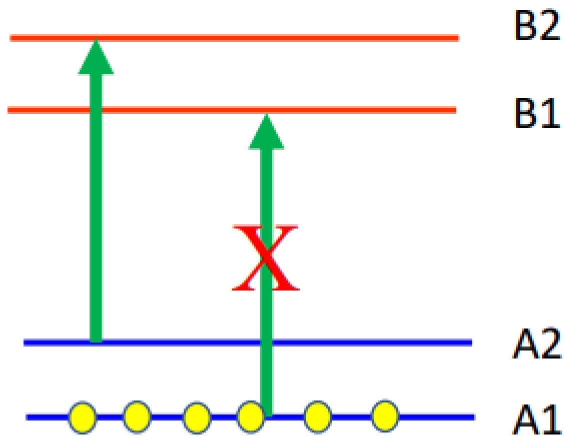
When an external magnetic field is applied, the atomic energy levels will separate. The figure below labels the separated energy levels as A1, A2, B1, and B2.



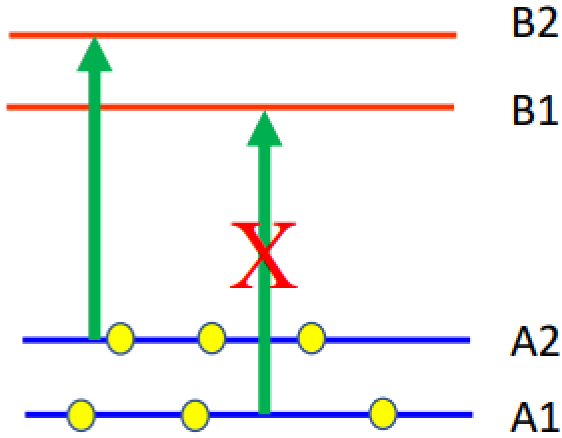
The energy gap between energy levels such as A1 and A2 is determined by the strength of the applied magnetic field. In an Akali Vapor Magnetometer, a beam of light with a frequency that corresponds to the energy gap between A2 and B2 can be used to irradiate the akali vapor. Electrons in the vapor at energy level A2 will absorb the energy from the beam and move to the higher energy level of B2. Additionally, the beam can be manipulated so that it does not contain any frequencies corresponding to the energy gap between A1 and B1. This process is called **polarization**.



In other words, after **polarization** electrons can only jump from A2 to B2 but cannot jump from A1 to B1. This means that electrons from A2 will all jump to B2, but B2 is at a higher energy state and therefore unstable. This means that energy will be released and the electrons will fall back down to either A1 or A2. This process of rising and falling atomic energy states will continue in a loop until all electrons eventually settle at A1. This process of overpopulating one energy level is known as **optical pumping**. At this point, the photons from the light beam will pass through the magnetometer's chamber with no energy loss. The vapor is therefore registered as transparent, and a photosensitive detector in the magnetometer will register a maximum current.



At this point, an RF (radio frequency) signal that has a frequency corresponding to the energy gap between A1 and A2 can be applied. This will lead some of the electrons to jump up to A2 which will absorb the photons from the light beam once again. The photosensitive detector will measure a decrease in current at this point. Unfortunately, the exact frequency of the desired RF frequency is unknown, so a varying RF signal is applied to sweep through a range of possible frequencies. This frequency is related to the energy gap which is determined by the strength of the magnetic field according to Zeeman splitting.



This means that by measuring the RF frequency accurately, an alkali vapor magnetometer can determine the strength of a magnetic field.

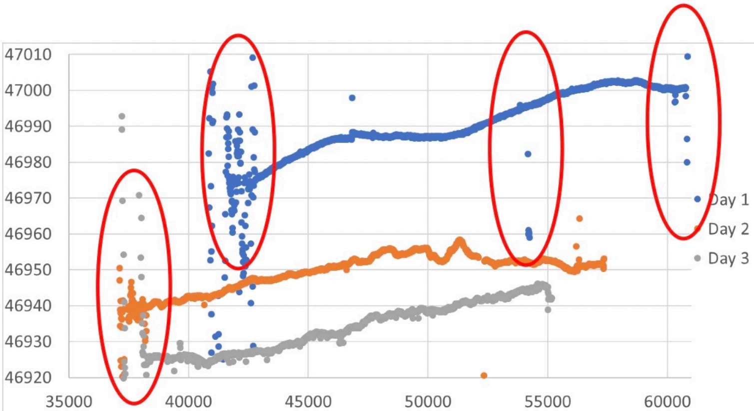
3.3.3.3. Alkali Vapor Magnetometer

When compared to the proton precession magnetometers mentioned previously, an alkali vapor magnetometer is an order of magnitude more sensitive. The sensitivities reported can range from as small as 0.001 to 0.01 nT. This is compared to a proton precession magnetometer which has an accuracy ranging from 0.1 to 1 nT. The most commonly used alkali vapor magnetometers are **cesium vapor** and **potassium vapor magnetometers**.

G-858 Cesium Vapor Magnetometer at a base station. This picture was taken at the UH 2018 Geophysics Field Camp.

It is important to note that the magnetometers that we have mentioned so far are **scalar magnetometers**. There are also **vector magnetometers**, such as **Fluxgate** and **SQUID (superconducting quantum interference devices)**. A **fluxgate magnetometer** has a sensitivity of about 0.1 to 1 nT like the proton precession magnetometer, but a **SQUID** magnetometer is far more sensitive than any of the other magnetometers that we have mentioned so far. It has a sensitivity of 10^{-5} nT, and can measure the three-components of magnetic field along with its gradients.

3.4. MAGNETIC DATA PROCESSING



Magnetic measurements at base station during 2018 UH Geophysical Field Camp.

Like raw gravity data collected in the field, magnetic data also needs some necessary corrections before being interpreted. As data are shown in the graph above, the blue, orange, gray dots are magnetic data recorded as the base station through the day 1, 2, and 3; the horizontal axis represents the time, the vertical axis denotes the magnetic data in nT. The red ellipses grouped data with a large number of noises.

The data measurements are taken with varied spatial locations in order to cover the study area, meanwhile, the recorded data are taken at different time periods of a day, which is reflected from the linear trend in the data shown in the above graph. Therefore, we need to remove the time variation effect, to left only the spatial variations due to the local geology we are interested in.

3.4.1 DIURNAL CORRECTION

The time variation effect removal from the magnetic data is called the **Diurnal correction**, which is to take consideration of the magnetic field changes due to the solar activities and their interaction with the ionosphere and magnetosphere. Because these changes have nothing to do with subsurface geology, therefore, they need to be removed from the raw measurements to enhance the signal from the geology features.

Usually, in data acquisition, one magnetometer was fixed at the base station to take continuous measurements of the magnetic field, say, every 10 seconds; while another magnetometer will be carried to the field to record at various locations.

3.4.2 IGRF SUBTRACTION

After subtracting the temporal variations of the magnetic field, what's left is the variation due to the normal Earth, and the materials deep in the core. Therefore, the **International Geomagnetic Reference Field (IGRF) subtraction** will be taken, in order to subtract the background magnetic field produced by the Earth's deep interior materials, so that, we can only focus on magnetic anomalies from the crust.

In general, the vertical gradient varies from approximately 0.03 nT/m at the poles to 0.01 nT/m at the magnetic equator, while the longitude variation is rarely greater than 6 nT/km. Therefore, elevation and latitude corrections are generally unnecessary.

To do this IGRF correction, we can input the longitude and the latitude of the measurement location into a mathematical model of the Earth magnetic field. After the subtraction, what's left in the data is from the local geology.

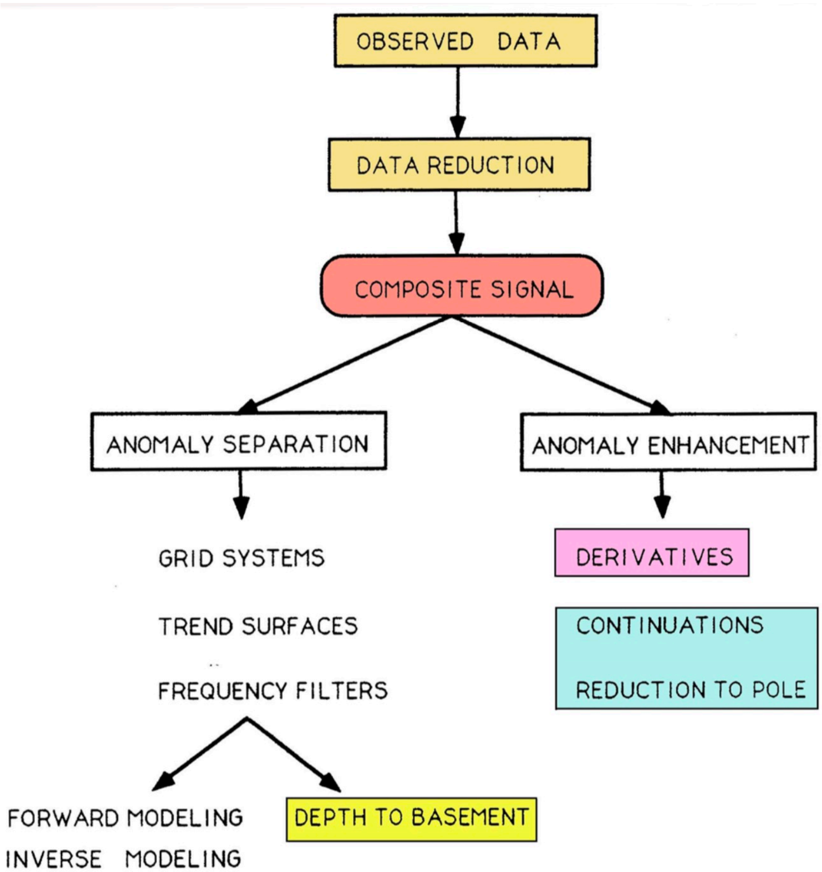
Moreover, besides the two corrections mentioned above, do we also need to consider the terrain effects in the magnetic data? The answer is no, in general. However, in some special cases, terrain effects can be significant. For example, terrain effects can be as large as 700 nT at steep slopes (e.g., 45 degrees) on only 10 m extent in formations containing 2% magnetite. In such cases, terrain correction should be done. However, this correction requires us to know the magnetic source bodies in the terrain and be able to model their magnetic effects, which is very hard to achieve! Moreover, crustal magnetization can vary by several orders of magnitudes at essentially all spatial scales. Thus, we generally do not do terrain correction. The effects of terrain are often left to the modeling and interpretation stage.

CHAPTER 4

Chapter 4: Anomaly Enhancement

4.1 DATA ENHANCEMENT

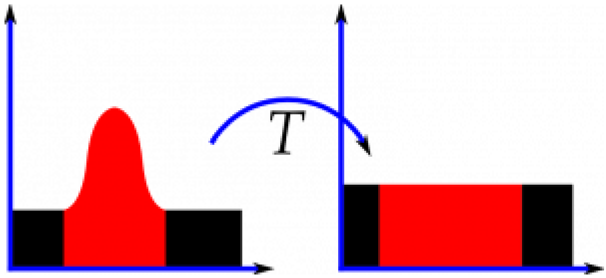
In previous chapters, we have discussed a lot about processing data which contains contributions from many sources, these sources have different physical properties and different scales or depths. Specifically, we have focused on anomaly separation from the observed data to remove unwanted contributions and single out the signal from targets. In practice, we would also apply various anomaly enhancement methods, in order to increase the visibility, or interpretability, or importance of the desired features, or signals with respect to others. Anomaly enhancement does not remove unwanted signals, instead, it helps to diminish or suppress their manifestation on a data map. There are many enhancement methods, such as the histogram equalization technique, and various types of derivative-based enhancement methods. We will discuss them in detail in this chapter.



Data processing flowchart. Image courtesy of Stuart Hall at UH.

4.1.1 HISTOGRAM EQUALIZATION TECHNIQUE

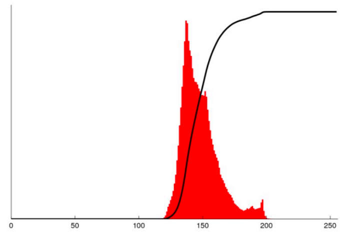
Histogram equalization technique is a common method in image processing to increase contrast in the intensities, such as in the pixel values intensity so that the adjusted values can be better distributed on the histogram, as illustrated in the image below.



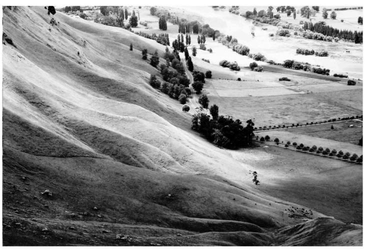
Histograms of an image before and after equalization. (https://en.wikipedia.org/wiki/Histogram_equalization)

The technique allows for areas of lower local contrast to gain a higher contrast, through spreading out the most frequent intensity values in the histogram. In the example of the figure shown below, the top row displays the original image and its pixel values histogram (in red) and its cumulative histogram (in black) before applying the histogram equalization method; The image is blurry, and the pixels values are concentrated within a narrow range. However, after implementing the equalization (shown in the bottom row), the pixel values are spread out in a wider range, and the resulting image is sharper and therefore more details could be visualized.

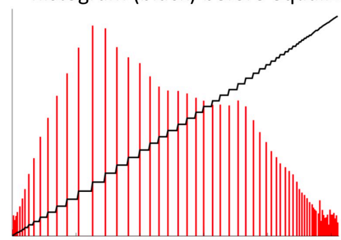
Before histogram equalization



Histogram (red) and cumulative histogram (black) before equalization



After histogram equalization

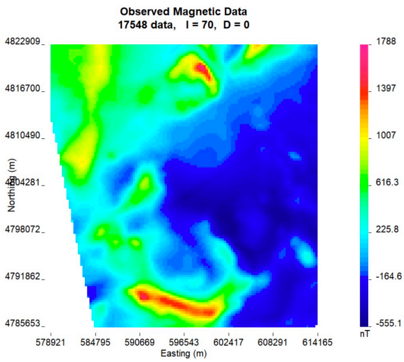


Histogram (red) and cumulative histogram (black) after equalization

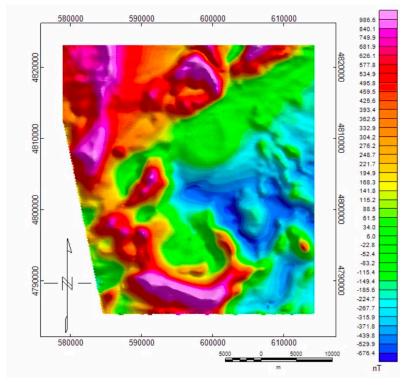
Implementation example. Image source from https://en.wikipedia.org/wiki/Histogram_equalization

Application to geophysics

In geophysical data, subtle features in data images with a high dynamic range, such as aeromagnetic data, are difficult to be visualized. Histogram equalization can help to bring out these subtle features. For example, the left figure below is the observed aeromagnetic data image without equalization, while the right figure displays the data after equalization was applied; We can observe that large-amplitude features are still there, while the smaller amplitude values (around 0) which are not clearly shown in the original figure now become more clearly visible to our naked eyes.



Aeromagnetic data without equalization



Aeromagnetic data with equalization

Application example to Geophysics.

4.1.2 DERIVATIVE-BASED ENHANCEMENT METHODS

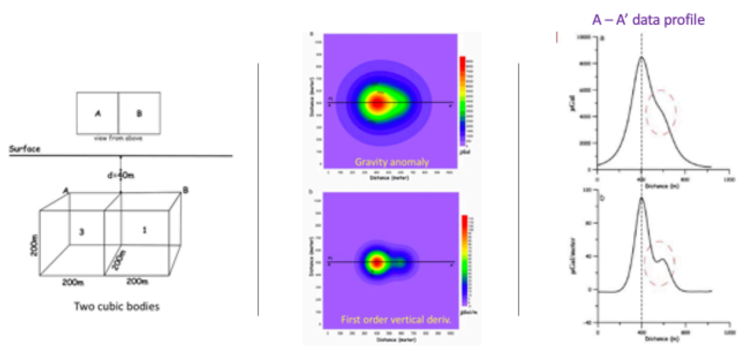
Usually, the anomalies to be enhanced are of small spatial scales than that of those we want to suppress. In general, features with shorter wavelengths are associated with steeper gradient and greater curvatures. Therefore, derivative-based methods are among the most widely used techniques for enhancing anomalies.

There are many derivative-based methods, such as the vertical derivative, total horizontal derivative, tilt derivative, and total horizontal derivative of the tilt derivative, which we will discuss in the following. Other enhancement methods are also popular, such as downward continuation, strike filtering, and phase preserving dynamic range compression filter (https://www.peterkovesi.com/projects/tonemapping/index.html).

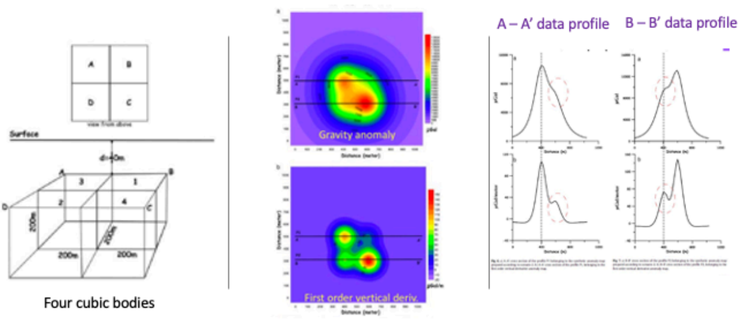
4.1.2.1 Vertical derivative

The vertical derivative of the data image could enhance the visibility and interpretability of the original data image.

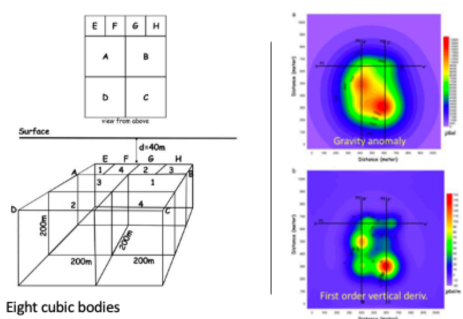
For example, as shown in the synthetic modeling figures below, the gravity gradient image (g_{zz}) is compared with the gravity anomaly image (g_z) in several synthetic models, with a different number of anomaly blocks in the subsurface. Generally, in comparison with the gravity images, the vertical derivative images could be used to better define and differentiate the spatial locations of the blocks, and more accurately to determine the number of blocks from the data images, while the gravity anomaly images cannot be used to clearly separate the signals from multiple blocks.



(Gonenc 2014, JAG) Two cubic blocks example.



(Gonenc 2014, JAG) Four cubic blocks example.



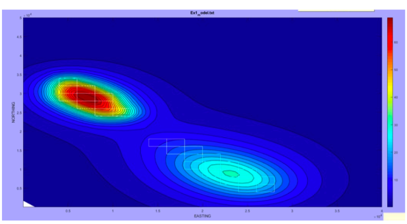
(Gonenc 2014, JAG) Eight cubic blocks example.

4.1.2.2 Total horizontal derivative

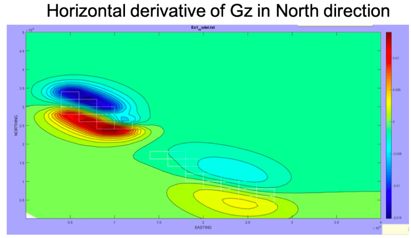
Similar to the vertical derivative of the data, the horizontal derivative of the data image could also be used to better enhance the anomaly interpretability.

For example, the figures below represent the gravity data g_z, the horizontal derivative of g_z in the North direction, and horizontal derivative of g_z in the East direction, respectively. All the derivative data images can be used to better define the boundaries of the target features in the subsurface. More specifically, the horizontal derivative in the North direction defines the north/south boundary of the target body, while the horizontal derivative in the East direction defines the west/east boundary of the target body. Altogether, the derivative images could help us determine the general boundary of the target feature.

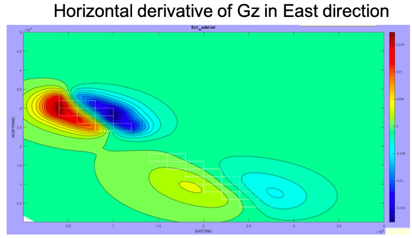
Gravity anomaly



Horizontal derivative of Gz in North direction



Horizontal derivative of Gz in East direction



The total horizontal derivative magnitude is the square root of the sum of the squares of the horizontal derivatives in x (or E/W) and y (or N/S) directions. Its math expression and an example are shown in the following.

$$h(x,y) = \left[\left(\frac{\partial g_z(x,y)}{\partial x} \right)^2 + \left(\frac{\partial g_z(x,y)}{\partial y} \right)^2 \right]^{\frac{1}{2}}$$

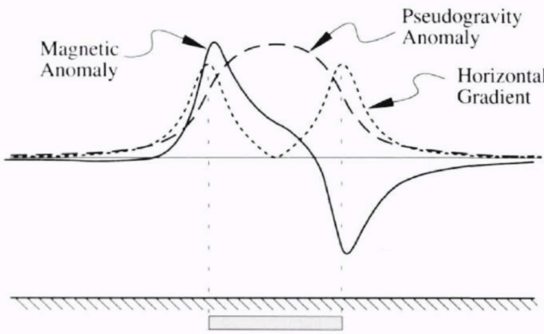
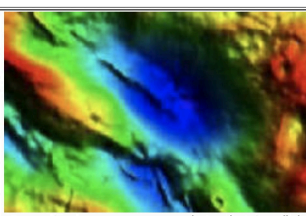
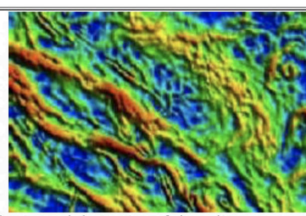


Fig. 12.13. The magnetic anomaly, pseudogravity anomaly, and magnitude of the horizontal gradient over a tabular body.

Blakely, 1996, p348

Another comparison example was shown in the images below, the total horizontal derivative of the gravity data shows much more detailed structures in comparison to the Bouguer gravity map.

	
Bouguer gravity: contour map plotted using "shading."	Total horizontal derivative of the adjacent gravity map.

Gravity map v.s. The total horizontal derivative of the gravity map. (https://www.eoas.ubc.ca/ubcgif/iag/methods/meth_4/gravgrads.htm)

4.1.2.3 Tilt derivative

Tile derivative is the arctangent of the ratio between the vertical derivative and the total horizontal derivative of the data. Its mathematical expression is shown in the following.

$$\mathrm{TILT} = \tan^{-1}\frac{\text{vertical component of gradient}}{\text{horizontal component of gradient}}$$

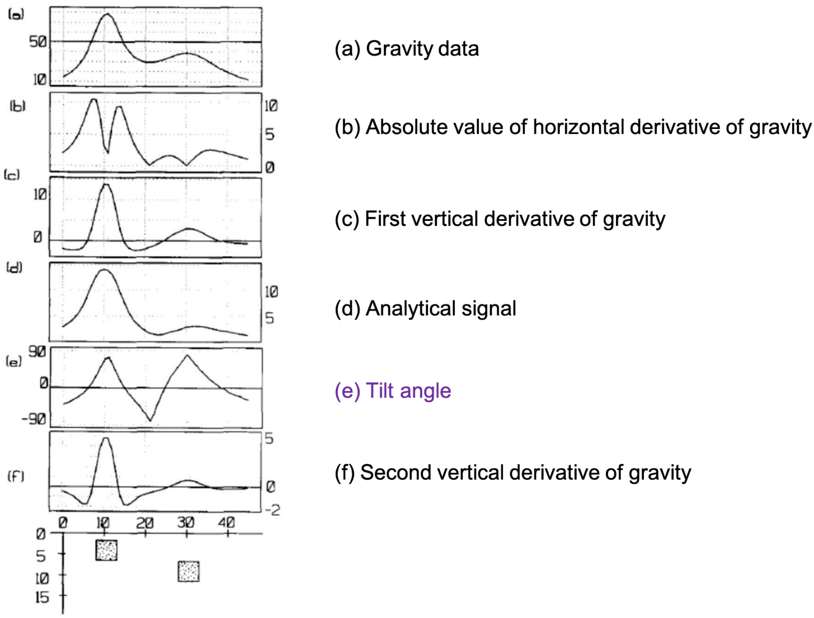
$$= \tan^{-1}\frac{\left(\frac{\partial f}{\partial z}\right)}{\left(\frac{\partial f}{\partial h}\right)}$$

where:

$$\frac{\partial f}{\partial h} = \left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right]^{1/2}$$

Tilt derivative. (Miller and Singh, 1994, JAG)

We can compare and contrast the characteristics of different derivatives of the data along with the tilt derivative of the gravity data simulated from two density blocks buried at different depths shown below.



Miller and Singh (1994, JAG)

Based on the data images above, we can make the following observations:

- in figure (b), the horizontal derivative shows 2 peaks corresponds to the boundary of the blocks;
- in figure (c), the vertical derivative shows a more compact pattern, which defines the shape and distribution of the source bodies;
- in figure (e), the **tilt angle image** shows two signal peaks, where the peak from the deeper body is comparable in amplitude with the shallower body signal peak! And the peaks are right above the body locations. That is the significant characteristic of the tilt angle enhancement method.

If we further take the horizontal derivative of the tilt angle, then the data profile will show the boundary of the source bodies, and this will be called the **horizontal derivative of the tilt derivative**.

CHAPTER 5

Chapter 5: Fourier-domain Modeling & Transformations

5.1. BACKGROUND

5.1.1. MOTIVATION

Up until this point, we have dealt with 2D gravity and magnetic data maps in the spatial domain. In this chapter, we will be looking at different way in dealing with these data: in the **frequency domain**. A lot of data processing techniques require the data to be in frequency domain due to its simpler and more efficient mathematical operators. The motivation to studying Fourier domain modeling for modeling any potential fields data are the following:

- The expression of potential field in frequency domain gives us a **different perspective** to look at our data
- Much of the **processing and analysis** of potential field data are done in frequency domain
- Potential field is related to the source distribution by a

convolution with the Green's function (will be explained later on)

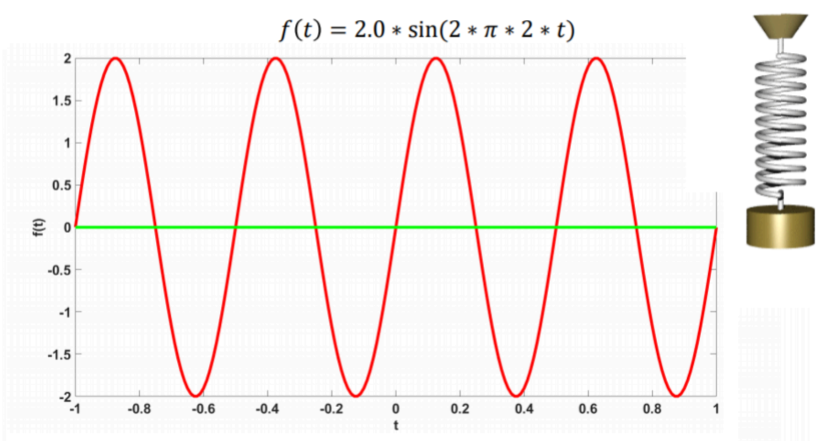
In this chapter, we will examine Fourier expression of simple potential functions and Fourier expression of field due to simple sources.

5.1.2. FREQUENCY AND PHASE

Before going over Fourier Transforms, we first need to know the basic theory behind it. There are several basic concepts that we will first cover before going through Fourier transforms.

5.1.2.1. Waveforms

A spring-mass system, as we see in the figure below, oscillates throughout time. The displacement of the mass from the oscillatory behavior can be graphed with respect to time as we see in the figure. It turns out this movement can be expressed as a sinusoidal function. We call this a **waveform.**

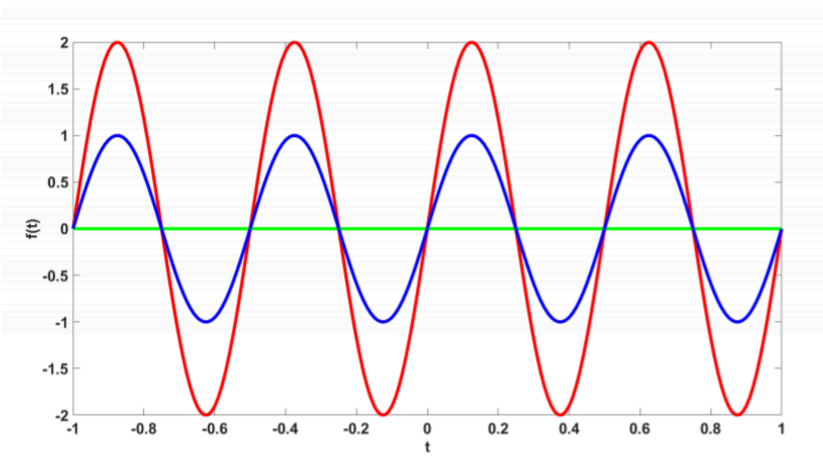


The **amplitude** of the waveform depends on the constant before the sin function. For example, the above figure shows an waveform with an amplitude of 2 units. The **blue** value in the below indicates the amplitude.

f(t)= 2.0 *sin (2* π*2*t)

If there is a function that has the form of *f(t) = 1.0*sin(2*π*2*t)*, then the waveform will be exactly the same as the aforementioned example, except that the amplitude would be halved, as we can see in the **blue** line in the following figure:

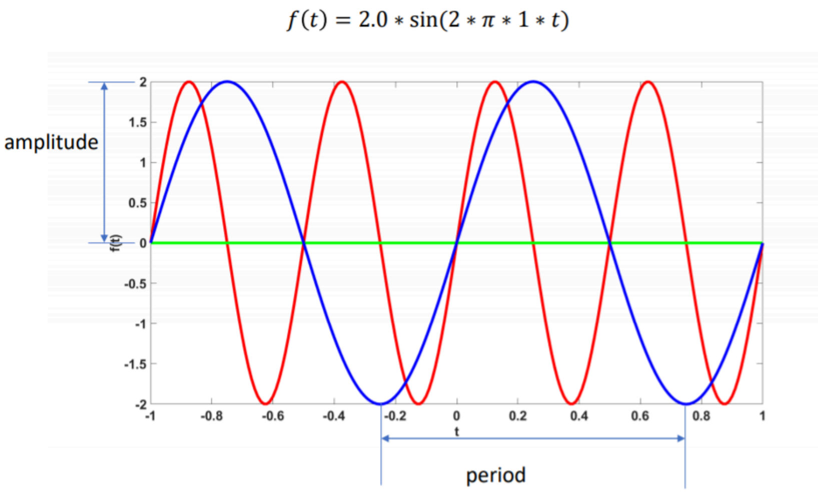
$$f(t) = 1.0 * \sin(2 * \pi * 2 * t)$$



The **period** of the waveform depends on the value preceding *t* inside the sin function . One period represents the amount of time it takes for the waveform to undergo one cycle (1 peak and 1 trough). For example, the red text in the following waveform determines the period:

f(t) = 2.0 * sin(2* π *2*t)

If there is a function that has the form of *f(t) = 2.0*sin(2*π*1*t)*, then the waveform will be exactly the same as the aforementioned example, except that the period is twice of that, as we can see in the **blue** line in the following figure:



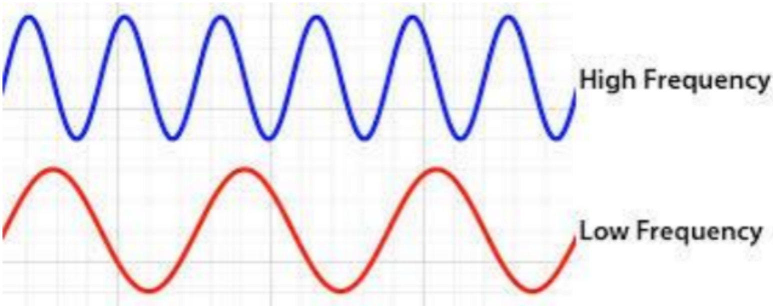
5.1.2.2. Frequency

The frequency of the waveform can be determined by the period. **Frequency** represents the number of cycles in one second, and the unit is **(/s)** or Hertz(**Hz). Frequency** can also be mathematically defined as:

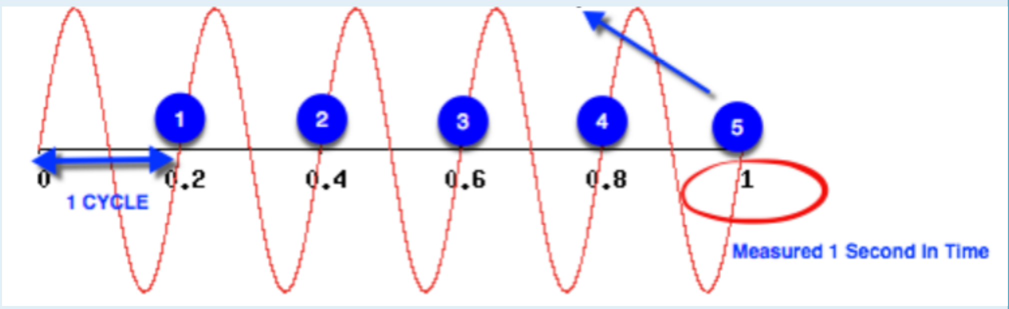
$$f = 1/T$$

Where T is the period.

If we have waveforms that have their peaks and troughs closer together, we say that the waveform has a **high frequency**. If they are far apart, then the waveform has **low frequency.**



Test your understanding: frequency



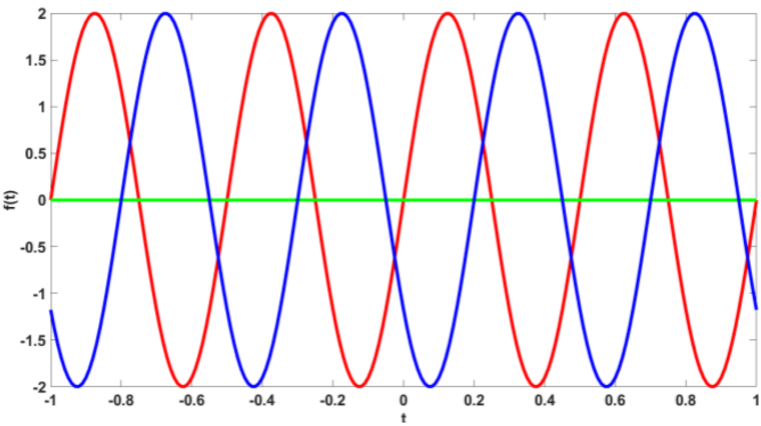
What is the **frequency** of this waveform? What about the **period?**

5.1.2.3. Phase

Phase is determined by the argument inside the sin function. For example, the following figure shows two waveforms of different phases. The first function below is for the red line in the graph, and the second one is for the blue line in the graph.

$$f(t) = 2.0 * \sin(2 * \pi * 2 * t)$$

$$f(t) = 2.0 * \sin(\boxed{2 * \pi * 2 * (t - 0.2)})$$

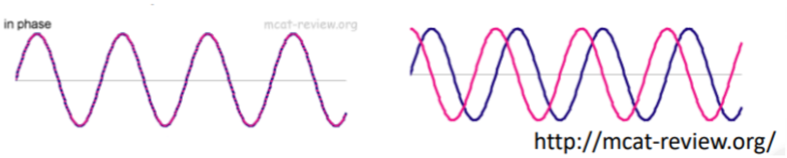


Notice that the two waveforms have the same amplitude and frequency, but one seems to be delayed compared to the other. This is because these two waveforms have different **phase**s.

Phase is the argument inside the sine(or cosine) function for a waveform. Given a fixed amplitude and frequency, phase determines when the peaks (or troughs) occur. Simply, it specifies where in the cycle is it oscillating at t=0. The unit for phase is in **degrees** or **radians.**

There are two terms that are related to phase: **in-phase** and **out-of-phase**.

If the two peaks of two signals with the same frequency are in the exact same alignment at the same time, then they are called **in-phase**. If the two peaks of two signals with the same frequency are not in an exact alignment at the same time, then they are **out-of-phase.** The following figure gives out a visualization of the two difference.



The important concept that we want to take home from this is the fact that **a sine wave and a cosine wave are 90 degrees out-of-phase with each other.** Thus, we can express a waveform in two ways: using a sine wave or a cosine wave.

5.1.2.4 General notation

The more general notation of a waveform can be expressed as:

$$f(t) = Asin(\omega t + \phi)$$

or

$$f(t) = Acos(\omega t + \phi)$$

Where **A** is the amplitude, and ω is the angual frequency in **radians**

per seconds. The angular frequency can be found from the equation:

$$\omega = \frac{2\pi}{T} = 2\pi f$$

Main takeaways from this frequency and phase

Any **periodically oscillating sinusoidal wave** can be characterized by three fundamental properties

- **Amplitude(A)**
- **Frequency($\omega = 2\pi f$)**
- **Phase(ϕ)**

A wave can be expressed using the following notation:

f(t) = A sin(ωt + φ)

or

f(t) = A cos(ωt + φ)

5.1.3. COMPLEX VARIABLE

A **complex variable** can be expressed as such:

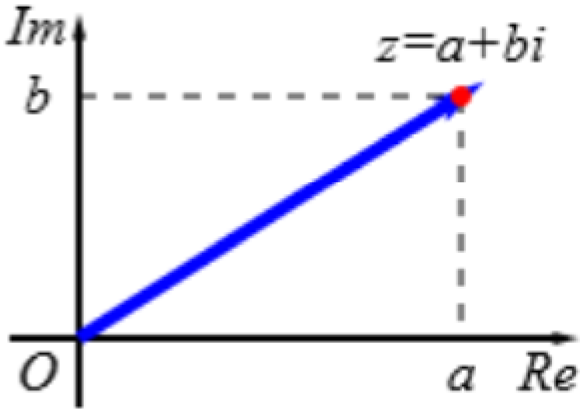
$$z = a + ib$$

where a and b are real numbers, and *i* is the imaginary unit that is equal to $\sqrt{(}-1)$, thus $i^2 = -1$

Thus, a complex number can be separated into two:

- a: real part
- b: imaginary part

A complex number can be viewed as a point in the **complex plane,** as the following figure:



https://en.wikipedia.org/wiki/Complex_number

5.1.3.1 Absolute value and argument

We know that a complex number can be expressed as such:

$$z = a + ib$$

The **absolute value(modulus)** of this complex number is thus:

$$|z| = r = \sqrt{(}a^2 + b^2)$$

The **argument(phase)** is the angle between the vector and the positive real axis. This makes more sense from the complex plane representation in 6.1.3.

The argument is thus expressed as:

$$arg(z) = \theta = atan2(y, x)$$

where y is the projection of the complex number in the imaginary axis, and x is the projection of the complex number in the real axis. Therefore:

$$a = rcos(\theta), b = rsin(\theta)$$

and thus:

$$z = rcos(\theta) + irsin(\theta)$$

Test your understanding: Complex Number

Given a complex number:

$$z = 1 + i\sqrt(3)$$

Calculate its **modulus** and **argument!**

5.1.3.2. Euler's formula

Euler's formula is the fundamental relationship between the **trigonometric functions** and the **complex exponential function**. This can be mathematically expressed as:

$$e^{i\theta} = cos(\theta) + isin(\theta)$$

From Euler's formula, we can express the complex number as:

$$z = rcos(\theta) + irsin(\theta) = re^{i\theta}$$

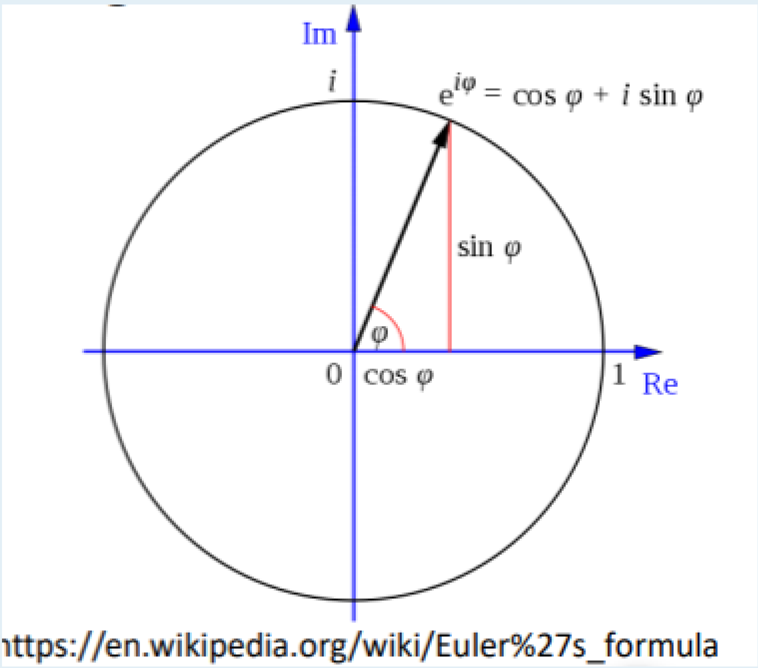
Therefore, z can also be expressed as:

$$z = re^{i\theta}$$

In other words, a complex number can be expressed as either with a **trigonometric function** or an **exponential function.**

Test your understanding: Euler's Formula

Let $\phi = \frac{\pi}{2}$
Calculate $e^{i\theta}$
Hint: Just plug in $\theta = \frac{\pi}{2}$ *to the function* $e^{i\theta} = cos(\theta) + irsin(\theta)$



Ans: i

Consider that a mathematical notation for a cosine wave is:

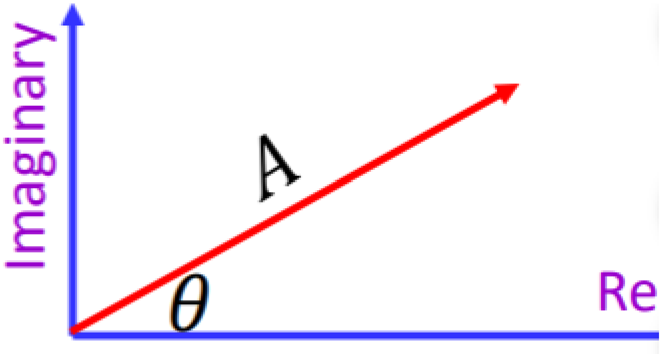
$$f(t) = A.cos(\theta)$$

where A is the **amplitude** and $\theta = \omega t + \phi$ is **phase.**

A different way to to express that wave is to use the following notation:

$$f(t) = Ae^{i\theta}$$

The representation of this in the complex plane is as follows;



(Note that if you were to see something that looks like $Ae^{i\theta}$, please always keep in mind that this represents sinusoidal waves with amplitude A and phase θ)

5.1.3.3. Deriving a complex number

Deriving an exponential function is much simpler than deriving a trigonometric function. Consider the following problem;

$$\frac{\partial e^{i\omega t}}{\partial t} = i\omega e^{i\omega t}$$

Now, remember that:

$$i = e^{i\frac{\pi}{2}}$$

Therefore:

$$\frac{\partial e^{i\omega t}}{\partial t} = i\omega e^{i\omega t} = e^{i\frac{\pi}{2}}\omega e^{i\omega t}$$

$$\frac{\partial e^{i\omega t}}{\partial t} = \omega e^{i(\omega t + \frac{\pi}{2})}$$

Notice that the argument in the exponential has changed from $i\omega t$ to $i(\omega t + \frac{\pi}{2})$

The conclusion to this is that **when you take a time derivative, the phase will change by** $\frac{\pi}{2}$.

5.1.3.4. Why bother with complex numbers?

The first question that you might ask is: why do we need to know about complex numbers? The reason is: It turns out, the whole Fourier theory was built upon complex variables.

Other than that, complex numbers offers a lot of mathematical conveniences, such as:

- No need to deal with **trigonometric functions**
- **Exponentials** are easier to manipulate
- Some problems can be solved **easier** using **complex number/complex analysis** (examples: quantum physics, conformal transformations, AC circuits, etc.)

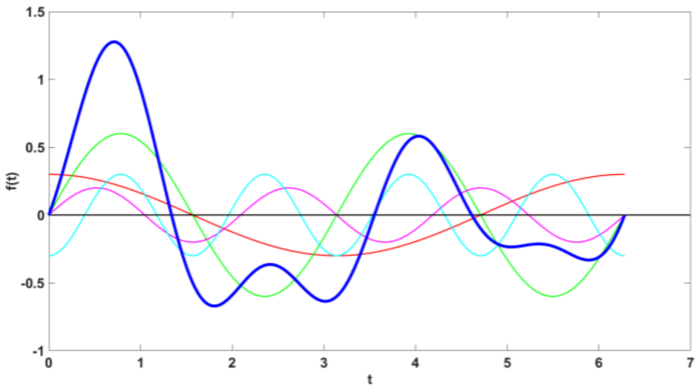
If you are interested to learn more about complex numbers, here are some resources that might be of interest:
• Complex Analysis Made Simple on youtube.
• MIT OpenCourseWare: Development of the
complex numbers
• Functions of complex variables in StackExchange
•Intuitive Arithmetic with Complex Numbers (Betterexplained.com)

5.2. FOURIER TRANSFORM

5.2.1. WAVEFORMS AND AMPLITUDE SPECTRUM

Virtually, any real world waveforms can be represented as a sum of sins, no matter how complicated they may look.

For example, take a look on this waveform:



Fourth component: $f(t) = 0.3\sin(\frac{8\pi}{T} * t - \frac{\pi}{2})$, where T = 6.28 s, t = 0:0.01:T

The **blue** waveform may look complicated, but it is actually the **sum** of four other sinusoidal functions in the above graph, with the equation for each line being;

Red line: $f(t) = 0.3sin(\frac{2\pi}{T}t + \frac{\pi}{2})$

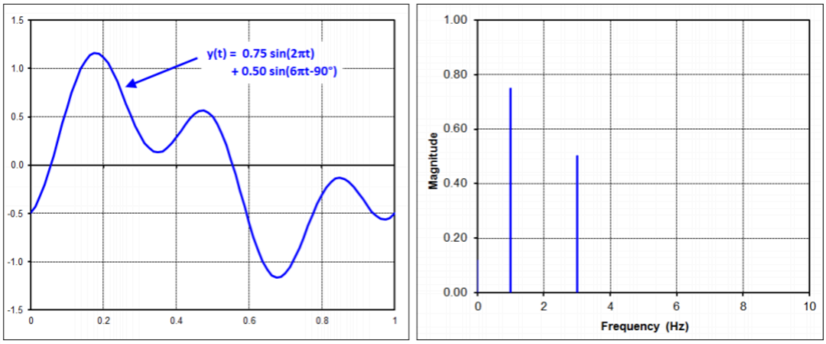
Green line: $f(t) = 0.6sin(\frac{4\pi}{T}t)$

Magenta line: $f(t) = 0.2sin(\frac{6\pi}{T}t)$

Cyan line: $f(t) = 0.3sin(\frac{8\pi}{T}t - \frac{\pi}{2})$

This proves that **no wonder how complex they are, waveforms are just a sum of sinusoidal waves.**

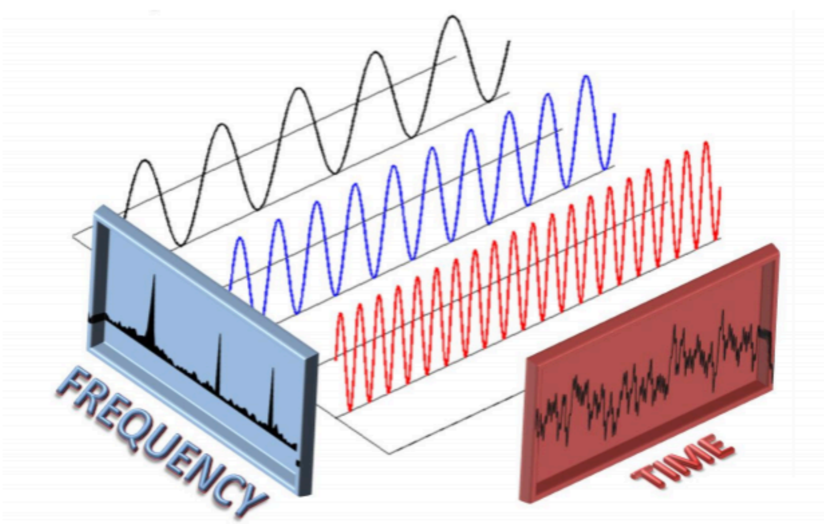
If we were to plot a waveform that is the sum of two sinusoidal waves, and plot the magnitude of each wave in terms of their frequency, we get something like this:



http://www.continuummechanics.org/fourierxforms.html

The signal results from the sum of two sine waves. These two waves have different amplitudes, frequencies, and phases, and this difference is made more clear in the **amplitude spectrum** on the right side of the above figure.

Another way to look at the difference between frequency domain and time domain is with the following figure:



http://visualizingmathsandphysics.blogspot.com/2015/06/fourier-transforms-intuitively.html

The figure above shows two different ways of viewing the combination of these 3 waveforms. In time domain, it is the summation of these three waveforms (with noise), and thus the perceived signal in the time domain looks more complicated. When viewing the three waveforms in the frequency domain, it is in the form of *three separate peaks* that tells you the frequency and amplitude of said waveforms.

5.2.2. DEFINITION

Mathematically speaking, any waveform function can be expressed as:

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}dw$$

The term $F(\omega)e^{i\omega t}$ is a sinusoidal wave, with the amplitude being $F(\omega)$. This essentially says that **any waveform is a sum of many sinusoidal waves with different amplitudes, phases, and frequencies.**

The **Fourier Transform** allows us to find the constituent frequencies given a signal in time domain. This is mathematically expressed as:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt$$

It decomposes a signal into a series of sinusoidal waves of different amplitudes, frequencies, and phases. This helps identify what the sine and cosine components that make up the signal are.

The **inverse Fourier transform** is thus:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} dw$$

In terms of notation, we refer the Fourier transform of a function $f(t)$ as the capital F, and the inverse Fourier transform as F^{-1}. Putting this in mathematical perspective:

Fourier Transform: $F[f(t)] = F(\omega)$

Inverse Fourier Transform: $F^{-1}[F(\omega)] = f(t)$

If you want to know more about Fourier transform(FT), here's some optional reading material on FT:

- https://betterexplained.com/articles/an-interactive-guide-to-the-fourier-transform/
- http://visualizingmathsandphysics.blogspot.com/2015/06/fourier-transforms-intuitively.html
- https://www.ritchievink.com/blog/2017/04/23/understanding-the-fourier-transform-by-example/
- https://learn.adafruit.com/fft-fun-with-fourier-transforms/background

5.2.3. FROM TIME DOMAIN TO SPATIAL DOMAIN

For now, you might notice that everything that we have talked about so far is in time domain. However, potential field data is in the **spatial domain**. This does not seem to be a problem because **Fourier transform also applies in the spatial domain**. The Fourier transform in time and spatial domain can be seen in the following figure:

Time domain

$$\mathcal{F}(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-iwt} dt$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathcal{F}(\omega)\, e^{iwt} dw$$

Spatial domain

$$\mathcal{F}(k) = \int_{-\infty}^{+\infty} f(x) e^{-ikx} dx$$

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathcal{F}(k)\, e^{ikx} dk$$

Notice that they are both the same, except for the fact that the Fourier transform of a in time(t) domain is a function of angular frequency (ω), and in spatial(x) domain, the Fourier transform is a function of wavenumber(k). Thus, everything in spatial domain is the same as in time domain, except we're dealing with space (x) instead of time (t), and the Fourier transform involves wavenumber (k) instead of angular frequency (ω).

As we mentioned before, frequency is a measure of how many cycles per unit of time. **Wavenumber** is the measure of how many cycles per unit of distance. The following equation applies for wavenumber:

$$k = \frac{2\pi}{\lambda}$$

$$f = \frac{1}{\lambda}$$

Where λ is the wavelength (one cycle in unit of distance).

Another concern is that the data maps that we have dealt with are 2-dimensional. Therefore, if we were to take the Fourier transform of such spatial data maps, then we need to do a **2D Fourier transform** in spatial domain. The following is such equation:

Spatial domain 2D

$$\mathcal{F}(k_x, k_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-i(k_x x + k_y y)}dxdy$$

$$f(x,y) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{F}(k_x, k_y)e^{i(k_x x + k_y y)}dk_x dk_y$$

Note that k_x is the wavenumber in the x axis and k_y is the wavenumber in the y axis respectively.

5.2.4. DC-COMPONENT

There is one interesting result that we can see from the Fourier transform:

$$F(k) = \int_{-\infty}^{\infty} f(x)e^{-ikx}$$

If we have **k=0**, that means the Fourier transform would be:

$$F(0) = \int_{-\infty}^{\infty} f(x)e^{-i.0.x} = \int_{-\infty}^{\infty} f(x)dx$$

Fourier transform of a function f(x) evaluated at k=0 is simply the integral of this function over the entire x axis. And it is always real if f(x) is real. We call this the **DC Component.** Thus, the **DC component** is the Fourier transform of a function f(x) evaluated at k=0.

5.2.5. AMPLITUDE AND PHASE

In general, Fourier transform of $F(k)$ is a complex function with a real and an imaginary part:

$$F(k) = Re(F(k)) + iIm(F(k))$$

Which also can be written as:

$$F(k) = |F(k)|e^{i\theta(k)}$$

where:

$$|F(k)| = \sqrt(Re(F(k))^2 + (Im(F(k)))^2)$$

and

$$\theta(k) = arctan\frac{Im(F(k))}{Re(F(k))}$$

5.2.6. PROPERTIES

There are several properties of Fourier transform. We will cover each one in this section.

5.2.6.1. Linearity

If we have a two functions in the spatial domain, with each having a Fourier transform:

$$f_1(x) \leftrightarrow F_1(k)$$
$$f_2(x) \leftrightarrow F_2(k)$$

Then the summation of the two functions in the spatial domain is also the summation in the Fourier domain as such:

$$a_1 f_1(x) + a_2 f_2(x) \leftrightarrow a_1 F_1(k) + a_2 F_2(k)$$

5.2.6.2. Scaling

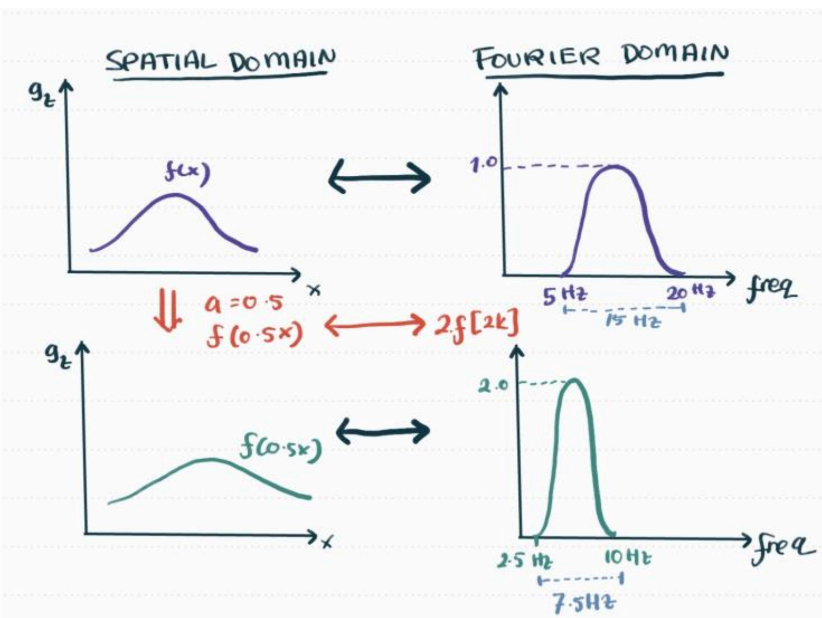
If we have a function $f(x)$ with its Fourier transform $F(k)$:

$f(x) \leftrightarrow F(k)$, then if you *stretch* your $f(x)$ by 0.5, we would have an $F(k)$ with narrower band and higher amplitude. This is mathematically expressed as:

$$f(ax) \leftrightarrow \frac{1}{|a|} F(\frac{k}{a})$$

- If a>1, we shrink the signal (the signal contains more high frequency content), then the Fourier transform of this new signal will contain higher frequencies (hence the $\frac{k}{a}$)
- If a<1, we stretch the signal (the signal contains more low-frequency contents) then the Fourier transform of this new signal will contain lower frequencies

An example of this property can be seen in the following graph.



Thus, a broad anomaly will have a narrower amplitude spectrum than a narrow anomaly. Because the width of an anomaly is directly related to the **depth** of its source, we can expect that the narrowness of a Fourier-transformed anomaly will also be related to the depth of the source.

Scaling Property

If you **stretch** in spatial domain $\longleftrightarrow$ you are **shrinking** in Fourier domain
If you **shrink** in spatial domain $\longleftrightarrow$ you are **stretching** in Fourier domain

5.2.6.3. Shifting

If you have a function such that $f(x) \leftrightarrow F(k)$, then:

$$f(x - x_0) \leftrightarrow F(k)e^{-ikx_0}$$

Shifting a function along the x axis in the spatial domain is equivalent to adding a linear **phase** factor to the function's Fourier transform. The amplitude spectrum is unaffected.

5.2.6.4. Differentiation

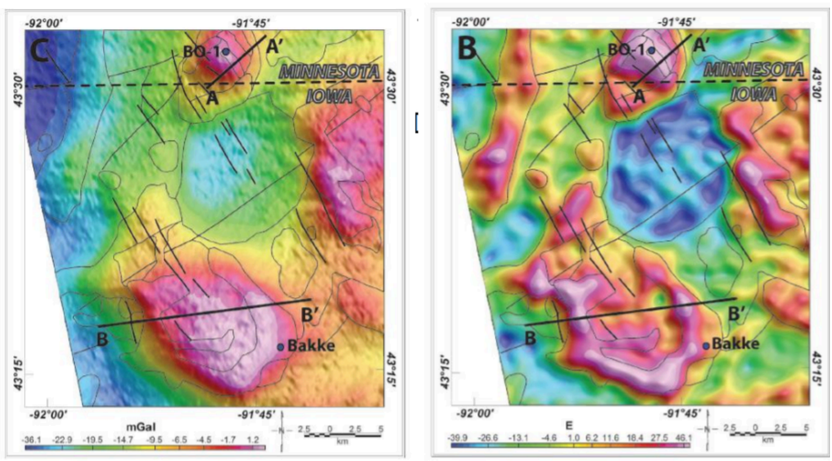
If we have a function such that: $f(x) \leftrightarrow F(x)$

then:

$$\frac{d}{dx}f(x) \leftrightarrow ikF(k)$$

As mentioned before in the previous section, the phase will change by $\frac{\pi}{2}$. Thus, the amplitudes for low-frequency components will be suppressed, and the amplitudes for high-frequency components will be amplified.

The concept of differentiation in the Fourier domain is applied in potential field data. Specifically, for **derivative-based anomaly enhancement.** For example, the following figure shows a bouguer anomaly and Gzz data map. We can see that by taking the derivative of the gravity data, the anomalies are more enhanced and thus we see the boundaries of the anomaly clearer.

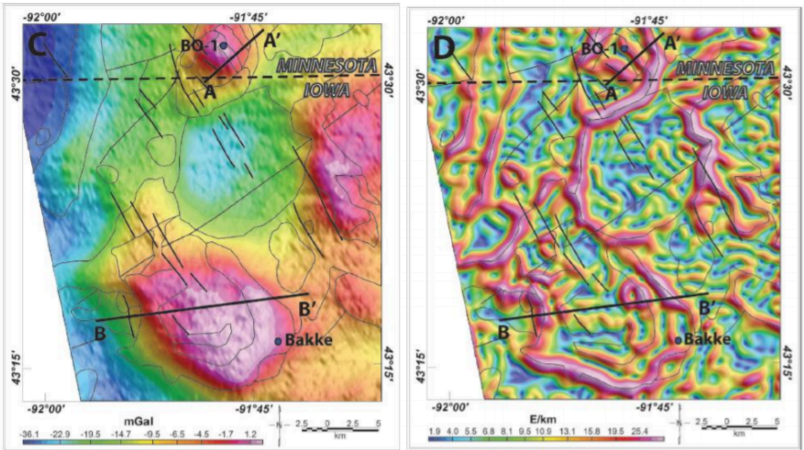


(Left) Bouguer anomaly over northeast Iowa. (Right) Gzz data.

Drenth et al. (2015)

Another way to enhance the anomaly is to take the horizontal gradient of the magnitude as seen in the following picture. We can see that the boundaries are more accentuated in the data map on the right.

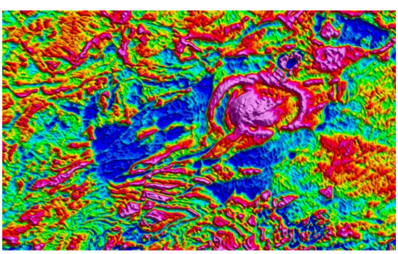
Bouguer anomaly



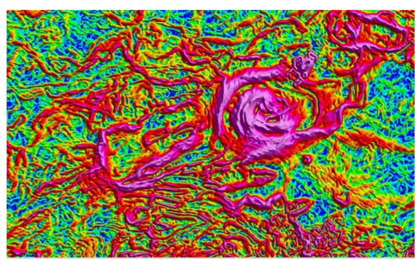
(Left) Bouguer anomaly over northeast Iowa. (Right) Horizontal gradient magnitude.

Drenth et al. (2015)

Tensor data



FTG: Gzz (https://www.bellgeo.com/what-is-ftg)



FTG: Total horizontal gradient (https://www.bellgeo.com/what-is-ftg)

With how well derivatives are in enhancing the anomalies, it is no surprise that we would want to do Fourier transform to these data maps. This is because, as mentioned before, the math involved in derivation in the fourier domain is much more simpler than in spatial domain.

Summary of Fourier Transform Properties

The properties of Fourier transform are:

- **Linearity:** Suppose that $f_1(x) \leftrightarrow F_1(k)$ and $f_2(x) \leftrightarrow F_2(k)$, then:

$$a_1 f_1(x) + a_2 f_2(x) \leftrightarrow a_1 F_1(k) + a_2 F_2(k)$$

- **Scaling**: Suppose that $f(x) \leftrightarrow F(k)$, then:

$$f(ax) \leftrightarrow \frac{1}{|a|} F(\frac{k}{a})$$

- **Shifting:** Suppose that $f(x) \leftrightarrow F(k)$, then:

$$f(x - x_0) \leftrightarrow F(k)e^{-ikx_0}$$

- **Differentiation:** Suppose that $f(x) \leftrightarrow F(x)$, then:

$$\frac{d}{dx} f(x) \leftrightarrow ikF(k)$$

5.2.6.5. Fourier transform of 1/r

If you still remember the gravity potential, the equation is shown:

$$U(p) = \gamma \frac{dm}{r}$$

where γ is the gravitational constant, dm is a small mass over the r.

Similar to the magnetic scalar potential:

$$V(p) = \frac{\mu_0}{4\pi} \frac{\mathbf{m} \cdot \widehat{\mathbf{r}}}{r^2} = -\frac{\mu_0}{4\pi} \mathbf{m} \cdot \nabla_p \frac{1}{r}$$

We observe that both gravity potential and magnetic scalar potential have the $\frac{1}{r}$. Before we apply Fourier transform to gravity potential and magnetic scalar potential, we need to figure out the Fourier transform of $\frac{1}{r}$. Specifically, we need to make use of Bessel function and Hankel transform. But the derivation is mathematically involved, so we will skip the detailed derivations. For those who are interested, please refer to Blakely (1996, p271-273).

After a series complicated derivation, we obtain:

$$\mathcal{F}(\frac{1}{r}) = 2\pi \frac{\mathrm{e}^{|\mathrm{k}|(\mathrm{z_0}-\mathrm{z}^{'})}}{|\mathrm{k}|}$$

where, $z^{'}>z_0$, $|\mathrm{k}| \neq 0$, k is wave number.

The equation above is the results for Fourier transform of $\frac{1}{r}$.

$$\mathcal{F}\left(\frac{1}{r}\right) = 2\pi \frac{e^{|k|(z_0-z')}}{|k|}, \text{ for } z' > z_0, |k| \neq 0$$

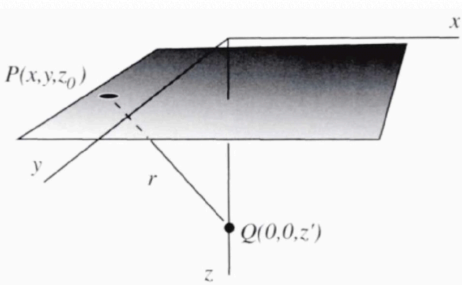


Fig. 11.2. Coordinate system for the derivation of Fourier transformed anomalies caused by point sources. Field is measured on a horizontal surface at z_0, and source is located on the z axis at z'.

Blakely, 1996, p272

5.3. TERMINOLOGY

5.3.1. AMPLITUDE AND PHASE, AMPLITUDE SEPCTRUM AND PHASE SPECTRUM

We can use an equation to do Fourier transform (shown in below).

$$\mathcal{F}(k) = \int_{-\infty}^{+\infty} f(x)e^{-ikx}\,\mathrm{d}\,x$$

where $f(x)$ is a signal. IN most general cases, Fourier transform is a complex function with a real and an imaginary part.

$$\mathcal{F}(k) = Re\mathcal{F}(k) + iIm\mathcal{F}(k)$$

Also, if we apply Euler's formulation, the Fourier transform can be expresses as:

$$\mathcal{F}(k) = |\mathcal{F}(k)|\, e^{i\Theta(k)}$$

where,

$$|\mathcal{F}(k)| = \sqrt{(Re\mathcal{F}(k))^2 + (Im\mathcal{F}(k))^2}$$

$$\Theta(k) = arc\tan\frac{Im\mathcal{F}(k)}{Re\mathcal{F}(k)}$$

The $|\mathcal{F}(k)|$ is called amplitude and $\Theta(k)$ is named phase. If we want to transform a signal from spatial domain to the Fourier domain, we need to know both amplitude and phase.

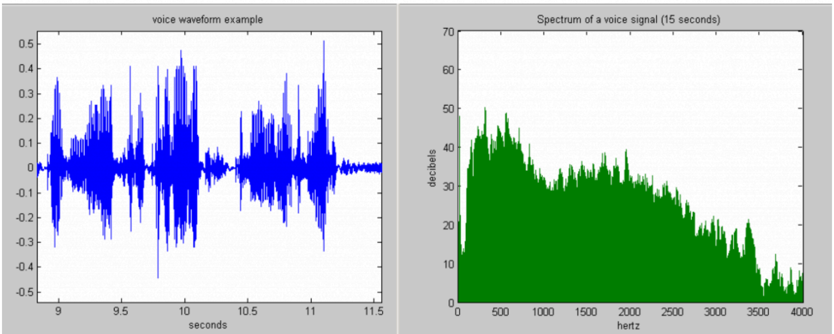
$|\mathcal{F}(k)|$ is simply called amplitude spectrum, which is a function of wave number k. In similar, we also called $\Theta(k)$ phase spectrum, which means how phase change with the wave number k.

5.3.2. ENERGY AND POWER SPECTRUM

The total energy of a real function f(x). It is an integral of entire real axis of the square of amplitude, which measures how much energy in function $\mathcal{F}(k)$.

$$E = \int_{-\infty}^{+\infty} |\mathcal{F}(k)|^2 \,\mathrm{d}\, k$$

where $|\mathcal{F}(k)|$ is Fourier transform which is a complex number. $|\mathcal{F}(k)|^2$ is also called energy density function over frequency, which tells us, how energy of this function distribute with different wave numbers over frequency band. Sometimes, $|\mathcal{F}(k)|^2$ is named energy spectral density, power spectral density (PSD) or power spectrum (PS).



Power spectrum: the voice waveform over time (left) has a broad audio power spectrum (right). https://en.wikipedia.org/wiki/File:Voice_waveform_and_spectrum.png

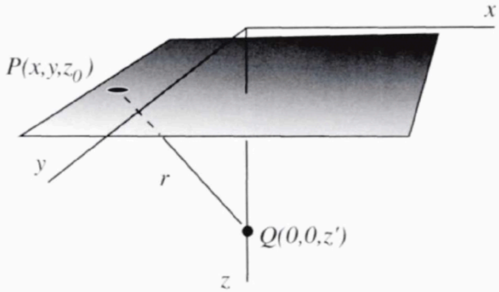
We can see some high amplitude and low amplitude in the figure above (left), if we do the integration of square of amplitude, we will obtain the power spectrum (right).

5.4. GRAVITY FOR A POINT MASS IN FOURIER

DOMAIN

Recall Fourier transform of 1/r

$$\mathcal{F}\left(\frac{1}{r}\right) = 2\pi \frac{e^{|k|(z_0 - z')}}{|k|}, \text{ for } z' > z_0, |k| \neq 0$$



- Fig. 11.2. Coordinate system for the derivation of Fourier transformed anomalies caused by point sources. Field is measured on a horizontal surface at z_0, and source is located on the z axis at z'.

Blakely, 1996, p272

Based on the Fourier transform of 1/r, we can calculate the 3D gravity modeling in Fourier domain with point mass.

$$U = \gamma \frac{\mu}{r}$$

where μ is point mass, U is gravitational potential. Then we apply Fourier transform, and will obtain:

$$\mathcal{F}(U) = \mathcal{F}(\gamma \frac{\mu}{r}) = \gamma\mu\mathcal{F}(\frac{1}{r})$$

After applying Fourier transform of 1/r, it is easily to get:

$$\mathcal{F}(\frac{1}{r}) = 2\pi\frac{\mathrm{e}^{|\mathrm{k}|(\mathrm{z_0}-\mathrm{z}^{'})}}{|\mathrm{k}|}$$

$$\mathcal{F}(U) = 2\pi\gamma\mu\frac{\mathrm{e}^{|\mathrm{k}|(\mathrm{z_0}-\mathrm{z}^{'})}}{|\mathrm{k}|}$$

where, $z^{'}$>z_0, $|\mathrm{k}| \neq 0$.

We currently measures the vertical component of gravity field Gz, which is derivative of gravity potential in z direction.

$$g_z = \frac{\partial U}{\partial z} = \gamma\mu\frac{\partial}{\partial z}\frac{1}{r}$$

Then, we apply Fourier transform,

$$\mathcal{F}(g_z) = \mathcal{F}(\gamma\mu\frac{\partial}{\partial z}\frac{1}{r}) = \gamma\mu\mathcal{F}(\frac{\partial}{\partial z}\frac{1}{r})$$

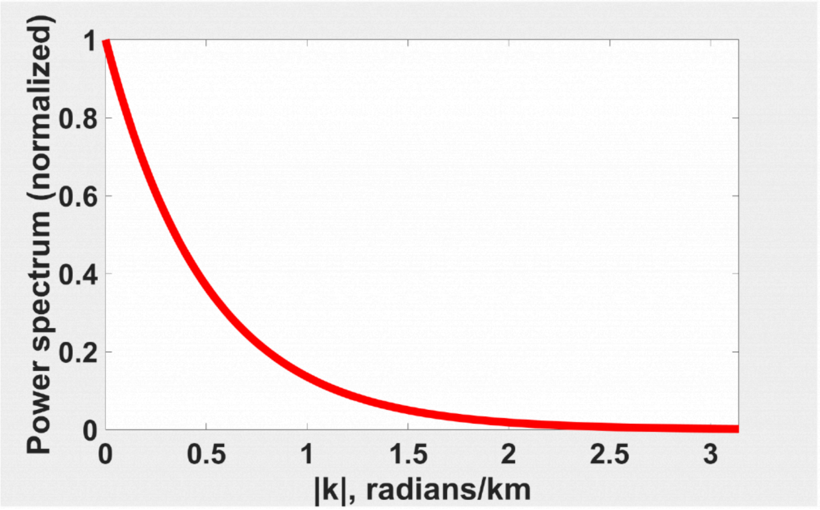
The derivative in z direction has nothing to do in x, y direction, so we can directly take the derivative outside, and obtain the following equation:

$$\mathcal{F}(g_z) = \gamma\mu\frac{\partial}{\partial z}\mathcal{F}(\frac{1}{r}) = 2\pi\gamma\mu e^{|\mathrm{k}|(\mathrm{z_0}-\mathrm{z}^{'})}$$

Thus, we obtain the gravity measure (z component) after Fourier transform:

$$\mathcal{F}(g_z) = 2\pi\gamma\mu e^{|\mathrm{k}|(z_0-z^{'})}$$

Next, we are going to make a few observations about the equation, which will help us better understand the gravity.



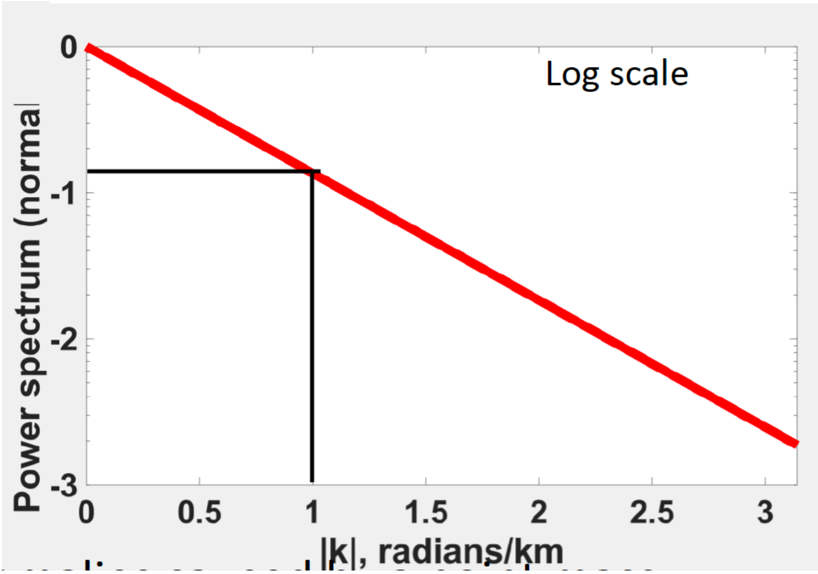
Power spectrum of gravity anomalies caused by a point mass of 1 kg at depth 1 km (assuming observation height 0 km).

The x axis is wave number, y axis is normalized power spectrum of gravity anomalies with a point mass. This figure thus tell us how much energy there is at each wavenumber. The most obvious thing here is exponential decay. Another observation is maximum energy occurs at k=0 (zero wavenumber). When k = 0, we can obtain:

$$\mathcal{F}(g_z) = 2\pi\gamma\mu$$

where, γ is a constant value, μ is mass.

We can easily observe that the power spectrum value at k=0 is proportional to the total mass.



Observation from the power spectrum in log scale, we noticed that it is a straight line with negative slope. Besides, the energy decreases exponentially with increasing wavenumber. The black lines show that the x axis is 1, y axis is -1, so the slope is -1. We knew that the depth is negative slope, so the depth for the anomaly in the figure above is 1 m.

The rate of decreases on log scale is equal to the elevation difference between the source body and observation plane.

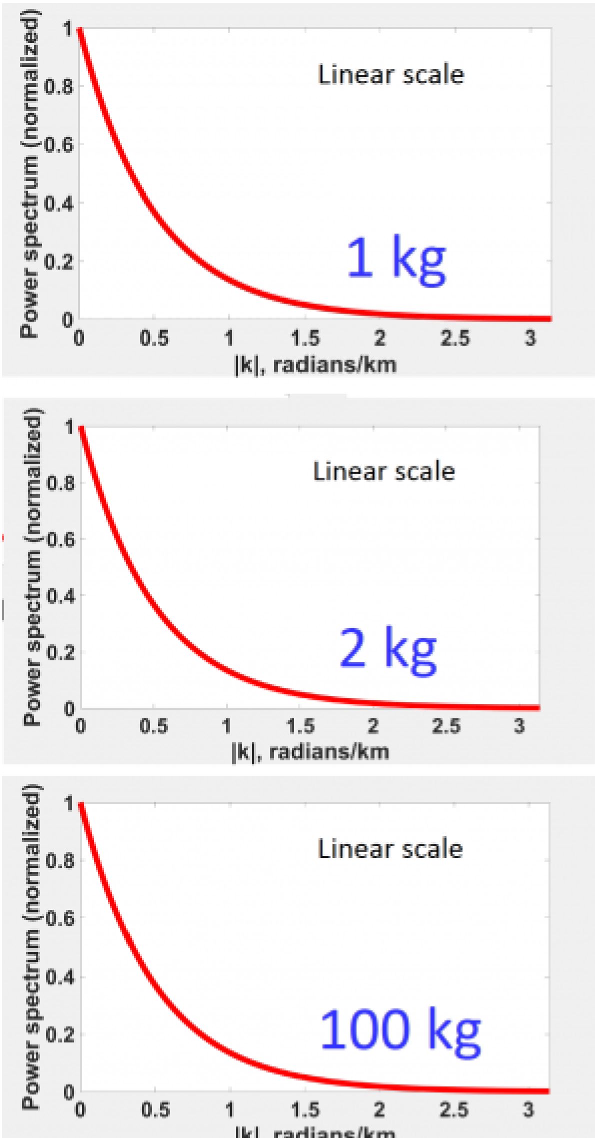
$$\mathcal{F}(g_z) = 2\pi\gamma\mu e^{|\mathrm{k}|(z_0 - z^{'})}$$

$$\log(\mathcal{F}(g_z)) = \log(2) + |k|\,(z_0 - z^{'})$$

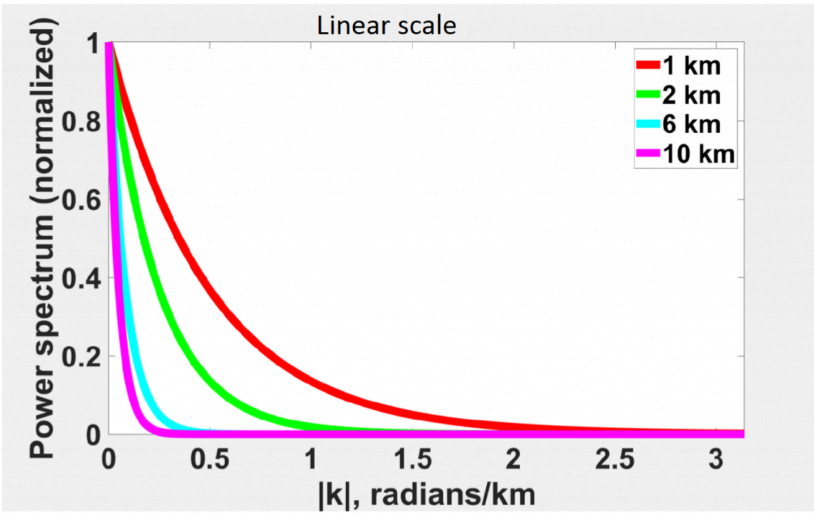
If we set $h = z_0 - z^{'}$,

$$\log(\mathcal{F}(g_z)) = \log(2) + |k|\,h$$

The equation above mathematically illustrates that the rate of decrease depends on the depth of the source body.



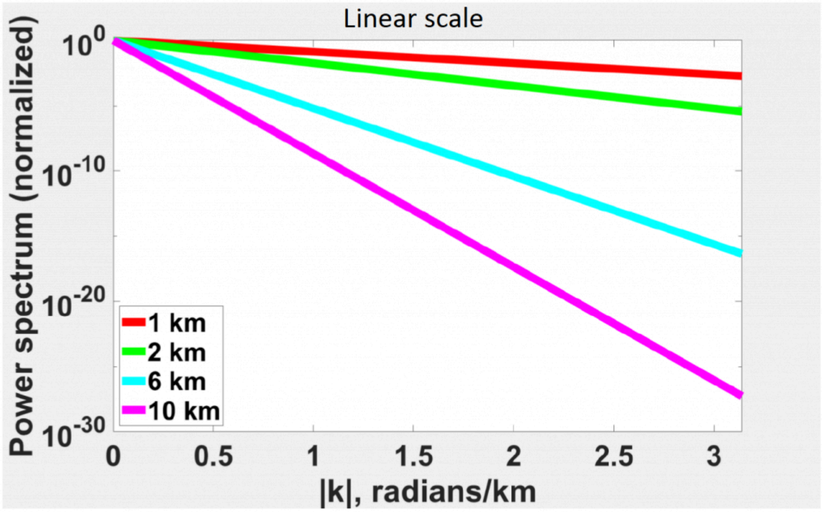
The figures above have same depth but different masses, we noticed the trend is completely same, which means the mass does not change the decay rate and the decay rate has only to do with h.



Power spectrum of gravity anomalies caused by a point mass of 1 kg at depths 1 km, 2km, 6km, and 10km (observation height 0 km).

Here we created a synthetic model with same mass but located in different depth. The red line is shallowest one. We noticed from the figure above is if the source body is deeper, the energy decay faster and faster. Specifically, look at the source body located at 10 km, the energy decay to 0 at 0.2 radians/km. In comparison, if source body located at 2 km (green) line, the energy decays to 0 when wavenumber is about 1.2 radians/km.

Then, we plot 4 cases in log scale, shown in below:



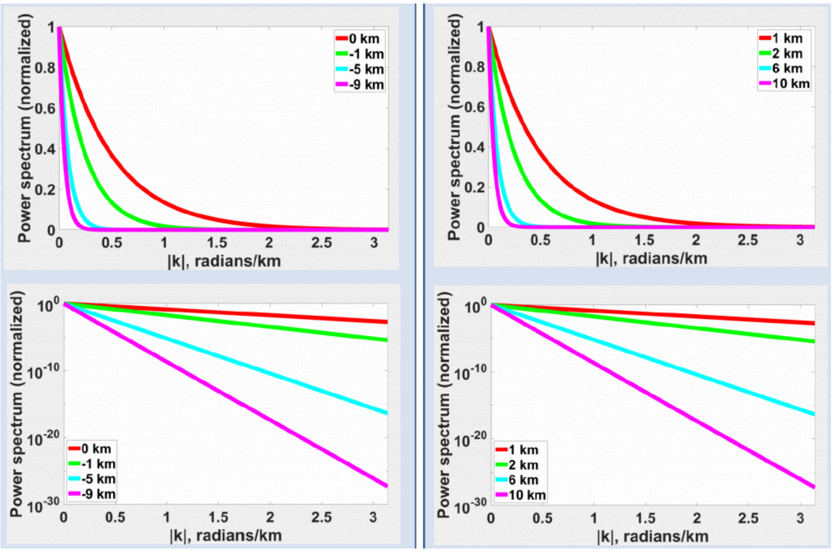
In similar trend, the source body is deeper, the energy decay is faster and faster, where energy decay is represented by the slope. The slope equal to -h, so the slope is smaller, the depth is deeper and the energy decay is faster.

The summaries of two figures above

- The decay rate depends upon depth.
- The deeper the source, the faster the energy decays.
- As source become deeper, energy at higher – wave numbers become smaller.
- The gravity anomaly is approximately band limited.

In previous experiment, we fixed observation at 0 km, and varied the depth of source bodies: 1km, 2km, 6 km and 10 km. If we fix the source body at a depth of 1 km and change the observation height:

0km, -1 km, -5 km, and -9km, we will observe the exact same thing (shown in the figure below).

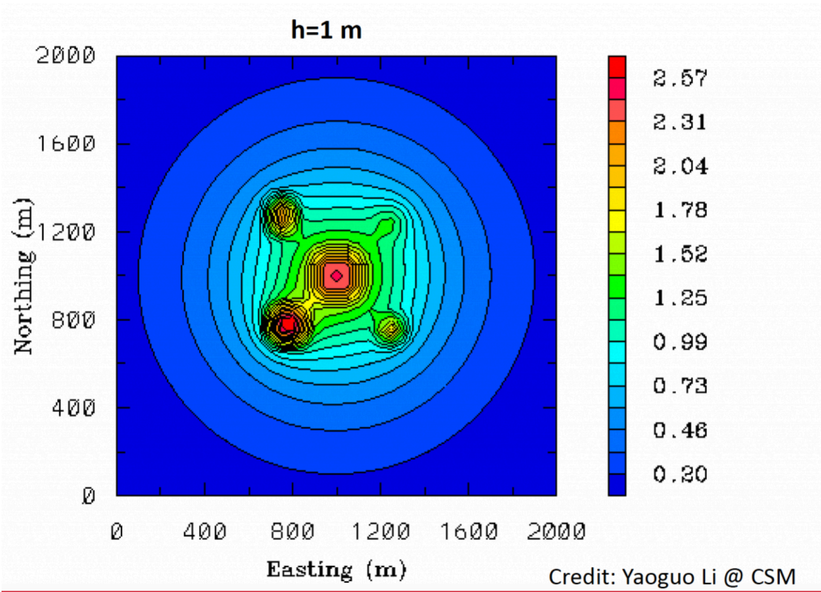


(top left) changing observation heights. (top right) changing source body depth. (bottom left) changing observation heights in log scale. (bottom right) changing source body depth in log scale.

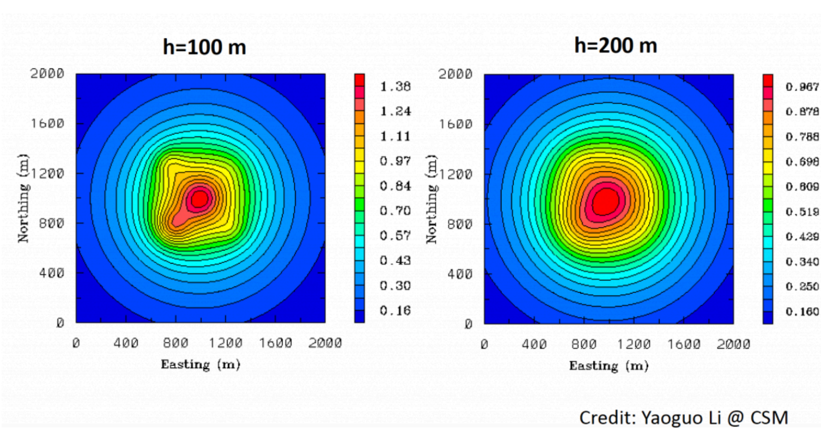
From the figure above (top left and bottom left), we noticed that when source depth fixed, the decay rate depends upon observation height. The higher the observation height, the faster the energy decays. As observation become higher, energy at higher, wave numbers become smaller.

This observation plays a critical role in data acquisition that is: if you want to resolve fine details (i.e., high wave number information) of your targets, you need to be closer to your target. If using the airborne platform, we have to fly as lower as we can.

Here are several examples.



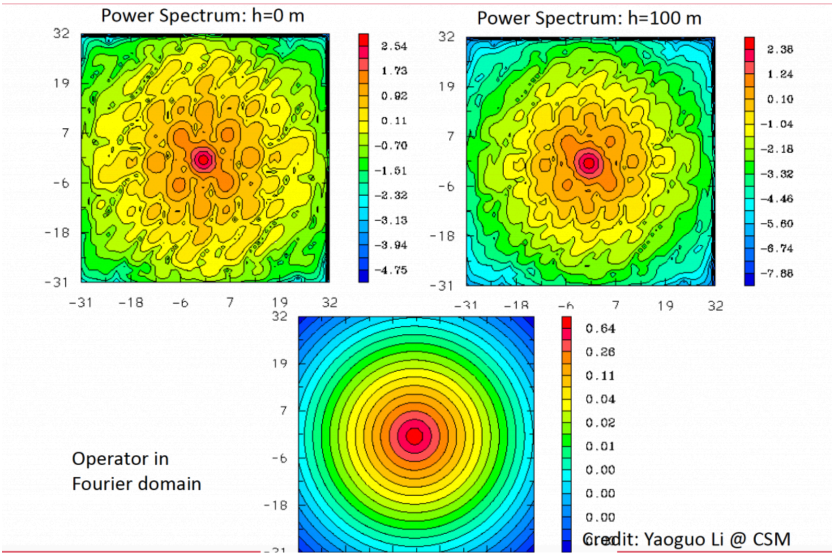
This is one field gravity data map simulated from five source bodies with fixed depth. The data map shown above is 1 meter above the surface. We can easily see the features at data maps.



In the top right figure, we keep source body shape, depth same and the only thing changed is observation height. The observation height is changed from 1 m to 100 m. Compared with previous map, we noticed that anomalies become much smooth. But we can still see a few features of source body. We can still tell there something subsurface. Keeping increasing observation height to 200 m, this map is smooth and it is always explained as one source body rather five.

The observation from the above figures is the height is higher, the gravity anomalies become smoother that means less high frequency information. This phenomenon can be explained as the higher the observation height, the faster the energy decays and wave numbers become smaller.

Another example is:



From the figure above, we noticed that the height of observation is increased, the power spectrum becomes more smooth. The more details about the figures above will be discussed in the following sections.

A quick summary

- We have made quite a few observations based on this simple equation.
- These insights help understand gravity well.
- This is one of the reasons why we look at gravity data in Fourier doamin.

5.5. FOURIER DOMAIN MODELING OF MAGNETIC DATA DUE TO A DIPOLE

Let's move on to the magnetic data for the simplest case of considering a magnetic dipole.

5.5.1. FOURIER TRANSFORM OF MAGNETIC POTENTIAL

For the magnetic scalar potential $V(P)$, which is the dot product of two vectors expressed in the following equation:

$$V(P) = -\frac{\mu_0}{4\pi}\mathbf{m} \cdot \nabla_P \frac{1}{r}$$

Where the dipole moment is expressed as its magnitude and directions in 3D by $\mathbf{m} = m[\hat{m}_x, \hat{m}_y, \hat{m}_z]^T$.

According to the definition of the dot product, the above magnetic scalar potential expression can be rearranged into the following equation:

$$V(P) = -\frac{\mu_0}{4\pi} m \left(\hat{m}_x \frac{\partial}{\partial x}\frac{1}{r} + \hat{m}_y \frac{\partial}{\partial y}\frac{1}{r} + \hat{m}_z \frac{\partial}{\partial z}\frac{1}{r} \right)$$

If we apply the Fourier transform to the magnetic potential in the above equation, through using properties such as the linearity and differentiation, we will then have the derivations as shown below. Since we are taking the Fourier transform in 2D x-y domain, the z-direction differentiation is not relevant here.

$$F[V] = -\frac{\mu_0}{4\pi} m \cdot \left(\hat{m}_x i k_x F\left[\frac{1}{r}\right] + \hat{m}_y i k_y F\left[\frac{1}{r}\right] + \hat{m}_z \frac{\partial}{\partial z} F\left[\frac{1}{r}\right] \right)$$

After above derivation, we are left with familiar expression of $F\left[\frac{1}{r}\right]$, which equals $F\left(\frac{1}{r}\right) = 2\pi\frac{e^{|k|\left(z_0 - z'\right)}}{|k|}$, with z’ > z0.

Therefore, after substituting it into the Fourier transform equation, the simplified expression for the Fourier transform of magnetic potential is as following, consisting of four parts, the constant number $-\frac{\mu_0}{2}$, the magnitude of dipole moment m, the complex number Θ_m, and the exponential expression $e^{|k|\left(z_0 - z'\right)}$ which only depends on the source body depth.

$$F[V] = -\frac{\mu_0}{2} m \Theta_m e^{|k|\left(z_0 - z'\right)}$$

Where the complex number

$$\Theta_m = \hat{m}_z + i\frac{\hat{m}_x k_x + \hat{m}_y k_y}{|k|}$$

depends only on the orientation of the dipole.

5.5.2. FOURIER TRANSFORM OF TOTAL-FIELD ANOMALY

Practically in the field survey, most likely we will collect total-field anomaly ΔT as the magnetic data. Therefore, we can derive the vector B field by taking the negative gradient of the potential V, which can be mathematically expressed as:

$$\mathbf{B} = -\nabla_p V$$

That is, any component (such as $B_x, B_y, B_z, \Delta T$) can be derived from the directional derivative of the magnetic potential.

As we have talked before, the total field anomaly ΔT is simply the projection of the **B**field onto the inducing field direction, which is $\hat{\mathbf{f}} = \left[\hat{f}_x, \hat{f}_y, \hat{f}_z\right]^T$. Therefore, mathematically, the total field anomaly is equal to the dot product of the inducing field direction with the **B**field, as shown below:

$$\Delta T = -\hat{\mathbf{f}} \cdot \nabla_p V = -\hat{f}_x \frac{\partial}{\partial x} V - \hat{f}_y \frac{\partial}{\partial y} V - \hat{f}_z \frac{\partial}{\partial z} V$$

Similarly, after applying the Fourier transform to the total field anomaly, we will get the expression as follows:

$$\mathbf{F}\left[\Delta T\right] = -\hat{f}_x i k_x \mathbf{F}\left[V\right] - \hat{f}_y i k_y \mathbf{F}\left[V\right] - \hat{f}_z \frac{\partial}{\partial z}\mathbf{F}\left[V\right]$$

After simplifying this expression, we will have the Fourier transform as shown below, which consists of four parts: the constant number, the magnitude of dipole moment m (i.e., it depends on the source strength), the two directional dependent complex numbers Θ_m andΘ_f , and the depth-relevant exponential expression $e^{|k|\left(z_0 - z'\right)}$.

$$\mathbf{F}\left[\Delta T\right] = \frac{\mu_0}{2} m \Theta_m \Theta_f \left|k\right| e^{|k|\left(z_0 - z'\right)}$$

Where another complex number is $\Theta_f = \hat{f}_z + i\frac{\hat{f}_x k_x + \hat{f}_y k_y}{|k|}$, and it depends only on inducing field direction;

And the complex number $\Theta_m = \hat{m}_z + i\frac{\hat{m}_x k_x + \hat{m}_y k_y}{|k|}$ depends only on the orientation of the dipole.

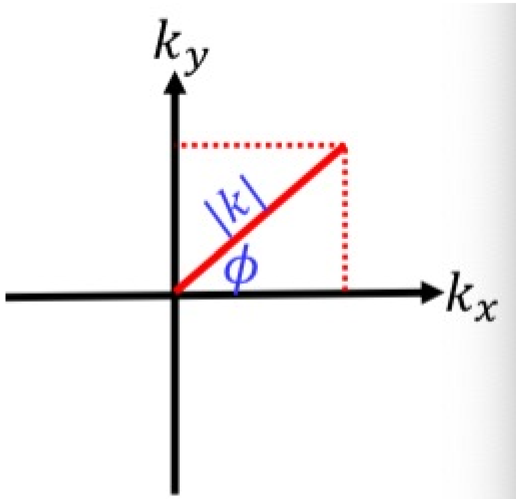
5.5.3. FOURIER TRANSFORM OF TOTAL-FIELD ANOMALY IN POLAR COORDINATES

In order to better understand the Fourier transform of the total field anomaly expression derived above, let's introduce its representation in polar system.

In the formulas of the complex number Θ_m,

$$\Theta_m = \hat{m}_z + i\frac{\hat{m}_x k_x + \hat{m}_y k_y}{|k|}$$

it consists of k_x and k_y, which are wavenumbers in x and y directions. If we express k_x and k_y in polar coordinates, graphically explained in the figure below, we will have:



$k_x = |k| \cos \phi$

$k_y = |k| \sin \phi$

where $|k|$ is the radial wavenumber.

After replacing the k_x and k_y in polar coordinates, the Θ_m can then be written as:

$\Theta_m = \hat{m}_z + i \left(\hat{m}_x \cos \phi + \hat{m}_y \sin \phi \right)$

Another way of representing Θ_m is by using the Euler's formula through amplitude and phase as a function of the angle ϕ, which can be re-written as

$\Theta_m = A_m \left(\phi \right) e^{iP_m(\phi)}$

Similarly, we can follow the same procedure to represent Θ_f, which we will get:

$\Theta_f = A_f \left(\phi \right) e^{iP_f(\phi)}$

Now, let us substitute the polar representations of Θ_m and Θ_f into the Fourier transform expression of the total field anomaly, we will get the following expression:

$\mathbf{F} \left[\Delta T \right] = \frac{\mu_0}{2} m A_m \left(\phi \right) e^{iP_m(\phi)} A_f \left(\phi \right) e^{iP_f(\phi)} \left| k \right| e^{|k| \left(z_0 - z' \right)}$

5.5.3.1. Interpretations of the first part of the amplitude spectrum

If we only focus on the **amplitude spectrum** term in the above expression, which is a function of wavenumber k and angle ϕ as shown below:

$$\mathbf{F}\left[\Delta T\right]=\frac{\mu_0}{2}mA_m\left(\phi\right)A_f\left(\phi\right)\left|k\right|e^{\left|k\right|\left(z_0-z^{'}\right)}$$

The first part of this amplitude spectrum $\frac{\mu_0}{2}mA_m\left(\phi\right)A_f\left(\phi\right)$ only varies with the angle ϕ, since once we fix the strength of the source body the magnitude m will become a constant. Therefore, given a fixed angle ϕ, along any ray originating from the origin, this part of the amplitude spectrum value is constant, that is, there is no decay on the ray since all values on this ray are equal.

Thus, we can express the first part as $C(\phi)$:

$$C\left(\phi\right)=\frac{\mu_0}{2}mA_m\left(\phi\right)A_f\left(\phi\right)$$

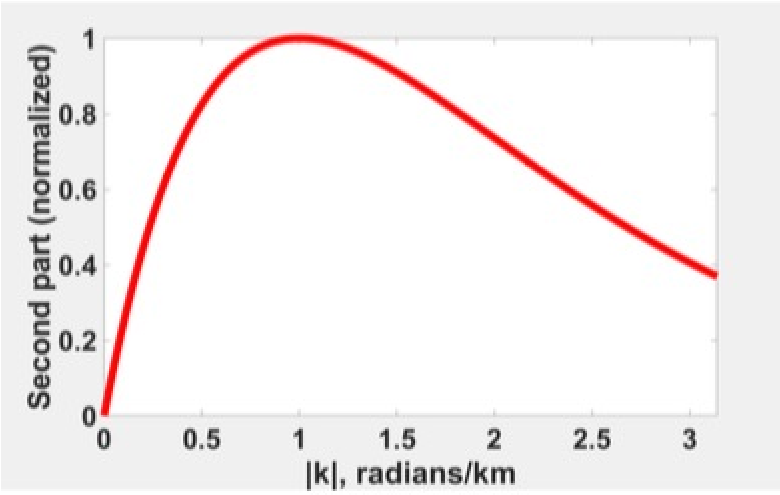
In brief summary so far, the **amplitude spectrum**term can be written as:

$$A\left(k,\phi\right)=C\left(\phi\right)\left|k\right|e^{\left|k\right|\left(z_0-z^{'}\right)}$$

5.5.3.2. Interpretations of the second part of the amplitude spectrum

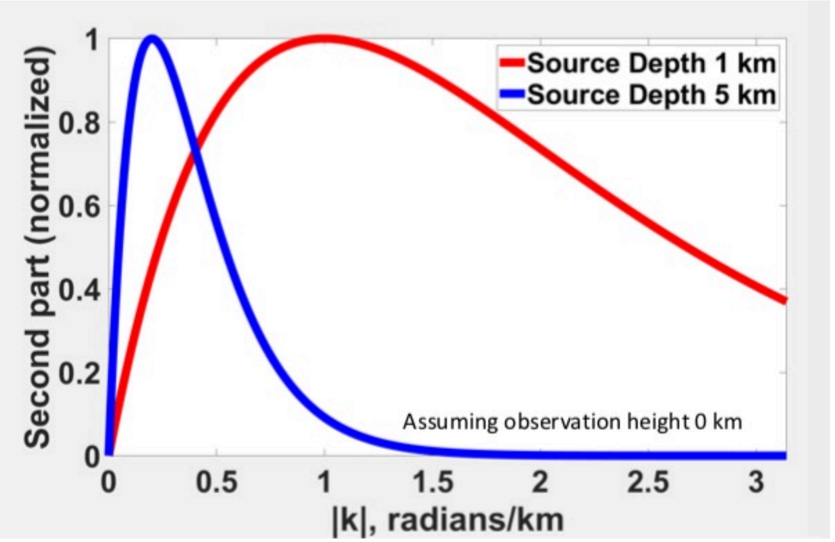
Now let us focus on the second part of the **amplitude spectrum**term $\left|k\right|e^{\left|k\right|\left(z_0-z^{'}\right)}$, which has nothing to do with angle ϕ, but it is related with the observation height z_0 and the depth $z^{'}$, and the radial wavenumber $|k|$. Then how does it look like?

The graph plotted below is what it looks like as a function of the radial wavenumber $|k|$. This figure implies that the energy increases to a peak value along with the increase of the wavenumber before starting decreasing.



Therefore, this second part of the amplitude spectrum determines how the amplitude spectrum changes along a ray and determines the shape of the energy change in that ray.

Moreover, since this part has nothing to do with angle ϕ, the shape of the amplitude spectrum along any ray is the same; and this shape is determined by the depth of the source body $z^{'}$.



Therefore, by looking at the shape of the amplitude spectrum along any ray, we can infer depth of the source body, no matter which ray we are looking at, since they all have the exact same shapes!

In addition, the maximum energy does not occur at wavenumber k=0, but at a larger radial wavenumber that is dependent on the depth of the source dipole; once the wavenumbers exceeding this peak energy wavenumber, energy starts decaying monotonically and approaching exponential decay at higher wavenumbers.

5.6. UPWARD CONTINUATION

5.6.1. CONCEPTS OF UPWARD CONTINUATION

Upward continuation is the process of using originally observed data to calculate the gravity or magnetic anomaly response that would be observed at locations above the original observation surface. Shifting the observation location to a higher elevation removes high-frequency signals from the near surface due to signal attenuation with distance. This allows for the focus of lower-frequency signals from deeper parts of a survey.

5.6.2. UPWARD CONTINUATION IN SPATIAL DOMAIN

Recall Green's third identity:

$$V\left(P\right) = \frac{1}{4\pi}\iint\limits_{Surf}\left(\frac{1}{r}\frac{\partial V}{\partial \hat{\mathbf{n}}} - V\frac{\partial}{\partial \hat{\mathbf{n}}}\frac{1}{r}\right)dS$$

We previously stated that this is the theoretical basis for the **upward continuation** and the **equivalent source technique**. If we consider V to be harmonic, Green's third identity is a **representation formula** which a potential field can be calculated at any point simply from the behavior of the field at its boundaries. This also means that no knowledge of the sources is required, except that none are located within the region. Unfortunately, the equation above requires not only the values of V on the surface, but also the values of its vertical derivative. These values are unlikely to be available in most practical applications. Fortunately, through some mathematical manipulation we can eliminate the derivative from the equation, but it is fairly complex so we won't show that here.

Eventually, we will obtain the following equation:

$$V(x, y, z_0 - \Delta z) = \frac{\Delta z}{2\pi}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\frac{V(x', y', z_0)}{[(x-x')^2+(y-y')^2+\Delta z^2]^{3/2}}dx'dy'$$

$\Delta z > 0$ 0" title="Rendered by QuickLaTeX.com" height="12" width="56" style="vertical-align: 0px;">

This equation is the upward-continuation integral. It shows how to calculate the value of a potential field at any point above a level, horizontal surface at z_0 from a complete knowledge of the field on the surface. This means upward continuation can be done in the spatial domain using this integral.

At this point, we will need to cover **convolution**. The **convolution** of two functions $f(x)$ and $g(x)$ is:

$$h(x) = \int_{-\infty}^{\infty} f(x')g(x - x')dx'$$

We can also convert this equation to the Fourier domain if we assume that $h(x) \leftrightarrow H(k)$, $f(x) \leftrightarrow F(k)$ and $g(x) \leftrightarrow G(k)$:

$$H(k) = F(k)G(k)$$

We will also need to examine **convolution** in 2D for **upward continuation**. In 2D, the **convolution** of two functions f(x) and g(x) in the spatial domain becomes:

$$h(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y')g(x - x', y - y')dx'dy'$$

If we assume that $h(x, y) \leftrightarrow H(k_x, k_y)$, $f(x, y) \leftrightarrow F(k_x, k_y)$ and $g(x, y) \leftrightarrow G(k_x, k_y)$, the Fourier domain equation is:

$$H(k_x, k_y) = F(k_x, k_y)G(k_x, k_y)$$

This is important as we go back to our previously established upward continuation equation:

$$V(x, y, z_0 - \Delta z) = \frac{\Delta z}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{V(x', y', z_0)}{[(x - x')^2 + (y - y')^2 + \Delta z^2]^{3/2}} dx' dy'$$

$\Delta z > 0$ 0" title="Rendered by QuickLaTeX.com" height="12" width="56" style="vertical-align: 0px;">

We can define

$$\psi(x, y, \Delta z) = \frac{\Delta z}{2\pi} \frac{1}{[x^2 + y^2 + \Delta z^2]^{3/2}}$$

to obtain the following equation:

$$V(x, y, z_0 - \Delta z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} V(x', y', z_0) \psi(x - x', y - y', \Delta z) dx' dy'$$

This equation is 2D convolution!

5.6.3. UPWARD CONTINUATION IN FOURIER DOMAIN

We can find the Fourier representation of upward continuation using the following equation:

$$F[V_U] = F[V] F[\psi]$$

All we need now is an analytical expression of F[\psi], and we will skip the derivations here to get the following equation:

$$F[\psi] = e^{-\Delta z |k|}$$

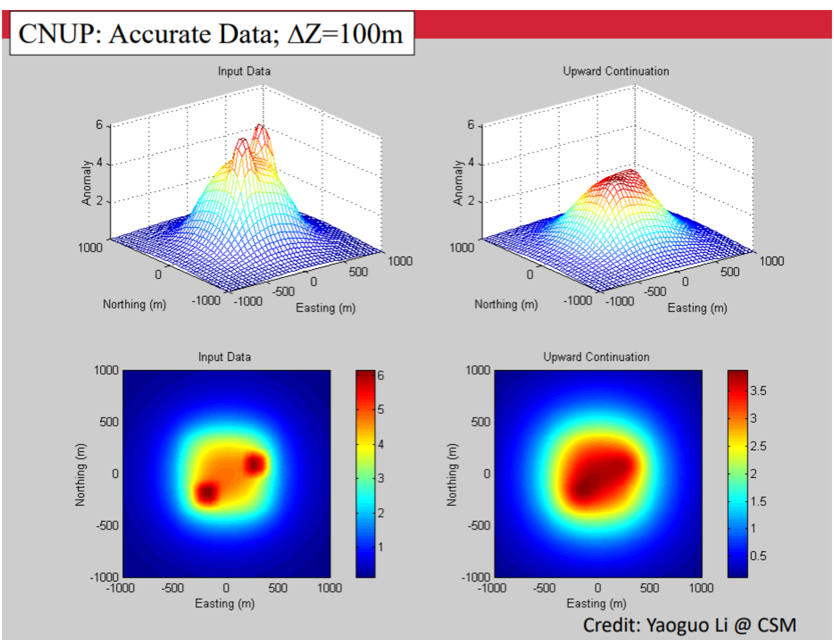
$\Delta z > 0$ 0" title="Rendered by QuickLaTeX.com" height="12" width="56" style="vertical-align: 0px;">

At this point, we can find the Fourier transform of the upward-continued field in three steps:

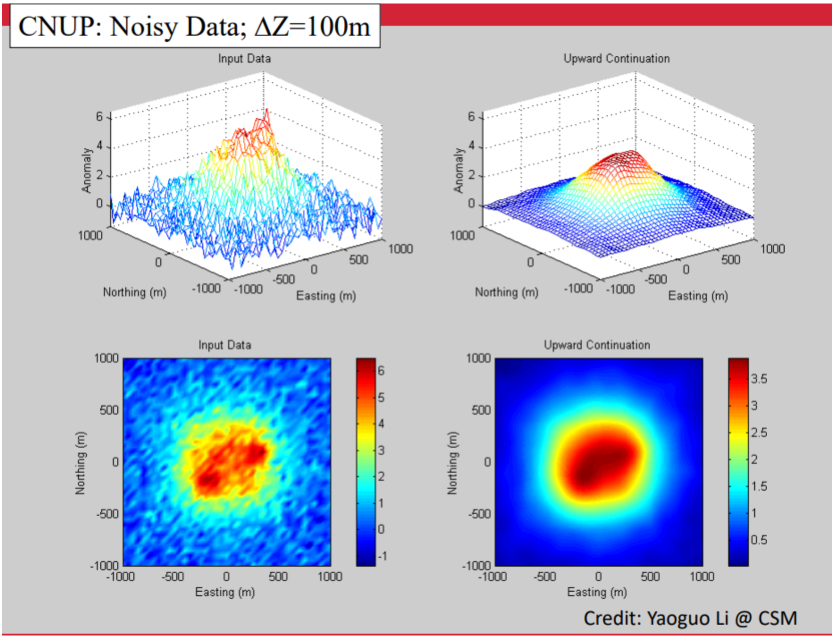
1. Transform the observed data via FFT (Fast Fourier Transform)
2. Multiply by the exponential operator ($e^{-h\sqrt{w_x^2 + w_y^2}}$)
3. Apply the inverse transform to obtain the upward-continued field

5.6.4. EXAMPLE

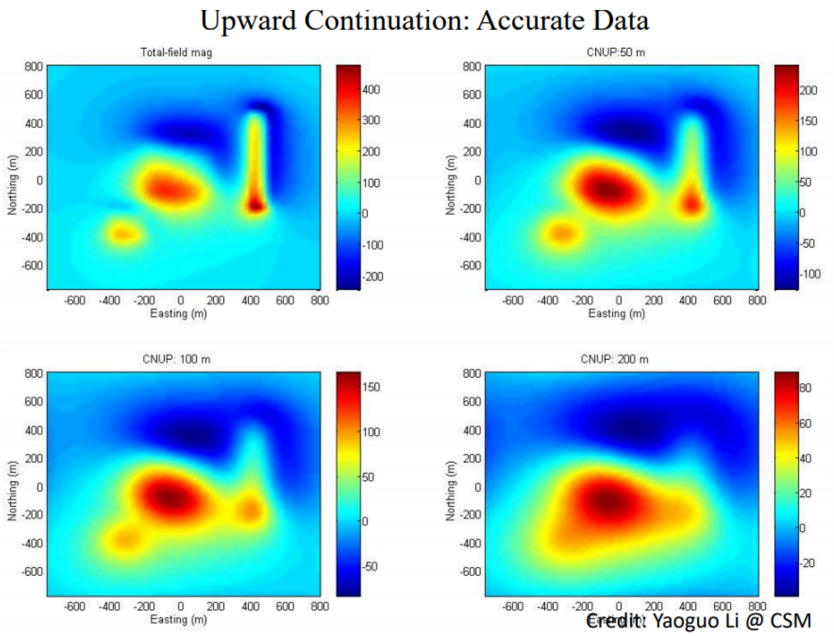
In the example below, we will be showing the influence that upward continuation has on clean and noisy data. The figure below shows the difference between the shape of the input data in 2D and 3D between its original elevation and its upward continuation.



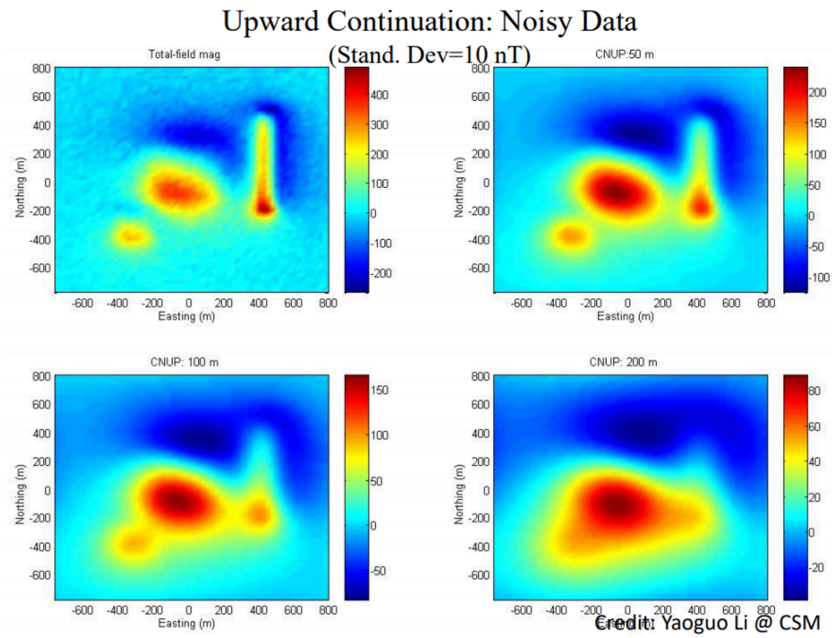
The following figure shows the same figures after noise has been introduced to the data. It is clear that the noise has a significant influence on the input data at its original data, but has been largely reduced by the upward continuation.



The figure below shows the magnetic response of the input data at various elevations. Pay attention to the smoothing of the overall feature along with the reduction in high-frequency data as the elevation increases.



The same figure was reproduced using noisy data. Notice the significant difference between the accurate and noisy data at the lower elevations, and the reduction in noise as the responses are recorded at higher elevations of upward continuation.



5.7. REDUCTION TO POLE

5.7.1. FUNDAMENTALS OF REDUCTION TO POLE

Reduction to Pole (RTP) is the process of adjusting recorded magnetic data so that it appears as if it was recorded if the Earth's inducing magnetic field was vertical. The figure below shows this adjustment, and note the difference in terms of symmetry between the unadjusted and adjusted data. It is important to note that RTP is difficult to do at low magnetic inclinations (next to the equator).

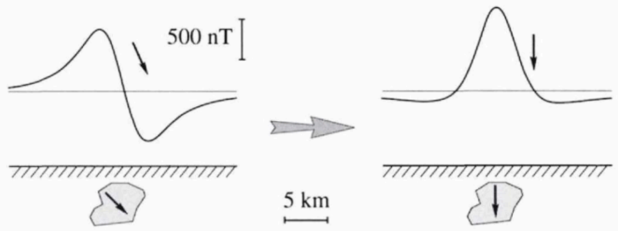


Fig. 12.7. A magnetic anomaly before and after being reduced to the pole.

Blakely, 1996, p330

The equation for RTP in the Fourier domain can be listed as the following:

$$F[\Delta T_r] = F[\Delta T]F[\psi_r]$$

$$F[\psi] = \frac{1}{\Theta_m \Theta_f} = \frac{|k|^2}{(\vec{R} \cdot \hat{f})(\vec{R} \cdot \hat{m})}$$

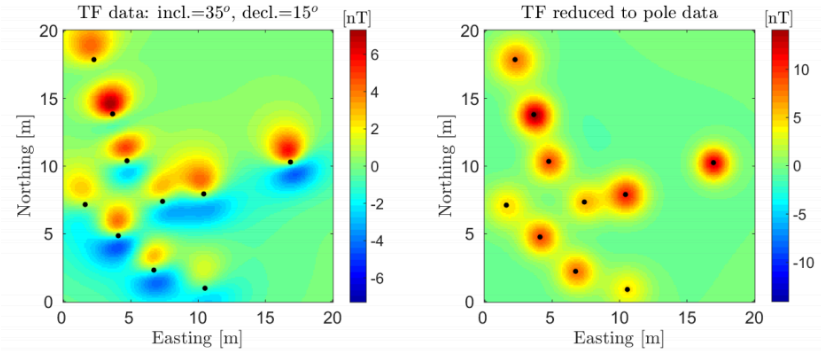
$$\vec{R} = (ik_x, ik_y, |k|)$$

where $\hat{f}$ is the inducing field direction and $\hat{m}$ is the magnetization direction

We can find the Fourier transform of RTP using the following three steps:

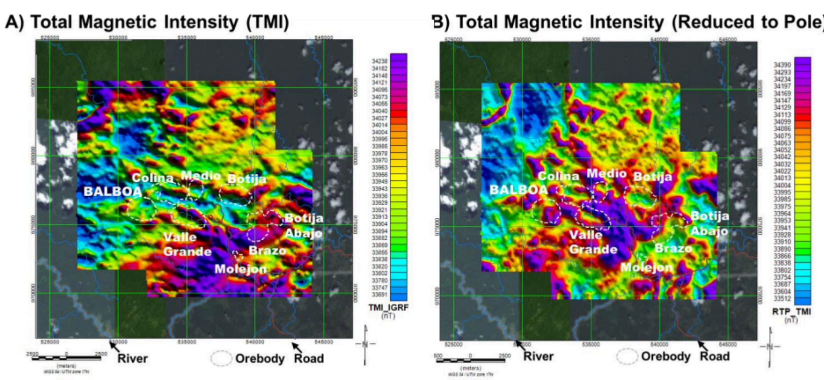
1. Transform the observed data via FFT (Fast Fourier Transform)
2. Multiply by $\frac{|k|^2}{(\vec{R} \cdot \hat{f})(\vec{R} \cdot \hat{m})}$
3. Apply the inverse transform to obtain RTP

5.7.2. EXAMPLES OF REDUCTION TO POLE



Total field magnetic data over UXO before (left) reduction to pole and after (right) reduction to pole. The true location of each object is labeled by a black dot.

https://gpg.geosci.xyz/content/magnetics/magnetics_processing.html



Magnetic data from a ZTEM survey over Cobre Panama deposit area: a) Raw total magnetic intensity (TMI), and b) TMI data after reduced to pole. Overall, all the known porphyry deposits are noticeably centered on magnetic lows, possibly representing demagnetized areas due to phyllic alteration. The only exception is Balboa that coincides with a positive magnetic anomaly.

https://em.geosci.xyz/content/case_histories/balboa/processing.html

5.8. ISSUES WITH FOURIER MODELLING

Throughout this chapter, we have gone through different methods in Fourier domain modelling using FFT (Fast Fourier Transform). However, FFT-based processing methods have two major **prerequisites** on data:

1. **The observation surface needs to be planar.**
2. **Data needs to be interpolated to uniform** grid.

These two prerequisites/assumptions is not made in practice because:

1. **Data are often acquired on undulating surface**. For example, in an airborne survey with a helicopter, the flight height will be made constant to 50 m above the underlying terrain. However, the underlying terrain has a lot of hills and troughs. As such, the data measured by the helicopter is not from a planar surface because it follows the topography.
2. **Majority of data are acquired along flight lines or scattered stations**. What this means is that the survey grid is not uniform in practice. For example, suppose an airborne survey using a helicopter. The helicopter will fly in one direction and obtain some measurements. After going through one survey line, it will move to the next survey line, which is around 100m away from the first survey line, and then it measures the data along that survey line. Now, the problem with this is that the data measured along the survey line (let's say y direction, measurement every 5m) is much finer than data measured adjacent to the survey lines (let's say x direction, measurement every 100m). Thus, you can have a grid of 5m x 100m. Clearly this grid is not uniform, and

thus FFT will not work with this survey grid. One way to solve this problem is to use interpolation method but this has other problems as well

3. **The process of interpolation is a heavy processing step.** Interpolation can be a dangerous process because it can create artifacts that does not exist in the data. Processed data maps would have many artifacts that would lead to misinterpretation.

In the next chapter, we will be going through a different method that does not produce artifacts in the data. This method is called the **equivalent source method.**

CHAPTER 6

Chapter 6: Interpretation methods

6.1. REVIEW OF GREEN'S FIRST & THIRD IDENTITY

Before going through equivalent source technique, it is best to review our understanding of Green's theorem to understand the key concepts to this method.

6.1.1. GREEN'S FIRST IDENTITY

Recall Green's first identity, which is:

$$\int\int\int_{Vol}(V\nabla^2 U+\nabla V\dot{\nabla}U)dv=\int\int_{Surf}V\frac{\partial U}{\partial \hat{n}}ds$$

Where **U, V** are continuous functions with continuous first order partial derivative. **U** also has a second order derivative. An arbitrary vector $A=V\nabla U$ is also defined.

Now, assume that: U is harmonic and U=V, then:

$$\int\int\int_{Vol} (\nabla U)^2 dv = \int\int_{Surf} U \frac{\partial U}{\partial \hat{n}} ds$$

Let's consider the above equation when U=0 on **S.** If that's the case, then the right hand side of the equation vanishes.

Because the $(\nabla U)^2$ is always positive and continuous, then, that means:

$(\nabla U)^2 = 0$

This implies that U is constant. We have already assumed that U=0 on the surface. Because U is continuous, then that constant must be zero.

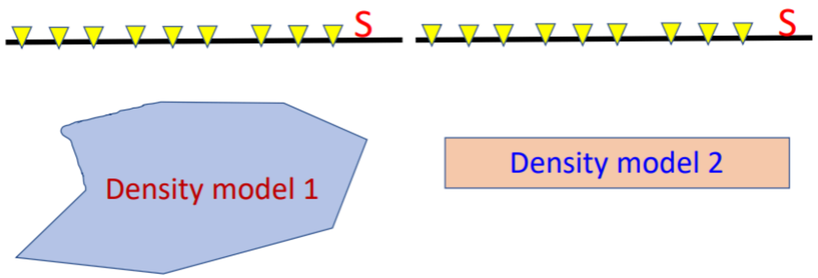
Therefore, if U is harmonic, U is continuously differentiable in R, and if U vanishes everywhere on the surface S, then **U must also vanish everywhere within the volumes.** Note that this only applies if the highlighted assumptions are true.

Now, consider two functions, U_1 and U_2, be harmonic and have identical boundary conditions ($U_1(S) = U_2(S)$). Thus, the function $U_1(S) - U_2(S)$ must be harmonic. Because $U_1 - U_2$ vanishes on S, then $U_1 - U_2$ must vanish at every point. Therefore, U_1 is identical to U_2. This implies that **a function that is harmonic and continuously differentiable in R is uniquely determined by its value on surface S.** Note that this only hold true if the highlighted assumptions are true.

What does it mean in potential field? Well, **gravity and magnetic scalar potentials** are harmonic and continously differentiable. Therefore, gravity and magnetic scalar potentials are uniquely determined by their values on S. As such: **gravity and magnetic data are uniquely determined by their values on surface S.**

Let's use the following illustration to understand this concept.

Suppose that there are two density distributions as such:



According to Green's first identity, if the measurements on S are the same, then the gravity data from these two density models are the same. The two sources are called **equivalent sources.**

6.1.2. GREEN'S THIRD IDENTITY

Recall Green's third identity. Considering the situation when V is harmonic:

$$V(P) = \frac{1}{4\pi}\int\int_{Surf}(\frac{1}{r}\frac{\partial V}{\partial \hat{n}} - V\frac{\partial}{\partial \hat{n}}\frac{1}{r})ds$$

This implies that a harmonic function can be calculated at any point simply from its value and derivatives on the boundary **(representation formula).** Recall that this is used as a theoretical basis for upward continuation and equivalent source technique.

Because gravity/magnetic potential is harmonic, thus a potential field can be calculated at any point simply from the behavior of the field on the boundary. In other words, no knowledge about sources is required, except that none may be located within the region.

Unfortunately, Green's third identity requires not only the values of V on the surface, but also the values of vertical derivative of ($\frac{\partial V}{\partial \hat{n}}$). This is unlikely to be available in most practical applications. Fortunately, by some mathematical manipulations, we can eliminate the derivative from the equation.

Through some mathematical manipulations, Green's third identity can be expressed as:

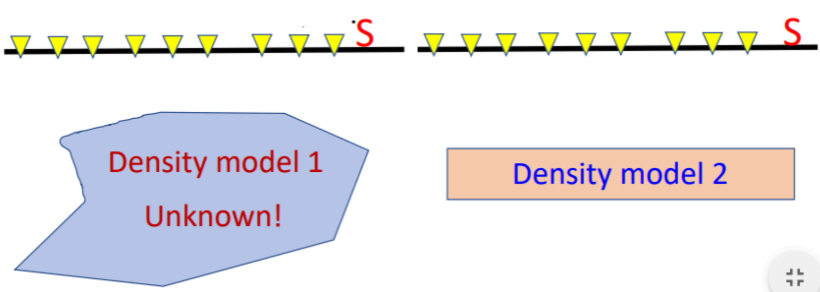
$$V(x, y, z_0 - \Delta z) = \frac{\Delta z}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{V(x', y', z_0)}{[(x - x')^2 + (y - y')^2 + \Delta z^2]^{3/2}} dx'dy'$$

$$\Delta z > 0$$

Blakely, 1996

The above equation is called the **upward-continuation integral**. It shows how to calculate the value of a potential field at any point above a level, horizontal surface at z_0 from a complete knowledge of the field on the surface. The function also still implies that the potential field in R is determined by its values on S. Thus, upward continuation can be done in spatial domain using the above integral.

Let's use the following example to understand this process.



Suppose that we do not know anything about density model 1. The anomaly from the source body is measured on S. Then, we can construct an equivalent source that reproduces the measurement on S using a known source body parameter, let's say a rectangular density model pictured in Density model 2. The equivalent source (density model 2) will have zero resemblace to the true source body (density model 1). But this is still applicable in processing (to do UC, RTP,, component conversion, etc.), as long as it produces the same measurements on S. We can use equivalent source to calculate potential field values at any other locations in R.

Since we do not care what the equivalent source looks like as long as it produces the same measurements on S, we typically do not bother constructing a 3D equivalent source. Thus, we typically construct an equivalent source layer (i.e. a thin sheet of density) as the equivalent source.

There are several advantages to Equivalent source layer (ESL):

1. Free of the instabilities with RTP at low latitudes
2. Can do uneven-to-uneven surface continuation
3. Can use ESL to do RTP (Reduction-To-Pole), UC (Upward Continuation), component conversion, as we can work with ESL in Fourier domain.

The caveat is that to construct an equivalent source layer, we need to do an **inversion**.

6.2. EQUIVALENT SOURCE

Equivalent source is a very powerful image processing technique in processing and interpreting the potential field data. Almost all the data processing techniques we have talked about so far in Fourier domain can be accomplished by using equivalent source technique.

6.2.1. INVERSION

After reviewing the Green's first and third identity, we have built a solid theoretical understanding of the equivalent source technique now. However, in order to construct an equivalent source layer, we do need to do an inversion, which can be computationally expensive. Therefore, let us quickly go through the inversion firstly before we move on to the examples of applying the equivalent source technique.

Assuming that we have collected N number of discretized data samples on the surface of the Earth, which are denoted as a vector array as shown below.

$$\vec{d} = (d_1, \cdots, d_N)^T$$

Then we need equivalent source that can be represented by point or piece-wise constant values (such as density or magnetization parameters) that can reproduce the collected data. The equivalent source can be mathematically denoted by a vector as follows:

$$\vec{m} = (m_1, \cdots, m_M)^T$$

These two elements, the data measurements on the boundary and the equivalent source, are linked by a kernel matrix, $\mathbf{G}$:

$$\vec{d} = \mathbf{G}\vec{m}$$

That is, given the measured data $\vec{d}$, we want to construct $\vec{m}$. Unfornately, we need to store the dense matrix $\mathbf{G}$ that requires large amount of memory and CPU time. For example, a 128 by 128 grid would require up to 1 Gb of memory to store this $\mathbf{G}$ matrix!

$$\begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1M} \\ g_{21} & g_{22} & & \\ \vdots & & \ddots & \\ g_{N1} & & & g_{NM} \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_M \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix}$$

Moreover, the construction of $\vec{m}$ is often an ill-posed problem (we will discuss it in more details later). One way of mitigating this ill-posed problem is by the Tikhonov regularization method. That is, we formulate a cost function ϕ that consists of data misfit function ϕ_d, and the model objective function ϕ_m, these two terms are linked by a regularization parameter μ. Through minimizing the cost functional value ϕ, it is ensured that the reconstructed equivalent source parameters $\vec{m}$ can reproduce the measured data.

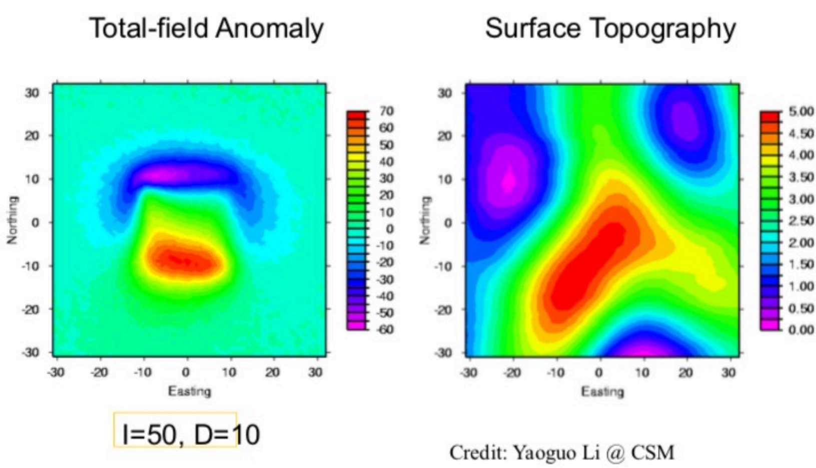
$$\phi = \phi_d + \mu\phi_m$$

6.2.2. NUMERICAL EXAMPLES

6.2.2.1. Upward continuation

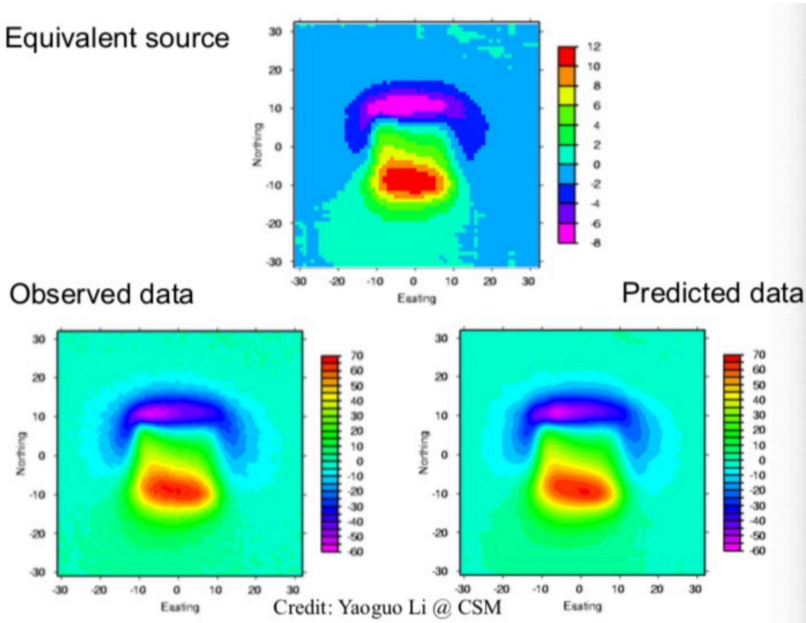
Here is a **synthetic example**of using equivalent source technique to do upward continuation.

The figure on the left is a synthetic total-field anomaly map that shows typical magnetic dipolar data pattern. This magnetic data is simulated by assuming the inclination equals 50 degrees, declination equals 10 degrees. The surface topography map displayed on the right can be interpreted as the data observation heights; that is, the observation data are not collected at an even surface.



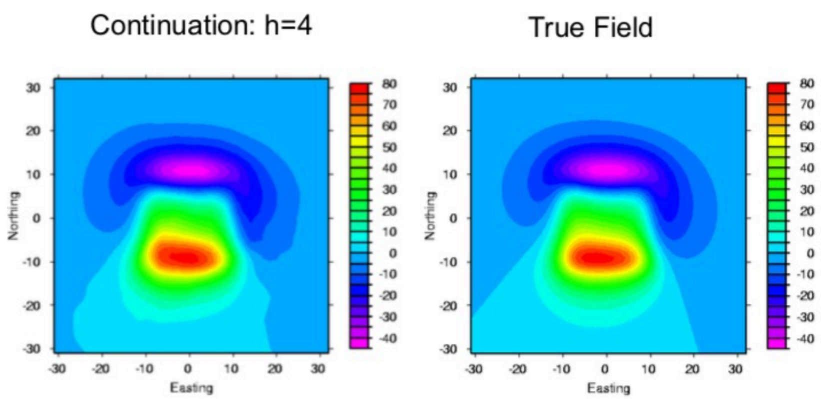
Synthetic magnetic total-field anomaly data and its surface topography. Image courtesy of Yaoguo Li @ CSM

The reconstructed equivalent source layer (top plot) is displayed below. The reproduced data (bottom right plot) are simulated based on this equivalent source, and it is compared with the observed data map (bottom left plot) on the surface of the Earth.



Constructed equivalent source layer and its corresponding data. Image courtesy of Yaoguo Li @ CSM

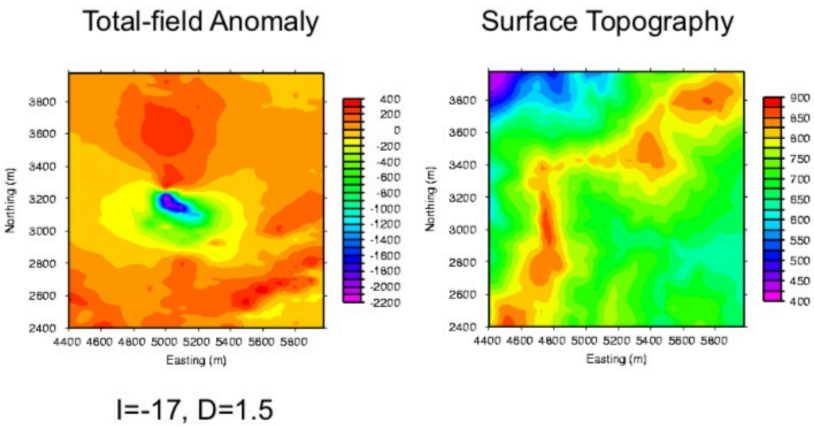
Since this recovered equivalent source can reproduce the measured data on the surface, therefore we will take this equivalent source layer to do upward continuation, as if it is the true source. The figure shown below is the upward continuation at 4 meters height (left plot), and it is compared with the true field measurement (right plot). These two plots look almost identical. Thus, this example proves that the equivalent source technique does work well.



Upward continuation in comparison with the true field. Image courtesy of Yaoguo Li @ CSM

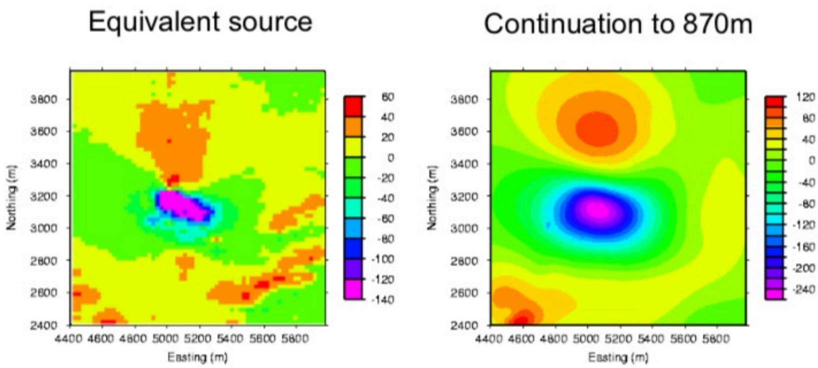
Here is another **field example** of using equivalent source technique to do upward continuation.

The magnetic map on the left figure below is collected from a low magnetic area with inclination of -17 degrees, declination of 1.5 degrees. The surface topography on the right map implies a very rough measurement height surface, with up to 500 meters variations. Since the data are not collected in a constant height, therefore we need to do upward continuation, and that requires an equivalent source layer.



Field measured total-field anomaly data, and the surface topography. Image courtesy of Yaoguo Li @ CSM

After constructing and minimizing the objective functional value, given the measured total-field anomaly data, the reconstructed equivalent source layer is shown on the figure below (left plot). Then this equivalent source layer is used to do a forward modeling at the height of 870 meters above the Earth surface, the upward continued data simulated from this equivalent source is shown on right plot below.

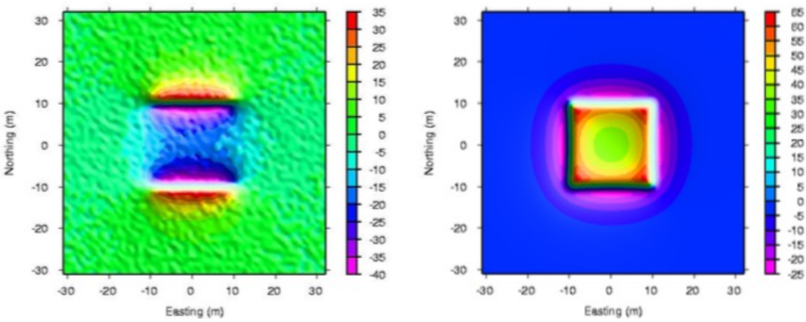


Constructed equivalent source layer, and the upward-continued data based upon it. Image courtesy of Yaoguo Li @ CSM

6.2.2.2. Stable RTP

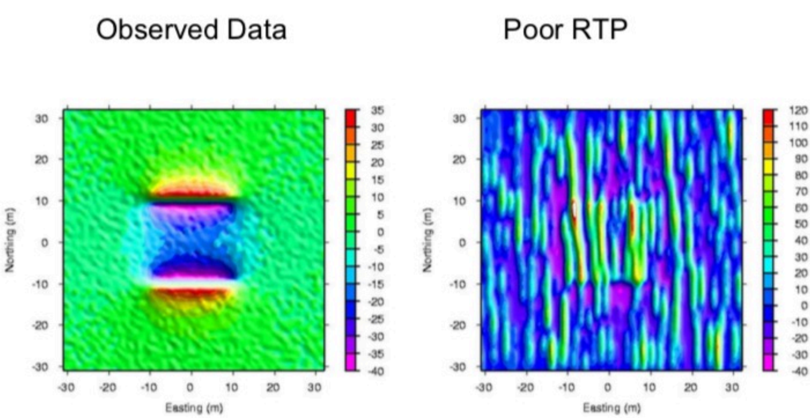
Equivalent source technique can also be used to do stable RTP even at low magnetic latitude.

Here is a **synthetic example**. The left figure is the synthetically simulated TMI total-field anomaly data at a low magnetic latitude area; the data shows a typical positive-negative-positive pattern for magnetic data collected at low latitude region. The RTP in Fourier domain will be very challenging. Since this is a synthetic modeling, so we can use the true source to calculate the true RTP, which is displayed in the right plot below.



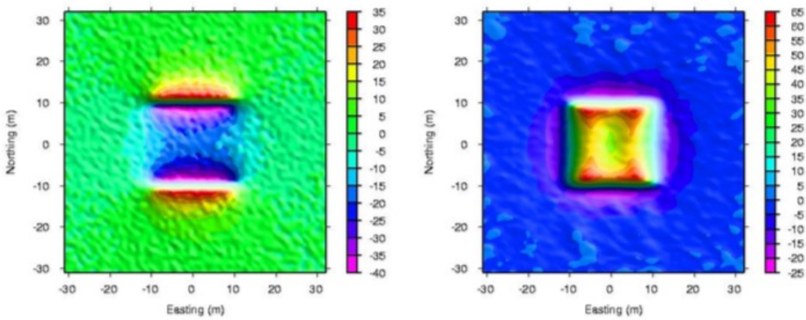
Synthetic TMI total-field anomaly data, and the true RTP. Image courtesy of Yaoguo Li @ CSM

If the noise is added into the clean total-field anomaly data (left plot below), then do the RTP by using the Fourier domain technique we talked about before, the resulted poor RTP figure shown on the right plot below has a lot of elongated strips artifacts along the direction of declination (0 degree in this example).



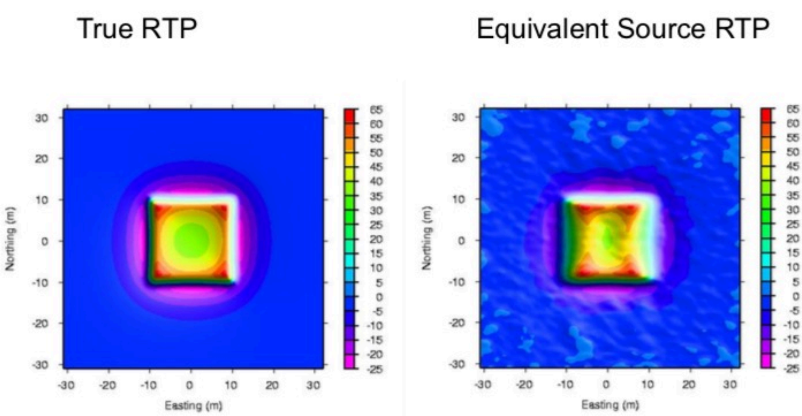
Observed data, and the computed poor RTP. Image courtesy of Yaoguo Li @ CSM

However, after using the magnetic data map (left plot below) to construct the equivalent source layer, then using the equivalent source layer to do the RTP (right plot below), the RTP quality is greatly improved.



Observed data, with regularized RTP with positivity. Image courtesy of Yaoguo Li @ CSM

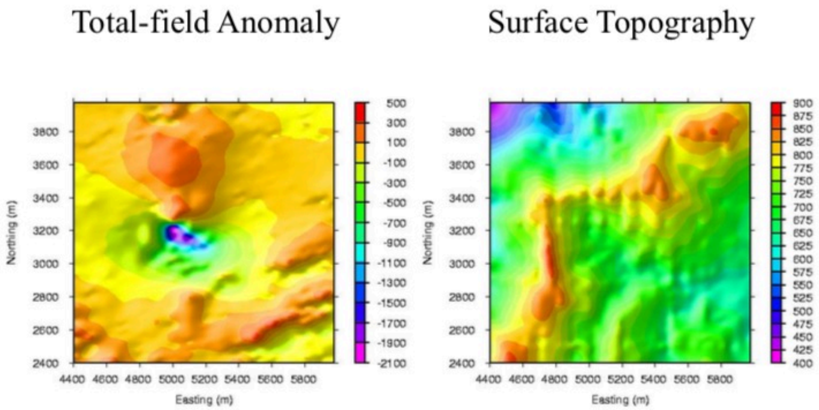
The equivalent source RTP is very similar with the true RTP, as shown in the comparison figure below, which is very interpretable.



Comparison of the true RTP with the Equivalent source RTP. Image courtesy of Yaoguo Li @ CSM

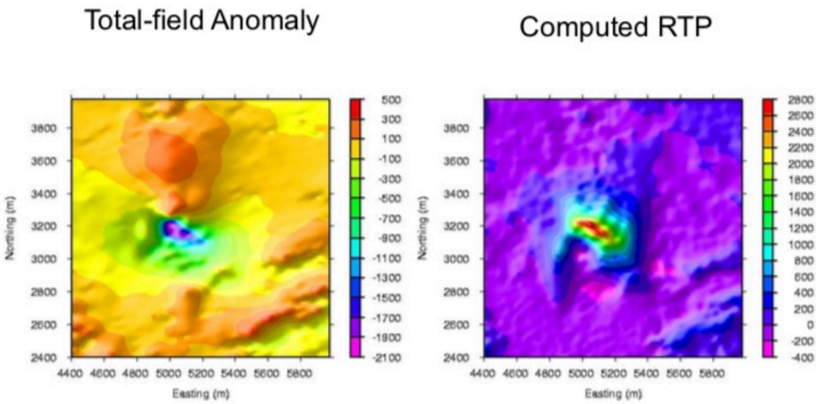
Another is a **field example**of using equivalent source technique to compute RTP.

The measure total-field anomaly at a low magnetic latitude is shown on the left figure below, and its data collection surface topography (right plot) shows a 500 meters topography relief.



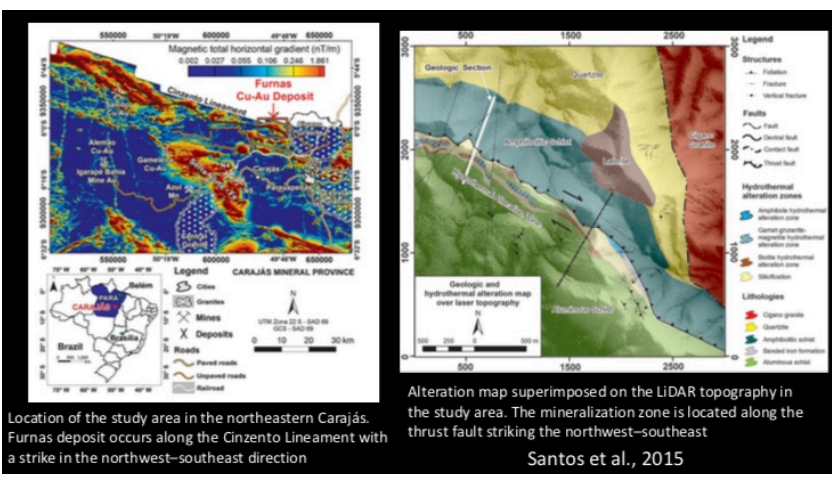
Measured data map, and the surface topography. Image courtesy of Yaoguo Li @ CSM

The computed RTP by using the equivalent source technique (right plot below) is very interpretable, since its peak anomaly lies on top of the anomaly body, and there are no elongated strips artifacts along the declination direction.



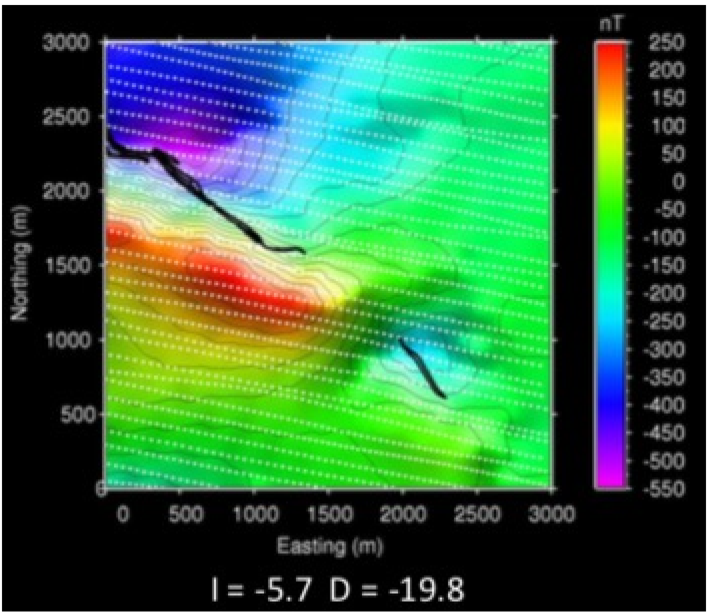
Measured data, and the computed Equivalent source RTP. Image courtesy of Yaoguo Li @ CSM

Here is another **field example** of using equivalent source technique to compute RTP at low latitude. The study area is in the northeastern Carajas, the interested mineralization zone is located along the major fault striking along the NW-SE direction. The simplified geologic unit map are shown in the figure below.



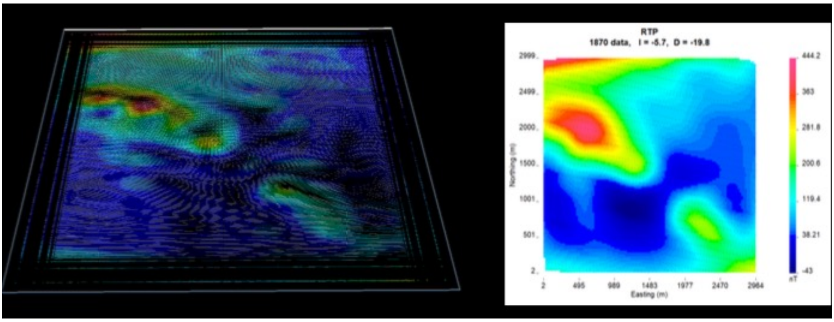
Study area location and the simplified geologic unit map. (Santos et al., 2015)

The collected magnetic data map was displayed in the figure below. Although it shows a typical positive-negative magnetic data pattern, it is very challenging to deal with due to various conditions. Firstly, the data are from very low latitude of -5.7 degrees; Secondly, the data has strong remanence; moreover, it is self-demag and affected by anisotropy.

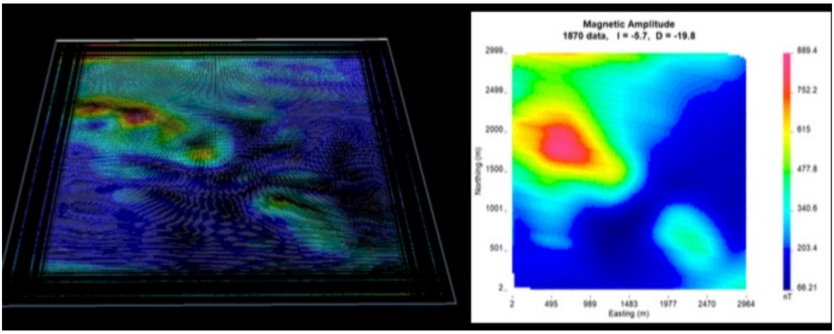


The measured magnetic data. (Leao-Santos et al., 2015)

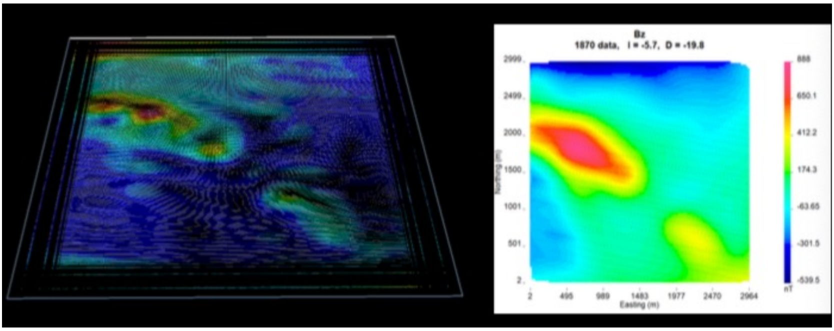
However, after applying the equivalent source technique, the data is much easier to work with. The equivalent source layer was constructed, and then it can be used to compute the **RTP**, the magnetic **amplitude**, the magnetic **component** Bx, By, Bz, or the **upward continuation**, since this equivalent source layer is able to reproduce the observed data on the surface of the Earth.



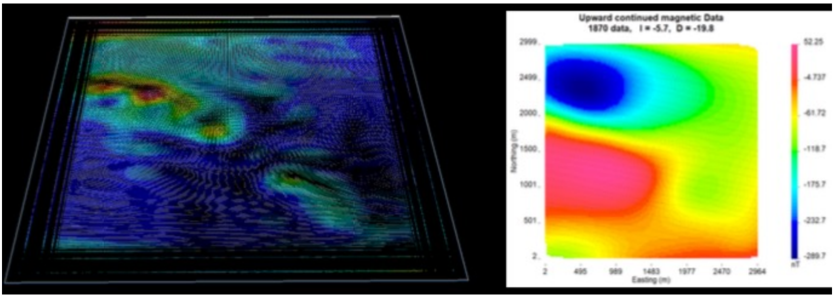
Constructed equivalent source layer, and the RTP.



Constructed equivalent source layer, and the magnetic amplitude.



Constructed equivalent source layer, and the Bz component.



Constructed equivalent source layer, and the upward continuation at 800 m.

6.3. DEPTH ESTIMATES

6.3.1. HALF WIDTH METHOD

Shown in the figure below, assuming there is a small ore body in the subsurface. It is the spherical shape with a radius of 10 m, located at a depth of 25m and the density contrast is 0.5 gm/cc. We are going to look at the gravity of the uniform sphere at different depth.

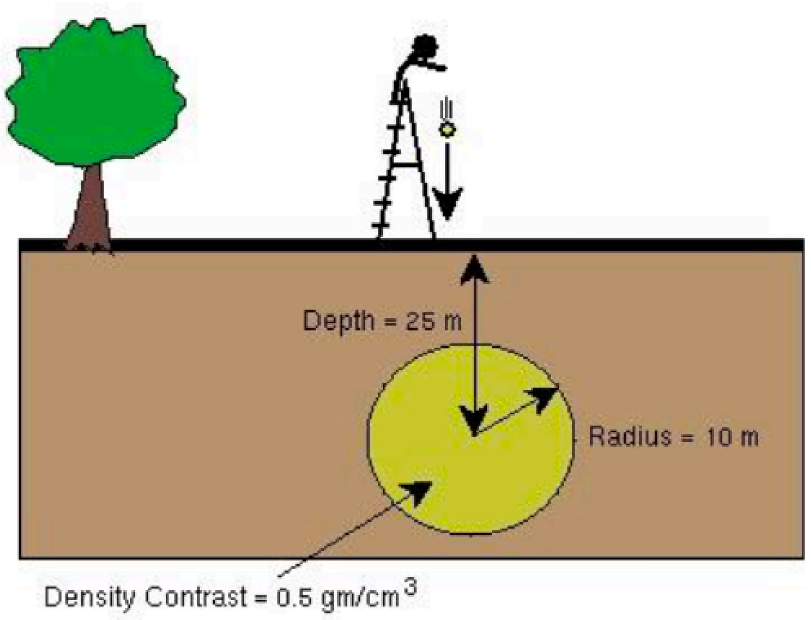
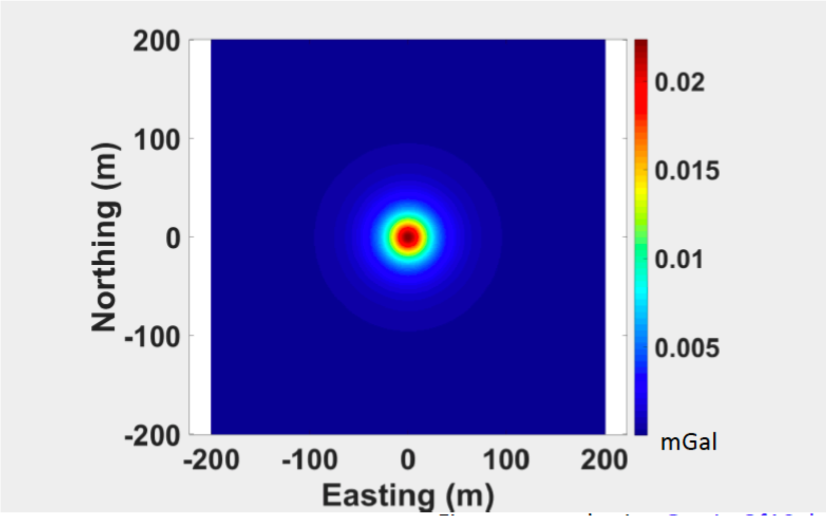
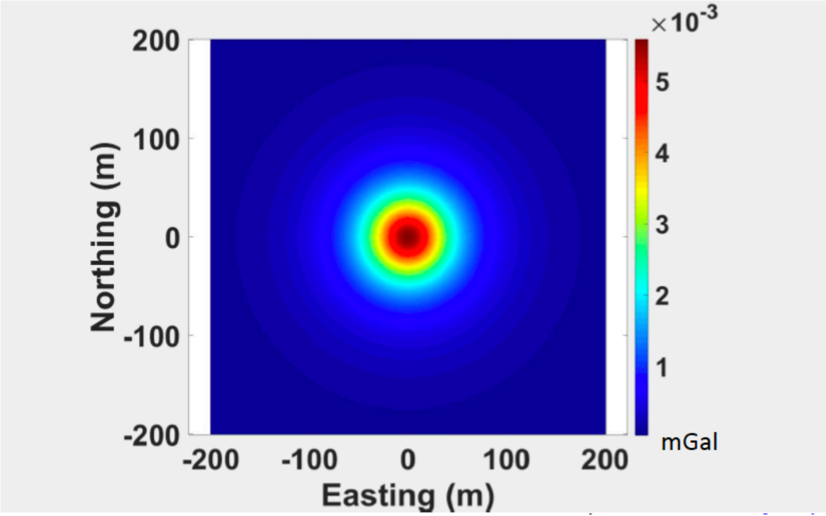


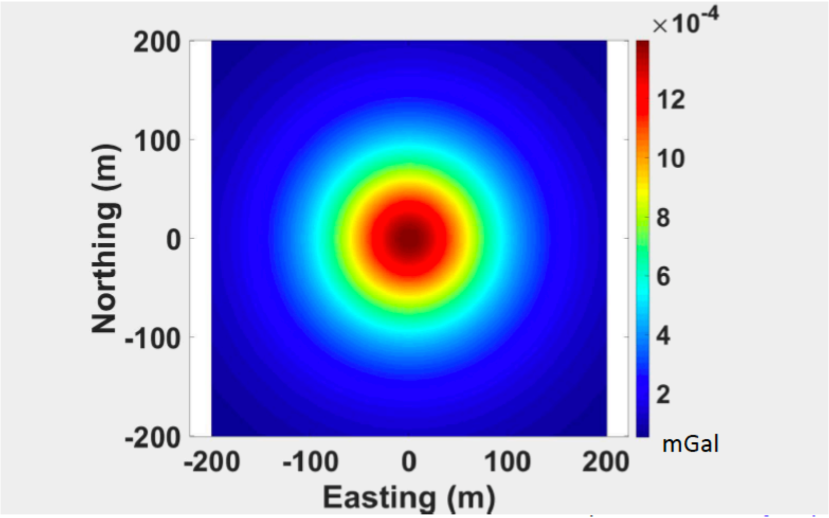
Image from t.ly/p75Wr



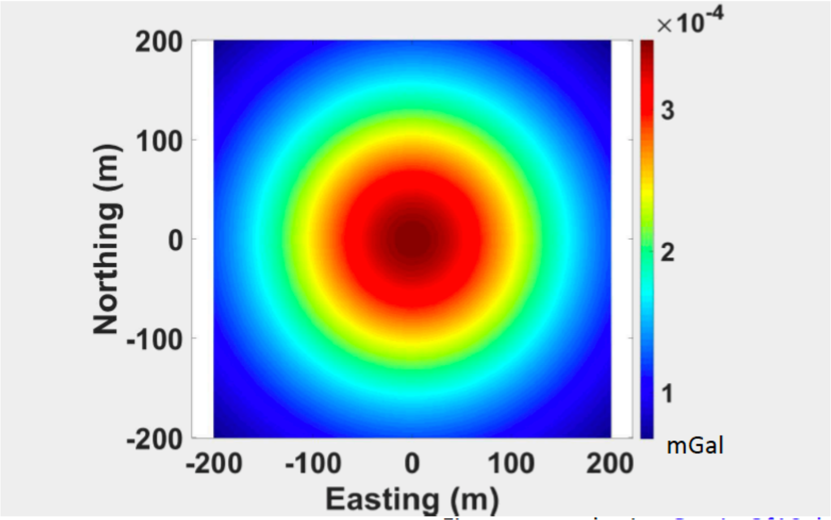
Gravity due to a uniform sphere where depth=25 m.



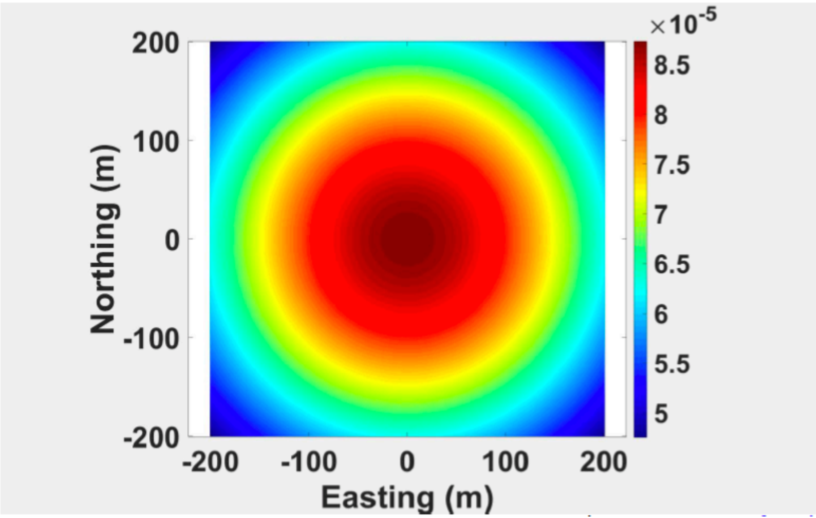
Gravity due to a uniform sphere where depth=50 m.



Gravity due to a uniform sphere where depth=100 m.



Gravity due to a uniform sphere where depth=200 m.

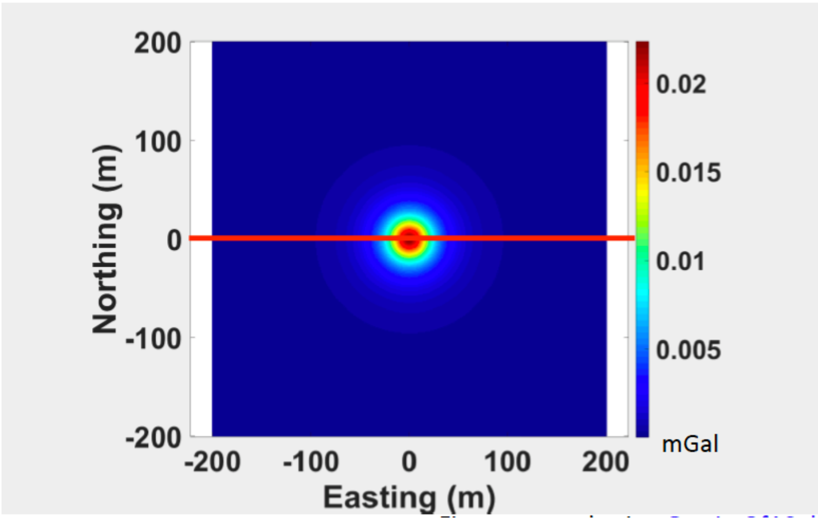


Gravity due to a uniform sphere where depth=400 m.

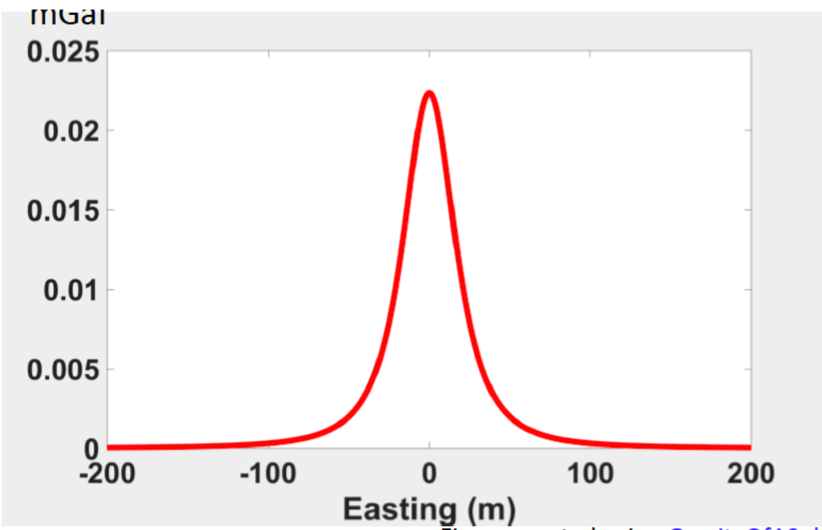
Everything is same in the above 5 figures, the radius is same, the geometry of anomaly is same. The only thing different is the depth changing from the shallower to deeper.

We can observe from the above figures that the maximum gravity measure decreases as depth increases. Besides, the width of the central peak region increases.

Let's look at a profile of gravity at depth = 25 m (seeing figure below), where the red line is profile.

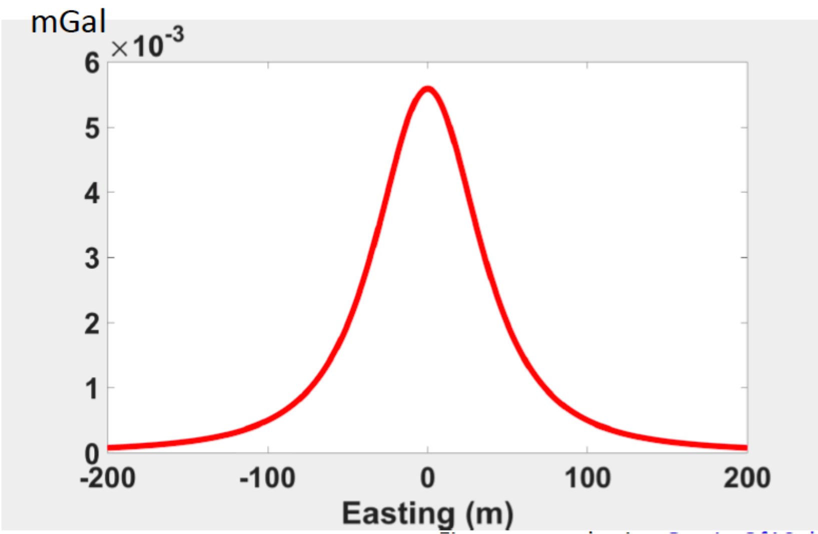


We plot the gravity of this profile shown in below:

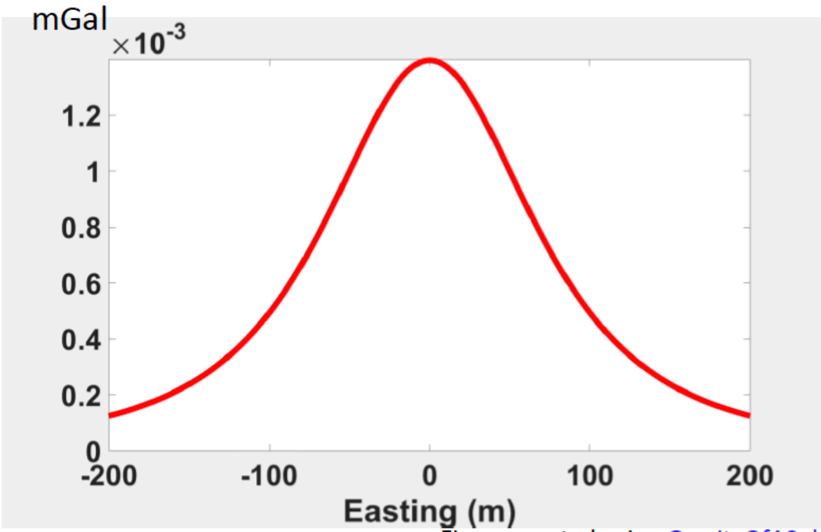


A profile (Northing = 0 m) of the gravity measures, where anomaly body located at depth = 25 m.

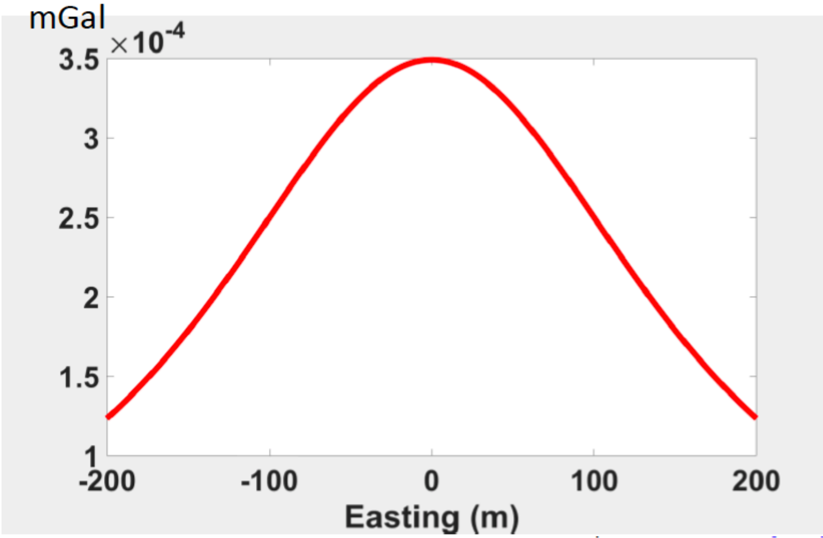
The observations from the figure above, the location at 0 corresponds to the location of source body and the peak located directly above the center of the anomaly body. The same things when we change the depth while keeping the same profile position (figures below).



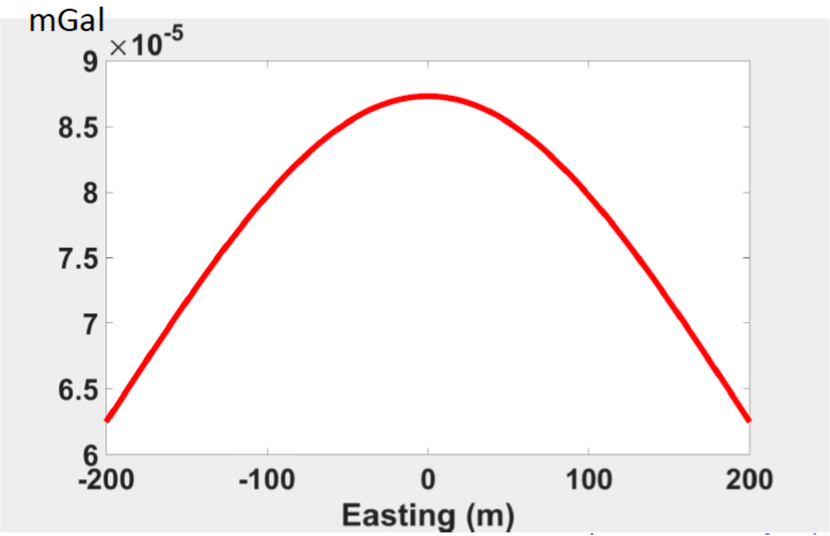
A profile (Northing = 0 m) of the gravity measures, where anomaly body located at depth = 50 m.



A profile (Northing = 0 m) of the gravity measures, where anomaly body located at depth = 100 m.



A profile (Northing = 0 m) of the gravity measures, where anomaly body located at depth = 200 m.



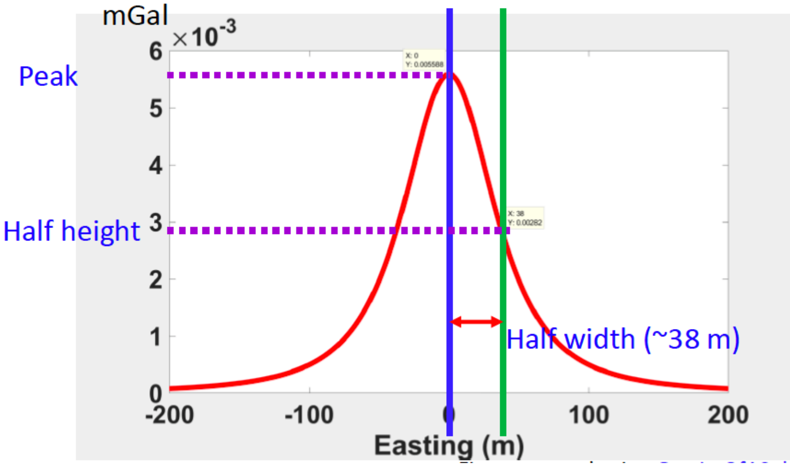
A profile (Northing = 0 m) of the gravity measures, where anomaly body located at depth = 400 m.

Here, we observe two thing from the figures above: Firstly, the maximum value (peak) decreases as depth increases. Secondly, width of the central peak region increases. These two trends are same with 2D maps of gravity measures.

Question

Can we use width to estimate depth???
Yes, we can!

We noticed there are some correlation between width of the peak and depth. The width seems to be depended on the depth. So logically speaking, we can use the width to estimate the depth of source body. Let's take a careful look at the profile at depth = 50 m.



Observation from the figure above, the peak value is 0.0055. The blue line is location of peak. The green line is the position of half of peak value, and at this point the value is 0.00282. The distance between blue line and green line is half width. In this particular case, the half width is 38 m.

There is a rule of thumb that we can use to estimate depth for spheres:

$$depth = 1.3 * half - width$$

In the equation, the depth is really the vertical distance between your observation and the center of the sphere. Thus, in our particular example, the depth = 1.3 * 38 = 49.4 m, which is really close to the true depth 50 m.

The half-width rule of thumb estimates depth for cylinders:

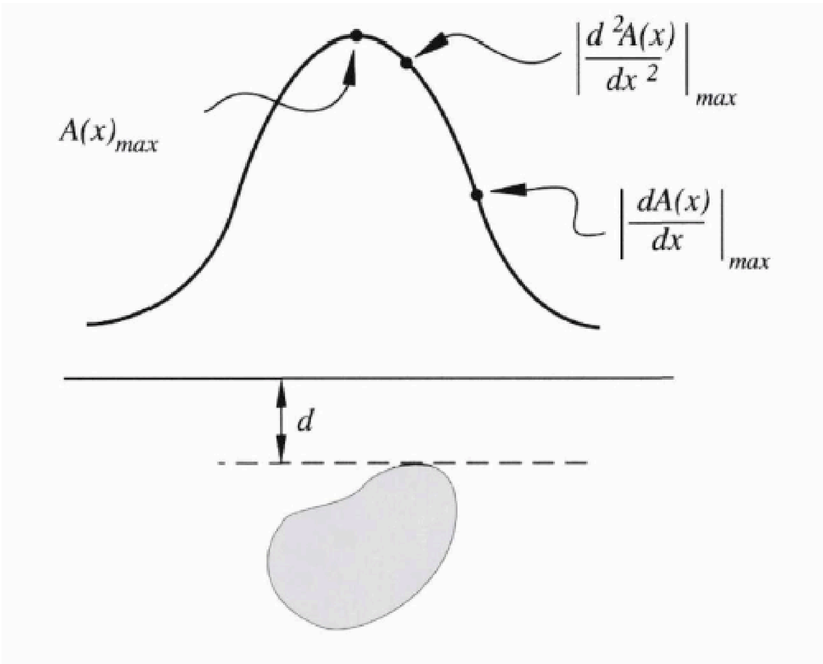
$$depth = 1.0 * half - width$$

The half-width rule of thumb is a practical rule, but there is a limitation that is we have to assume the shape of source body. If the source body is assumed to be sphere, we need to use 1.3 * half-width. In comparison, if it is a cylinder, we need to use 1.0 * half-width.

6.3.2. RULES OF THUMB FOR MAXIMUM DEPTH

Firstly, we make clear of a term: non-uniqueness. Many geologically reasonable density or magnetization solutions may perfectly satisfy the observed anomaly, which means the many potential correct models can fit our observed gravity or magnetic data equally well. For example, gravity due to s sphere is the same as a point mass. Because of non-uniqueness, actually we can create multiple source bodies and each of created source body can fit our observed data. However, some parameters about the source bodies can be uniquely determined from the observed anomalies, without assumptions about the source distribution (Blakely, 1996, p239), eg., maximum depth.

We are going to talk how to estimate maximum depth. The first question is what is maximum depth.



Maximum depth to causative sources based on first, second, and third derivatives fo their anomalies. Profile A(x) represents either a magnetic or gravity anomaly. (Blakely, 1996, p240)

The black curve is gravity or magnetic measures. The depth to a plane (the dash line shown above) below which the entire source distribution lies, which is represented by 'd' in the figure above. The maximum depth 'd' can be determined based on first, second or third derivatives of gravity or magnetic anomalies along profiles. We can use the information from derivative to help us estimate the depth.

We can estimate maximum depth for 3D gravity anomalies using the following equations (Blakely, 1996, p240):

$$d \leqslant 5.40 \; \frac{\gamma \rho_{max}}{|A''|_{max}}$$

$$d^2 \leqslant 6.26 \; \frac{\gamma \rho_{max}}{|A'''|_{max}}$$

where, the density of both signs that means the gravity anomalies could be both positive and negative. γ is gravitational constant. ρ_{max} is maximum density value. $|A''|_{max}$ is maximum value of the second order derivative. $|A'''|_{max}$ is maximum value of the third order derivative.

For both above equations, the upper limitation of maximum depth is right hand side. If we wanna use it to estimate maximum depth, we have to roughly estimate the maximum density value ρ_{max}. For example, we can estimate the maximum density value through measures of crops, borehole gravity measurements or our basic understanding of rocks. Here is a specific example to show what's the meaning of these two equations. If we measures the maximum depth is 400 meters, which means the depth of source body cannot be deeper than 400 m.

If the density entirely positive or entirely negative rather density of both signs, there are many other extra rules we can use to estimate the maximum depth.

$$d \leqslant 1.5 \; \frac{A}{|A'|}$$

where the rule is for all x.

$$d^2 \leqslant -3 \; \frac{A}{A''}$$

where the rule is for all x and A'' is negative.

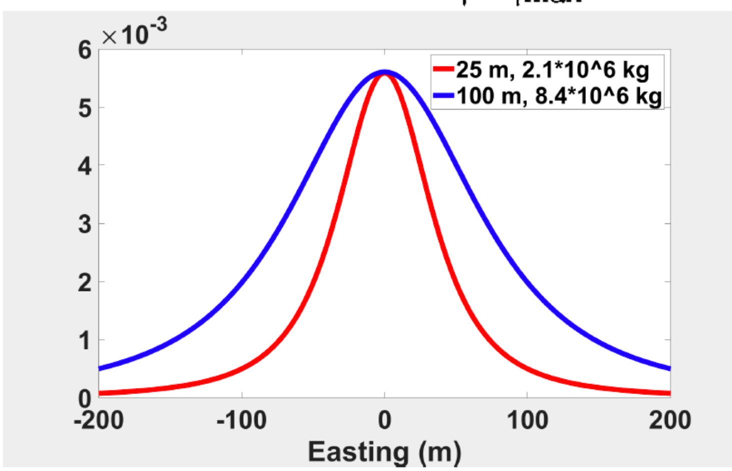
$$d \leqslant 0.86 \; \frac{A_{max}}{|A'|_{max}}$$

$$d \leqslant 2.70 \, \frac{\gamma \rho_{max}}{|A''|_{max}}$$

$$d^2 \leqslant 3.13 \, \frac{\gamma \rho_{max}}{|A'''|_{max}}$$

All these five rules above can help us estimate the maximum depth. We are going to talk about one rule specifically.

Understanding $d \leq 0.86 \, \dfrac{A_{\max}}{|A'|_{\max}}$



In the rule $d < 0.86 \, \frac{A_{max}}{|A'|_{max}}$, we don't need estimate maximum density value, which is more convenient to use. In the figure, we have two gravity profiles. These two profile have same maximum value because of peak is overlap with each other. The red curve decay faster than blue curve that means the first order derivative of red curve is larger than blue curve. Therefore, the maximum depth of blue curve is deeper than red curve, because the denominator of red curve is larger than blue cure. Another way to understand this is based on the first method we talked before, the half-width of the blue curve is larger than red, so the depth of blue curve is deeper than red curve.

The rules to estimate the maximum depth for 2D gravity anomalies are shown below:

$$d \leqslant \frac{A}{|A'|}$$

where the rule is for all x.

$$d^2 \leqslant -2\ \frac{A}{A''}$$

where the rule is for all x and A'' is negative.

$$d \leqslant 0.65\ \frac{A_{max}}{|A'|_{max}}$$

$$d \leqslant \triangle\omega$$

where the rule is for symmetric anomaly.

The maximum depth estimation for 3D magnetic anomalies are following:

No restrictions on magnetization:

$$d \leqslant 6.28\left(4\widehat{r}_x^2 + 3\widehat{r}_y^2 + 3\widehat{r}_z^2\right)^{0.5}\ \frac{M_{max}}{|A'|_{max}}$$

$$d^2 \leqslant 9.73\left(3\widehat{r}_x^2 + 2\widehat{r}_y^2 + 2\widehat{r}_z^2\right)^{0.5}\ \frac{M_{max}}{|A''|_{max}}$$

Magnetization everywhere parallel and same sense:

$$d \leqslant 3.14\left(4\widehat{r}_x^2 + 3\widehat{r}_y^2 + 3\widehat{r}_z^2\right)^{0.5}\ \frac{M_{max}}{|A'|_{max}}$$

$$d^2 \leqslant 4.87\left(3\widehat{r}_x^2 + 2\widehat{r}_y^2 + 2\widehat{r}_z^2\right)^{0.5}\ \frac{M_{max}}{|A''|_{max}}$$

Vertical $\widehat{r}$ and vertical $\mathbf{M}$:

$$d \leqslant 5.18 \, \frac{M_{max}}{|A'|_{max}}$$

$$d^2 \leqslant 6.28 \, \frac{M_{max}}{|A''|_{max}}$$

Vertical $\widehat{r}$ and vertical $\mathbf{M}$; $\mathbf{M}$ everywhere of same sense:

$$d \leqslant 2.59 \, \frac{M_{max}}{|A'|_{max}}$$

$$d^2 \leqslant 3.14 \, \frac{M_{max}}{|A''|_{max}}$$

where $\mathbf{M}_{max}$ is maximum value of magnetization. $\widehat{r} = (\widehat{r}_x, \widehat{r}_y, \widehat{r}_z)$ is direction in which magnetic field is measured.

The maximum depth estimation for =2D magnetic anomalies are following:

No restrictions on magnetization:

$$d \leqslant 8 \left(\widehat{r}_x^2 + \widehat{r}_z^2\right)^{0.5} \, \frac{M_{max}}{|A'|_{max}}$$

$$d^2 \leqslant 9.42 \left(\widehat{r}_x^2 + \widehat{r}_z^2\right)^{0.5} \, \frac{M_{max}}{|A''|_{max}}$$

$\mathbf{M}$ everywhere parallel and same sense:

$$d \leqslant 4 \left(\widehat{r}_x^2 + \widehat{r}_z^2\right)^{0.5} \, \frac{M_{max}}{|A'|_{max}}$$

$$d^2 \leqslant 4.71 \left(\widehat{r}_x^2 + \widehat{r}_z^2\right)^{0.5} \, \frac{M_{max}}{|A''|_{max}}$$

6.3.3. EULER DECONVOLUTION

The previous methods of depth estimation that we covered are best suited for anomalies caused by single and isolated bodies. However, there is another class of techniques that consider magnetic or gravity anomalies caused by multiple and relatively simple sources. One of these methods is **Euler Deconvolution**. In order to understand Euler Deconvolution, we need to recall Harmonic functions.

6.3.3.1 Harmonic Functions

$\nabla^2 V = 0$

If the Laplacian of a vector V is equal to zero, then V is a harmonic function. Additionally, any spacial derivative of a harmonic function is also harmonic. The spatial derivatives and Laplacian operator are also commutative. E.g., $\nabla^2(\frac{\partial}{\partial x}V) = \frac{\partial}{\partial x}(\nabla^2 V)$. This means that we can generate a host of harmonic functions from a given harmonic function. For example, if we consider the vector (\frac{1}{r}), we know that it is harmonic as $\nabla^2 \frac{1}{r} = 0$. It therefore follows that $\frac{\partial}{\partial z}\frac{1}{r}$, $\nabla \frac{1}{r}$, and $m \cdot \nabla \frac{1}{r}$ are all harmonic. The last harmonic function listed is important to remember, as it describes the magnetic potential of a dipole. Here, we can note that the scalar magnetic potential is harmonic, which also means that any component of the Earth's magnetic field is also harmonic. This is because any component of the magnetic field is simply a spacial derivative of the potential.

6.3.3.2 Homogeneous Functions

A function V is homogeneous of degree n if it satisfies Euler's equation:

$$x\frac{\partial V}{\partial x} + y\frac{\partial V}{\partial y} + z\frac{\partial V}{\partial z} = nV$$

In a simple example, if $V = xyz$, we can state that V is homogeneous of degree 3. Similarly, $V = xy$ is homogeneous of degree 2 and $V = x$ of degree 1. In a more complex example, we can analyze $V = log\frac{r+z}{r-z}$ which is homogeneous of degree 0. Similarly, $V = \frac{1}{r}$ is homogeneous of degree -1 and V=$\frac{z}{r^3}$ of degree -2. For any vector V, if it is homogeneous then any derivative of V is also homogeneous. As we established previously, $\frac{1}{r}$ is homogeneous and therefore the potential of a magnetic dipole is too.

If a homogeneous function is also harmonic, it can be represented in spherical coordinates as a sum of spherical surface harmonics.

6.3.3.3 Euler's Equation

We can express Euler's equation using the following general form:

$$r \cdot \nabla f = -nf$$

In this form, it is easy to show that $f = \frac{1}{r}$ satisfies Euler's equation with $n = 1$. Therefore, the potential of a point mass must also be homogeneous with $n = 1$. Next, we will take a look at the total-field anomaly of a dipole:

$$\Delta T = \frac{mu_0}{4\pi}\hat{b} \cdot \nabla(m \cdot \frac{1}{r})$$

This equation satisfies Euler's equation with $n = 3$.

Next, let us consider a magnetic survey. We will set ΔT_i as the ith measurement at location (x,y,z). Additionally, we will assume the center of the source body is at location (x_0, y_0, z_0). We can represent this survey according to Euler's equation:

$$[\frac{\partial}{\partial x}\Delta T_i \quad \frac{\partial}{\partial y}\Delta T_i \quad \frac{\partial}{\partial z}\Delta T_i]\begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix} = n\Delta T_i$$

Similarly, other measurements at other locations can be represented by expanding the previous equation:

$$\begin{bmatrix} \frac{\partial}{\partial x}\Delta T_1 & \frac{\partial}{\partial y}\Delta T_1 & \frac{\partial}{\partial z}\Delta T_1 \\ \frac{\partial}{\partial x}\Delta T_2 & \frac{\partial}{\partial y}\Delta T_2 & \frac{\partial}{\partial z}\Delta T_2 \\ \vdots & \vdots & \vdots \end{bmatrix}\begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix} = \begin{bmatrix} n\Delta T_1 \\ n\Delta T_2 \\ \vdots \end{bmatrix}$$

We can solve this equation using a least-squares method for (x_0, y_0, z_0) if we assume a value of n.

6.3.3.4 Example of Euler's Deconvolution

The example used below comes from a Geophysics Tutorial by Leonardo Uieda, Vanderlei C. Oliveira Jr, and Valéria C. F. Barbosa.

Uieda, L., V. C. Oliveira Jr, and V. C. F. Barbosa (2014), Geophysical tutorial: Euler deconvolution of potential-field data, The Leading Edge, 33(4), 448-450, doi:10.1190/tle33040448.1

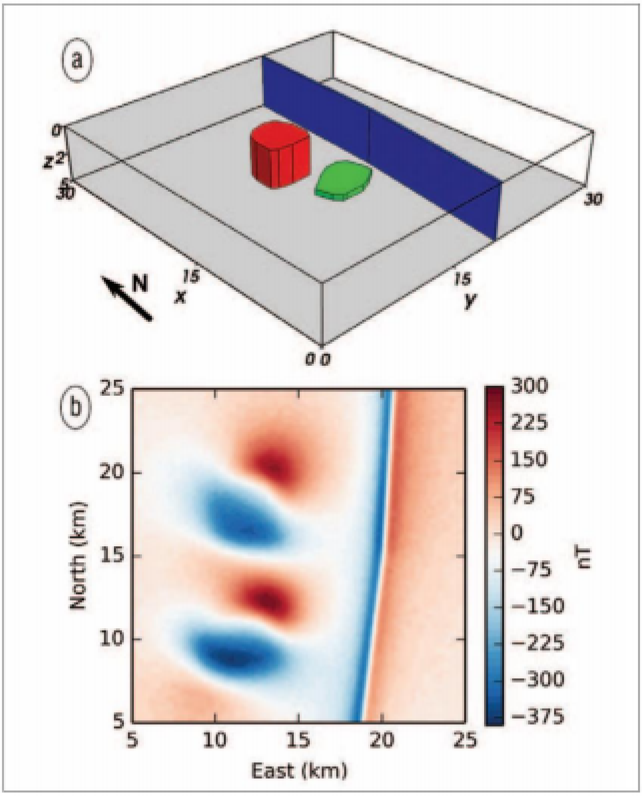


Figure 1. (a) Our model. The model simulates a dike (blue), a sill (green), and an intrusive body (red). (b) Synthetic total-field anomaly data. Data are corrupted with 5-nT pseudo-random Gaussian noise.

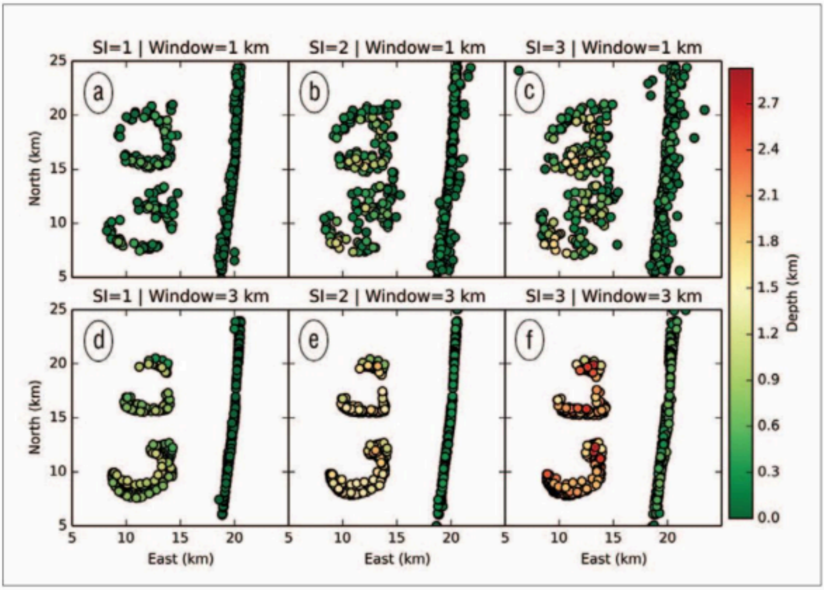


Figure 2. Euler deconvolution solutions for varying structural index (SI) and moving-window size.

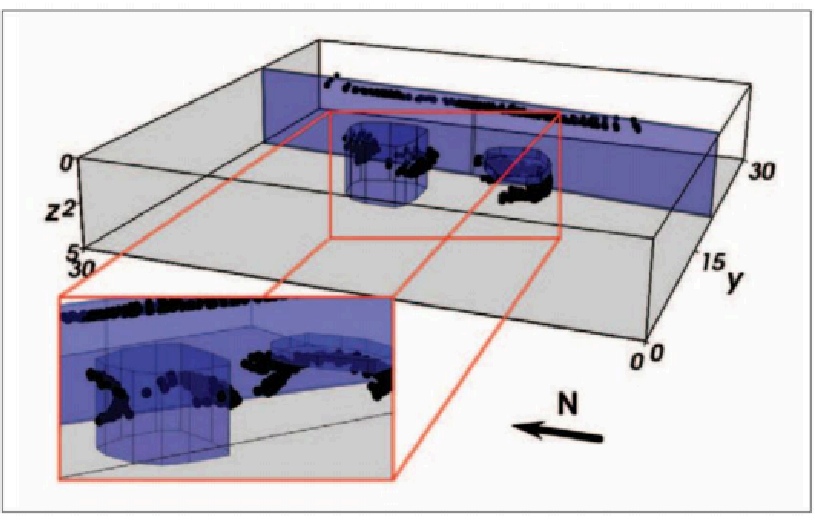
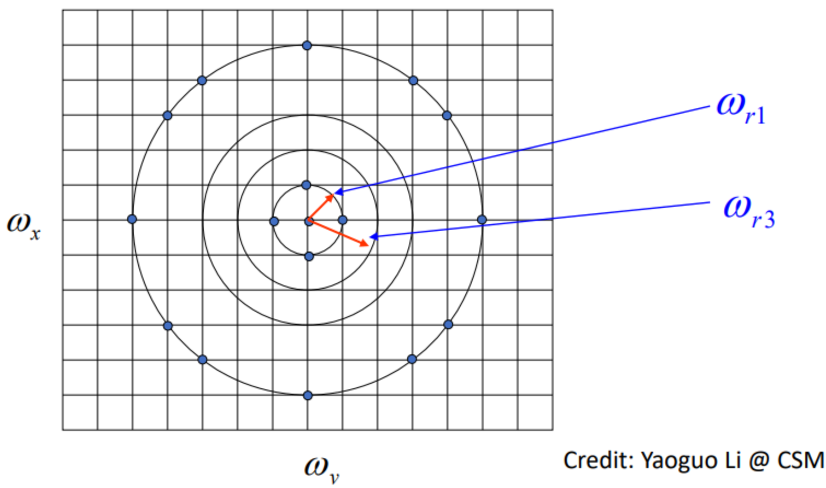


Figure 3. *3D view of Euler deconvolution solutions (black dots) for a moving window of 3 km and structural index of 3.*

6.3.4. SPECTRAL DEPTH ESTIMATION

Spectral depth estimation is done by calculating the radially averaged power spectrum of potential field data using the Fourier Transform. It is found by taking the average over points on concentric circles followed by smoothing along radial wave numbers. An example is shown in the figure below.



There a couple of important properties related to the radially averaged power spectrum for single source bodies that we will discuss. In regards to the decay of the power spectra of gravity and magnetic data, this decay depends primarily on the source depth. Deeper source depths result in faster decay of the power spectra, also leading to narrower recorded bandwidth. A semi-log plot can be used to plot the radially averaged power spectra, as parts of the spectra will appear as straight line segments. The slope of these line segments is proportional to the source depth, and can be described using this equation:

$$ln[p(w_r)] \propto 2lnw_r - 2hw_r$$

Note that in the intermediate wavenumber band, the slope is proportional to twice the source depth.

We will briefly discuss a paper by Spector and Grant (1970) on the Ensemble Average and power spectra. In this paper, an ensemble of rectangular prisms were used to represent parameters that were randomly distributed, uncorrelated, and had uniform probability. They used this to examine the expected power spectra of a magnetic field, and reached some surprising conclusions.

$$p_B(w_r) \propto e^{-2\bar{h}w_r}$$

if $\Delta h << \bar{h}$

In this case, $\bar{h}$ is the average central depth of the ensemble and 2 is the depth interval within which h is uniformly distributed. The authors concluded that the power spectrum of the ensemble behaves the same as an "average" member from within the ensemble.

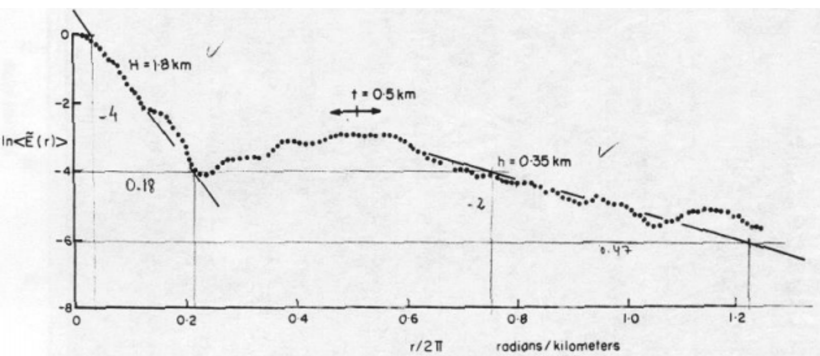


FIG. 4b. Power spectrum analysis of an aeromagnetic map of Figure 4a.

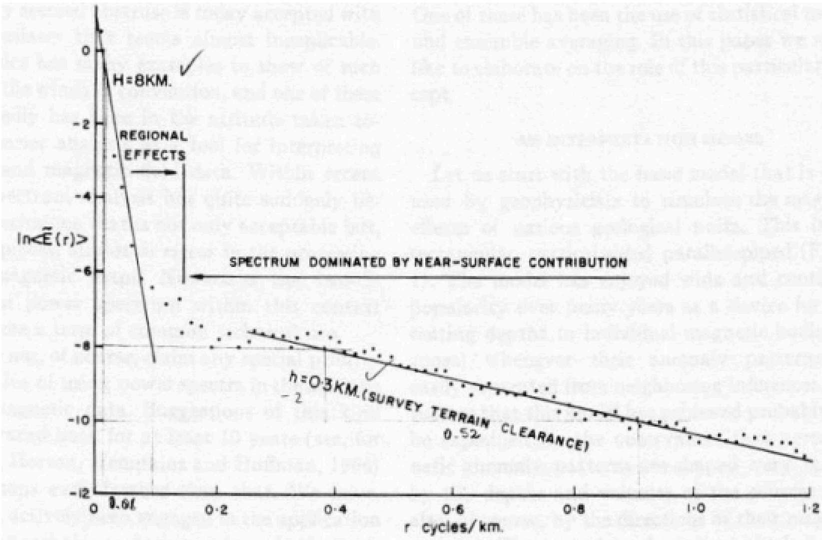
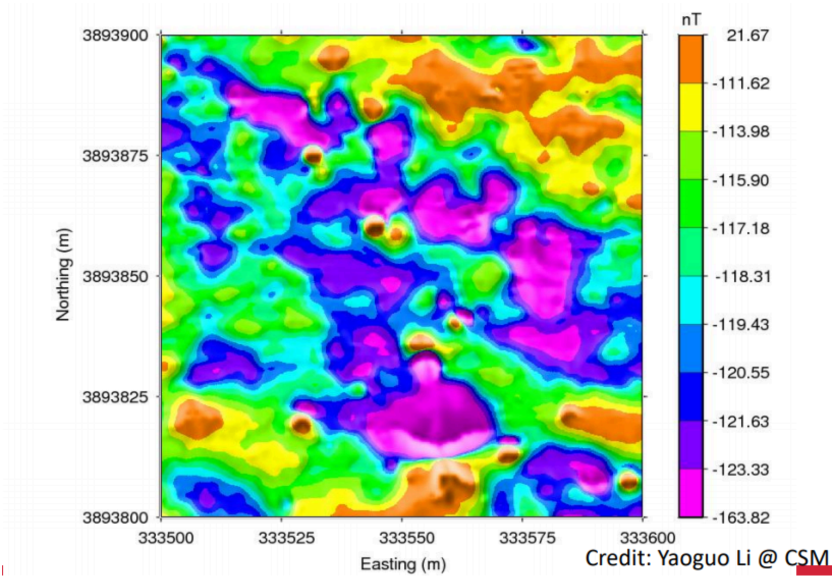
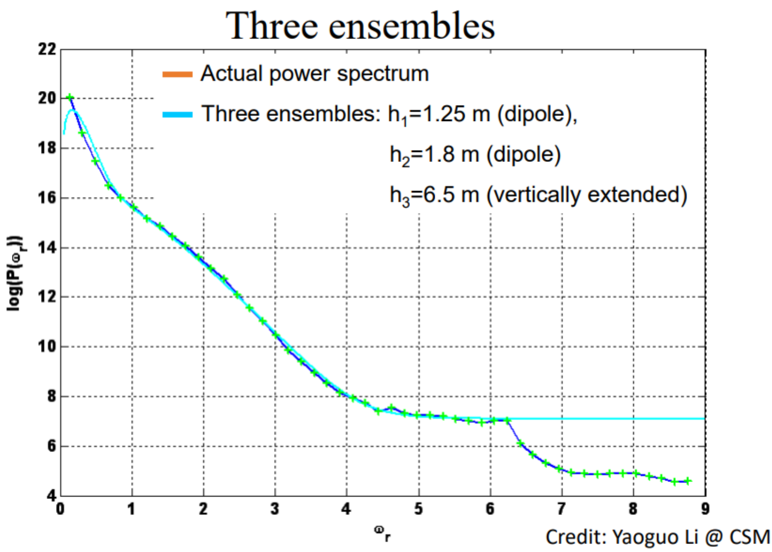


FIG. 6b. A power spectrum analysis of an aeromagnetic map of Figure 6a.

Next, we will take a look at field data collected from Kirtland.



Airborne magnetic data collected over Kirtland



The radially averaged power spectrum of the Kirtland data, with the power spectra of three rectangular ensembles compared to the actual power spectrum

Note that with the radially averaged power spectrum of the Kirtland data, the majority of the rectangular ensembles match the actual power spectrum almost perfectly, confirming the results of the Spector and Grant paper.

This is where you can add appendices or other back matter.