Harmonic and modal analysis

1 Systems, signals, stochastic processes, data and noise

We are becoming ever more careless in our analysis of geophysical (and probably geological) data. We carelessly avoid asking ourselves the relevant question that we might want to answer. This can be extremely problematical when carried into harmonic and modal analysis. We might, for example, want to ask ourselves what inherent periodicity is there in the natural system that we are studying. We might ask ourselves what is the inherent periodicity of tides or temperature. We know that there are forcings of tides and temperature that relate closely to well known geodynamic and astronomical periods. If we are concerned about the statistical significance of the possibly periodic forcings, we should be quite careful about how it is that we view or model the data that we obtain from the natural system. And, then, we must be very careful not to bring to bear worthless analytical tools in determining the frequencies and amplitudes of the periodicities. Some workers sluff off the responsibility to do statistically valid science by attaching to utterly simplistic models.

Properly, one might regard a geological survey or the acquistion of geophysical data as taking a piece from a process. The character of the process which provides the survey (say map) or data sequence is most important in obtaining information about the process. In typical time-series or space-series analysis, we would like our process to be characterized by some of several properties.

We might find a geology that appears to be a *truly periodic process*; for example, you might have geologically mapped the granite-tiled floor of an imposing building and found that there is an inherent periodicity in the tiling. Still, one must ask oneself a question concerning such a map: "Does this periodic 'geology' endlessly repeat itself in all directions to cover the planet? Or, is it merely a 'sample' of floors of imposing buildings?" That is, we should be concerned with understanding whether our map represents "geology" (in the whole) or is merely a small sample of a particular local geology of the whole which might not be measured by the same rules everywhere. The issue here is that there is a difference between the process and the sample of the process. Most of you know that one can come closer to understanding the larger process by ever more elaborately sampling the process. In principle we approach a complete knowledge of the process as we accumulate towards infinitely detailed and complete mapping of it. In simply idealized cases[1] we can learn twice as much about a process through increasing our sampling by 4 times. If the process is characterized by an infinity of detail, to know it completely would require an infinity **[2]** of samples.

[1] What we often describe as a Gaussian random process.

The "geological process" is infinitely detailed. We typically look to find the essential elements of the process.

One might also ask another important question: "Is the process limited to a finite 'interval'?" Clearly, the surface of the Earth is nearly a finite interval. It doesn't extend forever. A plane or a line can be unconstrained in area or length (well, to 13.7 billion light years in any direction in our universe) and so might be regarded as an open or infinite interval. What processes play on finite intervals and infinite intervals are inherently and conceptually different. The main difference is that on a closed or finite interval, there is the possibility of an exact harmonic decomposition of a functional variation. That means that over a closed interval, there are fixed modal components that are constrained by (and periodic in) the interval. Over the spherical Earth, we describe an infinity $^{\mathbf{2}}$ of modal components. Over, for example, an infinite plane, we would require an infinity $^{\mathbf{4}}$ of modal components for an exact model. Of course, there is no way to even conceive of dealing with exact modelling over an infinite plane which is fundamentally aperiodic. Over a finite-dimensioned plane, we can. For example, we could find all the periodic vibrational modes of a rectangular drum skin... simple 2-dimensional Fourier analysis is the appropriate tool. Over the (exactly) spherical Earth, we can in principle describe all the possible periodic modes required to build any functional relationship on the sphere... spherical harmonic analysis is the appropriate tool[2].

Returning to the problem of modal analysis of process described over open intervals like an infinite line, we might approximate or estimate some of the harmonic modes based on a finite-length sample of the process. We should understand, now, that we only have an estimate of the process' harmony; rather than accomplishing a "harmonic analysis" we seek "harmonic estimation" which is necessarily faced with an error due to the insufficiency of the sample interval which we deal with as statistical errors on our harmonic estimates. For data constrained on a line, say a time series which might have begun before our first sample of it and might continue on past our last sample, we responsibly obtain "spectral estimates" rather than "Fourier components". Depending on the conceptual model of the infinite or semi-infinite process generating the time series, we employ spectral analyses appropriate to our model. Classically, in time-series analysis, we might separate models of a restricted class of processes, the stationary-ergodic processes, into two fundamental models with some derived variants.

- **Moving-average model:** The data are modelled as a moving-averaging across a white Gaussian (sometimes called purely random) excition. That is, given such an excitation, say $\boldsymbol{e_i}$ which might range infinitely, the process from which

[2] This is the core topic of the remaining lectures

our data are extracted as a sample is described as

$$\boldsymbol{d_i} = \sum_{k=0}^{K} \boldsymbol{s_k e_{i-k}}.$$

Our data might further be corrupted by noise or measurement error:

$$\boldsymbol{d_i} = \sum_{k=0}^{K} \boldsymbol{s_k e_{i-k}} + \boldsymbol{n_i}.$$

- **Autoregressive model:** The data are modelled as a recursion on past process, weighted by a (typically) short regression operator:

$$\boldsymbol{d_i} = \sum_{m=0}^{M} \boldsymbol{r_m d_{i-m}} + \boldsymbol{p_i},$$

where, now, $\boldsymbol{p_i}$ continually innovates the data generating process; again with measurement error or noise,

$$\boldsymbol{d_i} = \sum_{m=0}^{M} \boldsymbol{r_m d_{i-m}} + \boldsymbol{p_i} + \boldsymbol{n_i}.$$

You might note that if the statistical characteristics of the innovation, $\boldsymbol{p_i}$, and the noise or measurement error, $\boldsymbol{n_i}$, are not clearly distinct, we cannot know what is measurement error and what is continuing innovation.

- **Combined autoregressive-moving-average model:** We might easily combine these models:

$$\sum_{m=0}^{M} \boldsymbol{r_m d_{i-m}} + \boldsymbol{p_i} = \sum_{k=0}^{K} \boldsymbol{s_k e_{i-k}} + \boldsymbol{n_i}.$$

In determining the models, we assign a statistical model to $\boldsymbol{e_i}$, $\boldsymbol{n_i}$ and/or $\boldsymbol{p_i}$ and then fit our data sequence, $\boldsymbol{d_i}$ with the best sequences $\boldsymbol{r_m}$ and $\boldsymbol{s_k}$ according to some prior-chosen objective function. This objective function in the simplest analyses is usually the least sum of squares of misfit between our modelled estimates $\widehat{\boldsymbol{d_i}}$ and the data measured, $\boldsymbol{d_i}$.

Not wanting to carry this story into a full course on time-series analysis, I shall now present some image-based explanations of some important issues.

1.1 Harmonic content of "known" processes:

I have generated a pseudo-uncorrelated Gaussian random sequence of 10000 points. An uncorrelated Gaussian random process (infinite length) is spectrally "white" in the sense that all harmonic components that describe the process have equal squared amplitudes; nothing is demanded of their phase relationships, or equivalently of the division of the amplitudes into real and imaginary parts. I note that the process is only "pseudo-uncorrelated" because no rule-generated recursions obtained by computer can be uncorrelated. For the "linear-congruent generator" used in my code, the minimum period of correlation is, theoretically, $\mathbf{2^{31}/6}$ points. We estimate the harmonic composition of the process based on various sample lengths while well knowing that, it should be "white" or of even squared amplitude in all harmonics.

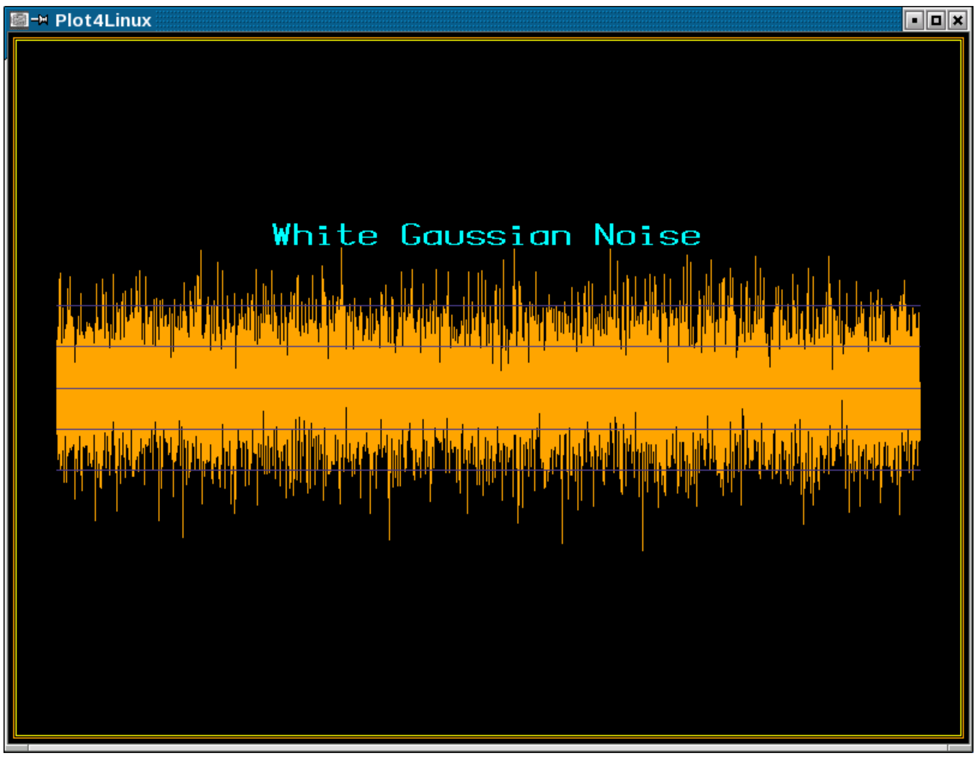


Figure 1 Our "white Gaussian noise process – sample of 10000 points.

What harmonic composition we might discover or estimate in the process from which this 10000-point sample is taken depends on how we look at it or on how it is that we model it for estimation. I show four different modellings and the results of estimations.

- **The Periodogram:** This is the crudest and least statistically significant with respect to its estimates of all the harmonic methodologies when applied to random processes that are not constrained to a fixed interval. Given a fixed interval and data sampled through that interval, it might be seen as the very best estimator.

The periodogram is calculated as the square of the Fourier transform components of the windowed sample. Depending on the length of the Periodogram one chooses to use, we migh obtain a “power density spectrum” of its harmonic composition with seemingly high resolution. I show two calculated periodograms, one based on the 10000-point Fourier transform and another on only the first 100 points of the 10000-point series.

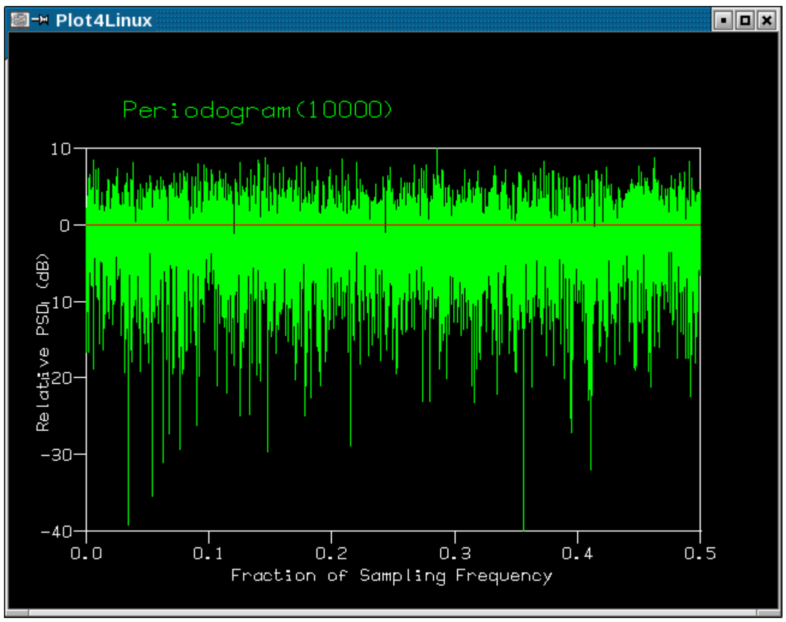


Figure 2 Periodogram estimates of the “white Gaussian noise process – sample of 10000 points.

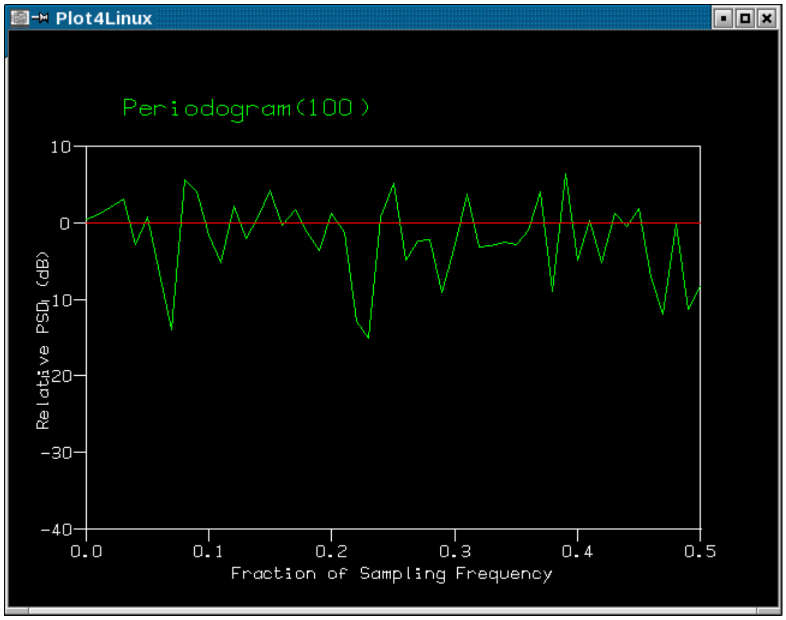


Figure 3 Periodogram estimates of the "white Gaussian noise process – sample of 100 points.

With respect to the statistical validity of these estimates as being representative of the infinite process from which the 10000-point or 100-point sample is taken, you should note the PSD ("power spectral density") of any component (peak or trough) is constrained at $\mathbf{1}\boldsymbol{\sigma}$ level to anywhere between $\mathbf{0}$ and $\sim \mathbf{1.7}\times$ its apparent value. We can improve the statistical significance of any estimate by averaging across a window of values. If we were to window-average $\pm\mathbf{50}$ points on our 10000-point periodogram, our resulting spectrum becomes ever more as one would expect: an even distribution across the entire spectral band. The effect of smoothing is really to add ever more periodogram components into our estimate so that the variance on our harmonic estimates is reduced according to $\sqrt{\boldsymbol{N}}$ where $\boldsymbol{N}$ is the number of raw periodogram components in our estimate. By averaging over 100 components, we reduce the variance on our estimates by a factor of 10. There are, however, several ancilliary issues that arise according to how we weight the components in our averaging. This, again, is beyond this short story.

Still, if you are to accept that our input "white Gaussian noise process" does, in principle, have a perfectly even distribution of harmonic component power, Figure 4 shows that we can approach the "known" correct power density spectrum through smoothing.

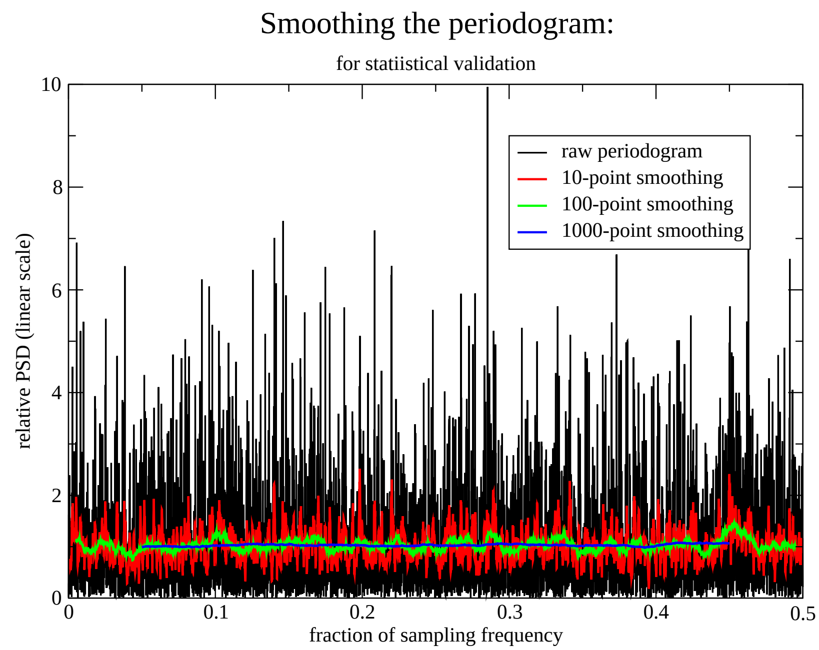


Figure 4 Smoothing improves our PSD estimates.

One can regard the raw periodogram as the spectrum described by a moving-average data model with as many coefficients as there are data points. By smoothing, we reduce the order (number of coefficients) in the model and in so-doing we improve the statistical validity of the estimates.

There are dozens of data models (the three simplest are described in the previous section) that lead us to resolving particular spectra. As you might recognize, I favour purely autoregressive models because the systematic formation of geophysical processes tends toward resonances rather than anti-resonances. That is, geophysical processes (time series) tend to be developed as the equivalent of multiple causal pendulums. Geophysical space-series data are not often directionally causal and so one might resist using such models in their analysis.

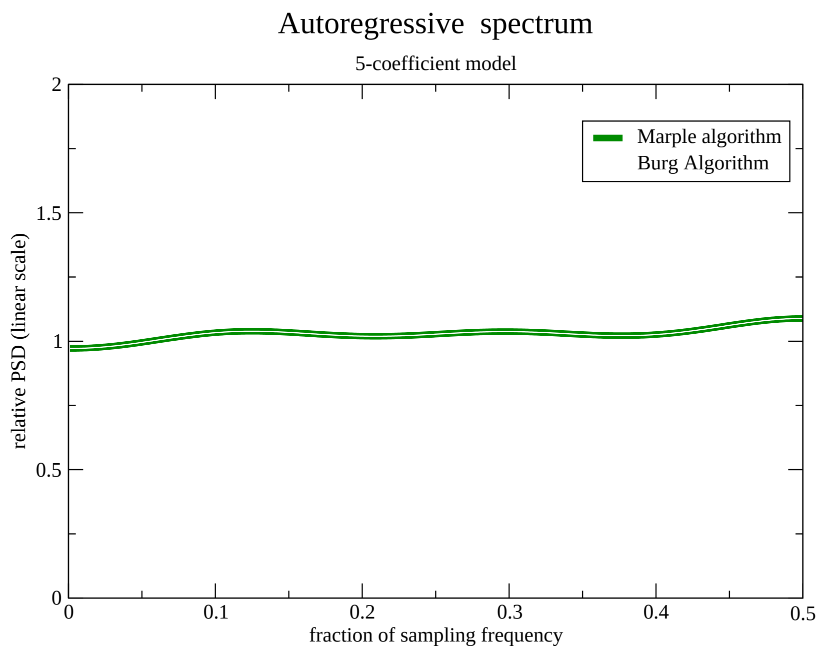


Figure 5 AR PSD estimates by Marple's and Burg's algorithms.

1.2 Data and measurements over a closed (finite) interval

While it might not be clear to you, on a closed interval, it is implicitly possible to determine a unique harmonic decomposition. The harmonic decomposition is obtained in terms of the special-function set that "fits" the interval. For example, if the interval is a line of length, L, we can explicitly fit a continuous sequence of data measurements exactly with a Fourier series. For example, suppose on the closed interval $[0,\ 1]$ we have a continuous function, $f(x)$ such that $f(x) = 0, x < 0, x > 1.$ We can decompose this function into its distributed Fourier components:

$$F(k_x) = \int_{-\infty}^{+\infty} f(x) e^{i k_x x} dx,$$

a continuous function of "wavenumber" k_x. More commonly, when dealing with data, we do not have the function $f(x)$ sampled continuously (ie., with infinitesmal interval). Suppose we have it sampled "*sufficiently*[3] with N steps of interval Δx so that the function data samples are equal to the function values at all points $x = n\Delta x; n = 0, N-1$ where $N = 1/\Delta x$ is the total number of samples. In

[3] Shannon's sampling theory

principle, this sampling is explicitly impossible to accomplish as no function over a finite interval can be "band-limited" within the "Nyquist" band of wavenumbers, that is $k_{nyq} = \pm 1/(2\Delta x)$. Still, if if Fourier components beyond the Nyquist frequecies have sufficiently low amplitude, $F(\pm k_x) \to 0; k_x > k_{nyq}$, we can well (enough) approximate our function over the interval with an $N-$finite sequence. Given a discretely sampled "version" of our function, $f(x)$, we can, then, reasonably obtain an discrete samples of independent Fourier components via the "discrete Fourier transform".

$$F_m = \sum_{n=0}^{N-1} f_n e^{-i2\pi mn/N}$$

from which we can reform our input sequence

$$f_n = \frac{1}{N} \sum_{m=0}^{N-1} F_m e^{+i2\pi mn/N}.$$

You might note that these forms describe sequences that are periodic outside the interval $[0, N-1]$; F_m and f_n can be formed for any value of m or n, not just those constrained within the interval. In discretizing our function $f(x)$ through sampling we have, in principle, also made the discretized version periodic with period, N. In trying to be careful about the effects this periodicity might have on our determination of the DFT (discrete Fourier transform), we sometimes embed our interval $[0, N-1]$ into a much longer interval $[0, N'-1]$ where $N' > 2N$. This eliminates a wrap-around effect on the ends of the interval. All of this is perhaps a time-series or geophysical-data analyst's fussiness but it does matter in the doing of a proper job.

The DFT (discrete Fourier transform) is almost always calculated using an algorithm rediscovered by Cooley and Tukey in 1964[4], usually called the FFT (fast Fourier transform). The algorithm had actually been "invented" in 1805 by Carl Friedrich Gauss. Until the rediscovery, most of us were calculating DFT's by numerical integrations which were incredibly slower. What I regard as being, by far, the very best realization of the FFT algorithm is the strict **FORTRAN** code written by Norman Brenner of MIT in 1967: FOURT.F[5]. I have another version which I rewrote in ***C*** which might be interesting to serious numerical analysts and computer coders.

What have I argued? I have tried to explain of the difference between

- 1. data analysis based on a short sample of an extended process, and

[4] The Cooley-Tukey FFT
[5] The FOURT.F code

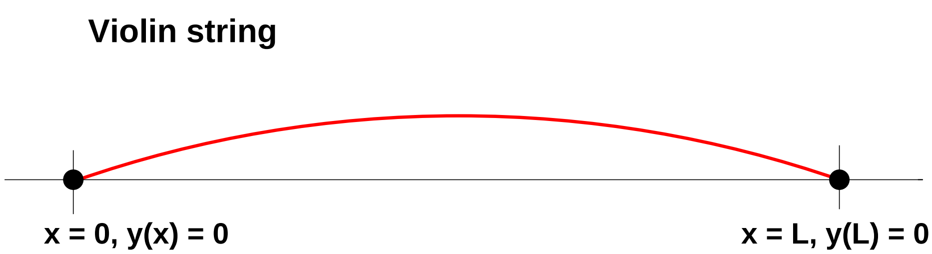
- 2. the possibility of exact harmonic analysis of a function which is precisely constrained to a fixed and finite interval.

The first is properly a problem in statistical estimation; the second is a problem is re-representing the data according to some linear transformation which is appropriate to the interval. For a finite line, the Fourier transform is the appropriate tool. For a function defined over a circular area, the Bessel transform is the appropriate tool; for a function defined over the surface of a sphere, the Spherical harmonic transform is the appropriate tool. We now, by example analysis of a much simplified problem in seismology, introduce the Spherical harmonic analysis.

2 Free Oscillations of an elastic fluid "Earth"

The Earth is an elastic body and responds to transient excitations such as earthquakes by "ringing" with its characteristic normal or free modes much as does a violin string respond with vibration when plucked. One might note that the tension, mass/length and style of excitation choose the vibrational modes with which the string vibrates.

- **The violin string:** To start, let us look at the physics of the violin string. The string has length, L, and at any point $0 < x < L$, $y(x, t)$ represents its momentary deflection from the equilibrium state.



A wave equation describes the oscillation

$$c^2 \frac{\partial^2 y}{\partial x^2} = \frac{\partial^2 y}{\partial t^2}$$

which (by Sturm-Liouville theory) must have standing wave solutions of the Fourier form

$$y(x, t) = \sum_{n=0}^{\infty} a_n Y_n(x, t)$$

where

$$Y_n(x, t) = \sin({}_n\omega x/c) e^{i_n \omega t}$$

are the "free modes" of vibration of the string. The "eigenfrequencies" of the oscillations are then

$${}_n\omega = (n + 1)\pi c/L.$$

The speed in this wave equation is related to the tension, say T, on the violin string and the mass/unit-length, M_l, of the string. Not surprisingly, the wave speed increases as we tension the string and a thick or massive string's speed is clearly lower and so its tone is lower.

$$c = \sqrt{\frac{T}{M_l}}.$$

Also note, the longer the string the lower the tone.

The Earth is more complex a spherically layered self-gravitating sphere; it possesses an extremely rich spectrum of vibrational modes.

- **A spherical, fluid, elastic Earth:** We shall discuss an extremely simple Earth model because we can obtain fully analytic solution (at least, analytic in special functions) that is most instructive about the character of the free modes. As our Earth model has no rigidity, $\mu = 0$ everywhere in a fluid, we shall find no torsional oscillations. These do exist in rich measure for "real Earth models". We ignore the surface tension, choosing to describe a stress-free surface. This approximation is valid for large oscillations. While our Earth is self-gravitating, we only include its self-gravity in terms of the pressure increasing with depth. Second-order self-gravity due to the density changes associated with compressions and dilations in oscillation is ignored.

  We shall use classical vector formalism rather than full tensor formalism in our following description as none of the elements of our simple model require a 2-tensor formalism. We shall describe our "vibrations" in a spherical coordinate system.

  In this much simplified case, the equation of particle motion at any point within the sphere of the "Earth" is easily described

  $$\rho \frac{\partial^2 \vec{u}}{\partial t^2} = -\nabla P$$

  where P is the pressure deviation from equilibrium and ρ, the fluid's density. Hooke's law of linear elasticity for the fluid also relates pressure and fluid displacements:

  $$P = -\kappa \nabla \cdot \vec{u}$$

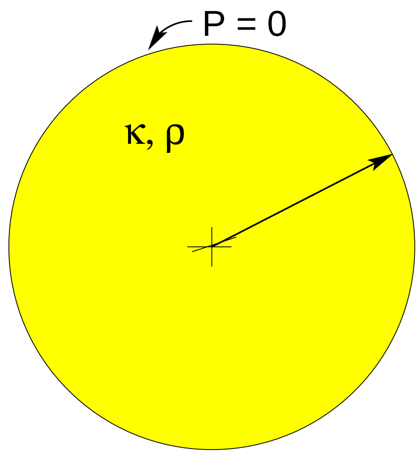
  and

  $$\rho \frac{\partial^2 (\nabla \cdot \vec{u})}{\partial t^2} = -\nabla \cdot \nabla P$$

  so

  $$\frac{\rho}{\kappa} \frac{\partial^2 P}{\partial t^2} = \nabla^2 P$$

  which you will recognize as a wave equation characterized by a speed of $c = \sqrt{(\kappa/\rho)}$. κ measures the fluid's bulk incompressibility.



We have the equation; what of boundary conditions? Let us assign the Earth's surface as a "free boundary" meaning that it is pressure or stress free. We ignore the trivial fluid surface tension. So, as we are seeking $P(\vec{r}, t)$, assign $P(\vec{r_o}, t) = 0.$ To the extent that our coordinate system is well matched to our functional dependence, we shall seek solution via a separation of variables. This wouldn't work if we were to use, say, Cartestian coordinates to describe $P(\vec{r}, t)$. We expect separable solutions of the form

$$P(\vec{r}, t) = R(r)\Theta(\theta)\Phi(\phi)T(t).$$

Expecting solutions to the wave equation will be oscillatory, we expect

$$T(t) \sim e^{\pm i\omega t}.$$

In spherical polar coordinates,

$$\nabla^2 P = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2\frac{\partial P}{\partial r}) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}(\sin\theta\frac{\partial P}{\partial \theta}) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 P}{\partial \phi^2} = \frac{1}{c^2}\frac{\partial^2 P}{\partial t^2}.$$

Now, substituting for $P = R\Theta\Phi T$ in this equation, noting that the functional subforms are independent of each other and so are only differentiable by their particular vector/time component, dividing through by P, we find

$$\frac{\sin^2\theta}{R}\frac{d}{dr}(r^2\frac{dR}{dr}) + \frac{\sin\theta}{\Theta}\frac{d}{d\theta}(\sin\theta\frac{d\Theta}{d\theta}) + \frac{\omega^2 r^2}{c^2}\sin^2\theta = -\frac{1}{\Phi}\frac{d^2\Phi}{d\phi^2}.$$

You might note that we have selected the negative exponent sign for the time dependence, $T(t) \sim e^{-i\omega t}$ because it forms a "nicer" previous equation.

– **Solving this equation:** You might recognize, further, that each term of this equation depends on only one of the spherical coordinates or on time. That leads us to this simple, iterative approach to solution. The only way that the left and right sides of the previous equation can hold is if each side is exactly $\mathbf{0}$ which is essentially uninteresting or a constant value, say $-m^2$. Then, we form the equation which describes separably the $\phi-$ dependence of $P(\vec{r}, t)$:

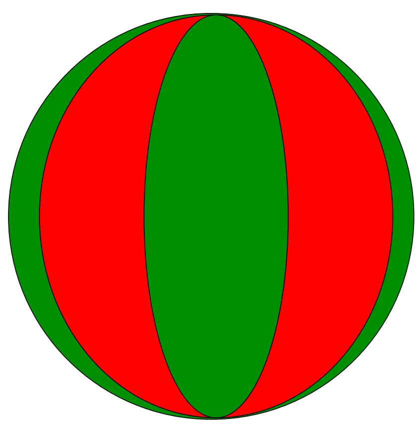
$$\frac{d^2\Phi}{d\phi^2} + m^2\Phi = 0.$$

We immediately recognize that

$$\Phi(\phi) = e^{im\phi}, m = 0, \pm 1, \pm 2, ...$$

We have our first set of *special functions* addressing our problem: the Fourier functions $\cos m\phi, \sin m\phi$.

Referenced to our chosen orientation of the spherical coordinate system, $\Phi(\phi)$ represents waves wrapped around our sphere varying only in apparent *longitude*. If we are applying such analysis to the Earth and if we choose the geographical coordinate system to accord with our spherical coordinate system, $\Phi(\phi)$ represents waves in *geographical longitude*.



$m = \pm 4$

We now deal with the left-hand side of the equation in all variables (above) by first dividing through by $\sin^2 \theta$

$$\frac{1}{R}\frac{d}{dr}(r^2\frac{dR}{dr}) + \frac{\omega^2 r^2}{c^2} = \frac{m^2}{\sin^2\theta} - \frac{1}{\sin\theta}\frac{1}{\Theta}\frac{d}{d\theta}(\sin\theta\frac{d\Theta}{d\theta}).$$

Again as the right side of this equation is independent of coordinate r, it can only take value 0 or some constant value, say, K so

$$\frac{d}{dr}(r^2\frac{dR}{dr}) + (\frac{\omega^2 r^2}{c^2} - K)R = 0,$$

$$\frac{d}{d\theta}(\sin\theta\frac{d\Theta}{d\theta}) = (\frac{m^2}{\sin^2\theta} - K)\sin\theta\Theta.$$

We shall first solve the equation with $\theta-$dependence.

- **Working through to the solution of the equation in Θ:** Cleaning up the equation a little,

$$\sin\theta\frac{d^2\Theta}{d\theta^2} + \cos\theta\frac{d\Theta}{d\theta} = (\frac{m^2}{\sin^2\theta} - K)\Theta\sin\theta.$$

Let $\cos\theta = x$ so that $dx = d(\cos\theta) = -\sin\theta d\theta$. $d\theta = -dx/\sin\theta, d\theta^2 = dx^2/\sin^2\theta.$ Substituting the appropriate form as we work toward an equation that varies in x rather than in θ, directly,

$$\sin^3\theta\frac{d^2\Theta}{dx^2} + \cos\theta\sin\theta\frac{d\Theta}{dx} = (\frac{m^2}{\sin^2\theta} - K)\Theta\sin\theta$$

and factoring out $\sin\theta$ while noting that $\sin^2\theta = 1 - \cos^2\theta = 1 - x^2$, we come to the *Legendre equations* in two cases.

$$(1-x^2)\frac{d^2\Theta}{dx^2} + x\frac{d\Theta}{dx} + (K - \frac{m^2}{(1-x^2)})\Theta = 0.$$

  * **Case I, $m = 0$:** Note that this case implies no variation in longitude according to the previous Fourier equation for $\Phi(\phi)$.

$$(1-x^2)\frac{d^\Theta}{dx^2} + \frac{d\Theta}{dx} + K\Theta = 0.$$

This is the Legendre equation which has, for solutions, an orthogonal set of polynomial defined on the the interval $[-1 \leq x \leq 1]$ or equivalently $[0 \leq \theta \leq \pi]$.

$$\Theta(x) = \sum_{l=0}^{\infty} b_l P_l(x),$$

where for these, $K \equiv l(l+1)$ defines l. The P_l?

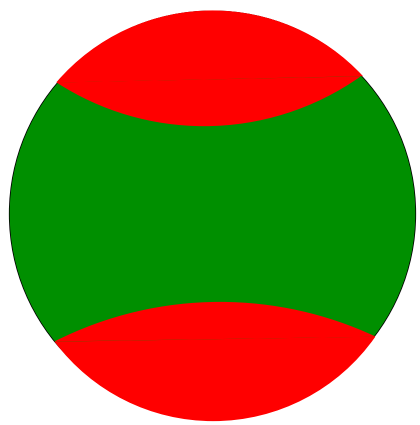
$P_0(x) = 1$ $\quad$ $P_1(x) = x$ $\quad$ $P_2(x) = (3x^2 - 1)/2$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x) \qquad P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)\ldots$$

and, in general, by *Rodrigues' formula*

$$P_l(x) = \frac{1}{2^l l!}\frac{d^l}{dx^l}(x^2 - 1)^l.$$

Note that $P_0(x = \cos\theta) = 1$, which tells us that this functional form contributes no "co-latitude" dependence over the sphere. We'll come back to this but, in the moment, you might recognize that P_0- dependence characterizes the so-called *breathing mode* of seismic oscillations. You might also note that the $P_2(x = \cos\theta) = 2\cos^2\theta - \sin^2\theta$, varying from value $+2$ on our nominal poles to -1 on our nominal equator. That is, it describes a *low* equatorial band and *high* polar regions.



$m = 0, l = 2$

You might recognize that this represents one phase of oscillation of the so-called seismic *football mode*. We shall see that a periodic oscillation is formed so that one half period later, the equatorial region becomes *high* and the polar regions *low*. The sphere, in our nominal coordinates (which we chose to align with the geographical coordinates for ease of our description), is either extended along the polar axis when the coefficient b_2 is positive or flattened along the polar axis when the coefficient is negative.

* **Case II, $m \neq 0$:** What I would like to say is this: "Recall from your course in Partial Differential Equations that the solution of the equation has the form:"

$$\Theta_m(x) = \sum_{l=0}^{\infty} b_l^m (1 - x^2)^{m/2} \frac{d^m}{dx^m} P_l(x)$$

and reforming,

$$= \sum_{l=0}^{\infty} b_l^m P_l^m(x)$$

where again $K = l(l+1)$ and the now-defined $P_l^m(x)$ are the *associated Legendre functions.*

$$P_l^m(x) = (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x)$$

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x)$$

where $-l \leq m \leq l$.

We now have determined the full spatial dependence in θ and ϕ over the surface of the sphere:

$$\Theta(\theta)\Phi(\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} b_l^m P_l^m(\cos\theta) e^{im\phi}.$$

Still, there is something more to do to take us to a more standard view of the *spherical harmonic decomposition*; we shall obtain the *normalized* coefficients, B_l^m.

We form

$$b_l^m = B_l^m (-1)^m [\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}]^{1/2}$$

so that, now,

$$Y_l^m(\theta,\phi) = (-1)^m [\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}]^{1/2} P_l^m(cos\theta) e^{im\phi},$$

the **normalized surface harmonics**. These manipulations allow us to write, more simply

$$\Theta(\theta)\Phi(\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} B_l^m Y_l^m(\theta,\phi).$$

Now what did we mean by *normalization*? The Y_l^m are normalized as follows; if we integrate the product of any two over the sphere the result is either 0 or 1. It becomes 1 if the indices l and m are the same for the two and 0 otherwise:

$$\int_0^{2\pi} d\phi \int_0^{\pi} Y_l^{m*}(\theta,\phi) Y_j^n(\theta,\phi) \sin\theta d\theta = \delta_{lj} \delta^{mn}.$$

That is the surface product of the the two Y_l^m integrates to unity if $l = j$ and $m = n$ but to 0 otherwise. Note that the Y_l^m are complex-valued through the $e^{im\phi}$ and the $*$ notes the complex-conjugate of the first of the two in the product. It could be the second without any loss of generality. The $Y_l^m(\theta, \phi)$ are said to be **orthonormal** over the sphere.

2.1 Orthonormality on a simple line interval and sphere:

Fourier harmonics: For a function $f(x)$ described over the interval $[0 \le x \le L]$, we decompose it to find the coefficients of the Fourier orthonormal set

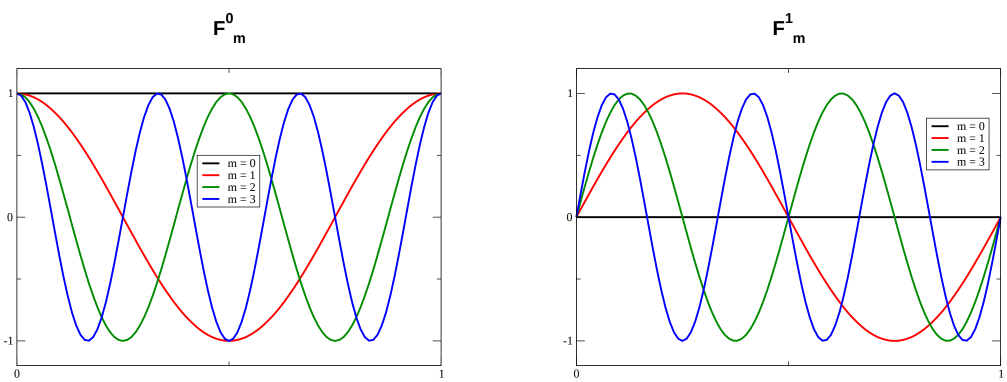
$$F_m^n(x), m = 0, 1, 2..., n = 0, 1.$$

We expand the function as

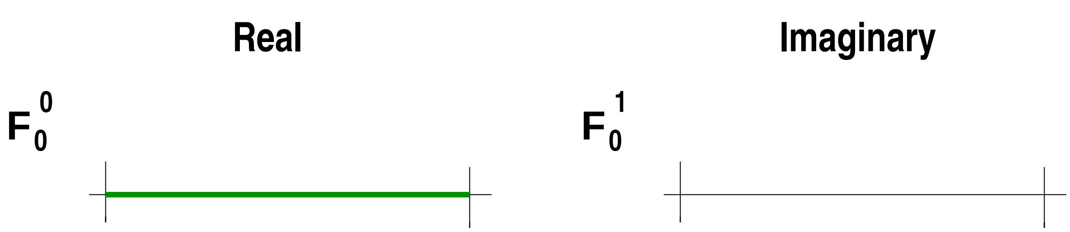
$$f(x) = \sum_{m=0}^{\infty} \sum_{n=0}^{1} a_m^n F_m^n(x),$$

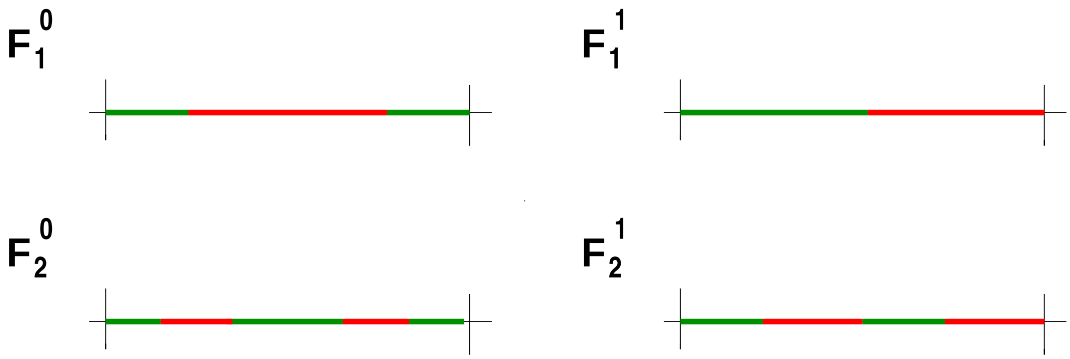
where

$$F_m^0(x) = \cos(2\pi m x/L) \qquad F_m^1 = i \sin(2\pi m x/L).$$



The Fourier type functions





The Fourier type functions (second perspective)

How do we find the appropriate coefficient set a_m^n that fully describe our function $f(x)$?

$$a_m^n = \frac{1}{L}\int_0^L f(x) F_m^n(x) dx.$$

The orthonormality of our special function set is easily shown. Take $f(x)$ to be $F_j^l(x)$ and then seek the coefficient set

$$a_m^n = \frac{1}{L}\int_0^L F_j^l(x) F_m^n(x) dx = \delta_{jm}^{ln}.$$

That is $a_m^n = 1$ if both $j = m$ and $l = n$ but is 0 otherwise.

Spherical harmonic decomposition of a known function, $\Gamma(\theta, \phi)$:

We can obtain the value of each coefficientf B_l^m through recognition of the orthonormality properties of the normalized surface harmonics. If

$$\Gamma(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} B_l^m Y_l^m(\theta, \phi),$$

choosing a specific surface harmonic distribution, $Y_j^n(\theta, \phi)$,

$$\int_0^{2\pi} d\phi \int_0^{\pi} \Gamma(\theta, \phi) Y_j^{n*}(\theta, \phi) \sin\theta d\theta =$$

$$\int_0^{2\pi} d\phi \int_0^{\pi} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} B_l^m Y_l^m(\theta, \phi) Y_j^{n*}(\theta, \phi) \sin\theta d\theta = B_{j=l}^{n=m},$$

the coefficient that scales the contribution of the form $\boldsymbol{Y_j^n(\theta,\phi)}$ to our $\boldsymbol{\Gamma(\theta,\phi)}$. Practically, how do we accomplish this integration to find the $\boldsymbol{B_l^m}$? Rather than calculating $\boldsymbol{Y_j^n(\theta,\phi)}$ on the fly, what is often (normally?) done is to create a surface map of $\boldsymbol{Y_j^n(\theta,\phi)}$ at all points on a regular angular grid (say, every $\boldsymbol{1^\circ}$ in $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$) and multiply the grid value into the function $\boldsymbol{\Gamma(\theta,\phi)}$ measurement and integrate numerically. The numerical integration is easily accomplished. Another approach is to use some rather fancy codes that are specific to spherical harmonic analysis such as ***shaeC*** or ***shagC*** both of which are part of the graphical package *NCL, NCAR Command Language*[6].

To this point in the story, we have only discovered the $\boldsymbol{\theta-\phi}$-dependence on the surface. The Earth has depth and that depth, as well as the temporal variations, is dealt with through the *Radial Wavefunctions*. We have already assumed a temporal dependence of the form $\boldsymbol{T(t)\sim e^{i\omega t}}$ but we haven't found the particular $\boldsymbol{\omega}$ for which this form will hold. That is, we haven't yet found the temporal normal modes. The radial wavefunction is determined by the equation

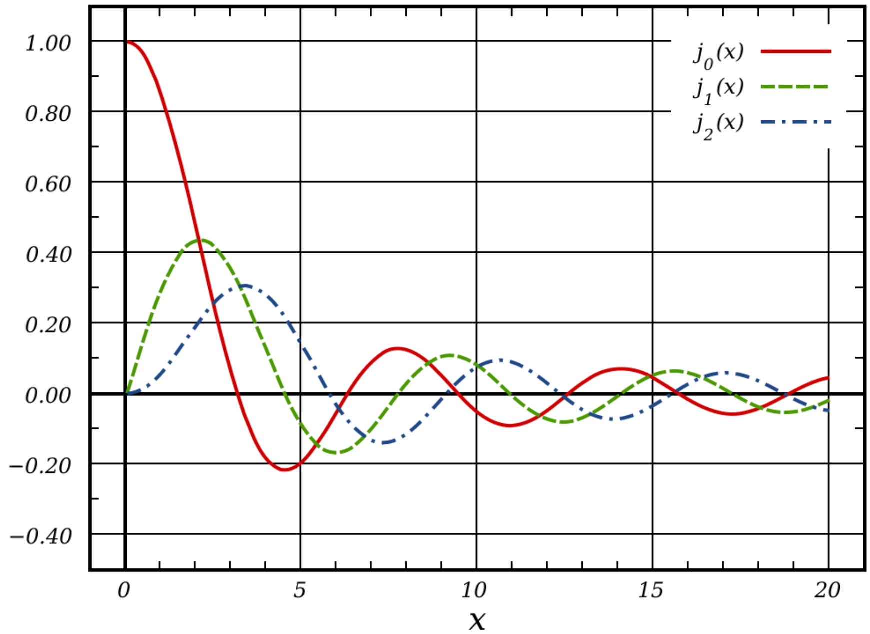
$$\frac{1}{r^2}\frac{d}{dr}(r^2\frac{dR}{dr})+[\frac{\omega^2}{c^2}-\frac{l(l+1)}{r^2}]R=0,$$

for $\boldsymbol{l=0,\pm1,\pm2,...\pm m}$. This equation is satisfied by superpositions of any of three special function sets – the "spherical Bessel functions" of the first, second and third kinds.

- **Spherical Bessel function of the first and second kinds**: The spherical Bessel functions are formed from the circular Bessel functions by a weighting in radius and a normalization.

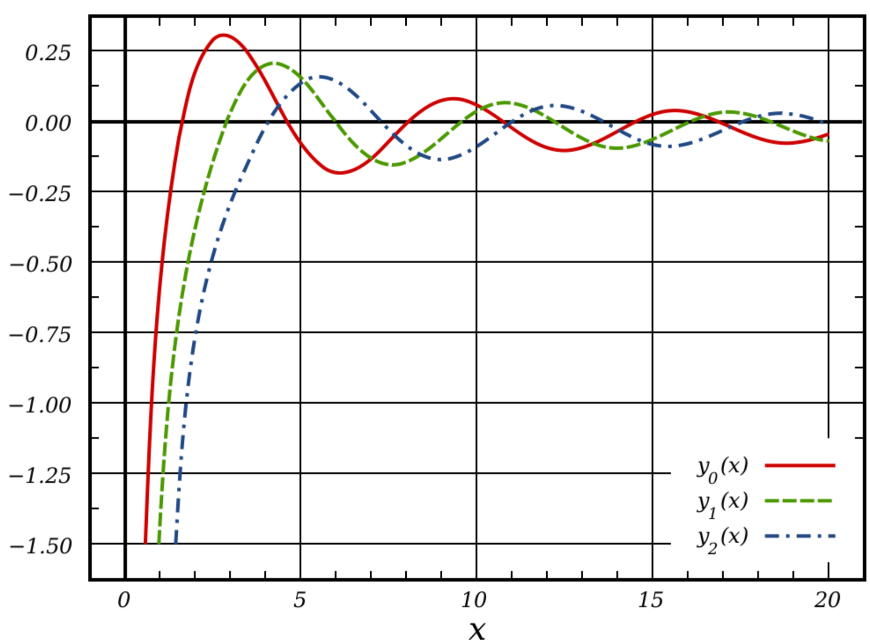
$$j_l(z)=\sqrt{\frac{\pi}{2z}}J_{l+\frac{1}{2}}(z),\quad z=\frac{\omega r}{c},$$

[6] NCAR NCL



Spherical Bessel function of the first kind

$$y_l(z) = \sqrt{\frac{\pi}{2z}} Y_{l+\frac{1}{2}}(z), \quad z = \frac{\omega r}{c}.$$



Spherical Bessel function of the second kind

- **Spherical Bessel functions of the third kind form from complex sums of the first two:**

$$h_l^{(1)}(z) = j_l(z) + iy_l(z)$$
$$h_l^{(2)}(z) = j_l(z) - iy_l(z).$$

Note the analogy of the Bessel function of the first kind to the Fourier $\cos(\alpha)$ and of the second kind to $\sin(\alpha)$ and then the Bessel functions of the third kind might be seen as analogous to the combined Fourier form $e^{\pm i\alpha}$.

Looking for the complete solution for $P(\vec{r}, t)$ while recognizing that pressure is not a complex-valued physical measure, we can immediately discount solutions based upon spherical Bessel functions of the third kind.

$$R(r) = \sum_{l=0}^{\infty} u_l j_l(\frac{\omega r}{c}) + v_r y_l(\frac{\omega r}{c}).$$

As we now pay attention to the radial boundary conditions, note that

$$\lim_{z \longrightarrow 0} y_l(z) \longrightarrow -\infty,$$

so that the $v_l \equiv 0$, and

$$R(r) = \sum_{l=0}^{\infty} u_l j_l(\frac{\omega r}{c}).$$

The spherical Bessel functions, like the associated Legendre functions, form a series

$$j_0(z) = \frac{\sin z}{z} \qquad j_1(z) = \frac{\sin z}{z^2} \qquad j_2(z) = (\frac{3}{z^3} - \frac{1}{z}) \sin z - \frac{3 \cos z}{z^2} ...$$

Rayleigh's generating formulae

$$j_l(z) = z^l (-\frac{1}{z}\frac{d}{dz})^l (\frac{\sin z}{z}),$$

$$y_l(z) = -z^l (-\frac{1}{z}\frac{d}{dz})^l (\frac{\cos z}{z}).$$

We apply the stress (pressure) free surface boundary condtion

$$P(r = r_0, \theta, \phi, t) = 0,$$

$$R(r = r_0) = 0,$$

to find the "eigenfrequencies" of oscillation. Let us first look at the simplest case for $\boldsymbol{l = 0}$,

$$R(r) \sim j_0(\frac{\omega r}{c}),$$

so that

$$\frac{\omega r}{c} R(r) \propto \sin(\frac{\omega r}{c}).$$

$R(r_0) = 0$, which requires that ω be such that

$$R(r_0) \propto \frac{\sin(\frac{\omega r_0}{c})}{(\frac{\omega r_0}{c})}.$$

Recalling l'Hopital's rule for ratios of zero-valued functions, for $\omega = 0$, the ratio for $R(r_0, \omega = 0) \equiv 1$ (i.e. we have no possible solution for $\omega = 0$. but for

$$\frac{\omega r_0}{c} = (n+1)\pi, \quad n = 0, 1, 2, ...$$

$$R(r_0) \equiv 0.$$

The values

$$_n\omega_0 = \frac{(n+1)\pi c}{r_0}, \quad n = 0, 1, 2, ...$$

determine the eigenfrequencies of oscillation of the "Earth's" body which obey the boundary conditions. Here, $n = 0$ determines the "fundamental" mode of oscillation for the $l = 0$ spatial harmonic; $n = 1, 2, ...$ determine the "overtones".

In past courses in *Earthquakes and Earth Structure* I used to leave students with an exercise:

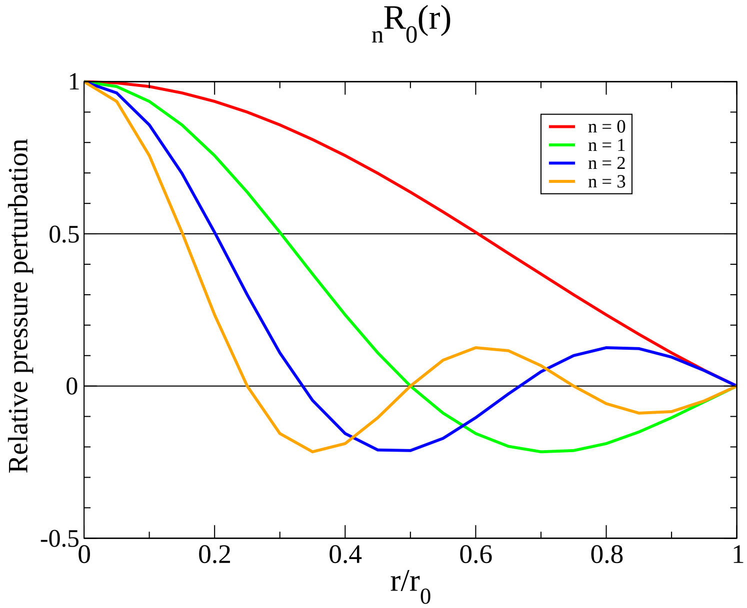
Find the eigenmodes $_n\omega_2$.

Each eigenfrequency is associated with a *radial eigenfunction* to complete the description of the *eigenmode*. Assign the eigenfuntions as

$$_nR_l(r) \propto j_l(\frac{_n\omega_l r}{c}).$$

In order to determine actual amplitude scales, we would have to address an intial condition of excitation. So, here, we restrict our interest to the relative amplitudes as a function of radius among the eigenmodes.

The "radial eigenfunctions" describe the relative perturbation in pressure below the stress-free (pressure-free) surface for each of the multiple infinity of possible "eigenfrequencies".



The $_n\boldsymbol{R}_0$ radial eigenfuntion set.

Note that we have obtained our solution for the radial oscillations in terms of pressure. We might be more interested in the perturbation on the surface position. In the simple model just obtained, this would not be terribly difficult. We know the surface harmonic distribution over the sphere. We need only integrate the volume change on the sphere due to the change in the distribution in pressure from the "Earth's" centre to its surface and then scale by the surface harmonic distributions. This might be seen as a somewhat sloppy approach to finding realistic whole body oscillations. While I don't intend to lecture the next section, it describes one incremental step in taking our description to more realistic models of the Earth. Euphemistically, these are called "Real"-Earth models.

3 Free Oscillations of "Real"-Earth Models

From your courses in Earth Physics and possibly from Earthquakes and Earth Structure, you might recall the relationship between internal particle motions in an elastic solid and the spatial variations in stresses that drives them:

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial}{\partial x_j} p_{ij},$$

where ρ is the local material density, u_i, the $i-$direction component of the particle motion, p_{ij}. The differentiation $\partial / \partial x_j$ obtains the variation in local stress according to direction. You might note that if the material is a fluid, rather than an elastic solid, the material might "flow"; in accounting for flow, we could replace the second partial derivative with respect to time with the second ordinary derivative and then note that

$$\frac{d^2 u_i}{dt^2} \approx \frac{\partial v_i}{\partial t} + v_j \frac{\partial v_i}{\partial x_j}$$

where, now, $v_i = \partial u_i / \partial t$. In our following example analysis of the Earth's free elastic oscillations, even though we model the Earth as an elastic fluid, we shall ignore any flow. This is valid at relatively short periods of oscillation as the fluid materials of the Earth have little time for flow adjustments. It is a reasonably valid simplification for periods of oscillation shorter than a few hours; it begins to seriously fail when we try to analyse the forced oscillations of Earth tides. We won't go there.

Assuming a Hookean solid,

$$p_{ij} = \lambda \frac{\partial u_k}{\partial x_k} \delta^i_j + \mu \left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j} \right).$$

Smylie and Mansinha (1971) noted that the Earth is gravitationally pre-stressed; they included this pre-stress, P_{ij} as independent of the perturbation displacement, u_i, and perturbation stress, p_{ij} and then formed a total stress

$$\sigma_{ij} = P_{ij} - u_k \frac{\partial P_{ij}}{\partial x_k} + p_{ij}.$$

We might regard $P_{ij} = p_o \delta_{ij}$ where p_o is the equilibrium hydrostatic pressure which varies in place according to local density and the gravitational force acting on it:

$$\frac{\partial p_o}{\partial x_i} = \rho_o g_{oi}$$

where ρ_o is the reference local density and g_{oi} the reference local gravitational state which might not be purely radial. In the deformed state

$$\frac{\partial \sigma_{ij}}{\partial x_j} = -\rho g_i - f_i$$

where $\boldsymbol{g_i}$ is the gravitational force and $\boldsymbol{f_i}$ any additional body force per unit volume. Density and gravitational force vary with perturbation of our "locale" as $\boldsymbol{\rho} = \boldsymbol{\rho_o} + \boldsymbol{\rho_1}$ and $\boldsymbol{g_i} = \boldsymbol{g_{oi}} + \boldsymbol{g_{1i}}$ where the **1**s represent the perturbation due to elastic displacements.