SPECTRAL ANALYSIS

8.3 The Periodogram

We determined that the Fourier series expansion of our observed time-series d_i could be written

$$\hat{d}_i = \sum_{j=0}^{\leq n/2} [a_j \cos \omega_j t_i + b_j \sin \omega_j t_i]. \tag{8.54}$$

Remember that (8.2) started out by trying to fit a cosine of arbitrary amplitude A_j and phase ϕ_j, but that we could rewrite this single term as a sum of a cosine and sine components with different amplitudes and zero phases. We found

$$a_j = A_j \cos \phi_j, \quad b_j = A_j \sin \phi_j. \tag{8.55}$$

From these expressions we readily find a component's full amplitude and phase. Dividing the b_j by a_j gives

$$\tan \phi_j = b_j / a_j \quad \Rightarrow \quad \phi_j = \tan^{-1}(b_j / a_j). \tag{8.56}$$

Squaring a_j and b_j and adding them gives

$$A_j^2 = a_j^2 + b_j^2. \tag{8.57}$$

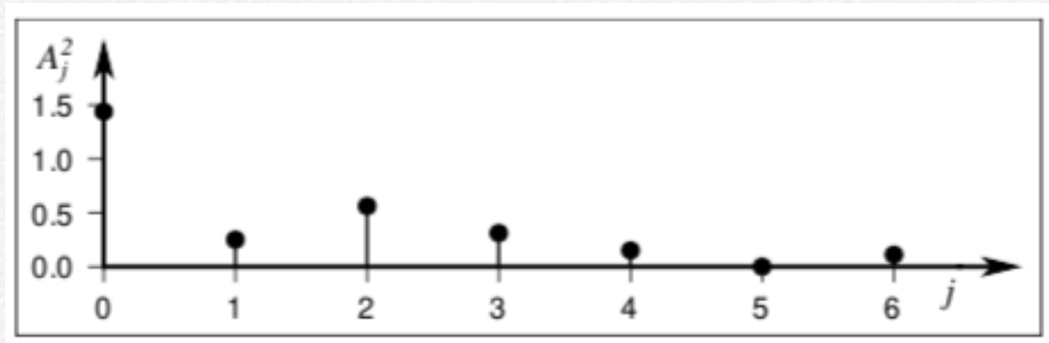


Figure 8.8: Raw periodogram of the function given in (8.58). The peak corresponds to the $A_{20} = a_{20}$ term defined to be twice the mean (-0.6) squared.

The *periodogram* is constructed by plotting A_j^2 versus j, f_j, ω_j, or P_j. While often called the *power spectrum*, it is strictly speaking a raw, discrete periodogram. The true spectrum is a smoothed periodogram showing frequency components of statistical regularity. However, the periodogram is the most common form of output of a Fourier transform. Figure 8.8 shows the periodogram for the function

$$d(t) = \frac{1}{2}\cos \omega_1 t + \frac{3}{4}\cos \omega_2 t + \frac{1}{2}\sin \omega_3 t + \frac{1}{4}\cos \omega_3 t + \frac{1}{3}\cos \omega_4 t + \frac{1}{5}\sin \omega_4 t + \frac{1}{3}\sin \omega_6 t - \frac{3}{5}. \tag{8.58}$$

SPECTRAL ANALYSIS

Let us look, for a moment, at the variance of the time series expansion. Recall, the variance is given by

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\hat{d}_i - \bar{d})^2. \tag{8.59}$$

We shall write the Fourier series as

$$\hat{d}_i = \bar{d} + \sum_{j=1}^{\leq \frac{n}{2}} (a_j \cos \omega_j t_i + b_j \sin \omega_j t_i), \tag{8.60}$$

by pulling the constant (mean) term out separately. Since the two means cancel, we find

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left\{\left[\sum_{j=1}^{\leq \frac{n}{2}} (a_j \cos \omega_j t_i + b_j \sin \omega_j t_i)\right]\left[\sum_{k=1}^{\leq \frac{n}{2}} (a_k \cos \omega_k t_i + b_k \sin \omega_k t_i)\right]\right\}. \tag{8.61}$$

Also recall that, because of orthogonality, all the cross terms ($k \neq j$) resulting from the full expansion of the squared expression will be zero when summed over i, while the remaining terms will sum to $n/2$ (since $j,k > 0$). Hence, we are left with

$$s^2 = \frac{n}{2(n-1)}\sum_{j=1}^{\leq \frac{n}{2}} (a_j^2 + b_j^2) \sim \frac{1}{2}\sum_{j=1}^{\leq \frac{n}{2}} A_j^2. \tag{8.62}$$

Therefore, the power spectrum (periodogram) of (a_j^2 + b_j^2) versus ω_j is a plot showing the contribution of individual frequency components to the total variance of the signal. For this reason, the power spectrum is often called the variance spectrum. However, most of the time it is simply called “the spectrum.” Hence, the Fourier transform converts a signal from the time domain to the frequency domain (or wavenumber domain), where the signal can be viewed in terms of the contribution of the different frequency components of which it is made. The phase spectrum (ϕ_j versus ω_j) shows the relative phase of each frequency component. In general, phase spectra are more difficult to interpret than amplitude (or power) spectra.

SPECTRAL ANALYSIS

8.3.1 Aliasing of higher frequencies

We mentioned before that the highest frequency (or shortest period, or wavelength) that can be estimated from the data is called the Nyquist frequency (or period, or wavelength), given by

$$f_N = f_{n/2} = \frac{1}{2\Delta t}, \quad \omega_N = 2\pi f_N = \frac{\pi}{\Delta t} \quad P_{n/2} = 2\Delta t. \tag{8.63}$$

Higher frequencies, whose wavelengths are less than twice the spacing between sample points cannot be detected. However, when we sample a signal every Δt and the original signal has higher frequencies than $f_{n/2}$, we introduce *aliasing*. Aliasing means that some frequencies will leak power into other frequencies. This concept is readily seen by sampling a high-frequency signal at a spacing larger than the Nyquist interval.

Sampling of the high-frequency signal actually results in a longer-period signal (Figure 8.9). When Clint Eastwood's wagon wheels seem to spin backwards in an old Western movie — that's aliasing: The 24 pictures/sec rate is simply too slow to capture the faster rotation of the wheel.

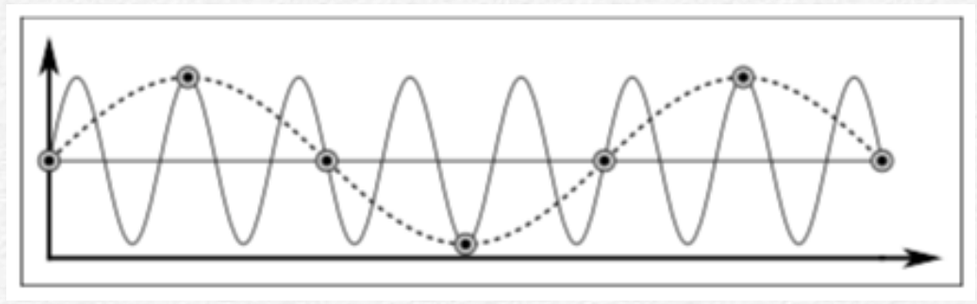


Figure 8.9: Aliasing: A short-wavelength signal that is not sampled at the Nyquist frequency or higher will instead appear as a longer-wavelength component that does not exist in the actual data.

8.3.2 Significance of a spectral peak

In some applications we may be interested in testing whether a particular component is dominant or if its larger amplitude is due to chance. The statistician R. A. Fisher devised a test that calculates the probability that a spectral peak s_j^2 will exceed the value σ_j^2 of a hypothetical time series composed of independent random points. We must evaluate the ratio of the variance contributed by the maximum peak to the entire data variance:

$$g = \frac{s_j^2}{2s^2}, \tag{8.64}$$

where s_j^2 is the largest peak in the periodogram (we divide by two to get its variance contribution) and s^2 is the variance of the entire series.

SPECTRAL ANALYSIS

For a prescribed confidence level, α, the critical value that we wish to compare to our observed *g is*

$$g_{\alpha,m} \approx 1 - \exp\left(\frac{\ln\alpha - \ln m}{m-1}\right), \tag{8.65}$$

with $m = n/2$ (for even n) or $m = (n-1)/2$ (for odd n). Should our observed g (obtained via 8.64) exceed this critical value we may decide that the dominant component is real and reflects a true characteristic of the phenomenon we are observing. Otherwise, s_j^2 may be large simply by chance.

8.3.3 Estimating the continuous spectrum

The power spectrum or periodogram obtained from the Fourier coefficients is discrete, yet we do not expect the power at frequency ω_j to equal the underlying continuous $P(\omega)$ at exactly ω_j, since the discrete spectrum must necessarily represent some average value of power at all frequencies between ω_{j-1} and ω_{j+1}. In other words, the computed power at ω_j also represents the power from nearby frequencies not among the chosen harmonic frequencies ω_j . Furthermore, the uncertainty in any individual estimate p_j^2 is very large; in fact, it is equal to ± p_j^2 itself.

Can we improve (i.e., reduce) the uncertainties in p_j^2 by using more data points or sample the data more frequently? The unpleasant answer is that the periodogram estimates do not become more accurate at all! The reason for this is that adding more points simply produces power estimates at a greater number of frequencies ω_j. The only way to reduce the uncertainty in the power estimates is to smooth the periodogram over nearby discrete frequencies. This can be achieved in one of two ways:

1. Use a time-series that is *M* times longer (so $f_1' = f_1/M$) and *sum* the *M* power estimates p_j^2 straddling each original ω_j frequency to obtain a smooth estimate $p_j^2 = \sum p_k^2$.

2. Split the original data into *M* smaller series, find the p_j^2 for each series, and take the *mean* of the *M* estimates for the same *j* (i.e., the same frequency).

SPECTRAL ANALYSIS

In both cases the variance of the power spectrum estimates drop by a factor of *M*, i.e., $s_j^2 = p_j^2 / M$. The exact way the smoothing is achieved may vary among analysts. Several different types of weights or spectral *windows* have been proposed, but they are all relatively similar. These windows arose because, historically, the power spectrum was estimated by taking the Fourier transform of the *autocorrelation* of the data; hence many windows operated in the lag-domain. The introduction of the Fast Fourier Transform made the FFT the fastest way to obtain the spectrum, which then is simply smoothed over nearby frequencies. The FFT is a very rapid algorithm for doing a discrete Fourier transform, provided *n* is a power of 2. It can be shown that one can always split the discrete transform into the sum of two discrete, scaled transforms of subsets of the data. Applying this result recursively, we eventually end up with a sum of transforms of data sets with one entry, whose transform equals itself. While mathematically equivalent, there is a huge difference computationally: While the discrete Fourier transform executes proportional to n^2, the FFT only takes $n \cdot \log(n)$. For a data set of 10^6 points, the speed-up is a factor of > 75, 000.

By doing a Fourier Analysis, we have transformed our data from one domain (time or space) to another (frequency or wavenumber). A physical analogy is the transformation of light sent through a triangular prism. White light is composed of many frequencies, and the prism acts as a frequency analyzer that separates the various frequency components, here represented by colors. Each color band is separated from its neighbor by an amount proportional to their difference in wavelength, and the intensity of each band reflects the amplitude of that component in the white light. We know that by examining the spectrum we can learn much about the composition and temperature of the source and the material the light passed through. Similarly, examining the power spectra of other processes may tell us something about them that may not be apparent in the time domain. Consequently, spectral analysis remains one of the most powerful techniques we have for examining temporal or spatial sequences.

8.4 Convolution

Convolution represents one of the most fundamental operations of time series analysis and is one of the most physically meaningful. Consider the passage of a signal through a linear filter, where the filter (a “black box”) will modify a signal passing through it (Figure 8.10). For instance, it may

1. Amplify, attenuate or delay the signal.
2. Modify or eliminate specific frequency components

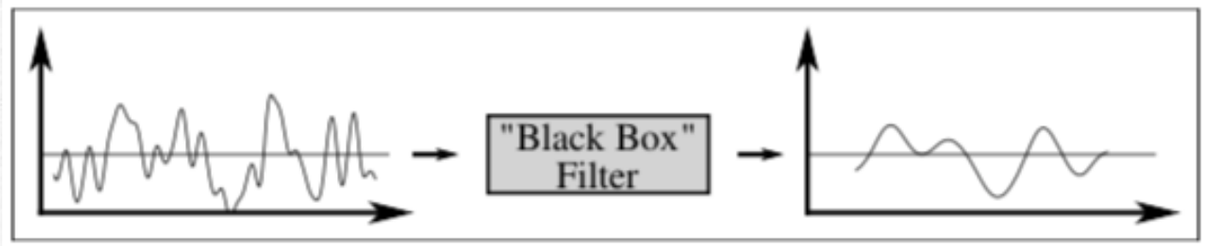


Figure 8.10: Example of convolution between an input signal and a filter.

SPECTRAL ANALYSIS

Consider the propagation of a seismic pulse through the upper layers of the Earth's crust, as illustrated in Figure 8.11. The generated pulse may be sharp and thus have high frequencies, yet the recorded signal that traveled through the crust may be much smoother and include repeating signals that reflect internal boundaries.

Figure 8.11: Convolving a seismic pulse with the Earth gives a seismic trace that may reflect changing properties of the Earth with depth.

Convolution is this process of linearly modifying one signal using another signal. In Figure 8.11 we convolved the seismic pulse with the "Earth filter" to produce the observed returned seismogram. Symbolically, we write the convolution of a signal $d(t)$ by a filter $p(t)$ as the integral

$$h(t) = d(t) * p(t) = \int_{-\infty}^{+\infty} d(u) \cdot p(t-u)du. \qquad (8.66)$$

Deconvolution, or *inverse filtering* is the process of unscrambling the convolved signal to determine the nature of the filter or the nature of the input signal. Consider these two cases:

1. If we knew the exact shape of our seismic pulse $d(t)$ and seismic signal received, $h(t)$, we could deconvolve the data with the pulse to determine the (filtering) properties of the upper layers of the Earth through which the pulse passed (i.e., $p(t) = d^{-1}(t) * h(t)$).
2. If we wanted to determine the exact shape of our pulse $d(t)$, we could pass it through a known filter $p(t)$ and deconvolve the output with the shape of the filter (i.e., $d(t) = p^{-1}(t) * h(t)$).

The hard work here is to determine the inverse functions $d^{-1}(t)$ or $p^{-1}(t)$, which is akin to matrix inversion. Other examples of convolution include:

1. Smoothing data — with running means, weighted means, removing specific frequency components, etc.
2. Recording a phenomenon with an instrument that responds slower than the rate at which the phenomenon changes, or which produces a weighted mean over a narrow interval of time, or which has lower resolving power than the phenomenon requires.

SPECTRAL ANALYSIS

3. Conduction and convection of heat.
4. Deformation and the resulting gravity anomalies caused by the flexural response of the lithosphere to a seamount load.

Convolution is most easily understood by examining its effect on discrete functions. First, consider the discrete impulse $d(t)$ sent through the filter $p(t)$, as illustrated in Figure 8.12:

The output $h(t)$ from the filter is known as the *impulse response function* since it represents the response of the filter to an impulse, $d(t)$. It represents a fundamental property of the filter $p(t)$. Next, consider a more complicated input signal convolved with the filter, as shown in Figure 8.13:

Since the filter is linear, we may think of the input as a series of individual impulses. The output is thus the sum of several impulse responses scaled by their amplitudes and shifted in time. Calculating convolutions is a lot like calculating cross-correlations, except that the second time-series must be reversed. Consider the two signals as finite sequences on separate strips of paper (Figure 8.14):

We obtain the zero lag output by aligning the paper strips as shown in Figure 8.15, after reversing the red strip.

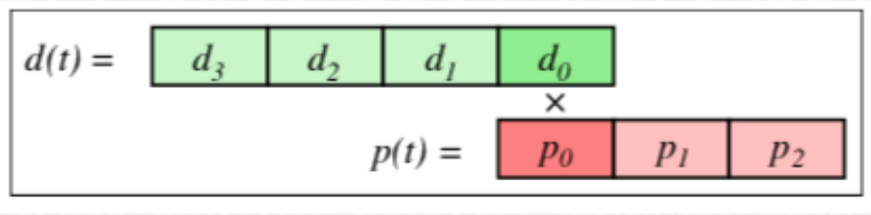


Figure 8.15: Convolution, zero lag. Reverse one strip and arrange them to yield a single overlap.

Figure 8.12: A filter's impulse response is obtained by sending an impulse $d(t)$ through the filter $p(t)$.

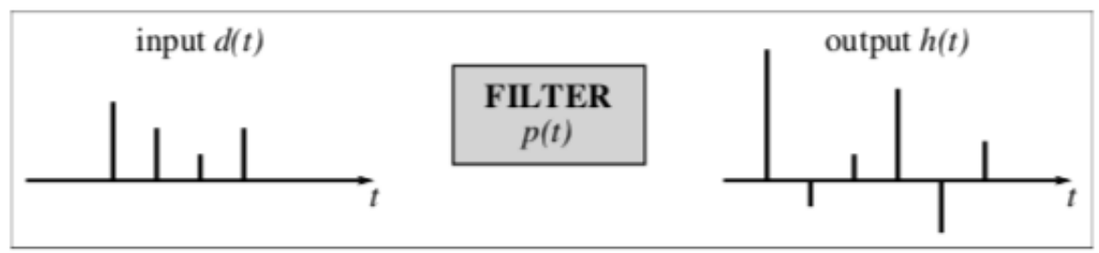


Figure 8.13: Filtering seen as a convolution.

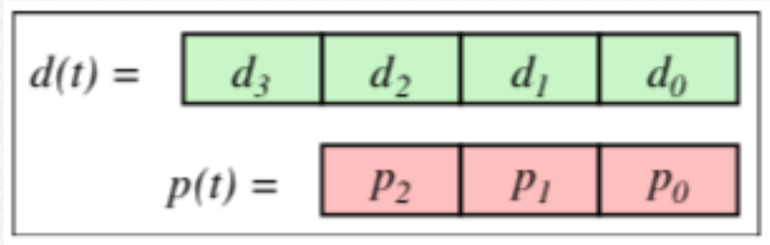


Figure 8.14: Graphical representation of a convolution. We write the discrete values of $d(t)$ and $p(t)$ on two separate strips of paper

SPECTRAL ANALYSIS

The zero lag result h_0 is thus simply $d_0 \cdot p_0$. Moving on, the first lag results from the alignment shown in Figure 8.16:

This simple process is repeated, and for each lag k we evaluate h_k as the sum of the products of the overlapping signal values. This is a graphic (or mechanical) representation of the discrete convolution equation (compare this operation to the integral in 8.66). Consider the convolution of the two functions shown in Figure 8.17.

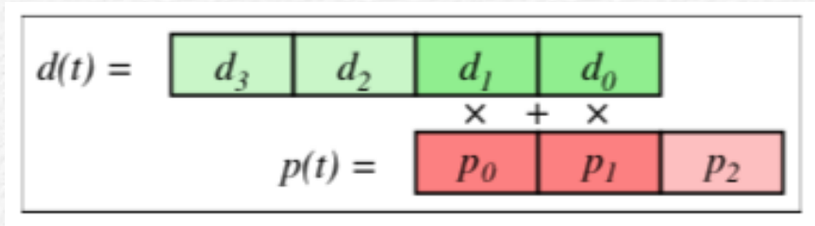


Figure 8.16: Convolution, first lag. We shift one strip by one to increase the overlap.

Given the simple nature of $p(t)$, we can estimate the values of h_k directly:

$$
\begin{aligned}
h_0 &= d_0/5 \\
h_1 &= \tfrac{1}{5}(d_0+d_1) \\
&\vdots \\
h_4 &= \tfrac{1}{5}(d_0+d_1+d_2+d_3+d_4) \\
h_5 &= \tfrac{1}{5}(d_1+d_2+d_3+d_4+d_5) \\
&\vdots \\
h_{18} &= d_{14}/5
\end{aligned}
\qquad (8.67)
$$

This is simply a five-point running (or moving) average of $d(t)$, and the result is shown in Figure 8.19. An n-point average would be the result if $p(t)$ consisted of n points, each with a value of $1/n$.

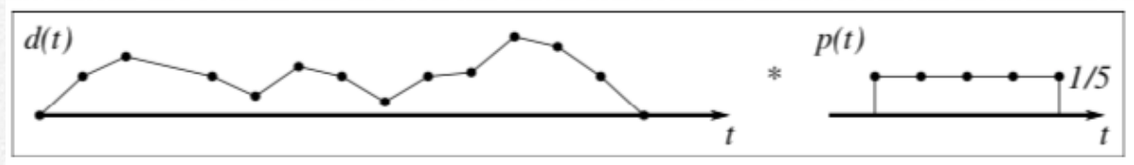


Figure 8.17: Moving averages is obtained by the convolution of data with a rectangular function of unit area.

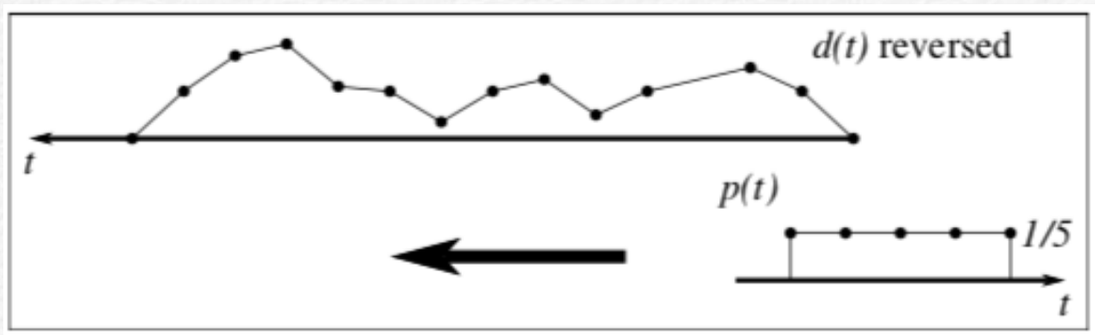


Figure 8.18: The mechanics of convolutions, this time without the paper strips.

SPECTRAL ANALYSIS

8.4.1 Convolution theorem

Although not shown here, it can be proven that a convolution of two functions $p(t)$ and $d(t)$ in the time-domain is equivalent to the product of $P(f)$ and $D(f)$ in the frequency domain (here, uppercase letters indicate the Fourier transforms of the lowercase, time-domain functions). The converse is also true, thus

$$\begin{aligned} p(t)*d(t) &= h(t) \quad \leftrightarrow \quad P(f)\cdot D(f) = H(f), \\ p(t)\cdot d(t) &= z(t) \quad \leftrightarrow \quad P(f)*D(f) = Z(f). \end{aligned} \tag{8.68}$$

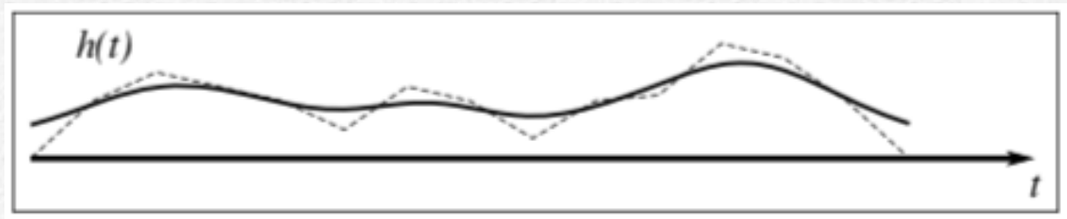


Figure 8.19: The final result of the convolution is a smoothed data set since any short-wavelength signal will be greatly attenuated.

Because convolution is a slow calculation it is often advantageous to transform our data from one domain to the other, perform the simpler multiplication, and transform the data back to the original domain. The availability of *fast Fourier transforms* (FFTs) makes this approach practical.

8.5 Sampling Theory

The *sampling theorem* states that if a function is *band-limited* (i.e., the transform is zero for all radial frequencies $f > f_N$), then the continuous function $d(t)$ can be uniquely determined from knowledge of its sampled values given a sampling interval $\Delta t \le 1/(2 f_N)$. From distribution theory, we have

$$d_t = \sum_{j=-\infty}^{+\infty} d(t)\delta(t - j\Delta t) = \sum_{j=-\infty}^{\infty} d(j\Delta t)\delta(t - j\Delta t) = d(t)\cdot\Delta(t), \tag{8.69}$$

where

$$\Delta(t) = \sum_{j=-\infty}^{+\infty} \delta(t - j\Delta t) \tag{8.70}$$

is the sampling or “comb” function in the time domain (Figure 8.20).

SPECTRAL ANALYSIS

Thus, *dt* is the continuous function *d*(*t*) sampled at the discrete times *j*Δ*t*. Consequently, it is true that the original signal *d*(*t*) can be reconstructed exactly from its sampled values *dt* via the *Whittaker-Shannon* interpolation formula

$$d(t) = \sum_{j=-\infty}^{+\infty} d_j \,\mathrm{sinc}\left(\frac{t - j\Delta t}{\Delta t}\right), \qquad (8.71)$$

where $d_j = d(j\Delta t)$ are the sampled data values and the sinc function is defined as

$$\mathrm{sinc}(x) = \frac{\sin \pi x}{\pi x}. \qquad (8.72)$$

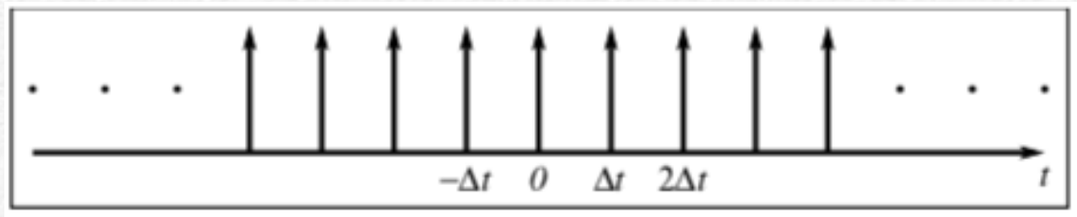


Figure 8.20: The sampling or "comb" function, Δ(*t*), represents mathematically what we do when we sample a continuous phenomenon *d*(*t*) at discrete, equidistantly spaced times.

Recall that the multiplication of two functions in the time domain is equivalent to the convolution of their Fourier transforms in the frequency domain, hence

$$d(t) \cdot \Delta(t) \leftrightarrow D(f) * \Delta(f). \qquad (8.73)$$

The time-domain expression is visualized in Figure 8.21.

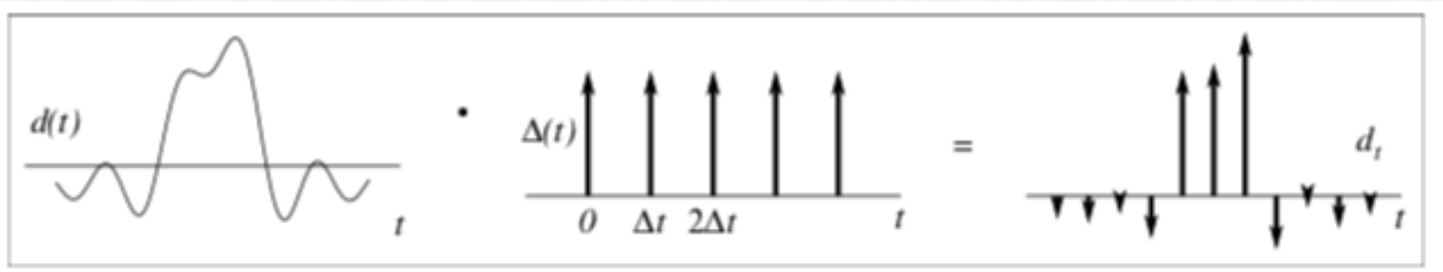


Figure 8.21: Sampling equals multiplication of a continuous signal *d*(*t*) with a comb function Δ(*t*) in the time-domain.

SPECTRAL ANALYSIS

The transformed function $\Delta(f)$ can be shown to be a series of impulses as well (Figure 8.22).

In the frequency domain, $d(t)$ is represented as $D(f)$ and illustrated in Figure 8.23. We note that while the time- domain comb function $\Delta(t)$ is a series of impulses spaced every Δt, the frequency-domain comb function $\Delta(f)$ is also a series of impulses but spaced every $1/\Delta t$. The time and frequency domain spacings of the comb functions are thus reciprocal: A finer sampling interval leads to a larger distance between the impulses in the frequency domain.

Given $D(f)$ and $\Delta(f)$,$D(f)*\Delta(f)$ is schematically shown in Figure 8.24.

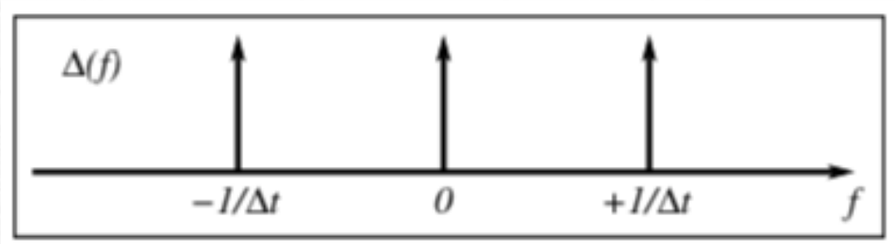


Figure 8.22: The Fourier transform of the comb function, $\Delta(t)$, is another comb function, $\Delta(f)$, with a spacing of $1/\Delta t$ between impulses.

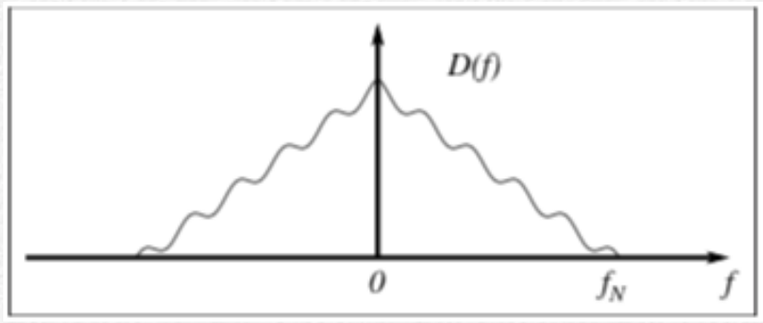


Figure 8.23: The Fourier transform of our continuous phenomenon, $d(t)$. We assume it is band-limited so that the transform goes to zero beyond the highest frequency, f_N .

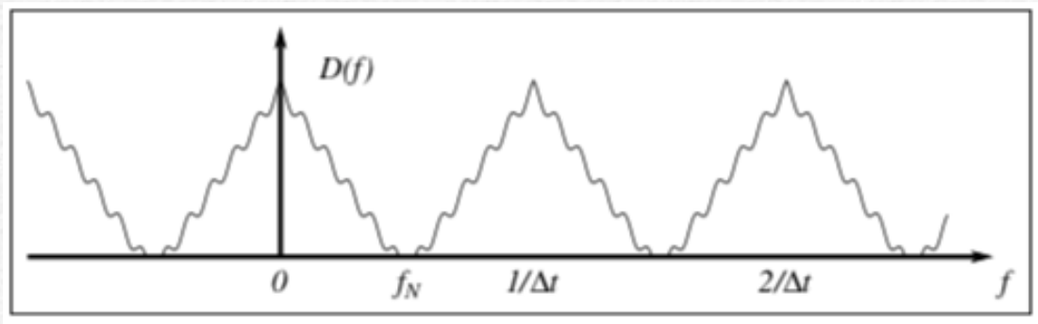


Figure 8.24: Replication of the transform, $D(f)$, due to its convolution with the comb function, $\Delta(f)$.

If the impulses in $\Delta(f)$ are spaced closer than $1/\Delta t$ then there will be some overlap between the $D(f)$ replicas that are centered at the location of each impulse (see Figure 8.25).

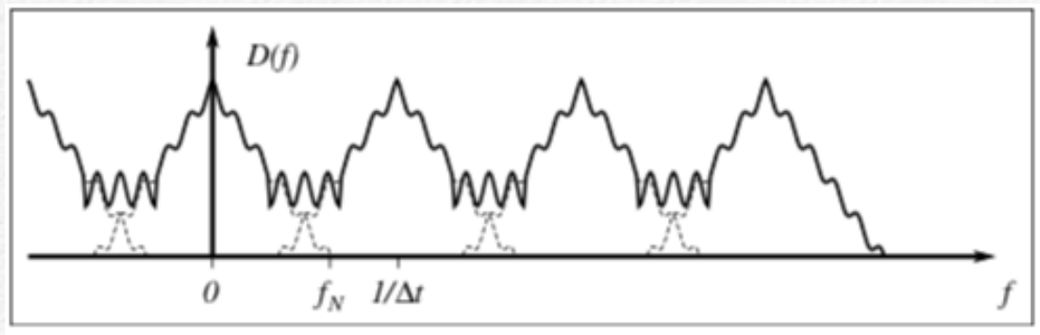


Figure 8.25: Aliasing in the frequency domain occurs when the sampling interval Δt is too large.

SPECTRAL ANALYSIS

This overlap introduces *aliasing* (which we shall discuss more later). To prevent aliasing, we must ensure $\Delta t \le 1/(2f_N)$, where f_N is the highest (radial) frequency component present in the time series. We call f_N the *Nyquist frequency* and the Nyquist sampling interval is $\Delta t = 1/(2 f_N)$, hence $f_N = 1/(2\Delta t)$. As long as we follow the sampling theorem and select $\Delta t \le 1/(2 f_N)$, with f_N being the highest frequency component, there will be no spectral overlap in $D(f) * \Delta(f)$ and we will be able to recover $D(f)$ completely. Therefore (and to prove the sampling theorem) we recover $D(f)$ by truncating the signal:

$$D(f) = [D(f) * \Delta(f)] \cdot H(f), \qquad (8.74)$$

which is illustrated in Figure 8.26 as a multiplication of the replicating spectrum with a *gate* function, $H(f)$.

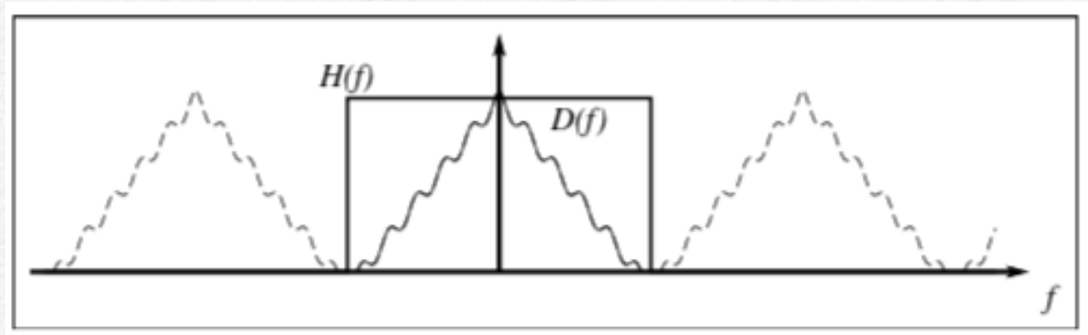


Figure 8.26: Truncation of the Fourier spectrum via multiplication with a rectangular gate function, $H(f)$.

8.5.1 Aliasing, again

Aliasing can be viewed from several angles. Conceptually, if $\Delta t > 1/(2 f_N)$ (where f_N is highest frequency component in phenomenon of interest) then a high frequency component will *masquerade* in the sampled series as a lower, artificial frequency component, as shown in Figure 8.27. If Δt is a multiple of P (e.g., see the squares in Figure 8.27), then this frequency component is indistinguishable from a horizontal line (i.e., a constant, with frequency $f = 0$). If $\Delta t = 5P/4$ (see circles in Figure 8.27) then this frequency component is indistinguishable from a component with frequency $1/5P$ (i.e., period of $5P$). Therefore, the under-sampled frequency components manifest themselves as lower frequency components (hence the word alias). In fact, every frequency *not* in the range

$$0 \le f \le 1/(2\Delta t) \qquad (8.75)$$

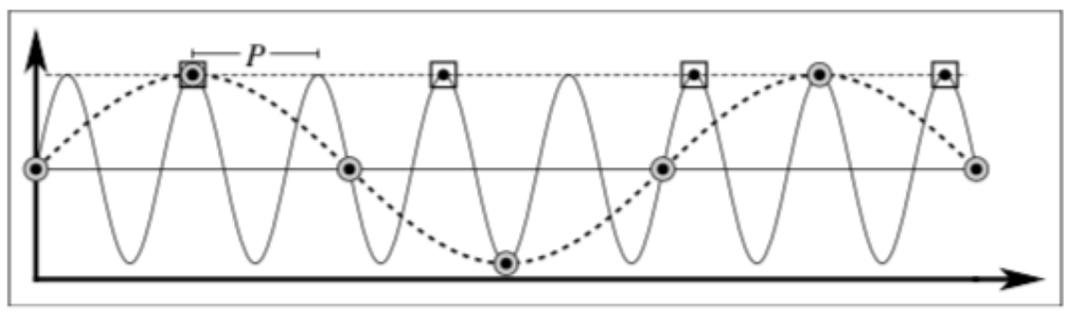


Figure 8.27: Aliasing as seen in the time domain. Thin line shows a phenomenon with period P. The circles and heavy dashed line show the signal obtained using a sampling rate of $1.25P$, while the squares and blue dashed line show a constant signal ($f = 0$) obtained with a sampling rate of $2P$.

has an alias in that range — this is its *principal alias*. Furthermore, any frequency $f_H > f_N$ will be indistinguishable from its principal alias. That is, the actual frequency $f_H = f_N + \Delta f$ will appear as the aliased frequency $f_L = f_N - \Delta f$.

SPECTRAL ANALYSIS

Because of this relationship, the Nyquist frequency (f_N) is often called the *folding frequency* since the aliased frequencies ($f > f_N$) have their principal aliases folded back into the range $\leq f_N$ (Figure 8.28). Therefore, when computing the transform of a data set, any frequency components in the phenomenon with true frequencies $f > f_N$ have been folded back into the resolved frequency range during sampling. Consequently, we must carefully choose Δt so that the powers at frequencies $f' > f_N$ are either small or nonexistent, or we must ensure that f_N is high enough so that the aliased part of the spectrum only affects frequencies higher than those of interest ($f \leq fI$, see Figure 8.29).

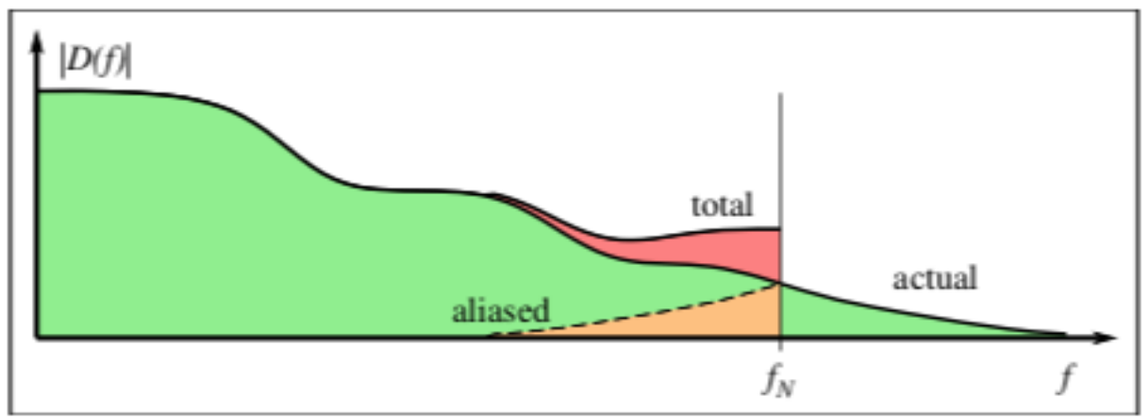


Figure 8.28: Aliasing and folding frequency. Power at higher frequencies than the Nyquist (f_N) will reappear as power at lower frequencies, "folded" around f_N . This extra power (orange) is then added to the actual power and the result is a distorted, total power spectrum (red).

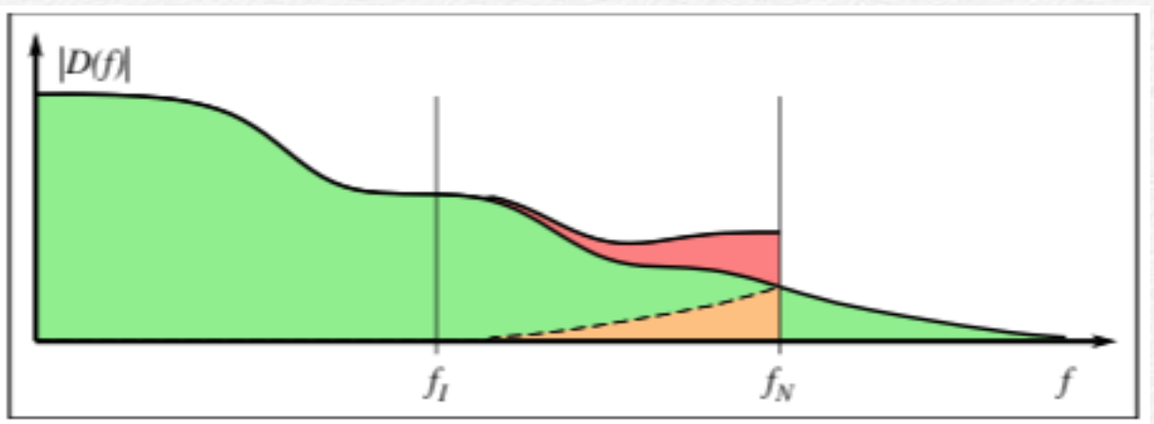


Figure 8.29: Selecting the Nyquist frequency so that aliasing only affects frequencies higher than the frequencies of interest ($f \leq f_I$). In this case, the extra power (orange) that is folded around f_N does not reach into the lower frequencies of interest, and consequently the total spectrum is unaffected for frequencies lower than f_I .

8.6 Aliasing and Leakage

We were exploring the relationship between the continuous and discrete Fourier transform and found that we could illustrate the process graphically. First, we found that we had to sample the time-series $d(t)$ (Figure 8.30a). The sampling of the phenomenon by the sampling function $\Delta(t)$ (Figure 8.30b) is a multiplication in the time-domain, which implies a convolution in the frequency domain. This sampling yields discrete observations in the time domain, but the corresponding convolution in the frequency domain enforces periodicity of the spectrum (Figure 8.30c). Depending on the chosen sampling interval we may or may not have spectral overlap (aliasing). This discrete infinite series must then be truncated to contain a finite number of observations. The truncation is conceptually performed by multiplying our infinite time series with a finite gate function.

SPECTRAL ANALYSIS

This truncation of the infinite and periodic signal amounts to a multiplication in the time domain with a gate function, $h(t)$, whose transform is

$$H(f) = \mathrm{sinc}(fT) = \frac{\sin \pi f T}{\pi f T}, \tag{8.76}$$

with both functions displayed in Figure 8.30d. This process results in the finite discrete observations shown in Figure 8.30e. It is this truncation that is responsible for introducing *leakage*.

Leakage arises because the truncation implicitly assumes that the time-series is periodic with period T (Figure 8.30f). Consequently, the discretization of frequencies is equivalent to enforcing a periodic signal (Figure 8.30g). Because both the time and frequency domain functions have been convolved with a series of impulses (by $\Delta(t)$ in time and $\Delta(f)$ in frequency), both functions are periodic in n discrete values, so the final discrete spectrum (for a real series as shown here) between 0 and f_N represents the discrete transform of the series on the left (which is periodic over T).

If the procedure in Figure 8.30 is followed mathematically, it is seen that the continuous Fourier transform is related to the discrete Fourier transform by the steps outlined graphically above. These show that a discrete Fourier transform can differ from the continuous one by two effects:

1. Aliasing — from discrete time domain sampling.
2. Leakage — from finite time domain truncation.

Aliasing can be prevented by choosing $\Delta t \leq 1/(2 f_N)$ or reduced as discussed previously. Leakage is always a problem for most observed (and hence truncated) time series. As discussed, leakage arises from truncation in the time domain, which corresponds to a convolution with a sinc function in the frequency domain. Conceptually, consider the effect of time domain truncation (Figure 8.31). Fourier analysis is essentially fitting a series of sines and cosines (using the harmonics of the fundamental frequency $1/T$) to the series $d(t)$. Since the Fourier series is necessarily periodic, it follows that

$$d(T/2 + \Delta t) = d(-T/2). \tag{8.77}$$

In other words, the transform is equivalent to that of a time series in which $d(t)$ is repeated every T (Figure 8.32).

SPECTRAL ANALYSIS

The leakage (conceptually) thus results from the frequency components that must be present to allow the discontinuity, occurring every *T*, to be fit by the Fourier series. If the series *d*(*t*) is perfectly periodic over *T* then there is no leakage because *d*(*T* +Δ*t*) = *d*(Δ*t*) and the transition will be continuous and smooth across *T*.
To minimize leakage we attempt to minimize the discontinuity (between *d*(0) and *d*(*T*)) or minimize the lobes of the sinc(*f T*) function convolving the spectrum. This is accomplished by truncating the time series with a more gently sloping gate function (called a taper, fader, window, etc.). In other words, we use a smoother function that has fewer high frequency components (Figure 8.33).

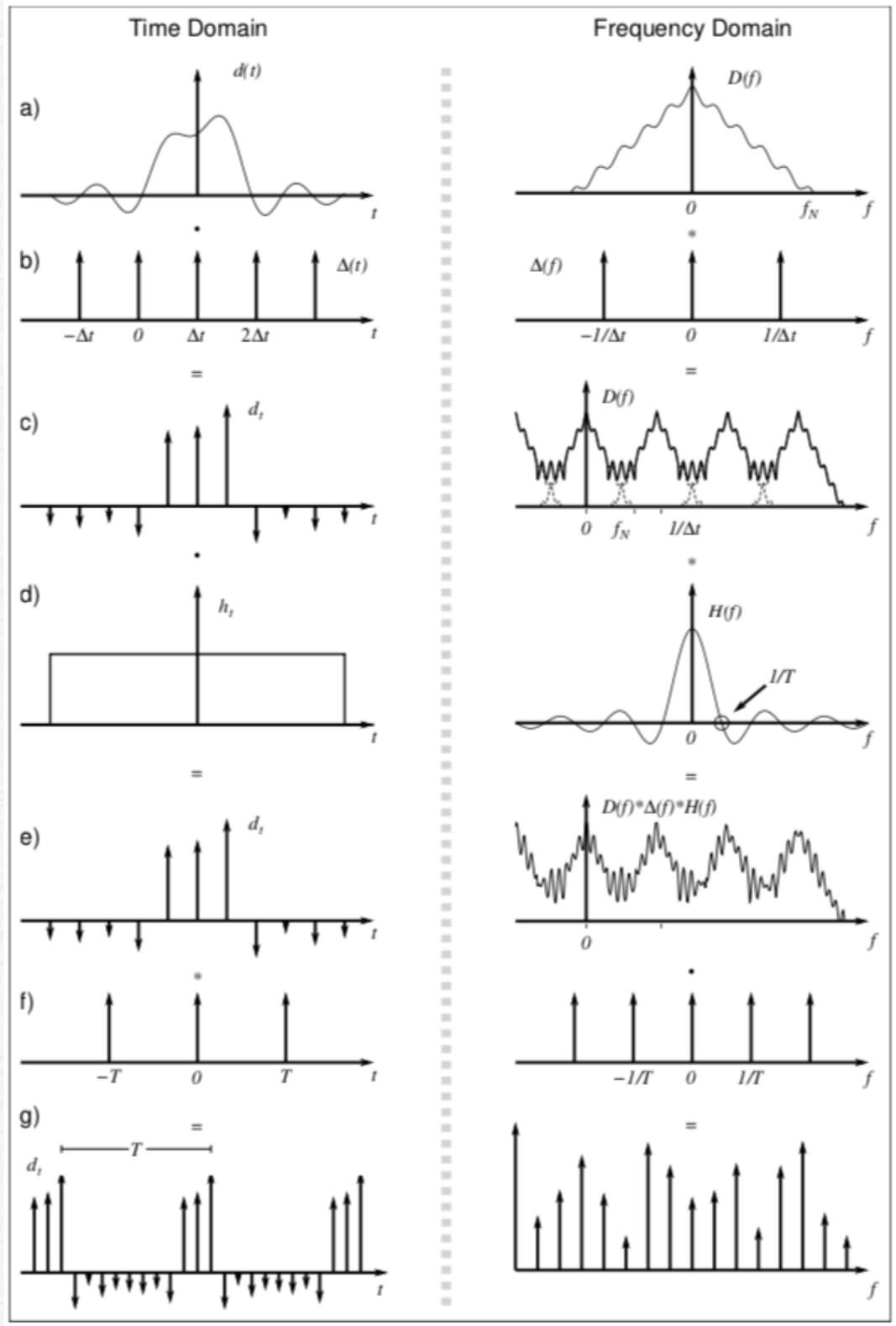


Figure 8.30: The continuous and band-limited phenomenon of interest, represented both in the time and frequency domains. Left column represents the time domain and the right column represents the frequency domain, separated by a vertical dashed gray line. The multiply, convolve, and equal signs indicate the operations that are being performed. a) Continuous phenomenon, b) Sampling function, c) Infinite discrete observations, d) Gate function, e) Truncated discrete observations, f) Assumed periodicity *T* , g) Aliasing and leakage of signal.

SPECTRAL ANALYSIS

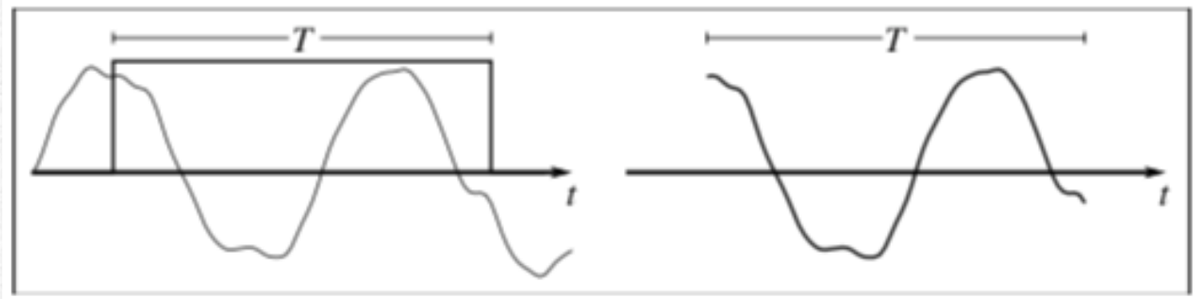


Figure 8.31: Truncation of a continuous signal, the equivalent of multiplying the signal with a gate function $h(t)$, determines the fundamental frequency, $f = 1/T$.

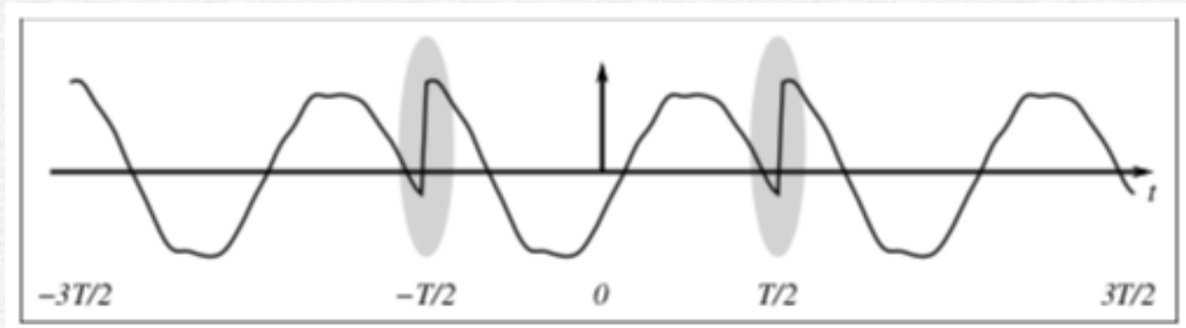


Figure 8.32: Artificial high frequencies are introduced due to the forced periodicity of a truncated time-series, which produces a discontinuous signal (highlighted by the gray regions).

The triangular function is the *Bartlett* window, which is the rectangle function convolved with itself (hence its transform is $\mathrm{sinc}^2(fT)$). The dashed line is the split cosine-bell window. Other windows include: *Hanning* (a cosine taper), *Parzen* (similar to Hanning but decays sooner and more steeply, *Hamming* (like Hanning), and *Bartlett- Priestley* (which is quadratic and has "optimal" properties, satisfying specific error considerations.) All of these tapers have transforms that are less oscillatory than the sinc function but they are also wider. Therefore, multiplication of the time series with one of these gate functions results in a convolution with its transform in the frequency domain that will smear spectral peaks more than the sinc function did. In return, it will not introduce ripples far away from these spectral peaks.

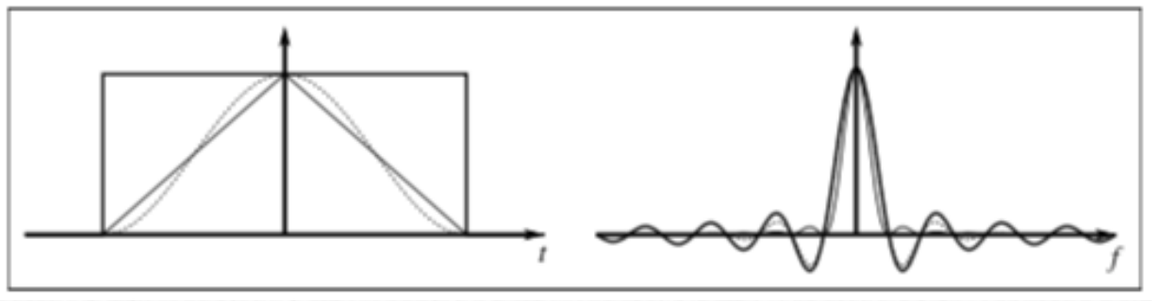


Figure 8.33: Alternative gate functions and their spectral representations. The less abrupt a gate function is in the time domain the less ringing it will introduce in the frequency domain.

Note that multiplying by, say, a Hanning window will make $d(T/2+\Delta t) \sim d(-T/2)$, so the bothersome discontinuity is eliminated — however damping of all $d(t)$ away from $d(T)$ acts like a modulation, which accounts for the smearing of spectral peaks. Hence, leakage is still not completely eliminated.

SPECTRAL ANALYSIS

8.7 Filtering

Filtering of data is typically performed in order to either smooth the signal or suppress power at particular frequencies or wavenumbers. So far, we have learned that filtering can be considered an example of a convolution between the data and the filter, i.e.,

$$h(t) = d(t) * p(t). \tag{8.78}$$

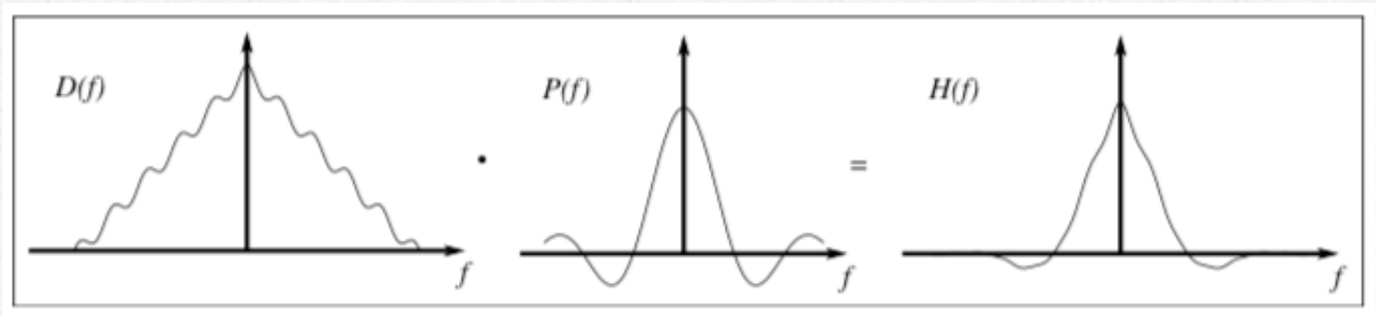
Thus, we can immediately take advantage of the convolution theorem and write

$$H(f) = D(f) \cdot P(f). \tag{8.79}$$

Hence, we can simply take the Fourier transform of $d(t)$ and multiply its spectral components $D(f)$ by $P(f)$. For example, we learned previously that a simple MA filter consists of convolving the signal with a rectangle function of width W (Figure 8.34). Figure 8.35 shows how this operation might look like in the frequency domain.

Figure 8.34: Moving average filter convolution seen in the time domain.

Figure 8.35: Moving average filtering seen in the frequency domain.



Thus, the MA filter output is obtained by multiplying the signal's Fourier coefficients by a sinc function. This means that some of the coefficients will have their signs reversed, which we know equates to a phase change of 180°. Also, while the filter coefficients do fall off with increasing f , they do so in a very oscillatory way and take a long time to approach zero. This fact would suggest that the sinc function is a poor choice for a filter if you really wanted to cut out, say, high frequency information.

SPECTRAL ANALYSIS

Perhaps we should instead design $P(f)$ directly so that it truncates all power at frequencies higher than f_{cut}? Surely, by simply removing all power from frequencies of no interest we should retain the part of the signal that we are most interested in. This alternative operation, which is simply carried out in the frequency domain, is illustrated in Figure 8.36. The filtered output is finally obtained by transforming $H(f)$ back to the time domain (Figure 8.37).

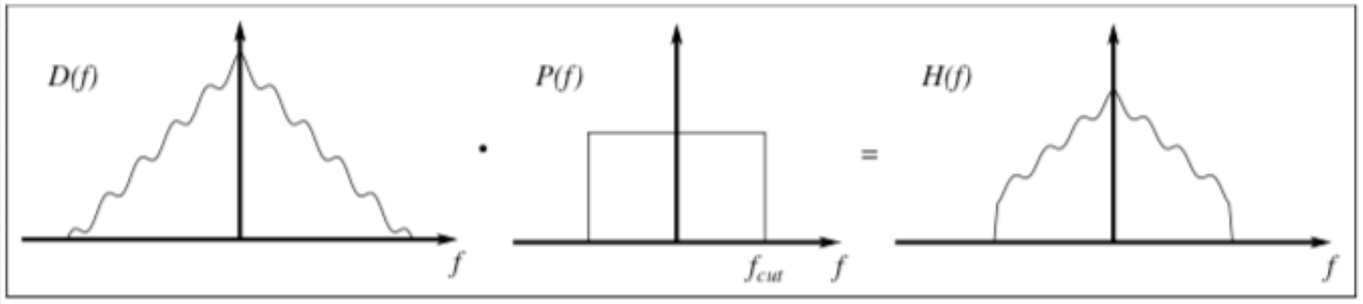


Figure 8.36: Truncation of high frequencies is straightforward in the frequency domain.

We discover that the output has a lot of "ripples" (here somewhat exaggerated) at about the period corresponding to $1/f_{cut}$. This "ringing" is caused by our convolving of $d(t)$ with $p(t)$, which in this case is a sinc function. Because all power at higher frequencies are completely eliminated, the power at the remaining highest frequency f_{cut} stands out. This effect is referred to as "Gibbs' phenomenon".

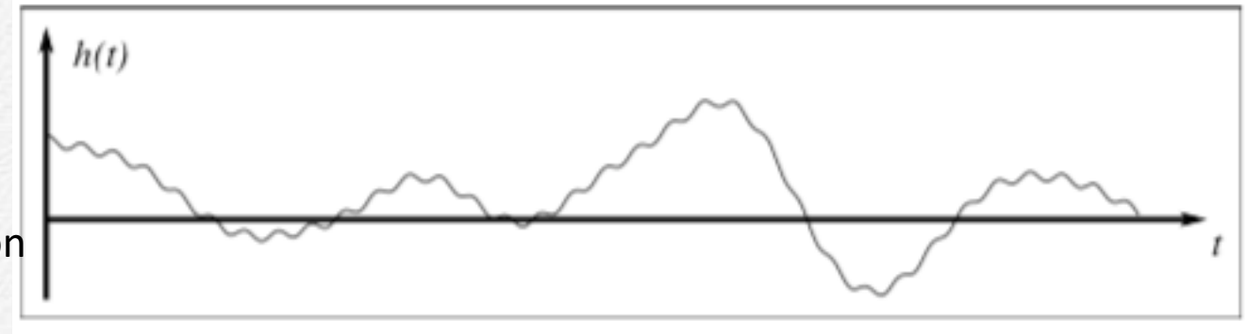


Figure 8.37: "Ringing" in the time domain due to truncation in the frequency domain.

It is clear from these two examples of the use of a rectangular-shaped function that the price we pay for having a sharp truncation of filter coefficients in one domain is excessive "ringing" in the other domain. This, of course, is simply the convolution theorem at work. A rectangle function is necessarily a poor choice for a filter because of the slowly decaying oscillatory nature of the sinc function.

Since time and frequency are inversely proportional, making a filter broader in one domain simply narrows it in the other. Clearly, somewhere in the middle there must be a function that looks similar in both domains. In fact, there is such a function. Here, we will just state that the Gaussian normal distribution behaves exactly this way. Filters based on the Gaussian shape are some of the most widely used filters in data analysis and data processing. Because of their smooth transform properties we know that they will minimize ringing in either domain.

ANALYSIS OF DIRECTIONAL DATA

"Essentially, all models are wrong, but some are useful."
George E. P. Box, Statistician

Much environmental data contain information about orientations in the plane where the *orientations* and not the lengths of the features are the important attributes. Examples of such data types abound: Strikes and dips of bedding planes, fault surfaces, and joints are familiar from structural geology, augmented by glacial striations, sole marks, lineaments in satellite images, directions of winds and currents, and much more. First, we must distinguish between *directional* and *oriented* data. Directional data can take on unique values in the entire 0–360° range, like drumlins and wind directions, while oriented data consist of "two-headed" vectors: there is a 180° ambiguity inherent in the data. Examples of oriented data include fault traces and other lineaments on maps. Such data require special care in the analysis. However, environmental data involve not only directions in the plane but spatial directions as well, which introduces another degree of complexity. Because 3-D vector data are very common, particularly in the Earth and environmental sciences, we need to examine such data in more detail. We find that much of the measurements used in structural geology, such as strike and dip of fault planes, can be expressed as a normal vector to the fault plane. Other examples include vector measurements of the geomagnetic field, palaeomagnetic measurements, stress directions, wind and current directions, and determinations of crystallographic axes for petrofabric studies.

9.1 Circular Data

9.1.1 Displaying directional distributions

Directional data can be displayed in circular diagrams. One can either plot each direction as a unit vector, or we may count the number of vectors within a given angular sector and draw a polar histogram of the distribution (Figure 9.1):

Unfortunately, the sector diagram as described here is biased in the way it presents the data. Nevertheless, it is still the most commonly used type of display for directional data. Where is this bias coming from?

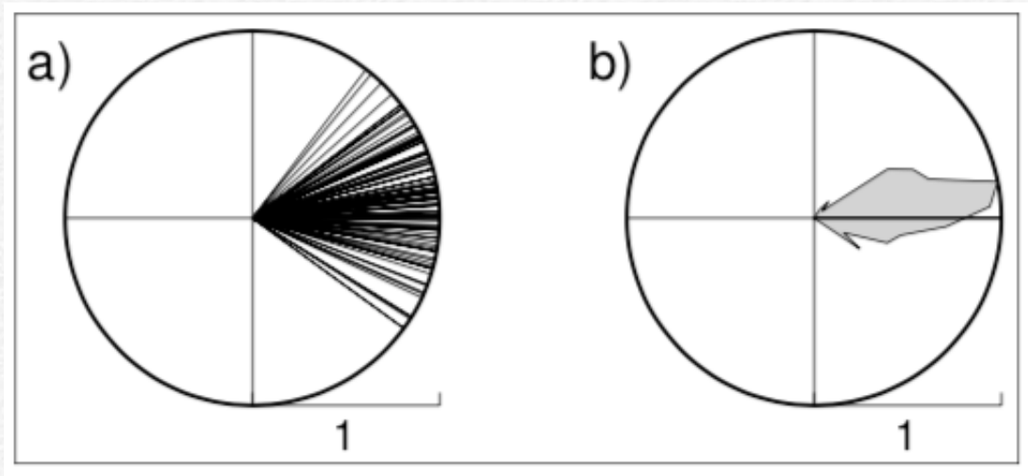


Figure 9.1: Two types of graphical presentations of the same directional data set. (a) Windrose diagram shows all individual directions, (b) Sector diagram shows distribution via a polar histogram (a so-called *rose* diagram).

ANALYSIS OF DIRECTIONAL DATA

Consider the area of a sector of width Δα, given by

$$A = \frac{\pi r^2 \Delta\alpha}{360} \propto r^2. \tag{9.1}$$

We see that the area of a sector is proportional to the radius squared, whereas for a conventional Cartesian histogram the column area is proportional to height, not height squared. Consequently, this leads to a visual distortion of sectors with high counts. Therefore, we should let the sector radius be proportional to the square root of the frequency (or count) to ensure that the final rose diagram will have an area proportional to frequency. If this is not done, the larger counts for some sectors will completely swamp smaller sectors due to the r^2 effect. Thus, small but significant trends may not be detected or may simply be considered noise.

Before we can statistically analyze orientation data, the 180° ambiguity must be accounted for. The simplest way to do this is to double all angles and analyze these doubled angles instead. For instance, if two fault traces are reported as having strikes of 45° and 225° (which means the two strikes have the same orientation), we double the angles and find 90° and 450° − 360° = 90° . Statistics derived from doubled angles can then be divided by two to recover their original meanings.

The dominant or *mean direction* can be found by computing the *vector resultant* (i.e., the vector sum) of the unit vectors that represent the various directions in the data. Since the coordinates of these vectors are given by

$$x_i = \cos\theta_i, \quad y_i = \sin\theta_i, \tag{9.2}$$

we find the coordinates of the resultant *R* to be

$$x_r = \sum_{i=1}^{n} \cos\theta_i, \quad y_r = \sum_{i=1}^{n} \sin\theta_i. \tag{9.3}$$

The mean direction $\bar{\theta}$ is then simply given by

$$\bar{\theta} = \tan^{-1}(y_r/x_r) = \tan^{-1}\left(\sum_{i=1}^{n} \sin\theta_i \Big/ \sum_{i=1}^{n} \cos\theta_i\right). \tag{9.4}$$

Computer implementations of (9.4) should take care to use the atan2 function rather than atan so that the mean direction is found in the correct quadrant.

ANALYSIS OF DIRECTIONAL DATA

The magnitude of the resultant should be normalized by the number of vectors before it can be used further. This gives us the *mean resultant length*

$$\bar{R} = \frac{R}{n} = \frac{\sqrt{x_r^2 + y_r^2}}{n}, \qquad (9.5)$$

Which ranges from 0 to 1 and is a measure of (normalized) dispersion analogous to the variance. Since it increases for focused distribution we often convert it to a *circular variance* via

$$s_c^2 = 1 - \bar{R} = (n - R)/n. \qquad (9.6)$$

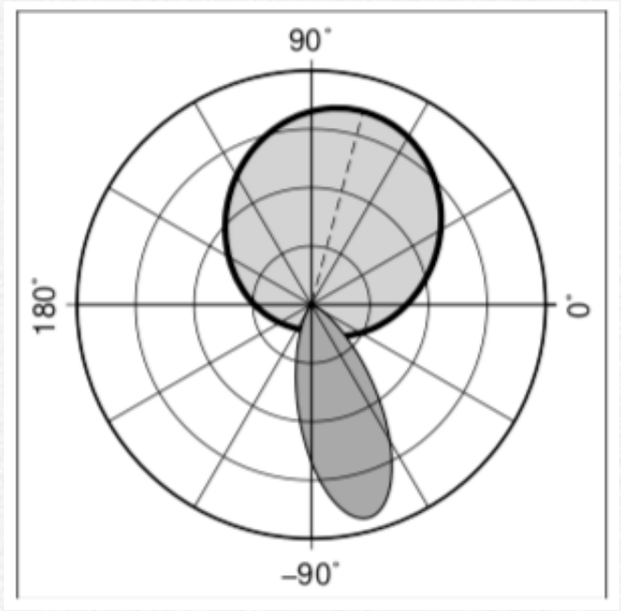


Figure 9.2: The circular von Mises distribution. One population is centered on $\mu = 75$ (dashed line), with very low directionality ($\kappa = 1$), and shows the effect of wrapping the wide distribution around the full circle. The other, at $\mu = -75$, is much more directional ($\kappa = 10$).)

To test various statistical hypotheses about circularly distributed data we need a probability density function that can be used to obtain critical values. A traditional distribution that has been used extensively in studies of directional data is the *von Mises* distribution, given by

$$p(\theta) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)}, \qquad (9.7)$$

where κ is a measure of the *concentration* of the distribution about the mean direction μ and I_0 is the modified Bessel function of the first kind and order zero (it normalizes the cumulative distribution of $p(\theta)$ over 360° to unity.) A large κ means we have a strongly preferred direction. The concentration parameter κ is obviously related to R (or s_c^2) and this relationship can be derived if we assume that the data represent a random sample from a population having a von Mises distribution. Statistical tables (e.g., Table A.16) report solutions to (A.2) that relate κ and $\bar{R}$.

ANALYSIS OF DIRECTIONAL DATA

9.1.2 Test for a random direction

The simplest test for circular data is to determine whether the directional observations are random or not, i.e., that there is no *preferred direction*. In that case we have a *uniform* distribution. In terms of the von Mises distribution (9.7), a uniform distribution means that κ must be zero. Hence, the hypothesis test becomes

$$H_0 : \kappa = 0 \quad \text{versus} \quad H_1 : \kappa > 0. \tag{9.8}$$

If the data came from a uniform distribution (i.e., κ = 0), we would expect a small but not necessarily zero value for $\bar{R}$. Lord Rayleigh devised this test, which gives critical values for $\bar{R}$ depending on the ν = *n* degrees of freedom and chosen level of significance, α. For ν ≥ 10 we can approximate the critical mean resultant via

$$\bar{R}_{\alpha,n} = \frac{1}{2n}\sqrt{1 + 4n(n+1) - (\log\alpha + 2n + 1)^2}, \tag{9.9}$$

while for smaller samples we need to consult exact tables (e.g., Table A.17). If $\bar{R}$ exceeds the critical $\bar{R}_{\nu,\alpha}$, then we must reject the null hypothesis.

Example 9–1. We have measured the strikes of fault planes in an area of the sea floor (called Area 1) imaged by a side-scan sonar device and found the following orientations:

$$\theta_i : 110°,300°,310°,135°,138°,320°,141°,145°,330°,335°,280°,160°,170° \; n = 13.$$

Since these are oriented features, we double the angles to remove any ambiguity in the orientations:

$$2\theta_i : 220°,240°,260°,270°,276°,280°,282°,290°,300°,310°,200°,320°,340°.$$

Summing up the sines and cosines via (9.2), we find

$$x_r = 1.2972,\ y_r = -10.349, \tag{9.10}$$

which gives us

$$2\bar{\theta}_1 = 277°,\ \bar{R}_1 = 0.802. \tag{9.11}$$

We convert back to original angles by dividing by 2 and find the mean orientation to be

$$\bar{\theta}_1 = 138.5°. \tag{9.12}$$

The null hypothesis is not affected by doubling the angles since we are simply testing if κ = 0. At the α = 0.05 level of significance we use Table A.17 to find critical $\bar{R}_{13,0.05} = 0.475$. Clearly, this value is greatly exceeded by the observed $\bar{R} = 0.802$ and we must conclude that the fault strikes are not randomly oriented.

ANALYSIS OF DIRECTIONAL DATA

9.1.3 Test for a specific direction

There are instances when we would like to test whether an observed trend equals a specified trend. The determination of critical values for such a scenario is more difficult and involves using complex equations or charts. As an alternative approach we find the *confidence angle* $\Delta\theta$ around the mean direction of the sample and investigate whether this angular interval is large enough to accommodate the specified trend we want to test against. The approximate standard error in θ is given (in radians) by

$$s_e \approx 1/\sqrt{n\bar{R}k} = 1/\sqrt{Rk}, \tag{9.13}$$

where k is our sample estimate of κ, the population parameter. If we assume that the estimation errors are normally distributed then we may use critical z-values for a given confidence level to construct the angular interval as

$$\bar{\theta} \pm \left| z_{\frac{\alpha}{2}} \right| s_e, \tag{9.14}$$

which will contain the true mean (double angle) orientation, 2μ, $100 \cdot \alpha$ % of the time.

Example 9–2. In our fault strike data case above, it is believed that the faults arose as tensional features in response to tectonic forces operating in the N30°E direction. Under such a stress regime we would theoretically expect the faults to trend at 90° to the direction of these tensile stresses, i.e., in the N120°E direction. From Table A.16 we find $\kappa = 2.897$, which yields

$$s_{e_2} = \frac{1}{\sqrt{13 \cdot 0.802 \cdot 2.897}} = 0.1820 = 10.4^\circ. \tag{9.15}$$

Note that this is the standard deviation about the *doubled* mean angle orientation of 277°. We divide this deviation by 2 to find the standard deviation associated with the original orientations, $s_{e1} = 5.2°$. The 95% confidence interval corresponds to $z_{0.95} = 1.96$, so the confidence interval around $\bar{\theta}_1$ becomes

$$\bar{\theta}_1 = 138.5^\circ \pm 10.2^\circ. \tag{9.16}$$

Since the predicted direction of 120° is outside this interval we conclude that the observed mean direction deviates from that predicted based on the orientation of current tectonic forces. It is possible that the orientation of stresses may have changed since the formation of the faults.

ANALYSIS OF DIRECTIONAL DATA

9.1.4 Test for equality of two mean directions

Another common situation where we may want to apply a statistical test is to determine if two sample mean directions are equal at some prescribed level of confidence. This situation, of course, is reminiscent of the two-sample mean test for scalars that we described in Section 4.2.2. For example, we may have obtained new side-scan sonar data for an area farther away and would like to test whether the mean direction of fault strikes at the second location is similar to what was found at the first location. We can carry out this test by comparing the vector resultants of the two groups to that produced by pooling the two data sets and recalculate a single grand resultant. The key principle here is that if the mean directions are different, then the pooled resultant should be *shorter* than the sum of the two individual resultants. This expectation can be examined using an *F*-test by comparing observed *F* to critical $F_{1,n-2,\alpha}$. Unfortunately, due to approximations needed to relate Cartesian and circular critical values, the procedure to evaluate the *F*-statistic changes with the dispersion parameter, κ, represented by our sample-based value, *k*. For large $k > 10$, we may compute

$$F = \frac{(n-2)(R_1 + R_2 - R_p)}{n - R_1 - R_2}, \tag{9.17}$$

with *n* being the combined number of observations, R_1 and R_2 are the resultants (not *mean* resultants) for each data set, and R_p is the resultant from the combined data. Here, *k* is estimated from $\bar{R}_p$, the mean resultant. However, if $2 < k < 10$ then a more accurate *F*-statistic is obtained via the adjustment

$$F^* = \left(1 + \frac{3}{8k}\right) F, \tag{9.18}$$

while for even smaller values of *k* special tables must be used.

ANALYSIS OF DIRECTIONAL DATA

Example 9–3. We will use the equality test to determine whether the mean strike of the faults in our second area (Area 2) is the same as what was found for the first area ($\theta_1 = 138.5°$, double angle = 277°). The new data are:

$$\theta_i = 91°,280°,111°,115°,118°,300°,122°,126°,130°,80°,320°,149° \ (n = 12),$$

which we double to get

$$2\theta_i = 182°,200°,222°,230°,236°,240°,240°,244°,252°,260°,160°,280°,298°.$$

Carrying out a similar analysis on our new data we now find $2\bar{\theta}_2 = 234.5°$ and $\bar{R}_2 = 0.805$. The hypothesis becomes

$H0 : \bar{\theta}_1 = \bar{\theta}_2$ versus $H_1 : \bar{\theta}1 \neq \bar{\theta}2$, at $\alpha = 0.05$. For the combined data set (with $n = 25$), we obtain

$$2\bar{\theta}_p = 256.7°, \ \bar{R}_p = 0.749. \tag{9.19}$$

We find observed k from Table A.16 to be 2.36. Hence, we use (9.18) to obtain

$$F = \left(1 + \frac{3}{8 \cdot 2.36}\right) \frac{23\,(13 \cdot 0.802 + 12 \cdot 0.805 - 25 \cdot 0.749)}{25 - 13 \cdot 0.802 - 12 \cdot 0.805} = 7.42. \tag{9.20}$$

Table A.5 gives the critical $F_{1,n-2,\alpha}$ value for 1 and 23 degrees of freedom as 4.28. Therefore, we must reject the null hypothesis and conclude that the fracture directions in the two areas appear to have different orientations at the 95% confidence level.

9.1.5 Robust directions

The concept of robustness is also applicable to directional data since such data may also contain outliers. However, directional outliers cause less harm than their Cartesian counterparts discussed earlier since directional data are forced to be periodic. It is clear from Figure 9.3 that an outlier causes most damage when it is oriented 90° away from the trend of the bulk of the data. We may find the circular analog of the median by finding the direction $\tilde{\theta}$ that minimizes the sum

$$\text{minimize} \sum_{i=1}^{n} d\left(\theta_i, \tilde{\theta}\right), \tag{9.21}$$

where $d(\theta_i, \tilde{\theta})$ is the arc distance between θ_i and $\tilde{\theta}$ measured along the perimeter of the unit circle.

ANALYSIS OF DIRECTIONAL DATA

Since this distance is proportional to $|\theta_i - \tilde{\theta}|$ it is the equivalent of minimizing the sum

$$\text{minimize} \sum_{i=1}^{n} |\theta_i - \tilde{\theta}|, \tag{9.22}$$

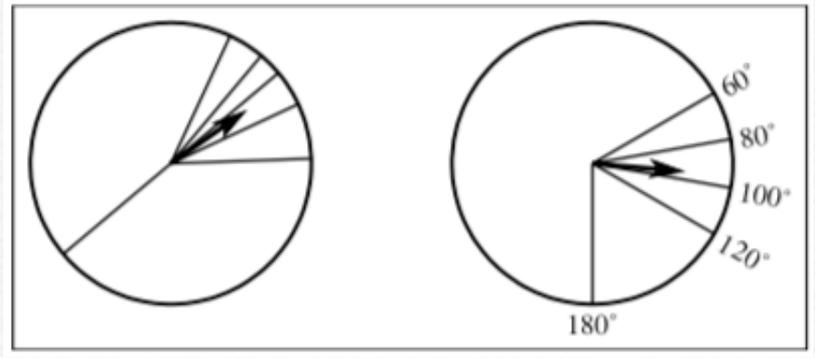
which we know gives the median of the θ_i data set. Similarly, a LMS mode estimate can be found by determining the midpoint of the shortest arc containing $n/2 + 1$ points. Again, this would involve sorting the directions first, possibly after doubling orientations.

Example 9–4. Consider the wind directions

$$\theta_i : 75°, 85°, 90°, 98°, 170° \quad \bar{\theta} = 100.7°.$$

The shortest arc over 5/2 + 1 = 3 points is located between 85° and 98° , giving the mode estimate $\hat{\theta} = 91.5°$. This estimate suggests that 170° is an outlier with respect to the rest of the atmospheric data.

Figure 9.3: The effect of outliers on the mean direction is most severe when the outlier is orthogonal to the bulk of the data directions. Headed vectors represent the mean resultants.



9.1.6 Data with length and direction

In the analysis so far we have only considered the direction (or orientation) of a feature and not its length. However in many cases, such as fracture data or fault traces, the features will have very different lengths. The analysis described above, when applied to such data, would give both a one km long and a 100 km long fault the same weight, which may not be appropriate. We can account for this bias by weighing the directions by the respective *lengths* of the features. By keeping track of the length of the faults (and not their numbers) per sector we can obtain a rose diagram that reflects the proportions of the various fracture directions. The rationale employed here is that large and/or long faults may be more representative of the tectonic stresses than a few short fractures. The rose diagram may then be normalized by the total length of the fractures to give overall proportions in percent.

ANALYSIS OF DIRECTIONAL DATA

Staying with fault strike data, it is clear that in many regions the faults are not entirely straight lines but may actually curve or bend. From a directional analysis point of view, such fractures must first be approximated by shorter straight line segments. It then becomes obvious that we must weigh the pieces by their lengths, otherwise the directional frequencies would depend on the number of pieces used. Fault traces must therefore be digitized prior to analysis on a computer. A typical fault trace is displayed in Figure 9.4. Depending on the angular width of the polar histogram,

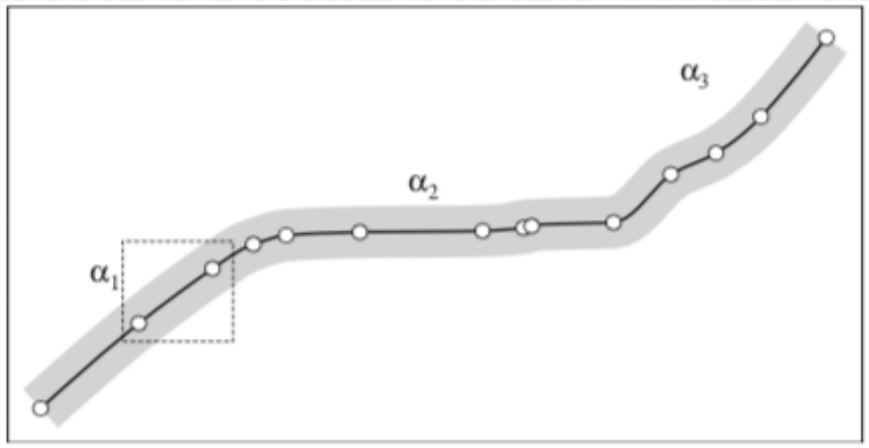


Figure 9.4: A digitized fault trace. Open circles indicate the digitized points. Digitizing too finely may lead to short-wavelength noise in the representation of the fault. Points within the dashed box are examined in Figure 9.5.

the digitized fault may end up in two bins corresponding to the orientations α_1 and α_2. We must also be aware of the fact that the digitizing process will introduce uncertainties in the digitized points. To see how this may affect the analysis, consider the line segment in Figure 9.5. We may assume that the exact position of each point is uncertain, here represented by the one-sigma uncertainty estimate s_r in radial position (gray circles) for each point. There are two points of interest here: First, the length of the segment, d, will have an uncertainty since it reflects a difference between two uncertain values. In Chapter 2, we found this to be

$$\Delta d = \sqrt{2} s_r, \quad d = d_0 \pm \Delta d, \tag{9.23}$$

assuming the uncertainties are *independent*. Second, we see that the direction (or orientation) α may be in error by $\pm\Delta\alpha$, given by

$$\Delta\alpha = \tan^{-1}(2s_r/d), \quad \alpha = \alpha_0 \pm \Delta\alpha. \tag{9.24}$$

Consequently, one should not digitize lines so frequently that $\Delta\alpha$ exceeds the desired polar histogram interval. For instance, if this interval is 10° then you are best served by making the average digitizing interval $d > 10 s_r$. Alternatively, one can *filter* the digitized track so that short-wavelength noise is smoothed out before binning the segments.

ANALYSIS OF DIRECTIONAL DATA

The length of all segments that have a direction within the width of a single bin is simply computed by adding up all the individual lengths. Note that since internal nodes are shared by adjacent line segments the uncertainty in the total length is independent of the number of line segments (e.g., Figure 2.2) and only depends on (9.23). However, the *sum* of the lengths of the individual segments will have an uncertainty that is cumulative since the segments are no longer connected. This method will provide a frequency distribution with error bars in which directions with many small segments will have higher uncertainty than directions with fewer and longer faults.

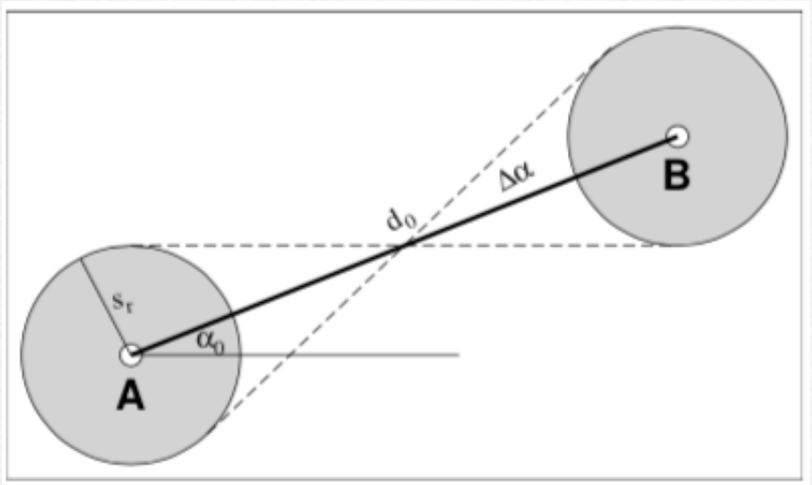


Figure 9.5: The uncertainty r_s in the locations of two digitized points (A and B) from Figure 9.4 introduce uncertainties Δd and $\Delta\alpha$ in the length (d_0) and orientation (α_0) of a line segment, respectively.

9.2 Spherical Data Distributions

The analysis of 3-D directions and orientations is an extension of the methods used for circular data. It is common to require that these 3-D vectors have unit lengths so that their endpoints all lie on the surface of a sphere with unit radius — hence the name spherical distributions. Similar to the 2-D case for circular data, there will be spherical data that only reflect orientations (i.e., *axes*) rather than directions. We need to use a 3-D Cartesian coordinate system to describe the unit vectors (Figure 9.6). Thus, any vector v is uniquely determined by the triplet (x,y,z). We could also use spherical angles θ (colatitude) and ϕ (longitude) to specify the vector direction, assuming the length $r = 1$.

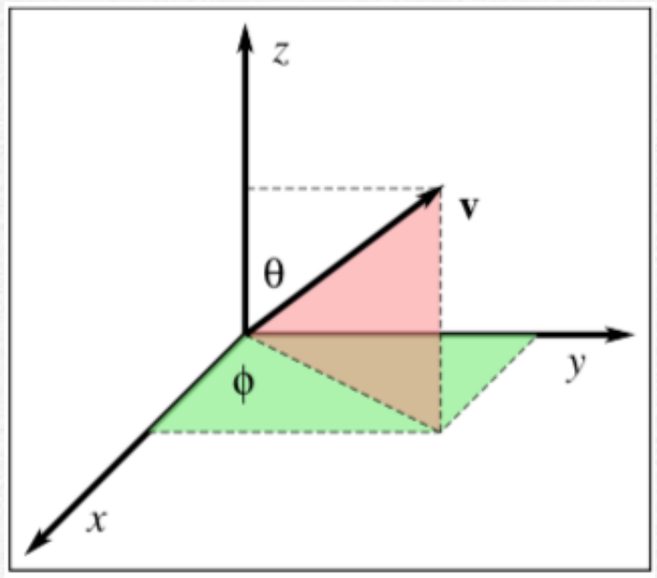


Figure 9.6: The relation between the Cartesian (x,y,z) and spherical (ϕ,θ,r) coordinate systems.

ANALYSIS OF DIRECTIONAL DATA

We relate the Cartesian coordinates and the spherical angles as follows:

$$\begin{aligned} x &= \sin\theta\cos\phi, \\ y &= \sin\theta\sin\phi, \\ z &= \cos\theta. \end{aligned} \qquad (9.25)$$

However, geological measurements like *strike* and *dip* are more commonly used than Cartesian coordinates and spherical angles and these follow their own convention. We define a new local coordinate system in which *x* points toward north, *y* points east, and *z* points vertically down (i.e., in order to maintain a right-handed coordinate system). In such a system, fault plane dips are expressed as positive angles. For the fault plane in Figure 9.7 we find that the angle *A* is the azimuth of the strike of the plane and *D* is the dip, measured positive down. The slip-vector *OP* is then given by its components

$$\begin{aligned} x &= -\sin A\cos D, \\ y &= \cos A\cos D, \\ z &= \sin D. \end{aligned} \qquad (9.26)$$

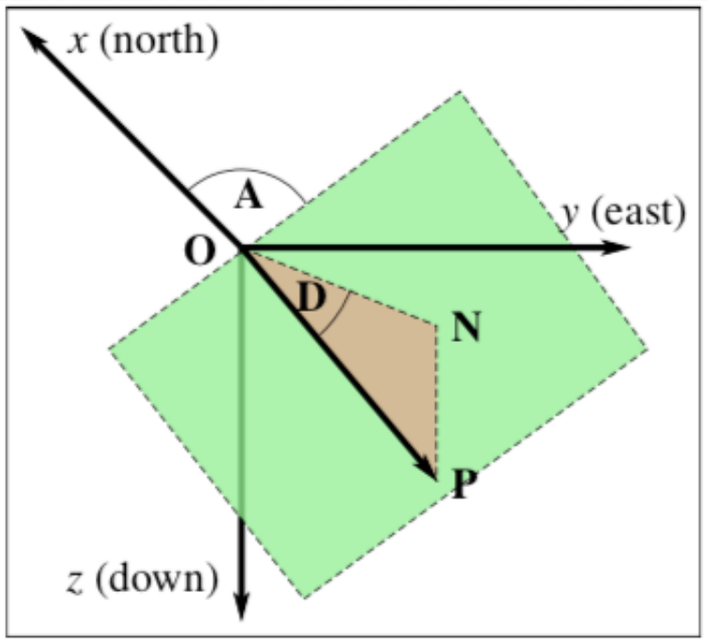


Figure 9.7: Local, right-handed coordinate system shows the convention used in structural geology. Here, *A* is the strike (measured from north over east) of the dipping plane (light green) and *D* is its dip, measured as the angle between the dip vector *OP* and its projection *ON* onto the horizontal *x* − *y* plane. The red plane containing the dip vector is orthogonal to the light green plane.

Once we have converted our (*A*, *D*) data to (*x*, *y*, *z*) we can compute such quantities as mean direction and spherical variance, which are simple extensions of the 2-D or directional analogs. The length of the resultant vector is simply

$$R = \sqrt{\left(\sum x_i\right)^2 + \left(\sum y_i\right)^2 + \left(\sum z_i\right)^2}, \qquad (9.27)$$

with the sum taken over all the *n* points. The resultant vector is usually normalized to give $\bar{R} = R/n$. The coordinates $\bar{x}$, $\bar{y}$ and $\bar{z}$ of the mean vector m are then obtained via

$$\bar{x} = \sum x_i/n, \quad \bar{y} = \sum y_i/n, \quad \bar{z} = \sum z_i/n, \qquad (9.28)$$

so that the mean slip-vector is given by its two components

$$\begin{aligned} \bar{D} &= \sin^{-1}\bar{z}, \\ \bar{A} &= \tan^{-1}(-\bar{x}/\bar{y}). \end{aligned} \qquad (9.29)$$

ANALYSIS OF DIRECTIONAL DATA

If all the vectors are closely clustered then the resultant R will approach n, but if the vectors are more randomly distributed then R will approach zero. As in 2-D, we can use R (or $\bar{R}$) as the basis for the *spherical variance*, s_s^2, given by

$$s_s^2 = (n - R)/n = 1 - \bar{R}. \tag{9.30}$$

9.2.1 Test for a random direction

We can perform simple hypothesis tests in a manner analogous to those we carried out for 2-D directional data. Unlike the case for 2-D we now require a *spherical* probability density function from which we can derive critical values for comparison with our observed statistics. In response to the need for 3-D statistical analysis, in particular for palaeomagnetic studies that started to explore "continental drift" in the 1950s, the famous British statistician Ronald Fisher developed a suitable theoretical distribution for spherical data. His probability density function has since been called the *Fisher* distribution on a sphere and is given by

$$P(\mathbf{x}) = \frac{\kappa}{4\pi \sinh \kappa} e^{\kappa(\mathbf{x}\cdot\boldsymbol{\mu})}, \tag{9.31}$$

where the dot product between the mean direction $\boldsymbol{\mu}$ and any other direction $\mathbf{x}$ equals the cosine of the angle ψ between them, κ is the precision parameter similar to the one we encountered for the von Mises distribution on the circle, and sinh is the *hyperbolic sine* function (which is needed to ensure the cumulative distribution of $P(\mathbf{x})$ over the unit sphere equals one).

Fisher showed one can estimate κ from the sample, provided $n > 7$ and $\kappa > 3$. The estimate is then given by

$$k \approx \kappa = \frac{n-1}{n-R}, \tag{9.32}$$

but more accurate tables also exist (e.g., Table A.18) which solve (A.3).

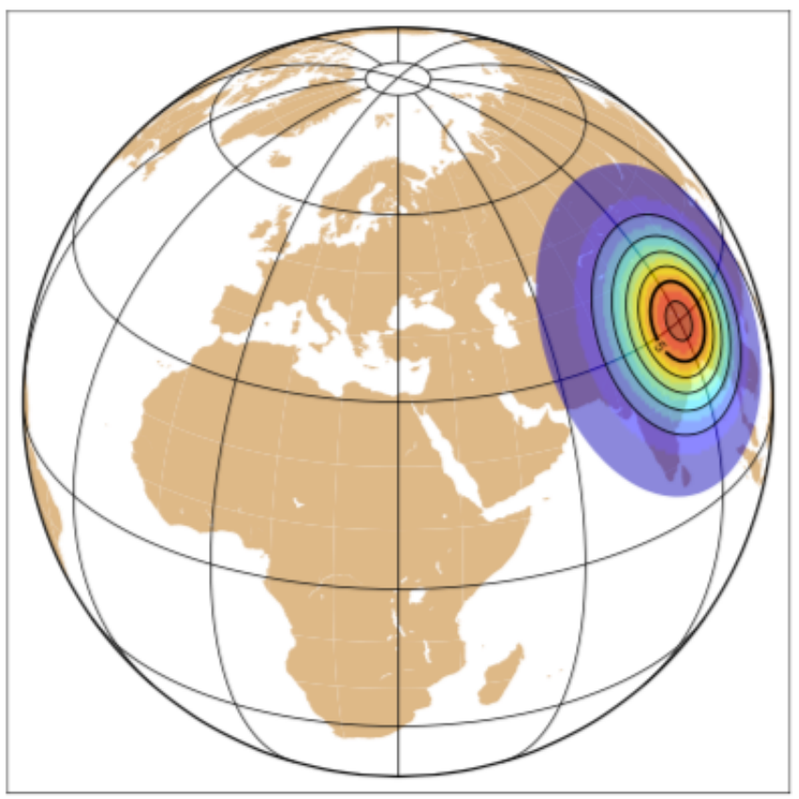


Figure 9.8: Well-focused Fisher distribution on a sphere, centered on a point in northern India, with κ = 40.

ANALYSIS OF DIRECTIONAL DATA

Testing a spherical distribution for randomness follows the same approach used for directional data: We must first evaluate $\bar{R}$, the mean resultant. Then, we state the null hypothesis to be

$$H_0 : \kappa = 0 \quad H_1 : \kappa > 0. \tag{9.33}$$

As before, this test is executed by establishing critical values of $\bar{R}$ given the prescribed level of confidence, α. Table A.19 shows such critical $\bar{R}$ values for selected values of α and n.

Example 9–5. Let us examine a set of palaeomagnetic measurements. We have been given six measurements of *declination* and *inclination*, reported as

Dec	105°	130°	115°	120°	118°	145°
Inc	40°	49°	57°	32°	55°	45°

First, we note that a dip-vector specified by strike A and dip D is not using the same geometry as a magnetic field vector given by its magnetic declination D and inclination I. Because the projection of the dip-vector onto the horizontal plane is 90° away from the strike, we must use a slightly modified conversion suitable for magnetic vectors to obtain the Cartesian coordinates:

$$\begin{aligned} x &= \cos D \cos I, \\ y &= \sin D \cos I, \\ z &= \sin I. \end{aligned} \tag{9.34}$$

We convert the n = 6 observed declinations and inclinations to x, y, z and find

x	−0.20	−0.42	−0.23	−0.42	−0.27	−0.58	$\bar{x}$	−0.354
y	0.74	0.50	0.49	0.73	0.51	0.41	$\bar{y}$	0.564
z	0.64	0.75	0.84	0.53	0.82	0.71	$\bar{z}$	0.715

ANALYSIS OF DIRECTIONAL DATA

Computing the mean direction and resultant gives

$$\bar{I}_1 = \sin^{-1}\bar{z} = 45.7^\circ, \tag{9.35}$$

$$\bar{D}_1 = \tan^{-1}\bar{y}/\bar{x} = 122.1^\circ, \tag{9.36}$$

$$\bar{R}_1 = \frac{1}{6}\cdot 5.86 = 0.977, \tag{9.37}$$

$$k_1 = \frac{6-1}{6-5.86} = 35.7. \tag{9.38}$$

It is evident that our distribution has a clear preferred direction since k_1 is so large (the critical $\bar{R}$ for $\alpha = 0.05$ is only 0.642).

9.2.2 Test for a specific direction

Often, we will be interested in testing whether the observed mean direction equals a prescribed direction, given the uncertainties due to random errors. As for circular data, such tests are best performed by constructing the α confidence region around the mean direction. This statistic is based on the Fisher distribution and gives a spherical cap radius for a *cone of confidence* around the mean direction. As usual, this radius, δ, is a function of both the confidence level α and R. We find

$$\delta_{1-\alpha} = \cos^{-1}\left\{1 - \frac{n-R}{R}\left[\left(\frac{1}{\alpha}\right)^{\frac{1}{n-1}} - 1\right]\right\}. \tag{9.39}$$

This fairly complicated expression simplifies considerably if we assume (or actually know) that $k > 7$ and standardize our tests for $\alpha = 0.05$. Then,

$$\delta_{95\%} \approx 140^\circ/\sqrt{kn} \tag{9.40}$$

given in spherical degrees. We may now say that there is a 95% probability that the true mean direction lies within the cone of confidence specified by the angular radius $\delta_{95\%}$.

ANALYSIS OF DIRECTIONAL DATA

The application of the cone of confidence is more involved than for circular data since we cannot directly compare the two angular measures defining the mean direction (e.g., strike, dip) to the comparable quantities of a specific direction to be tested, here called $\hat{\mathbf{v}}$. Instead, we need to utilize their Cartesian vector representations. The procedure is still relatively straightforward:

1. State the null hypothesis that the unit vectors are the same, i.e., $H_0 : \hat{\mathbf{m}} = \hat{\mathbf{v}}$.
2. Take the dot-product of the mean unit vector with the test unit vector, i.e, $c = \hat{\mathbf{m}} \cdot \hat{\mathbf{v}}$.
3. Since a dot product gives the cosine of the spherical angle between two vectors we find this distance to be given by $\psi = \cos^{-1} c$.
4. If ψ exceeds $\delta_{95\%}$ then $\hat{\mathbf{v}}$ is outside the cone of confidence and we can reject the null hypothesis; otherwise we must reserve judgment.

For other levels of confidence we would substitute the general radius specified in (9.39).

Example 9–6. Given the data we just analyzed we want to test if the mean vector we found is compatible with an hypothesis that says the inclination should be 45° and the declination should be 120°. First, we find the unit vector pointing in the same direction as the mean resultant, m. This vector is

$$\hat{\mathbf{m}} = \frac{\mathbf{m}}{|\mathbf{m}|} = \frac{(-0.354, 0.564, 0.715)}{\sqrt{0.354^2 + 0.564^2 + 0.715}} = (-0.3621, 0.5770, 0.7321). \tag{9.41}$$

We evaluate the test vector to be

$$\hat{\mathbf{v}} = (\cos 120^\circ \cdot \cos 45^\circ, \sin 120^\circ \cdot \cos 45^\circ, \sin 45^\circ) = (-0.3536, 0.6124, 0.7071). \tag{9.42}$$

Hence, the angle between the mean and the test vector is given by their dot product:

$$\psi = \cos^{-1}(0.3621 \cdot 0.3536 + 0.5770 \cdot 0.6124 + 0.7321 \cdot 0.7071) = \cos^{-1}(0.9990) = 2.53^\circ. \tag{9.43}$$

To determine the radius of the 95% confidence cone we use (9.40) and find

$$\delta_{95\%} \approx 140^\circ / \sqrt{35.7 \cdot 6} = 9.6^\circ. \tag{9.44}$$

This means that the true population mean direction probably (i.e., with 95% level confidence) lies within a spherical cone of radius 9.6° centered on the observed mean direction. Clearly, our test vector v also lies inside this cone. Therefore, we cannot reject the null hypothesis that they point in the same direction.

ANALYSIS OF DIRECTIONAL DATA

A.8 Relationship Between κ and $\bar{R}$ for 2-D Directional Data

Given a mean resultant ($\bar{R}$) we use Table A.16 to obtain the corresponding concentration parameter (κ) for directional data in the plane. Alternatively, one can solve the implicit equation for κ given by

$$\bar{R} = \frac{I_1(\kappa)}{I_0(\kappa)}. \tag{A.2}$$

$\bar{R}$	κ	$\bar{R}$	κ	$\bar{R}$	κ
0.00	0.00000	0.34	0.72356	0.68	1.89637
0.01	0.02000	0.35	0.74783	0.69	1.95357
0.02	0.04001	0.36	0.77241	0.70	2.01363
0.03	0.06003	0.37	0.79730	0.71	2.07685
0.04	0.08006	0.38	0.82253	0.72	2.14359
0.05	0.10013	0.39	0.84812	0.73	2.21425
0.06	0.12022	0.40	0.87408	0.74	2.28930
0.07	0.14034	0.41	0.90043	0.75	2.36930
0.08	0.16051	0.42	0.92720	0.76	2.45490
0.09	0.18073	0.43	0.95440	0.77	2.54686
0.10	0.20101	0.44	0.98207	0.78	2.64613
0.11	0.22134	0.45	1.01022	0.79	2.75382
0.12	0.24175	0.46	1.03889	0.80	2.87129
0.13	0.26223	0.47	1.06810	0.81	3.00020
0.14	0.28279	0.48	1.09788	0.82	3.14262
0.15	0.30344	0.49	1.12828	0.83	3.30114
0.16	0.32419	0.50	1.15932	0.84	3.47901
0.17	0.34503	0.51	1.19105	0.85	3.68041
0.18	0.36599	0.52	1.22350	0.86	3.91072
0.19	0.38707	0.53	1.25672	0.87	4.17703
0.20	0.40828	0.54	1.29077	0.88	4.48876
0.21	0.42962	0.55	1.32570	0.89	4.85871
0.22	0.45110	0.56	1.36156	0.90	5.30469
0.23	0.47273	0.57	1.39842	0.91	5.85223
0.24	0.49453	0.58	1.43635	0.92	6.53939
0.25	0.51649	0.59	1.47543	0.93	7.42572
0.26	0.53863	0.60	1.51574	0.94	8.61035
0.27	0.56097	0.61	1.55738	0.95	10.2717
0.28	0.58350	0.62	1.60044	0.96	12.7668
0.29	0.60625	0.63	1.64506	0.97	16.9289
0.30	0.62922	0.64	1.69134	0.98	25.2581
0.31	0.65242	0.65	1.73945	0.99	50.2506
0.32	0.67587	0.66	1.78953	1.00	∞
0.33	0.69958	0.67	1.84177		

Table A.16: Relationship between κ and $\bar{R}$ in 2-D.

ANALYSIS OF DIRECTIONAL DATA

A.9 Critical Values of $\bar{R}$ for 2-D Directional Data

Given the level of significance we determine the critical value for the mean resultant length under the null hypothesis of no preferred direction in the plane ($H_0 : \bar{R} = 0$), with the alternative hypothesis being that the data can be described via the von Mises distribution (9.7) with a preferred trend ($H_1 : \bar{R} \neq 0$).

α:	**0.10**	**0.05**	**0.025**	**0.01**
$n = 4$	0.768	0.847	0.905	0.960
5	0.677	0.754	0.816	0.879
6	0.618	0.690	0.753	0.825
7	0.572	0.642	0.702	0.771
8	0.535	0.602	0.660	0.725
9	0.504	0.569	0.624	0.687
10	0.478	0.540	0.594	0.655
11	0.456	0.516	0.567	0.627
12	0.437	0.494	0.544	0.602
13	0.420	0.475	0.524	0.580
14	0.405	0.458	0.505	0.560
15	0.391	0.443	0.489	0.542
16	0.379	0.429	0.474	0.525
17	0.367	0.417	0.460	0.510
18	0.357	0.405	0.447	0.496
19	0.348	0.394	0.436	0.484
20	0.339	0.385	0.425	0.472
21	0.331	0.375	0.415	0.461
22	0.323	0.367	0.405	0.451
23	0.316	0.359	0.397	0.441
24	0.309	0.351	0.389	0.432
25	0.303	0.344	0.381	0.423
30	0.277	0.315	0.348	0.387
35	0.256	0.292	0.323	0.359
40	0.240	0.273	0.302	0.336
45	0.226	0.257	0.285	0.318
50	0.214	0.244	0.270	0.301

Table A.17: Critical values for $\bar{R}$ in the plane. Note the use of n rather than ν.

ANALYSIS OF DIRECTIONAL DATA

A.10 Relationship Between κ and $\bar{R}$ for 3-D Directional Data

Given a mean resultant ($\bar{R}$) we use Table A.18 to obtain the corresponding concentration parameter (κ) for directional data in space. Alternatively, one can solve the implicit equation for κ given by

$$\coth \kappa - 1/\kappa = \bar{R}. \tag{A.3}$$

$\bar{R}$	κ	$\bar{R}$	κ	$\bar{R}$	κ
0.00	0.00000	0.34	1.09951	0.68	3.08456
0.01	0.03000	0.35	1.13739	0.69	3.19091
0.02	0.06001	0.36	1.17584	0.70	3.30354
0.03	0.09005	0.37	1.21490	0.71	3.42314
0.04	0.12012	0.38	1.25459	0.72	3.55051
0.05	0.15023	0.39	1.29497	0.73	3.68655
0.06	0.18039	0.40	1.33605	0.74	3.83232
0.07	0.21062	0.41	1.37789	0.75	3.98905
0.08	0.24093	0.42	1.42053	0.76	4.15819
0.09	0.27132	0.43	1.46401	0.77	4.34143
0.10	0.30182	0.44	1.50839	0.78	4.54076
0.11	0.33242	0.45	1.55372	0.79	4.75857
0.12	0.36315	0.46	1.60005	0.80	4.99772
0.13	0.39402	0.47	1.64745	0.81	5.26167
0.14	0.42503	0.48	1.69599	0.82	5.55463
0.15	0.45621	0.49	1.74573	0.83	5.88181
0.16	0.48756	0.50	1.79676	0.84	6.24971
0.17	0.51909	0.51	1.84915	0.85	6.66652
0.18	0.55083	0.52	1.90300	0.86	7.14279
0.19	0.58278	0.53	1.95842	0.87	7.69228
0.20	0.61497	0.54	2.01550	0.88	8.33333
0.21	0.64740	0.55	2.07437	0.89	9.09091
0.22	0.68009	0.56	2.13515	0.90	10.0000
0.23	0.71306	0.57	2.19799	0.91	11.1111
0.24	0.74632	0.58	2.26304	0.92	12.5000
0.25	0.77990	0.59	2.33049	0.93	14.2857
0.26	0.81381	0.60	2.40050	0.94	16.6667
0.27	0.84806	0.61	2.47331	0.95	20.0000
0.28	0.88269	0.62	2.54914	0.96	25.0000
0.29	0.91771	0.63	2.62825	0.97	33.3333
0.30	0.95315	0.64	2.71093	0.98	50.0000
0.31	0.98902	0.65	2.79751	0.99	100.000
0.32	1.02536	0.66	2.88836	1.00	∞
0.33	1.06218	0.67	2.98389		

Table A.18: Relationship between κ and $\bar{R}$ in 3-D.