

8. Fourier transform spectroscopy (cont.)

Complex representations of Fourier transformations

In cases where the time domain signal is neither perfectly even nor perfectly odd (as all of the previous examples except the filtered square wave have been) both sine and cosine Fourier coefficients are required to describe $f(t)$ in the frequency domain. For example, in interferometry, when the mirror alignment is not perfect the interferogram becomes asymmetric around the zero retardance point. (Notice in Marshall & Verdun, App D that symmetric functions have only cosine transform coefficients. This is because symmetric functions are even.) Attempting to compute $\hat{F}_c(\omega)$ from the asymmetric $f(t)$, with the simple cosine transform produces a product

$$|\hat{F}(\omega)| \cos(\phi(\omega)) = \sum_{t=0}^{\infty} f(t) \cos(\omega t), \quad (8.4)$$

where $|\hat{F}(\omega)|$ is the amplitude of the Fourier coefficient and $\phi(\omega)$ is the phase (we shall discuss these terms further shortly), rather than the transform $\hat{F}_c(\omega)$. So any odd function will have a cosine transform for which the phases are not zero and transformation yields the product $|\hat{F}(\omega)| \cos(\phi)$. It is clear that this situation is not optimal and Equations 8.2a, and 8.4 are inconvenient in this case because there is no way to separate the phase information from the coefficient information in these cases. It is always important to know both $\hat{F}(\omega)$ and $\cos(\phi)$ rather than just the product. This is why we use complex Fourier transforms; they allow us to determine the transform coefficients and phase separately. After a short review of complex number representations, we will see how they are used in Fourier transformations.

The imaginary number, i , is the square root of negative one, i.e., $(-1)^{1/2}$. (Sometimes $(-1)^{1/2}$ is represented by j to avoid confusion with the current.) Complex numbers have both a real and imaginary part: $z = a + bi$. The function Re extracts the real part of a complex number, while Im extracts the imaginary part:

$$\text{Re}(z) = a \quad \text{Im}(z) = b$$

One way to think about complex numbers is as the components of a vector stopped at some point as it travels around a circle of radius $|z| = \sqrt{a^2 + b^2}$. (I think this reflects the central role stargazing and astronomy have played in the development of our mathematics.) With this picture we also know the angle the vector makes with the x axis. In the picture it is θ , in our notation it is

$$\phi = \text{atan}(b / a)$$

This means we can re-express z using Euler's formula:

$$z = a + ib = |z|(\cos \phi + i \sin \phi) = |z|e^{i\phi}.$$

Now back to Fourier transforms. We use Euler's formula to combine cosine and sine transforms into a single function:

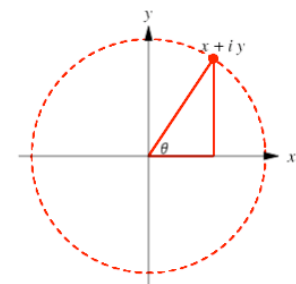


Figure 8.8:

Complex Number

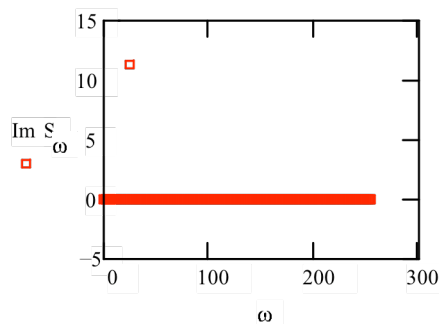
from <http://mathworld.wolfram.com>

$$\exp(-i\omega t) = \cos(\omega t) - i \sin(\omega t).$$

The real part of the complex transform is the cosine transform and the imaginary part of is the sine transform. So the complex Fourier transform separates a function into even and odd components. Even functions, which can be represented by sums of cosines, will have only real numbers in their complex Fourier transforms. Odd functions, which are represented by sums of sines (cosines shifted by $\pi/2$), will have only imaginary numbers in their complex transforms (Fig. 40). A function comprised of cosines shifted in phase by anything but an integer multiple of $\pi/2$ will have a complex transform.

Add phase shift of 90 deg. : $S_t = \cos(\omega t - \pi/2)$

$$S_\omega = \text{fft}(S_t)$$



Nonzero amplitude is now imaginary

$$s_{24} = -5.01 \cdot 10^{-15} + 2.26 \cdot 10^{-14}i$$

$$s_{25} = 11.314i$$

$$s_{26} = -4.37 \cdot 10^{-15} - 2.32 \cdot 10^{-14}i$$

Figure 8.9: Impact of Phase Shift on Transform

The key to understanding the impact of the odd components on the interferogram transform is a second form of the complex Fourier transform of $f(t)$ (see Eq. 1.3 for the first). We can write any function as a sum of the product of harmonics and an explicit phase term:

$$f(t) = \sum_{\omega=0}^{\infty} |F(\omega)| \exp(i\phi(\omega)) \exp(i\omega t)$$

where $|F(\omega)|$ is the amplitude of the combined cosine and sine coefficients. Technically the above expression is the inverse transform of $f(t)$. (FYI: The inverse transform has the same form as the transform, but with the sign in the exponential changed. The term $\exp(i\omega t)$ is the "complex conjugate" of $\exp(-i\omega t)$.) The transform of $f(t)$ is

$$\hat{F}(\omega) = \sum_0^{\infty} f(t) \exp(-i\omega t) = |F(\omega)| \exp(i\phi) = |F(\omega)| (\cos(\phi) + i \sin(\phi))$$

In this treatment the cosine transform of the asymmetric interferogram in the frequency domain is $|F(\omega)| \cos(\phi)$. Remember, the reason we've rewritten the transform is to separate the coefficient and phase information. The phase can be calculated from the complex transform components

$$\phi(\omega) = \tan^{-1} \left\{ \frac{\text{Im}[|F(\omega)| \exp(i\phi(\omega))]}{\text{Re}[|F(\omega)| \exp(i\phi(\omega))]} \right\} = \tan^{-1} \left\{ \frac{|F(\omega)| \sin(\phi(\omega))}{|F(\omega)| \cos(\phi(\omega))} \right\}$$

where real coefficients are given by $|F(\omega)|\cos(\phi)$ and the imaginary by $|F(\omega)|\sin(\phi)$. In this form, the phase can be removed from the amplitude using simple trigonometry, isolating the amplitude,

$$\begin{aligned}|F(\omega)| &= \left\{ \left(\operatorname{Re} [|F(\omega)| \exp(i\phi)] \right)^2 + \left(\operatorname{Im} [|F(\omega)| \exp(i\phi)] \right)^2 \right\}^{1/2} \\ &= \left\{ |F(\omega)|^2 \cos^2(\phi) + |F(\omega)|^2 \sin^2(\phi) \right\}^{1/2} \\ &= \sqrt{\hat{F}_c^2(\omega) + \hat{F}_s^2(\omega)}\end{aligned}$$

M&V call the amplitude the magnitude, $M(\omega)$, others call it the power spectrum. Clearly, it is analogous to $|z|$.

Let's reconsider the RC circuit using complex transforms. We know (from physics) that the charge in an RC circuit decays exponentially in time, consequently the frequency response of a low-pass filter can be described by taking the Fourier transform of a single-sided exponential decay. Using $f(t) = \exp(-t/\tau)$, $\tau=RC$ (only for $t>0$), the complex Fourier transform is

$$FT(\exp(-t/\tau)) = (2/\pi)^{1/2} \left\{ \frac{1}{1+(\omega\tau)^2} + i \cdot \frac{\omega\tau}{1+(\omega\tau)^2} \right\}.$$

It is not easy to see, but the filtered square wave is the convolution of the original function and the exponential decay that characterizes the filter. The sharp parts of the square wave rise and fall exponentially because passage through the filter attenuates the real and imaginary high frequency (fast) signal components by $(1+\omega\tau)^{-1}$ and $+\omega\tau/(1+\omega\tau)$, respectively.

A second example of this kind of qualitative analysis of Fourier transformation comes from considering the transformation of a single frequency interferogram for which the length of mirror travel is insufficient. The acquisition of the interferogram is stopped before the oscillation is completely acquired. This is tantamount to multiplying the time domain signal by a rectangular window function. Consequently, the Fourier transform of this incomplete interferogram is the convolution of the transform of the oscillation (a Dirac pulse) and the transform of the window (a sinc function). This is illustrated in Figure 8.13.

Numerical Fourier transformation

Fourier Tran. in NMR, Optical and Mass Spec.: Sections 3.1-3.5

Experimental data acquired by a computer is comprised of discrete points; therefore, the Fourier transforms used on experimental data are always discrete Fourier transforms. Before we look at the expressions for the discrete transform it is important to see how digitization impacts experimental signals. Figure 8.10 illustrates the digitization of a continuous, analog single frequency interferogram. This figure shows that digitization multiplies the analog signal by a sequence of evenly spaced Dirac pulses (the sampling function). The result is a sequence of N evenly spaced points that represent the analog signal: $\{f(t_0), f(t_1), \dots, f(t_{N-1})\}$. When a sequence is digitized, its

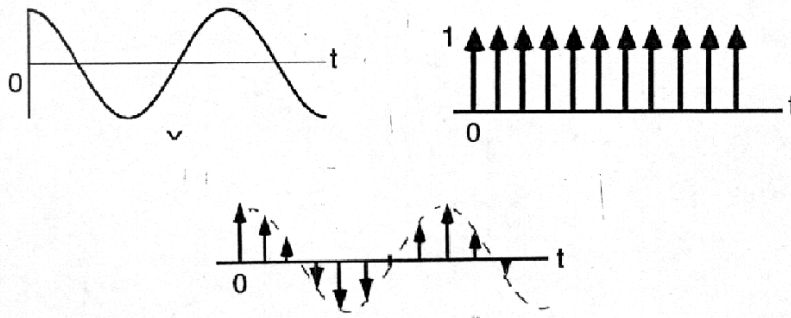


Figure 8.10: Digitization of Analog Signal (Marshall & Verdun, 1990.)

Fourier transform is very similar to its continuous counterpart. The principle difference is in the fact that the sequence and its transform are vectors and the indexing has important unique features. (There is a wide variety of commercial software for calculating the Fourier transform of a digitized waveform numerically. (Most of these programs are based upon the Cooley-Tukey algorithm, which is an efficient, timesaving way of programming the Fourier transform, and is often called the "Fast Fourier transform algorithm". Using any personal computer, a Fast Fourier transform (FFT) on 1024 data points takes just a few seconds.) The traditional FFT requires that the waveform (digital sequence) have 2^n data points, where n is any integer. Others do not, but in either case the calculated spectrum has only half the number of unique data points as the time-domain function. This arises from the fact that simple transformation does not distinguish between positive and negative frequencies. Remember that $\cos(\omega_0)$ and $\cos(-\omega_0)$ are identical so the positive and negative cosine (real) Fourier coefficients are identical. In the case of the sines, only the signs are different. Formally this is reflected in the Nyquist criterion, which states that a waveform of frequency, ω , can only be characterized by sampling the waveform at a rate of at least 2ω . Hence the number of unique frequencies in the spectrum is a factor of 2 smaller than the number of points in the time domain sequence. The discrete FT of a sequence $f(t_n)$ is

$$\hat{F}(\omega_m) = \sum_{n=0}^{N-1} f(t_n) \exp(i\omega_m t_n)$$

where the frequency index m steps from $-N/2$ to $N/2$ to reflect the reduction of the number of unique points in the transform.

The Nyquist criterion has another important consequence in computing transforms of discrete sequences: when the sampling rate is too low, signal components that have frequencies above the Nyquist frequency are "aliased" (folded) into lower frequencies. The example in Figure 8.11 shows what happens when the frequency of a single frequency interferogram exceeds $1/(2\Delta t)$, where $1/(2\Delta t)$ is the highest frequency the Fourier transform can compute and Δt is the time elapsed between sampled points. This re-states the Nyquist criterion: if $1/\Delta t$ is the sampling frequency, then $1/2\Delta t$ is the maximum frequency observed without aliasing. The highest frequency in the example transform is $2\pi(16/32)$ so a spectrum at $\omega = 2\pi(20/32)$, which should be outside the frame, is misrepresented at $\omega_{\text{apparent}} = \omega_{\text{sampling}} - \omega_{\text{true}} = 2\pi(12/32)$. When the frequency content of a waveform is high, the time axis has to have points sufficiently close

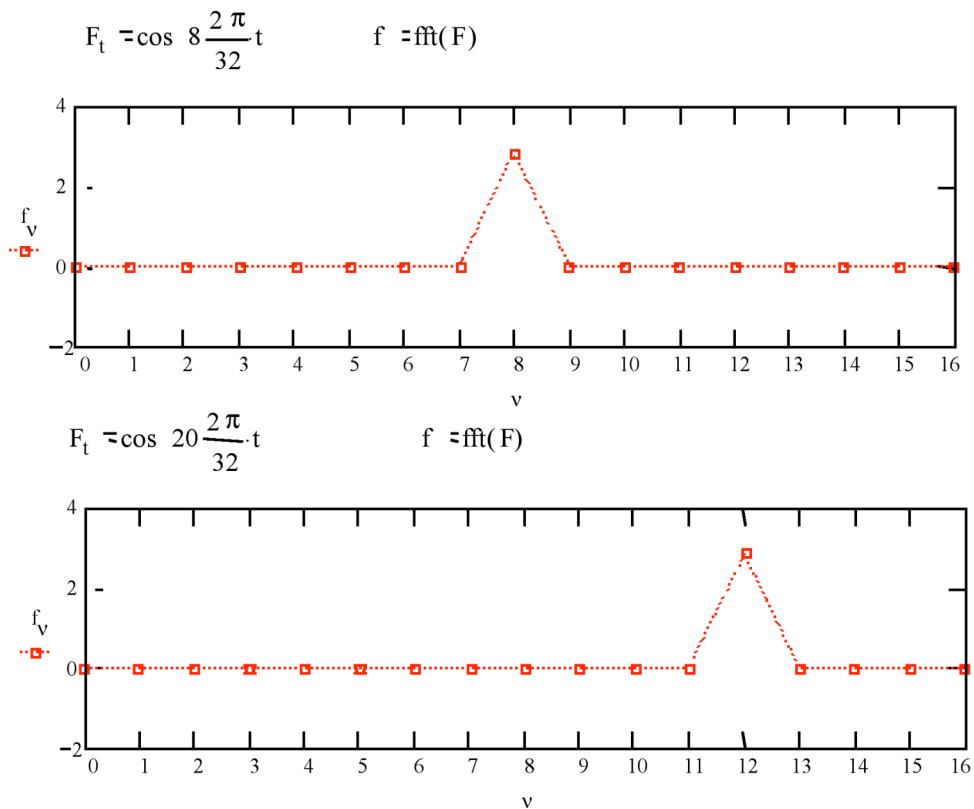


Figure 8.11: Aliasing (in the Frequency Domain)

together to capture the high frequency features. Otherwise each high frequency component appears as if it has a lower frequency in the transform. This is illustrated in Figure 8.12, where the same cosine wave is plotted with a time axis that has points close together (red) vs. far apart (blue). Sampling the high frequency wave with too few points makes it appear to be a lower frequency wave. Here the “free spectral range” (FSR) describes the range of frequencies that can be sampled without aliasing: $\text{FSR} = 1/2\Delta t$.

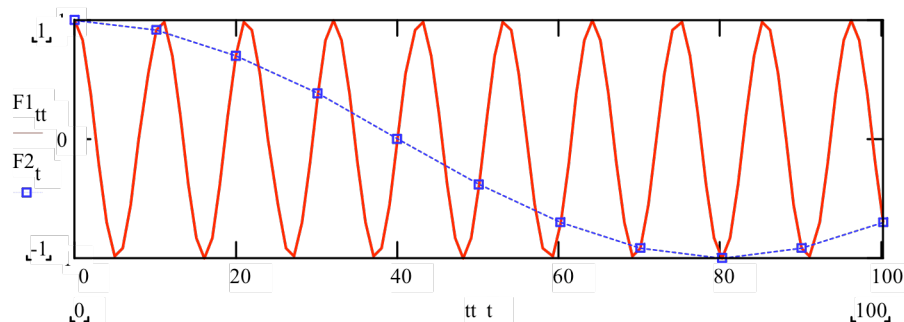


Figure 8.12: Aliasing (in the Time Domain)

Now that the signals have been digitized, we can discuss convolution as a series of sums rather than integrals. In practice most convolutions are carried out by computing the Fourier transforms, multiplying them, then computing the inverse transform of the product, but it is easiest to understand what convolution does in the

time domain. Convolution is a type of vector multiplication that blends the features of the two sequences. This operation occurs often in nature: when a simple camera captures an image of a large object (see http://www.vias.org/tmdatanaleng/...cc_convolution.html for a humorous, but effective cartoon), a telescope captures images of the sky, an AFM (atomic force microscopy) tip images a surface and when a source is dispersed into a spectrum by a monochromator equipped with entrance and exit slits. We will consider a simplified version of this last example. Consider the sequence $f=[0 \ 1 \ 2 \ 1 \ 0]$ to represent a spectrum and the sequence $h=[0 \ 1 \ 1 \ 1 \ 0]$ to represent the exit slit of a monochromator. The spectrum is recorded at the detector as the grating is turned and the wavelengths that pass through the slit change. The discrete convolution is given by the expression

$$g(i) = \sum_{k=1}^{k=N_f+N_h+1} h(k-i)f(k) = \left[\sum_{k-i=-2}^{k-i=2} h(k-i)f(k) \right]_{i=1}^{i=N_f+N_h+1}$$

The total number of points in a discrete convolution is one more than the combined number of points in the two sequences. The convolution of the two sequences is $g=[0 \ 0 \ 1 \ 3 \ 4 \ 3 \ 1 \ 0 \ 0]$. Technically, each element of the convolution is a sum of nine terms. Since most of them are zero, it is useful to only consider the indices that correspond to the spectrum points that are in front of the slit. We do this by changing the summation to cover the relative units around the center of the spectrum $k-i$. In this example $k-i$ steps from -2 to 2. So the first non-zero element of the convolution occurs when $i=3$. At this grating setting, the last non-zero element of the “spectrum” overlaps the first non-zero element of the “slit”, so the intensity recorded at the detector, $g(i)$, is 1. The spectrum was broadened by passing through the slit because the slit is not infinitely narrow as it would need to be to form a perfect 1:1 image.

In discrete processing, the truncation of signals by finite acquisition times is an important application of convolution too. Again we treat the data set as though it were an infinite function multiplied by a window (rectangle) function. (Remember the window function is defined as unity at the time points when the data set is measured and zero outside of the data set). As previously discussed, the Fourier transform of a rectangle is a sinc function. So when a 1024 point cosine wave is truncated to 300 data points, the transform is the convolution of the transform of the waveform (Dirac function) and the transform of the window function (sinc function). This is illustrated below in Figure 8.13.

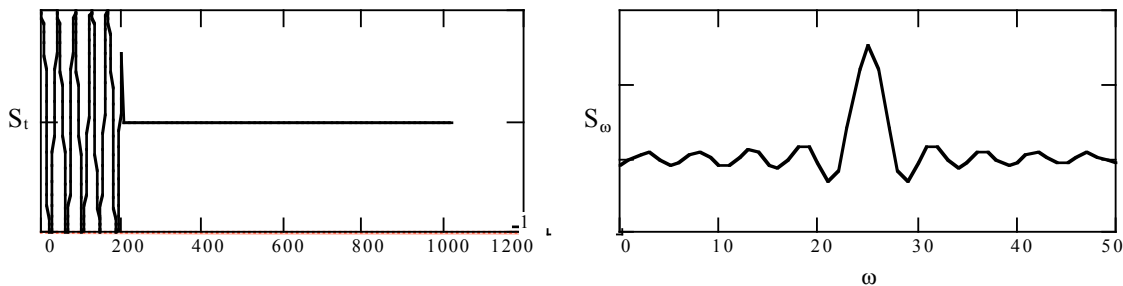


Figure 8.13: Truncated Oscillation & Transform (Marshall & Verdun, 1990.)

Now we see why even infinitely narrow spectral bands will appear as sinc functions that have widths inversely proportional to the mirror travel (the width of the window) in the spectrum. We also see why the spectral resolution of the instrument is controlled by the

mirror travel, Δz . The exact relation between the conjugate variables mirror travel, Δz , and spectral bandwidth (think FWHM), Δk , is $2\pi\Delta k = 1.2/\Delta z$. Most textbooks, including I&C, approximate 1.2 as 1 to simplify the calculation.

The oscillations in the baseline of the sinc function are undesirable in spectroscopy because they can distort neighboring resonances. The oscillation can be suppressed by “apodization”, in which the interferogram is multiplied by a monotonically decaying function such as a Gaussian or an exponential in order to attenuate the oscillation in the baseline. The example in Figure 8.14 shows that multiplication of the truncated cosine wave by an exponential gives a Lorentzian profile. Apodization always lowers the resolution, but its benefits often outweigh its disadvantages.

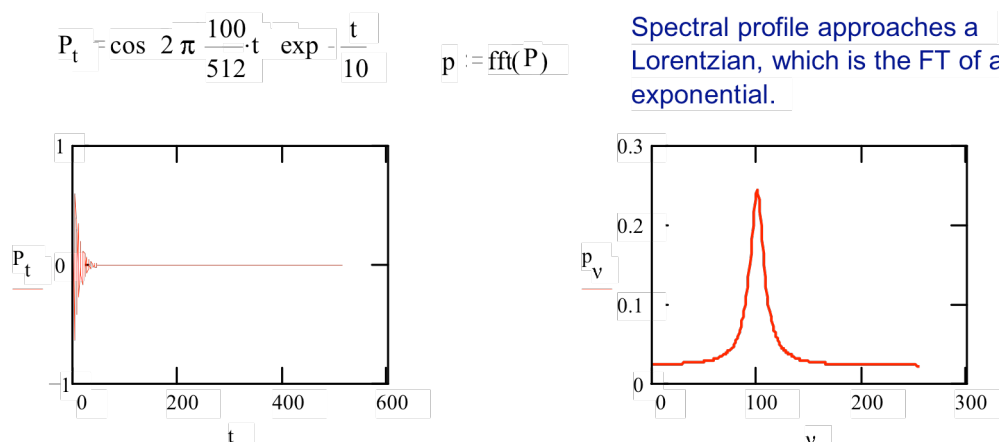


Figure 8.14: Apodization of Truncated Oscillation

This trade-off between resolution and baseline stability also is important when the FT is used to compute the spectrum from a time domain waveform. Since the number of waveform points determines the number of frequencies that can be represented in the spectrum. Since transformation cuts the number of points used to represent the spectrum, computing the transform of a waveform to which N zeros have been appended increased the number of unique frequencies to equal the number of waveform points. This does not increase the frequency content of the spectrum, remember this is determined by the Nyquist frequency ($1/2\Delta t$), but it increases the number of points used to represent the frequency axis. The addition of N points to the waveform can improve the resolution of a spectrum, as shown here in Figure 8,15.

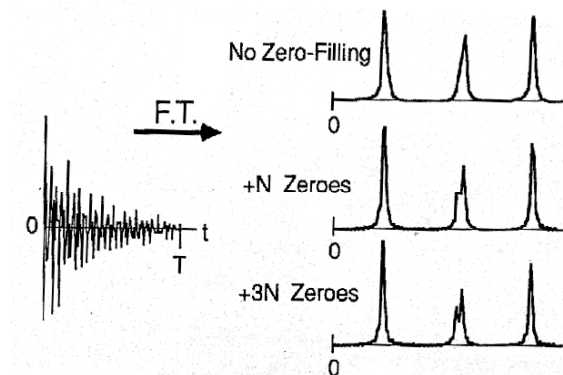


Figure 8.15: Impact of Zero-Filling
(Marshall & Verdun, 1990.)

Adding more than N zeros doesn't add any additional information to the spectrum, but it can make small shoulders more visible. However, this comes at the cost of decreased baseline stability. This trade-off is clearer in the original figure, which you can find in Figure 3.7 of Marshall & Verdun on page 81.

9. Noise

In experimentation, there is always noise. Noise consists of random fluctuations superimposed on the analytical signal. Noise is characterized by its magnitude, its frequency spectrum and/or its source. The magnitude of noise can be measured in terms of signal variance, σ^2 , when measurement errors are normally distributed or in terms of counts, n , in spectroscopic measurements described by counting (Poisson) statistics. In both cases, the signal to noise ratio is a convenient measure. Expressions for some of these parameters are

$$\sigma_{\Phi}^2 = \frac{\sum_i (\Phi_i - \bar{\Phi})^2}{n-1} \quad \sigma_N^2 = \bar{N} = \Phi_p t$$

$$\frac{S}{N} = \frac{\bar{\Phi}}{\sigma} = RSD^{-1}$$

$$\frac{S}{N} = \frac{\bar{N}}{\sigma_N} = \frac{\bar{N}}{\sqrt{\bar{N}}} = \sqrt{\bar{N}}$$

where Φ_i represents the radiant power measured at each point in the measurement, N_i represents the number of photons counted at each point in the measurement and n represents the total number of data points collected in the measurement.

How much noise is superimposed on the signal depends on several factors especially the frequency at which the measurements are made. Figure 9.1 depicts the relative amount of noise observed from various sources in a university laboratory. Notice in particular how noisy the low frequency range, i.e., where steady state measurements are made, is. Herein lies the value of modulation based strategies, such as lock-in

amplification, which move the measurement to a quiet frequency region where environmental noise is lower. Various instruments have similar noise spectra revealing the nature of the noise impacting the measurement process. For example the noise power spectrum of a fluorescence spectrometer, would look similar to Figure 9.1. The low frequency noise of such instruments is large because frequency dependent $1/f$ or flicker noise is a major noise source in fluorescence measurements. Of course, there should

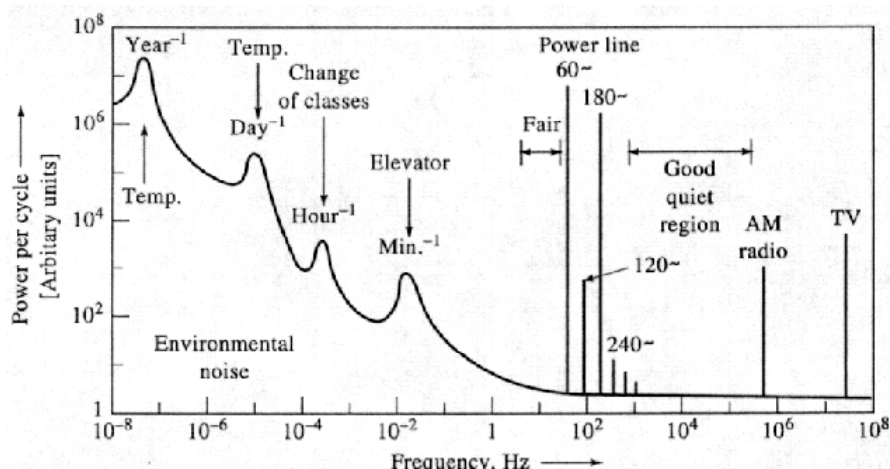


Figure 9.1: Environmental Noise Sources

Skoog, Nieman, Holler, Principles of Instrumental Analysis, 1998, 5th

spectrum of a fluorescence spectrometer, would look similar to Figure 9.1. The low frequency noise of such instruments is large because frequency dependent $1/f$ or flicker noise is a major noise source in fluorescence measurements. Of course, there should

not be such large interference bands since the instrument components should be insulated to minimize pick-up of environmental noise. However, there can be exceptions, for example large NMR spectrometers nearby (e.g., next door) can produce significant interference noise in the RF frequency range. The noise power spectrum of detector noise limited instruments, such as IR spectrometers, is essentially flat because it is dominated by frequency independent white noise such as shot and thermal (Johnson) noise.

Variances are additive so they can be partitioned according to their source. For example, Most spectroscopic signals are degraded by noise in the analytical signal, background signal and detector (dark) signal. Each of these variances can be partitioned further to reflect the circumstances of the measurement. For example, the noise in IR absorption measurements is dominated by fundamental and thermal noise from the detector because IR sources are typically weak so $\sigma_{IR}^2 \approx \sigma_{dark}^2 + \sigma_{ar}^2$ where σ_{dark}^2 is the detector dark noise and σ_{ar}^2 (ar stands for amplifier read-out) is the thermal noise impacting the detection electronics. On the other hand, fluorescence measurements are distorted by flicker (1/f) noise from the source in the signal and the background as well as fundamental (signal shot and quantum) noise in both these components, so the variance is

$$\sigma_F^2 \approx (\sigma_{signal}^2)_{shot} + (\sigma_{signal}^2)_{flicker} + (\sigma_{backgr}^2)_{shot} + (\sigma_{backgr}^2)_{flicker}.$$

In fact, the background can be so large in these measurements that it is more useful to measure and compare S/B rather than S/N. Ingle & Crouch give a very thorough treatment of this topic that can be difficult to read in some places. Read that material with a view to extract the definitions and descriptions of how noise is generated in instrument components.

In some cases it is enough to measure the noise and report its magnitude with the results, but more often we analyze the noise because we are looking for ways to reduce it. There are essentially two ways to do this. In the first, we use hardware or electronic components such as electronic (LRC) filters, photon counting electronics or lock-in amplifiers to suppress the high-frequency noise in the measured signal. In the other we use digital filters (computer algorithms) to reduce the noise in the signal after it has been collected. Fourier mathematics play a role in this second strategy because signals and noise are frequently characterized by their frequency content. In most measurements, spectral information primarily consists of low frequency components. If the noise is white noise, as is the case in detector noise limited measurements, it consists of (roughly) equal amounts of noise at all frequencies. In this case, filtering the higher frequency components of the signal improves the signal-to-noise ratio of the data because the signal is unaffected. Fourier transforms can facilitate this process, but the same concerns that guide choices in sampling and apodization apply here. The high frequency components of the signal are removed by multiplying the transform of the data (not the inverse transform, even if the data is a spectrum) by a smoothing function that attenuates the high frequency components. The simplest smoothing function is the rectangle or window function. All the frequencies above some user selected cut off frequency are set to zero. The problem with this is that a sinc function could introduce ripples to the baseline of the spectrum as the last panel of Figure 9.2 illustrates. More

often filters employ Gaussian or other smoothly decaying functions, which eliminate the high frequency components without introducing the unstable baseline.

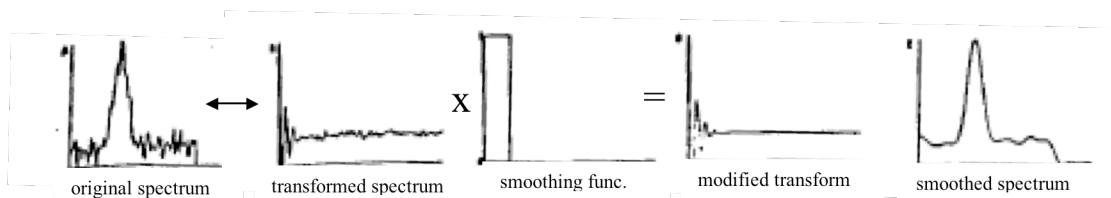


Figure 9.2: Digital filtering of Noisy Spectrum using Rectangle function

Filters do not always improve the quality of spectral data. Measurements that are dominated by noise from the source, e.g., fluorescence, are often distorted by flicker noise. Whereas white noise is frequency independent, flicker noise is larger at low frequencies. In fact, it is sometimes called pink noise (since low frequencies are associated with red rather than blue light). Filtering techniques are based on the idea that the noise does not, in general, consist of the same frequency components as the signal. In the case of flicker noise, this condition is not met. This is one reason that FT-UV-VIS spectrometers are not widely used. The improvement in signal to noise that comes from measuring the entire spectrum many times on an interferometer during the time a single spectrum can be measured on a dispersive (monochromator based) spectrometer is the square root of the number of times the spectrum was measured for shot noise limited (FT-IR) measurements. In other words, a ten point spectrum measured by the interferometer has a signal to noise ratio that is about three times ($\sqrt{10}$) larger than the same spectrum measured using a dispersive device because signal averaging reduces random noise by the square root of the number of acquisitions. (Remember this is called the multiplex or Fellgett advantage.) This improvement is always smaller in the presence of flicker noise, and in very crowded spectra, a multiplex disadvantage in which the signal-to-noise ratio in the data acquired by the interferometer is lower than in the dispersive instrument can be observed.