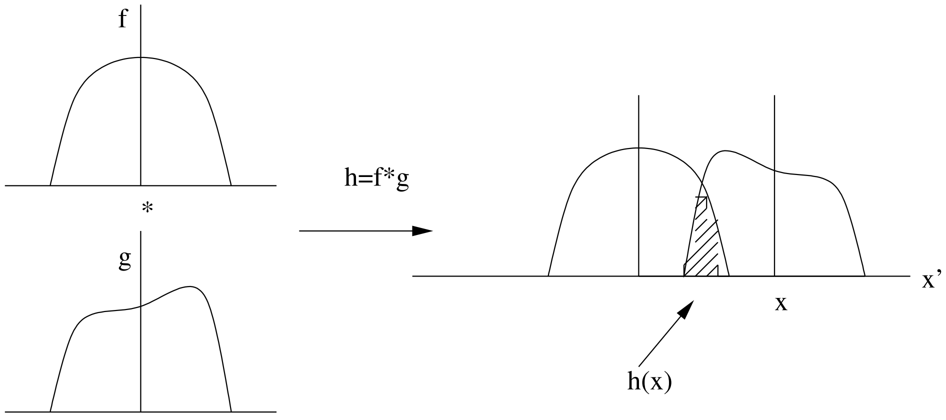

Figure 6.8: Illustration of the convolution of two functions, viewed as the area of the overlap resulting from a relative shift of x.

FOURIER ANALYSIS: LECTURE 11

6 Convolution

Convolution combines two (or more) functions in a way that is useful for describing physical systems (as we shall see). Convolutions describe, for example, how optical systems respond to an image, and we will also see how our Fourier solutions to ODEs can often be expressed as a convolution. In fact the FT of the convolution is easy to calculate, so it is worth looking out for when an integral is in the form of a convolution, for in that case it may well be that FTs can be used to solve it.

First, the definition. The convolution of two functions $f(x)$ and $g(x)$ is defined to be

$$f(x) * g(x) = \int_{-\infty}^{\infty} dx' \; f(x')g(x - x') \; , \tag{6.99}$$

The result is also a function of x, meaning that we get a different number for the convolution for each possible x value. Note the positions of the dummy variable x', especially that the argument of g is $x - x'$ and not $x' - x$ (a common mistake in exams).

There are a number of ways of viewing the process of convolution. Most directly, the definition here is a measure of *overlap*: the functions f and g are shifted relative to one another by a distance x, and we integrate to find the product. This viewpoint is illustrated in Fig. 6.8.

But this is not the best way of thinking about convolution. The real significance of the operation is that it represents a *blurring* of a function. Here, it may be helpful to think of $f(x)$ as a signal, and $g(x)$ as a blurring function. As written, the integral definition of convolution instructs us to take the signal at x', $f(x')$, and replace it by something proportional to $f(x')g(x - x')$: i.e. spread out over a range of x around x'. This turns a sharp feature in the signal into something fuzzy centred at the same location. This is exactly what is achieved e.g. by an out-of-focus camera.

Alternatively, we can think about convolution as a form of averaging. Take the above definition of convolution and put $y = x - x'$. Inside the integral, x is constant, so $dy = -dx'$. But now we are

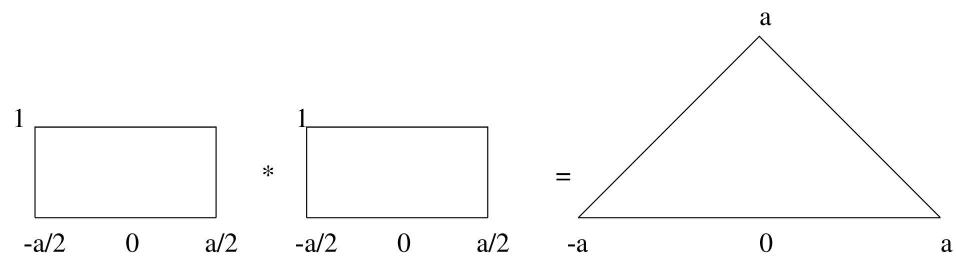

Figure 6.9: Convolution of two top hat functions.

integrating from $y = \infty$ to $-\infty$, so we can lose the minus sign by re-inverting the limits:

$$f(x) * g(x) = \int_{-\infty}^{\infty} dy\ f(x-y)g(y)\ . \tag{6.100}$$

This says that we replace the value of the signal at x, $f(x)$ by an average of all the values around x, displaced from x by an amount y and weighted by the function $g(y)$. This is an equivalent view of the process of blurring. Since it doesn't matter what we call the dummy integration variable, this rewriting of the integral showns that convolution is commutative: you can think of g blurring f or f blurring g:

$$f(x) * g(x) = \int_{-\infty}^{\infty} dz\ f(z)g(x-z) = \int_{-\infty}^{\infty} dz\ f(x-z)g(z) = g(x) * f(x). \tag{6.101}$$

6.1 Examples of convolution

1. Let $\Pi(x)$ be the top-hat function of width a.
   - $\Pi(x) * \Pi(x)$ is the triangular function of base width $2a$.
   - This is much easier to do by sketching than by working it out formally: see Figure 6.9.
2. Convolution of a general function $g(x)$ with a delta function $\delta(x-a)$.

   $$\delta(x-a) * g(x) = \int_{-\infty}^{\infty} dx'\ \delta(x'-a)g(x-x') = g(x-a). \tag{6.102}$$

   using the sifting property of the delta function. This is a clear example of the blurring effect of convolution: starting with a spike at $x = a$, we end up with a copy of the whole function $g(x)$, but now shifted to be centred around $x = a$. So here the 'sifting' property of a delta-function has become a 'shifting' property. Alternatively, we may speak of the delta-function becoming 'dressed' by a copy of the function g.

   The response of the system to a delta function input (i.e. the function $g(x)$ here) is sometimes called the *Impulse Response Function* or, in an optical system, the *Point Spread Function*.
3. Making double slits: to form double slits of width a separated by distance $2d$ between centres:

   $$[\delta(x+d) + \delta(x-d)] * \Pi(x)\ . \tag{6.103}$$

   We can form diffraction gratings with more slits by adding in more delta functions.

6.2 The convolution theorem

States that the Fourier transform of a *convolution* is a *product* of the individual Fourier transforms:

$$FT[f(x) * g(x)] = \tilde{f}(k)\, \tilde{g}(k) \tag{6.104}$$

$$FT[f(x)\, g(x)] = \frac{1}{2\pi} \tilde{f}(k) * \tilde{g}(k) \tag{6.105}$$

where $\tilde{f}(k)$, $\tilde{g}(k)$ are the FTs of $f(x)$, $g(x)$ respectively.

Note that:

$$\tilde{f}(k) * \tilde{g}(k) \equiv \int_{-\infty}^{\infty} dq\, \tilde{f}(q)\, \tilde{g}(k-q)\ . \tag{6.106}$$

We'll do one of these, and we will use the Dirac delta function.

The convolution $h = f * g$ is

$$h(x) = \int_{-\infty}^{\infty} f(x')g(x-x')\, dx'. \tag{6.107}$$

We substitute for $f(x')$ and $g(x-x')$ their FTs, noting the argument of g is not x':

$$\begin{aligned} f(x') &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx'}\, dk \\ g(x-x') &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{g}(k) e^{ik(x-x')}\, dk \end{aligned}$$

Hence (relabelling the k to k' in g, so we don't have two k integrals)

$$h(x) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx'}\, dk \int_{-\infty}^{\infty} \tilde{g}(k') e^{ik'(x-x')}\, dk' \right) dx'. \tag{6.108}$$

Now, as is very common with these multiple integrals, we do the integrations in a different order. Notice that the only terms which depend on x' are the two exponentials, indeed only part of the second one. We do this one first, using the fact that the integral gives 2π times a Dirac delta function:

$$\begin{aligned} h(x) &= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \tilde{f}(k) \int_{-\infty}^{\infty} \tilde{g}(k') e^{ik'x} \left(\int_{-\infty}^{\infty} e^{i(k-k')x'} dx' \right) dk' dk \\ &= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \tilde{f}(k) \int_{-\infty}^{\infty} \tilde{g}(k') e^{ik'x} \left[2\pi\delta(k-k') \right] dk' dk \end{aligned}$$

Having a delta function simplifies the integration enormously. We can do either the k or the k' integration immediately (it doesn't matter which you do – let us do k'):

$$\begin{aligned} h(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(k) \left[\int_{-\infty}^{\infty} \tilde{g}(k') e^{ik'x} \delta(k-k')\, dk' \right] dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(k) \tilde{g}(k)\, e^{ikx}\, dk \end{aligned}$$

Since

$$h(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{h}(k)\, e^{ikx}\, dk \tag{6.109}$$

we see that

$$\tilde{h}(k) = \tilde{f}(k)\tilde{g}(k). \tag{6.110}$$

Note that we can apply the convolution theorem in reverse, going from Fourier space to real space, so we get the most important key result to remember about the convolution theorem:

$$\begin{aligned} \text{Convolution in real space} &\Leftrightarrow \text{Multiplication in Fourier space} \\ \text{Multiplication in real space} &\Leftrightarrow \text{Convolution in Fourier space} \end{aligned} \tag{6.111}$$

This is an important result. Note that if one has a convolution to do, it is often most efficient to do it with Fourier Transforms, not least because a very efficient way of doing them on computers exists – the *Fast Fourier Transform*, or FFT.

CONVENTION ALERT! Note that if we had chosen a different convention for the 2π factors in the original definitions of the FTs, the convolution theorem would look differently. Make sure you use the right one for the conventions you are using!

Note that convolution commutes, $f(x) * g(x) = g(x) * f(x)$, which is easily seen (e.g. since the FT is $\tilde{f}(k)\tilde{g}(k) = \tilde{g}(k)\tilde{f}(k)$.)

Example application: Fourier transform of the triangular function of base width $2a$. We know that a triangle is a convolution of top hats:

$$\Delta(x) = \Pi(x) * \Pi(x) \ . \tag{6.112}$$

Hence by the convolution theorem:

$$FT[\Delta] = (FT[\Pi(x)])^2 = \left(\operatorname{sinc}\frac{ka}{2}\right)^2 \tag{6.113}$$

FOURIER ANALYSIS: LECTURE 12

6.3 Application of FTs and convolution

6.3.1 Fraunhofer Diffraction

Imagine a single-slit optics experiment (see Fig. 6.10). Light enters from the left, and interferes, forming a pattern on a screen on the right. We apply *Huygens' Principle*, which states that each point on the aperture acts as a source. Let the vertical position on the source be x, and the transmission of the aperture is $T(x)$. We will take this to be a top-hat for a simple slit,

$$T(x) = \begin{cases} 1 & -\frac{a}{2} < x < \frac{a}{2} \\ 0 & |x| > \frac{a}{2} \end{cases} , \tag{6.114}$$

but we will start by letting $T(x)$ be an arbitrary function (reflecting partial transmission, or multiple slits, for example).

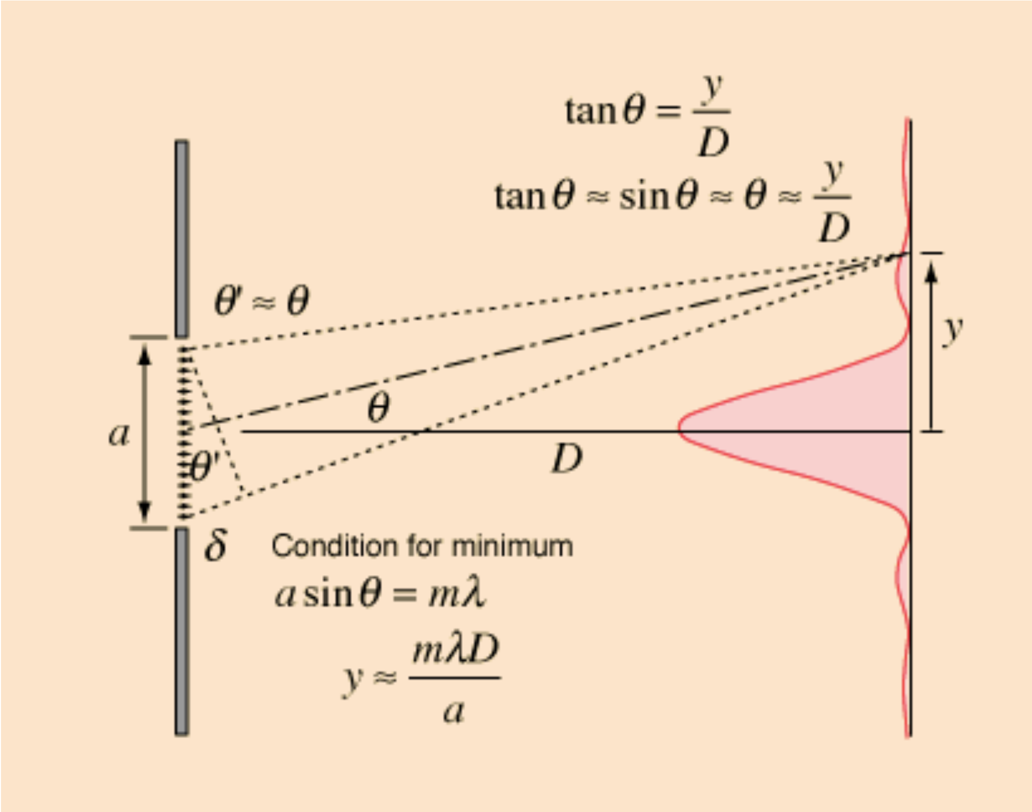

Figure 6.10: A slit of width a permitting light to enter. We want to compute the intensity on a screen a distance D away. Credit: Wikipedia

From a small element dx, the electric field at distance r (on the screen) is

$$dE = E_0 \frac{T(x)dx}{r}\, e^{i(kr-\omega t)}, \tag{6.115}$$

for some source strength E_0. To get the full electric field, we integrate over the aperture:

$$E(y) = E_0 \int_{-\infty}^{\infty} \frac{T(x)dx}{r} e^{i(kr-\omega t)}. \tag{6.116}$$

If the screen is far away, and the angle θ is small, then $\sin\theta = y/D \simeq \theta$, and $r \simeq D$ *in the denominator*. In the exponent, we need to be more careful.

If D is the distance of the screen from the origin of x (e.g. the middle of the slit), then Pythagoras says that

$$r = \left(D^2 + (y-x)^2\right)^{1/2} = D\,\left(1 + (y-x)^2/D^2\right)^{1/2} \simeq D + (y-x)^2/2D \tag{6.117}$$

(where we assume a distant screen and small angles, so that both x & y are $\ll D$). The terms in r that depend on x are $-(y/D)x + x^2/2D = -\theta x + (x/2D)x$; if we are interested in diffraction at fixed θ, we can always take the screen far enough away that the second term is negligible ($x/D \ll \theta$). To first order in θ, we then have the simple approximation that governs *Fraunhofer diffraction*:

$$r \simeq D - x\theta. \tag{6.118}$$

As a result,

$$E(y) \simeq \frac{E_0 e^{i(kD-\omega t)}}{D} \int_{-\infty}^{\infty} T(x) e^{ixk\theta} \, dx. \tag{6.119}$$

So we see that the electric field is proportional to the FT of the aperture $T(x)$, evaluated at $k\theta$:

$$E(y) \propto \tilde{T}\,(k\theta)\,. \tag{6.120}$$

Note the argument is not k!

The intensity $I(y)$ (or $I(\theta)$) is proportional to $|E(y)|^2$, so the phase factor cancels out, and

$$I(\theta) \propto \left|\tilde{T}\,(k\theta)\right|^2. \tag{6.121}$$

For a single slit of width a, we find as before that

$$I(\theta) \propto \mathrm{sinc}^2\left(\frac{ka\theta}{2}\right) \propto \mathrm{sinc}^2\left(\frac{\pi a\theta}{\lambda}\right), \tag{6.122}$$

where $\lambda = 2\pi/k$ is the wavelength of the light. Note that the first zero of the diffraction pattern is when the argument is π, so $\pi a\theta = \pi$, or $\theta = \lambda/a$.

If the wavelength is much less than the size of the object (the slit here), then the diffraction pattern is effectively confined to a very small angle, and effectively the optics is 'geometric' – i.e. straight-line ray-tracing with shadows etc).

Diffraction using the convolution theorem

Double-slit interference: two slits of width a and spacing $2d$. In terms of $k' \equiv \pi\theta/\lambda$,

$$\begin{aligned}
FT[\{\delta(x+d)+\delta(x-d)\} * \Pi(x)] &= FT[\{\delta(x+d)+\delta(x-d)\}]\ FT[\Pi(x)] && (6.123)\\
&= \{FT[\delta(x+d)] + FT[\delta(x-d)]\}\ FT[\Pi(x)] && (6.124)\\
&= (e^{ik'd} + e^{-ik'd})\ \mathrm{sinc}\left(\frac{k'a}{2}\right) && (6.125)\\
&= 2\cos(k'd)\,\mathrm{sinc}\left(\frac{k'a}{2}\right). && (6.126)
\end{aligned}$$

Hence the intensity is a sinc^2 function, modulated by a shorter-wavelength $\cos^2$ function. See Fig. 6.11.

6.3.2 Solving ODEs, revisited

Recall our problem

$$\frac{d^2z}{dt^2} - \omega_0^2 z = f(t). \tag{6.127}$$

Using FTs, we found that a solution was

$$z(t) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\tilde{f}(\omega)}{\omega_0^2 + \omega^2} e^{i\omega t} \, d\omega. \tag{6.128}$$

Now we can go a bit further, because we see that the FT of $z(t)$ is a product (in Fourier space), of $\tilde{f}(\omega)$ and

$$\tilde{g}(\omega) \equiv \frac{-1}{\omega_0^2 + \omega^2} \tag{6.129}$$

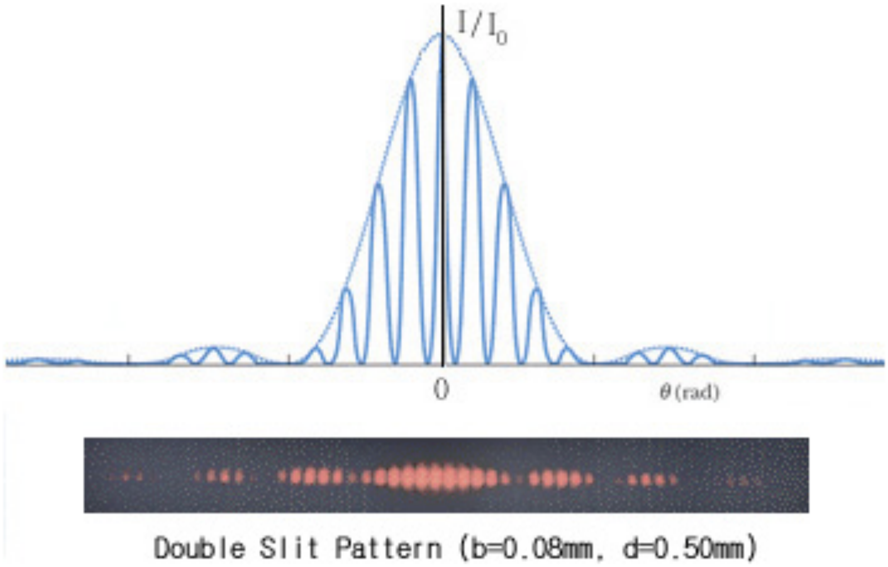

Figure 6.11: Intensity pattern from a double slit, each of width b and separated by $2d$ (Credit: Yonsei University) .

hence the solution is a convolution in real (i.e. time) space:

$$z(t) = \int_{-\infty}^{\infty} f(t')g(t-t')\,dt'. \tag{6.130}$$

An exercise for you is to show that the FT of

$$g(t) = -\frac{e^{-\omega_0|t|}}{2\omega_0} \tag{6.131}$$

is $\tilde{g}(\omega) = -1/(\omega_0^2 + \omega^2)$, so we finally arrive at the general solution for a driving force $f(t)$:

$$z(t) = -\frac{1}{2\omega_0}\int_{-\infty}^{\infty} f(t')e^{-\omega_0|t-t'|}\,dt'. \tag{6.132}$$

Note how we have put in $g(t-t') = e^{-\omega_0|t-t'|}/2\omega_0$ here, not $g(t)$ or $g(t')$, as required for a convolution.

FOURIER ANALYSIS: LECTURE 13

7 Parseval's theorem for FTs (Plancherel's theorem)

For FTs, there is a similar relationship between the average of the square of the function and the FT coefficients as there is with Fourier Series. For FTs it is strictly called *Plancherel's theorem*, but is often called the same as FS, i.e. Parseval's theorem; we will stick with Parseval. The theorem says

$$\int_{-\infty}^{\infty} |f(x)|^2\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\tilde{f}(k)|^2\,dk. \tag{7.133}$$

It is useful to compare different ways of proving this:

(1) The first is to go back to Fourier series for a periodic $f(x)$: $f(x) = \sum_n c_n \exp(ik_n x)$, and $|f|^2$ requires us to multiply the series by itself, which gives lots of cross terms. But when we integrate over one fundamental period, all oscillating terms average to zero. Therefore the only terms that survive are ones where $c_n \exp(ik_n x)$ pairs with $c_n^* \exp(-ik_n x)$. This gives us Parseval's theorem for Fourier series:

$$\frac{1}{\ell}\int_{-\ell/2}^{\ell/2} |f(x)|^2\, dx = \sum_n |c_n|^2 \Rightarrow \int_{-\ell/2}^{\ell/2} |f(x)|^2\, dx = \ell \sum_n |c_n|^2 = \frac{1}{\ell}\sum_n |\tilde{f}|^2, \tag{7.134}$$

using the definition $\tilde{f} = \ell c_n$. But the mode spacing is $dk = 2\pi/\ell$, so $1/\ell$ is $dk/2\pi$. Now we take the continuum limit of $\ell \to \infty$ and $dk \sum$ becomes $\int dk$.

(2) Alternatively, we can give a direct proof using delta-functions:

$$|f(x)|^2 = f(x) f^*(x) = \left(\frac{1}{2\pi}\int \tilde{f}(k)\exp(ikx)\, dk\right) \times \left(\frac{1}{2\pi}\int \tilde{f}^*(k')\exp(-ik'x)\, dk'\right), \tag{7.135}$$

which is

$$\frac{1}{(2\pi)^2}\iint \tilde{f}(k)\tilde{f}^*(k')\exp[ix(k-k')]\, dk\, dk'. \tag{7.136}$$

If we new integrate over x, we generate a delta-function:

$$\int \exp[ix(k-k')]\, dx = (2\pi)\delta(k-k'). \tag{7.137}$$

So

$$\int |f(x)|^2\, dx = \frac{1}{2\pi}\iint \tilde{f}(k)\tilde{f}^*(k')\,\delta(k-k')\, dk\, dk' = \frac{1}{2\pi}\int |\tilde{f}(k)|^2\, dk. \tag{7.138}$$

7.1 Energy spectrum

As in the case of Fourier series, $|\tilde{f}(k)|^2$ is often called the *Power Spectrum* of the signal. If we have a field (such as an electric field) where the energy density is proportional to the *square* of the field, then we can interpret the square of the Fourier Transform coefficients as the *energy* associated with each frequency. i.e. Total energy radiated is

$$\int_{-\infty}^{\infty} |f(t)|^2\, dt. \tag{7.139}$$

By Parseval's theorem, this is equal to

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} |\tilde{f}(\omega)|^2\, d\omega. \tag{7.140}$$

and we interpret $|\tilde{f}(\omega)|^2/(2\pi)$ as the energy radiated per unit (angular) frequency, at frequency ω.

7.1.1 Exponential decay

If we have a quantum transition from an upper state to a lower state, which happens spontaneously, then the intensity of emission will decay exponentially, with a timescale $\tau = 1/a$, as well as having a sinusoidal dependence with frequency ω_0:

$$f(t) = e^{-at}\cos(\omega_0 t) \qquad (t > 0). \tag{7.141}$$

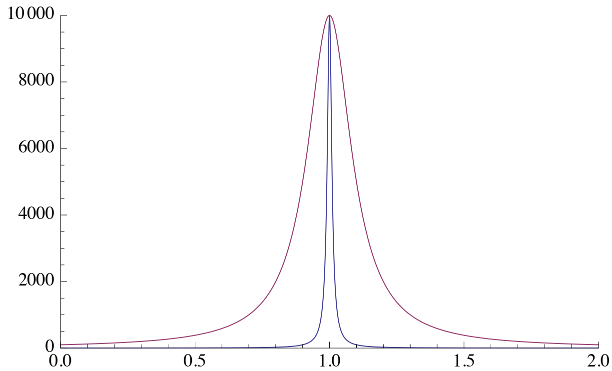

Figure 7.12: Frequency spectrum of two separate exponentially decaying systems with 2 different time constants τ. (x axis is frequency, y axis $\propto |\tilde{f}(\omega)|^2$ in arbitrary units).

Algebraically it is easier to write this as the real part of a complex exponential, do the FT with the exponential, and take the real part at the end. So consider

$$f(t) = \frac{1}{2}e^{-at}(e^{i\omega_0 t} + e^{-i\omega_0 t}) \qquad (t > 0). \tag{7.142}$$

The Fourier transform is [2]

$$\begin{aligned}
\tilde{f}(\omega) &= \frac{1}{2}\int_0^\infty (e^{-at-i\omega t+i\omega_0 t} + e^{-at-i\omega t-i\omega_0 t})\, dt & (7.143)\\
\Rightarrow 2\tilde{f}(\omega) &= \left[\frac{e^{-at-i\omega t+i\omega_0 t}}{-a-i\omega+i\omega_0} - \frac{e^{-at-i\omega t-i\omega_0 t}}{-a-i\omega-i\omega_0}\right]_0^\infty \\
&= \frac{1}{(a+i\omega-i\omega_0)} + \frac{1}{(a+i\omega+i\omega_0)} & (7.144)
\end{aligned}$$

This is sharply peaked near $\omega = \omega_0$; near this frequency, we therefore ignore the second term, and the frequency spectrum is

$$|\tilde{f}(\omega)|^2 \simeq \frac{1}{4\,[a+i(\omega-\omega_0)]}\frac{1}{[a-i(\omega-\omega_0)]} = \frac{1}{4\,[a^2+(\omega-\omega_0)^2]}. \tag{7.145}$$

This is a Lorentzian curve with width $a = 1/\tau$. Note that the width of the line in frequency is inversely proportional to the decay timescale τ. This is an example of the Uncertainty Principle, and relates the *natural width* of a spectral line to the decay rate. See Fig. 7.12.

7.2 Correlations and cross-correlations

Correlations are defined in a similar way to convolutions, but look carefully, as they are slightly different. With correlations, we are concerned with how similar functions are when one is displaced

[2]Note that this integral is *similar* to one which leads to Delta functions, but it isn't, because of the e^{-at} term. For this reason, you can integrate it by normal methods. If $a = 0$, then the integral does indeed lead to Delta functions.

by a certain amount. If the functions are different, the quantity is called the *cross-correlation*; if it is the same function, it is called the *auto-correlation*, or simply *correlation*.

The cross-correlation of two functions is defined by

$$c(X) \equiv \langle f^*(x)g(x+X)\rangle \equiv \int_{-\infty}^{\infty} f^*(x)g(x+X)\,dx. \tag{7.146}$$

Compare this with convolution (equation 6.99). X is sometimes called the *lag*. Note that cross-correlation does not commute, unlike convolution. The most interesting special case is when f and g are the same function: then we have the *auto-correlation function*.

The meaning of these functions is easy to visualise if the functions are real: at zero lag, the auto-correlation function is then proportional to the variance in the function (it would be equal if we divided the integral by a length ℓ, where the functions are zero outside that range). So then the *correlation coefficient* of the function is

$$r(X) = \frac{\langle f(x)f(x+X)\rangle}{\langle f^2\rangle}. \tag{7.147}$$

If r is small, then the values of f at widely separated points are unrelated to each other: the point at which r falls to $1/2$ defines a characteristic width of a function. This concept is used particularly in random processes.

The FT of a cross-correlation is

$$\tilde{c}(k) = \tilde{f}^*(k)\,\tilde{g}(k). \tag{7.148}$$

This looks rather similar to the convolution theorem, which is is hardly surprising given the smilarity of the definitions of cross-correlation and convolution. Indeed, the result can be proved directly from the convolution theorem, by writing the cross-correlation as a convolution.

A final consequence of this is that the FT of an auto-correlation is just the power spectrum; or, to give the inverse relation:

$$\langle f^*(x)f(x+X)\rangle = \frac{1}{2\pi}\int |\tilde{f}|^2 \exp(ikX)\,dk. \tag{7.149}$$

This is known as the *Wiener-Khinchin theorem*, and it generalises Parseval's theorem (to which it reduces when $X=0$).

7.3 Fourier analysis in multiple dimensions

We have now completed all the major tools of Fourier analysis, in one spatial dimension. In many cases, we want to consider more than one dimension, and the extension is relatively straightforward. Start with the fundamental Fourier series, $f(x) = \sum_n c_n \exp(i2\pi nx/\ell_x)$. $f(x)$ can be thought of as $F(x,y)$ at constant y; if we change y, the effective $f(x)$ changes, so the c_n must depend on y. Hence we can Fourier expand these as a series in y:

$$c_n(y) = \sum_m d_{nm} \exp(i2\pi my/\ell_y), \tag{7.150}$$

where we assume that the function is periodic in x, with period ℓ_x, and y, with period ℓ_y. The overall series is than

$$F(x,y) = \sum_{n,m} d_{nm} \exp[2\pi i(nx/\ell_x + my/\ell_y)] = \sum_{n,m} d_{nm} \exp[i(k_x x + k_y y)] = \sum_{n,m} d_{nm} \exp[i(\boldsymbol{k}\cdot\boldsymbol{x})]. \tag{7.151}$$

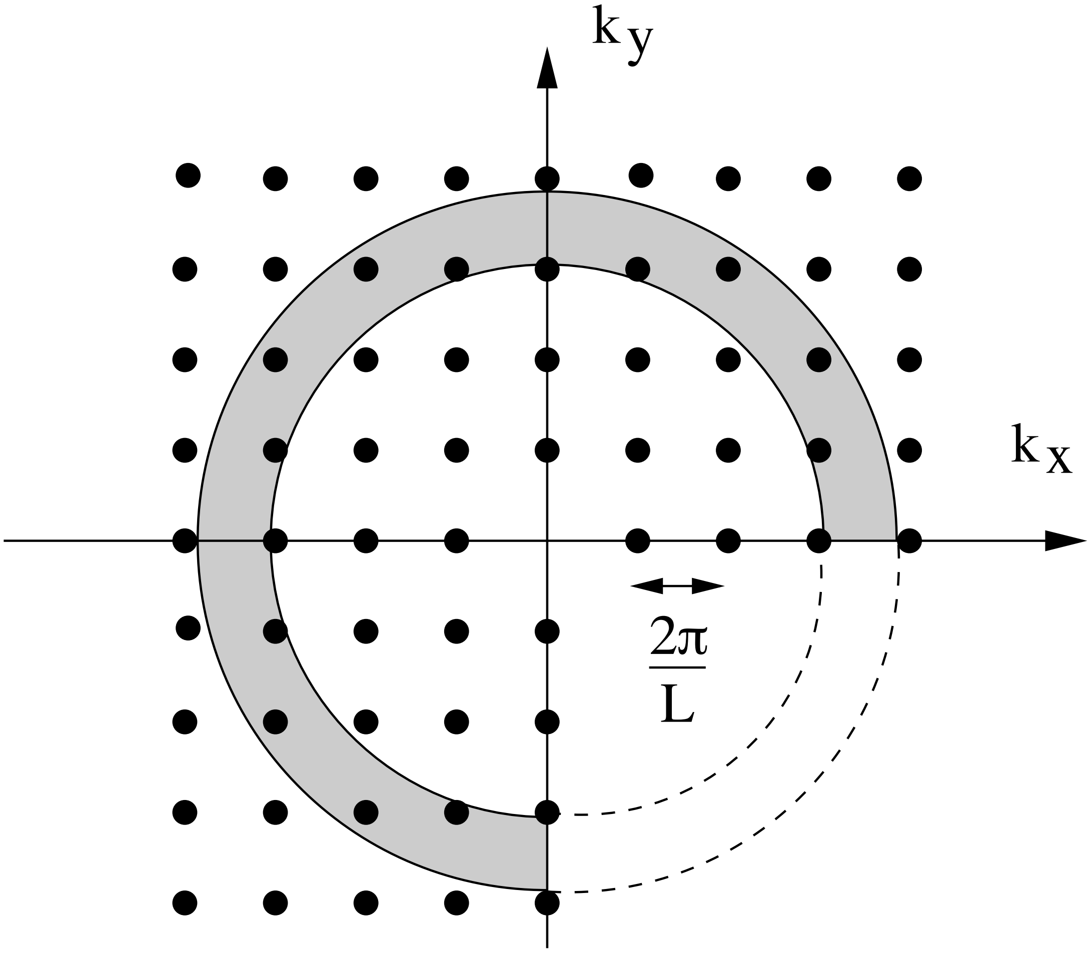

Figure 7.13: Illustrating the origin of the density of states in 2D. The allowed modes are shown as points, with a separation in k_x and k_y of $2\pi/\ell$, where ℓ is the periodicity. The number of modes between $|k|$ and $|k| + d|k|$ (i.e. inside the shaded annulus) is well approximated by $(\ell/2\pi)^2$ times the area of the annulus, as $\ell \to \infty$, and the mode spacing tends to zero. Clearly, in D dimensions, the mode density is just $(\ell/2\pi)^D$.

This is really just the same as the 1D form, and the extension to D dimensions should be obvious. In the end, we just replace the usual kx term with the dot product between the position vector and the wave vector.

The Fourier transform in D dimensions just involves taking the limit of $\ell_x \to \infty$, $\ell_y \to \infty$ etc. The Fourier coefficients become a continuous function of $\boldsymbol{k}$, in which case we can sum over bins in k space, each containing $N_{\rm modes}(\boldsymbol{k})$ modes:

$$F(\boldsymbol{x}) = \sum_{\rm bin} d(\boldsymbol{k}) \, \exp[i(\boldsymbol{k} \cdot \boldsymbol{x})] \, N_{\rm modes}. \tag{7.152}$$

The number of modes in a given k-space bin is set by the period in each direction: allowed modes lie on a grid of points in the space of k_x, k_y etc. as shown in Figure 7.13. If for simplicity the period is the same in all directions, the *density of states* is $\ell^D/(2\pi)^D$:

$$N_{\rm modes} = \frac{\ell^D}{(2\pi)^D} \, d^D k. \tag{7.153}$$

This is an important concept which is used in many areas of physics.

The Fourier expression of a function is therefore

$$F(\boldsymbol{x}) = \frac{1}{(2\pi)^D} \int \tilde{F}(\boldsymbol{k}) \, \exp[i(\boldsymbol{k} \cdot \boldsymbol{x}) \, d^D k], \tag{7.154}$$

Where we have defined $F(\boldsymbol{k}) \equiv \ell^D d(\boldsymbol{k})$. The inverse relation would be obtained as in 1D, by appealing to orthogonality of the modes:

$$\tilde{F}(\boldsymbol{k}) = \int F(\boldsymbol{x}) \, \exp[-i(\boldsymbol{k} \cdot \boldsymbol{x})] \, d^D x. \tag{7.155}$$

FOURIER ANALYSIS: LECTURE 14

8 Digital analysis and sampling

Imagine we have a continuous signal (e.g. pressure of air during music) which we sample by making measurements at a few particular times. Any practical storage of information must involve this step of *analogue-to-digital conversion*. This means we are converting a continuous function into one that is only known at discrete points – i.e. we are throwing away information. We would feel a lot more comfortable doing this if we knew that the missing information can be recovered, by some from of interpolation between the sampled points. Intuitively, this seems reasonable if the sampling interval is very fine: by the definition of continuity, the function between two sampled points should be arbitrarily close to the average of the sample values as the locations of the samples gets closer together. But the sampling interval has to be finite, so this raises the question of how coarse it can be; clearly we would prefer to use as few samples as possible consistent with not losing any information. This question does have a well-posed answer, which we can derive using Fourier methods.

The first issue is how to represent the process of converting a function $f(x)$ into a set of values $\{f(x_i)\}$. We can do this by using some delta functions:

$$f(x) \to f_s(x) \equiv f(x) \sum_i \delta(x - x_i). \tag{8.156}$$

This replaces our function by a sum of spikes at the locations x_i, each with a weight $f(x_i)$. This representation of the sampled function holds the information of the sample values and locations. So, for example, if we try to average the sampled function over some range, we automatically get something proportional to just adding up the sample values that lie in the range:

$$\int_{x_1}^{x_2} f_s(x) \, dx = \sum_{\text{in range}} f(x_i). \tag{8.157}$$

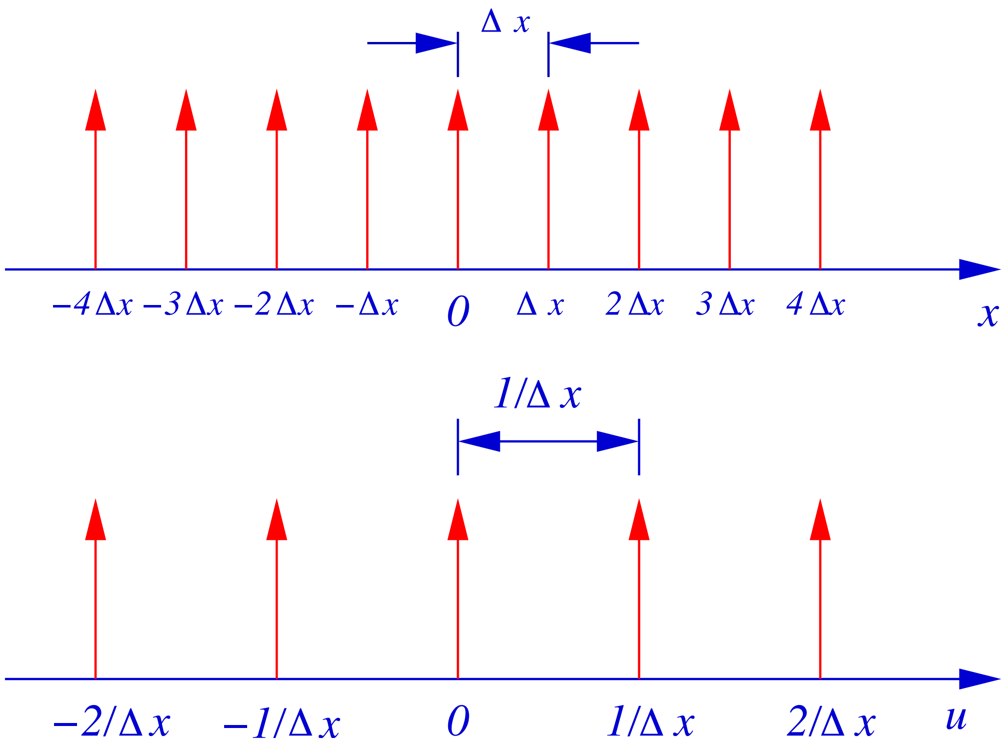

Figure 8.14: Top: An infinite comb in real space. This represents the sampling pattern of a function which is sampled regularly every Δx. Bottom: The FT of the infinite comb, which is also an infinite comb. Note that u here is $k/(2\pi)$.

8.1 The infinite comb

If we sample regularly with a spacing Δx, then we have an 'infinite comb' – an infinite series of delta functions. The comb is (see Fig. 8.14):

$$g(x) = \sum_{j=-\infty}^{\infty} \delta(x - j\Delta x) \tag{8.158}$$

This is also known as the *Shah function*.

To compute the FT of the Shah function, we will write it in another way. This is derived from the fact that the function is periodic, and therefore suitable to be written as a Fourier *series* with $\ell = \Delta x$:

$$g(x) = \sum_n c_n \exp(2\pi nix/\Delta x). \tag{8.159}$$

The coefficients c_n are just

$$c_n = \frac{1}{\Delta x} \int_{-\Delta x/2}^{\Delta x/2} \delta(x)\, dx = \frac{1}{\Delta x}, \tag{8.160}$$

so that

$$g(x) = \frac{1}{\Delta x} \sum_n \exp(2\pi nix/\Delta x) = \frac{1}{2\pi} \int \tilde{g}(k) \exp(ikx)\, dx. \tag{8.161}$$

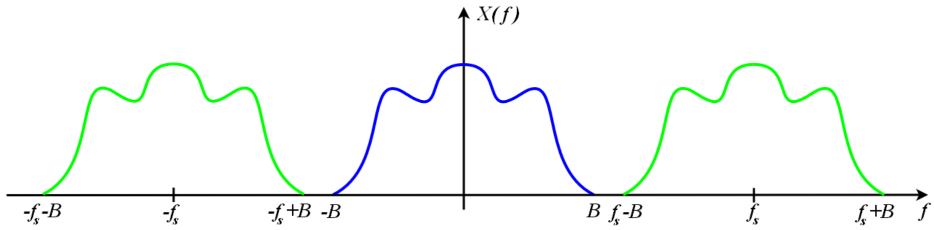

Figure 8.15: If the sampling is fine enough, then the original spectrum can be recovered from the sampled spectrum.

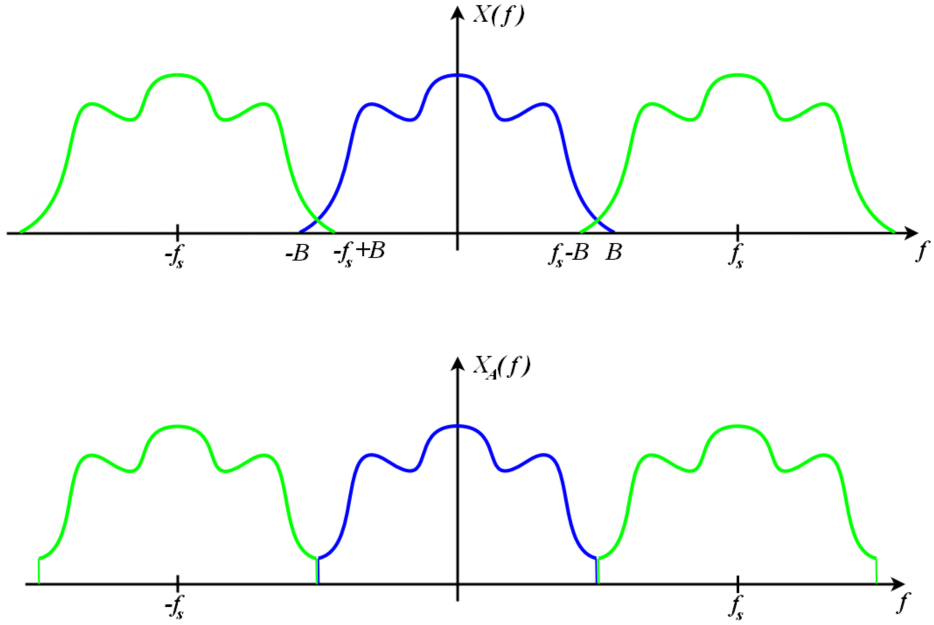

Figure 8.16: If the sampling is *not* fine enough, then the power at different frequencies gets mixed up, and the original spectrum cannot be recovered.

From this, the Fourier transform is obvious (or it could be extracted formally by integrating our new expression for $g(x)$ and obtaining a delta-function):

$$\tilde{g}(k) = \frac{2\pi}{\Delta x} \sum_{n=-\infty}^{\infty} \delta(k - 2\pi n/\Delta x). \tag{8.162}$$

which is an infinite comb in Fourier space, with spacing $2\pi/\Delta x$.

The FT of a function sampled with an infinite comb is therefore ($1/2\pi$ times) the convolution of this and the FT of the function:

$$\tilde{f}_s(k) = \frac{1}{2\pi} \tilde{f}(k) * \tilde{g}(k) = \frac{1}{\Delta x} \sum_{n=-\infty}^{\infty} \tilde{f}(k - 2\pi n/\Delta x). \tag{8.163}$$

In other words, each delta-function in the k-space comb becomes 'dressed' with a copy of the transform of the original function.

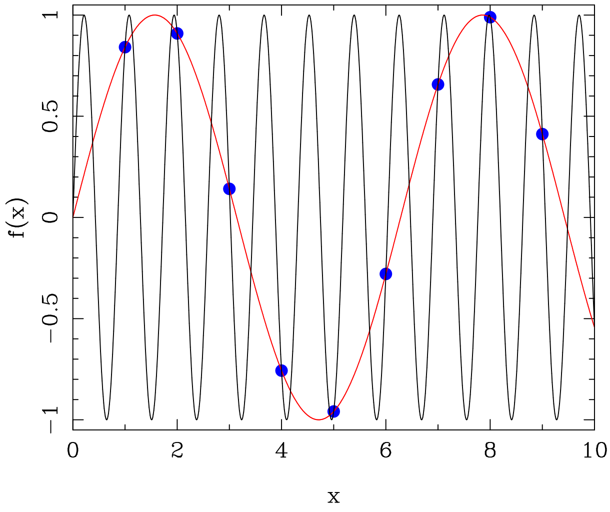

Figure 8.17: If $\sin t$ is sampled at unit values of t, then $\sin(t + 2\pi t)$ is indistinguishable at the sampling points. The sampling theorem says we can only reconstruct the function between the samples if we know that high-frequency components are absent.

8.2 Shannon sampling, aliasing and the Nyquist frequency

We can now go back to the original question: do the sampled values allow us to reconstruct the original function exactly? An equivalent question is whether the transform of the sampled function allows us to reconstruct the transform of the original function.

The answer is that this is possible (a) if the original spectrum is *bandlimited*, which means that the power is confined to a finite range of wavenumber (i.e. there is a maximum wavenumber k_{max} which has non-zero Fourier coefficients); and (b) if the sampling is fine enough. This is illustrated in Figs 8.15 and 8.16. If the sampling is not frequent enough, the power at different wavenumbers gets mixed up. This is called *aliasing*. The condition to be able to measure the spectrum accurately is to have a sample at least as often as the *Shannon Rate*

$$\Delta x = \frac{1}{\pi k_{\rm max}}. \tag{8.164}$$

The Nyquist wavenumber is defined as

$$k_{\rm Nyquist} = \frac{\pi}{\Delta x} \tag{8.165}$$

which needs to exceed the maximum wavenumber in order to avoid aliasing:

$$k_{\rm Nyquist} \geq k_{\rm max}. \tag{8.166}$$

For time-sampled data (such as sound), the same applies, with wavenumber k replaced by frequency ω.

There is a simple way of seeing that this makes sense, as illustrated in Figure 8.17. Given samples of a Fourier mode at a certain interval, Δ, a mode with a frequency increased by any multiple of $2\pi/\Delta$ clearly has the same result at the sample points.

8.3 CDs and compression

Most human beings can hear frequencies in the range 20 Hz – 20 kHz. The sampling theorem means that the sampling frequency needs to be at least 40 kHz to capture the 20 kHz frequencies. The CD standard samples at 44.1 kHz. The data consist of stereo: two channels each encoded as 16-bit integers. Allowing one bit for sign, the largest number encoded is thus $2^{15} - 1 = 32767$. This allows signals of typical volume to be encoded with a fractional precision of around 0.01% – an undetectable level of distortion. This means that an hour of music uses about 700MB of information. But in practice, this requirement can be reduced by about a factor 10 without noticeable degradation in quality. The simplest approach would be to reduce the sampling rate, or to encode the signal with fewer bits. The former would require a reduction in the maximum frequency, making the music sound dull; but fewer bits would introduce distortion from the quantization of the signal. The solution implemented in the MP3 and similar algorithms is more sophisticated than this: the time series is split into 'frames' of 1152 samples (0.026 seconds at CD rates) and each is Fourier transformed. Compression is achieved by storing simply the amplitudes and phases of the strongest modes, as well as using fewer bits to encode the amplitudes of the weaker modes, according to a 'perceptual encoding' where the operation of the human ear is exploited – knowing how easily faint tones of a given frequency are masked by a loud one at a different frequency.

8.4 Prefiltering

If a signal does not obey the sampling theorem, it must be modified to do so before digitization. Analogue electronics can suppress high frequencies – although they are not completely removed. The sampling process itself almost inevitably performs this task to an extent, since it is unrealistic to imagine that one could make an instantaneous sample of a waveform. Rather, the sampled signal is probably an average of the true signal over some period.

This is easily analysed using the convolution theorem. Suppose each sample, taken at an interval τ, is the average of the signal over a time interval T, centred at the sample time. This is a convolution:

$$f_c(t) = \int f(t')g(t - t')\, dt', \tag{8.167}$$

where $g(t - t')$ is a top hat of width T centred on $t' = t$. We therefore know that

$$\tilde{f}_c(\omega) = \tilde{f}(\omega) \sin(\omega T/2)/(\omega T/2). \tag{8.168}$$

At the Nyquist frequency, π/τ, the Fourier signal in f is suppressed by a factor $\sin(\pi T/2\tau)/(\pi T/2\tau)$. The natural choice of T would be the same as τ (accumulate an average signal, store it, and start again). This gives $\sin(\pi/2)/(\pi/2) = 0.64$ at the Nyquist frequency, so aliasing is not strongly eliminated purely by 'binning' the data, and further prefiltering is required before the data can be sampled.