The Discrete Fourier Transform and Sampling Theory

C. Godsalve
email:seagods@hotmail.com

June 29, 2009

Contents

1 Introduction **1**

2 Introducing the Discrete Fourier Transform **2**

3 Using the DFT for Data Analysis **3**

4 The N'th Root of Unity and the Fast Fourier Transform **5**

5 Some Theorems **7**

5.1 Parseval's and Plancherel's theorem's 7

5.2 Periodicity and Shifting 7

5.3 Convolution and Correlation 8

6 The Sampling Theorem **9**

7 Resampling in Two Dimensions **12**

1 Introduction

In this short article, we look at two separate subjects which are closely related. First we look at the Discrete Fourier transform, and then we look introduce sampling using the *continuous* form of the Fourier transform. Most undergraduate courses cover the continuous case quite well, but might miss out sampling theory. Mathematics courses for electronics degrees will cover the Discrete Fourier transform well, but the Discrete transform may be passed over in other numerate degree mathematics courses.

So, this brief introduction to the subjects may be useful to "plug a gap" in the reader's mathematical education. We recommend [1] as an accessible and thorough book on the subject as a whole. In the discrete Fourier transform we only mention windowing (a subject in itself) in passing). At the end of the section on the continuous transform, we mention the possibility of de-blurring.

The author is used to spatial data rather than time series, so has used x as a variable rather than t, also we have stuck to the frequency variable f rather than the angular frequency $\omega = 2\pi f$.

2 Introducing the Discrete Fourier Transform

Suppose we have some function $f(x)$ that takes the form

$$f(x) = \sum_{p=0}^{N-1} F_p \exp(ipx). \tag{1}$$

We impose the limitation $0 \leq x \leq 2\pi$. There is no loss of generality here, since we can map the interval $(0{,}2\pi)$ onto any other interval we please. We choose N values of x at equal intervals, so we have a set of points x_n such that

$$x_n = \left(\frac{2\pi n}{N}\right), \; n = 0, 1, 2, \ldots N-1. \tag{2}$$

We might think that we can make a rough approximation that, given some integer k, then

$$\int_0^{2\pi} f(x)e^{-ikx}\, dx \approx \frac{2\pi}{N} \sum_{n=0}^{N-1} f(2\pi n/N)\; \exp[-i \times (2\pi kn/N)]. \tag{3}$$

We have just an average of $f(x)\exp(-ikx)$ at equal intervals on the $(0, 2\,\pi)$. However, *this expression is exact* if $f(x)$ has the form of of eqn.1. Using eqn.1,

$$\int_0^{2\pi} f(x)e^{-ikx}\, dx = \frac{2\pi}{N} \sum_{n=0}^{N-1} \sum_{p=0}^{N-1} F_p \exp\left[\frac{2\pi i}{N}(pn - kn)\right].$$

$$= \frac{2\pi}{N} \sum_{p=0}^{N-1} F_p \sum_{n=0}^{N-1} \exp\left[\frac{2\pi in}{N}(p-k)\right]. \tag{4}$$

We need to examine S, where

$$S = \sum_{n=0}^{N-1} \exp\left[\frac{2\pi in}{N}(p-k)\right]. \tag{5}$$

We also need to examine

$$\exp\left[\frac{2\pi i}{N}(p-k)\right] S = \sum_{n=1}^{N} \exp\left[\frac{2\pi in}{N}(p-k)\right]. \tag{6}$$

We find that on subtracting eqns 5. and 6. that

$$S = \frac{1 - \exp(2\pi i(p-k))}{1 - \exp(2\pi i(p-k)/N)}. \tag{7}$$

As stated, p and k are integers, so S is zero if $p \neq k$, because $\exp(2\pi i \times integer) = 1$. However, if $p = k$ eqn.7 gives zero divided by zero. Looking back to eqn.5, we see that if $p = k$, then $S = N$.

So, given $f(x)$ of the form in eqn.1, then we have

$$\int_0^{2\pi} f(x)e^{-ikx}dx = \frac{2\pi}{N}\sum_{n=0}^{N} f\left(\frac{2\pi n}{N}\right)\exp\left(-ik\frac{2\pi n}{N}\right) = 2\pi F_k. \tag{8}$$

This motivates us to *define* the Discrete Fourier Transform (DFT) of $f(x)$ to be

$$\mathscr{F}_k(f) = \frac{2\pi}{N}\sum_{n=0}^{N} f_n \exp\left(-ik\frac{2\pi n}{N}\right) = 2\pi F_k,$$

$$k = 0, 1, 2, \ldots N-1. \tag{9}$$

We note that we have introduced a level of abstraction here in that f is no longer necessarily a function of the form of eqn.1, and it is not necessarily evenly sampled. It is, in this form, just a map of N (perhaps complex) numbers onto another set of N numbers. However, in data analysis, we assume the f_k to be an evenly spaced sampling of some function $f(x)$ on the interval (0,2π). In fact, if the function values are real, then we have $F_k = F^*_{N=k}$ so about half the DFT coefficients are redundant.

3 Using the DFT for Data Analysis

So, what does all this mean? If we are given N data values f_n at equal intervals on (0,2 π), and f *is* of the form of eqn.1, the Fourier series expansion coefficients are

$$F_k = \mathscr{F}_k(f) = \frac{1}{N}\sum_{n=0}^{N-1} f_n \exp\left(-\frac{2\pi ikn}{N}\right) \tag{10}$$

(The 2π factors cancel.) As we do this for $k = 0$ to $k = N-1$, we project the N points in real space x to N points in frequency space. To reconstruct the function at $x = 2n\pi/N$, we use the inverse discrete transform

$$f_n = \mathscr{F}^{-1}(F_k) = \sum_{k=0}^{N-1} F_k \exp\left(\frac{2\pi ink}{N}\right) \tag{11}$$

Note that we may have chosen to set this up is different to the standard definition. In the standard definition the Fourier Transform is not scaled to by $1/N$ which of course means that the inverse transform is scaled by $1/N$ instead. Also, we could define the transform as to have greater symmetry by defining both the transform and its inverse with a scaling of $1/\sqrt{N}$.

It seems that we can get the frequency spectrum of our periodic function from a set of function values evaluated at equal intervals from this simple summation. The function need not be known, just the function values (or measurements). If we can, this frequency spectrum can enable us to reconstruct exactly the continuous function $f(x)$ given the finite set of measurements. We do this by using eqn.11 with values of x not equal to $2\pi n/N$. Typically, it is required to reconstruct the function at a different resolution or sample spacing of the function. That is to say we may use this technique to *resample* the data.

So, what's the snag? (And yes, there is a snag.) Eqn.1 has a definite upper limit on the (spatial) frequency of $(N-1)$. In the parlance of the DFT, it is *band limited.* The *bandwidth* of f is $N-1$.

(If you have come across this subject before you might think there is a factor of 2 missing in the following. There isn't. We are not using the one sided base bandwidth over a frequency range $(-\infty, +\infty)$, that is all.)

It may well be the case that the data is band limited and the bandwidth is known. For instance the data may have been filtered in some way in the measuring process, or it is known by some other information about what it is that is being measured. Given that maximum value of N, we know the sampling interval required to reconstruct the function. In other words we must sample at *at least* $N-1$ evenly spaced intervals to reconstruct a function that has a bandwidth of $N-1$.

If we can only sample at a given frequency we must filter the data so that no higher frequency parts of $f(x)$ show up. Then we can only reconstruct the filtered function. If we have sampled the function at a particular (spatial) frequency, but don't know what the bandwidth of f is, we have no idea what we are reconstructing.

Suppose we are mistaken somehow. Suppose for instance that the filter isn't as good as we think and extra frequencies will show up. What happens then? Look at the denominator in eqn.7 we see the term like $(1-\exp(2\pi r(p-q)/N))$. We see the denominator vanishes if $p=q$ and $p+N-q$=0, and $p+2N-q$ =0. So, higher frequency terms will contribute to what we think are lower frequency components.

This phenomenon is known as *aliasing.* For instance, if p stopped at N instead of $N-1$, we would calculate an F_0 for our our data with a value that is really F_0+F_N.

So, we know that if the function is band limited, and we know the bandwidth, we know the minimum sample frequency (the *Nyquist* frequency) with which we can reconstruct the function. If we have, for instance, white noise added to the function, then there is no upper limit. What we do then, is to say, well we cannot reconstruct the function exactly, but we can apply a filter which "blurs" the data slightly, and reconstruct the blurred function.

This is better than having no idea what the reconstruction in between the data points actually means. In the following, we swap over to the more usual notation, where the $1/N$ term appears in the inverse DFT rather than the forward DFT.

Oh yes, we have not mentioned the *other snag*. The function which we are sampling will, in general, have components that are not in any of our f_k. It may have have a continuous Fourier Transform, or it may have just a single frequency between f_k and f_{k+1}. In this case, the DFT will have a spike, but it will also have other peaks called *side lobes*. This phenomena is called *spectral leakage*. There are measures that can reduce this problem. These methods involve what is known as *windowing*. There are many different kinds of window functions, and these have different merits for different applications. So, we shall just mention that they exist in this short article.

4 The N'th Root of Unity and the Fast Fourier Transform

There are N roots of $x^N = 1$. For instance, if N=8, these are $\zeta_0 = 1$, $\zeta_1 = e^{2\pi/8}$, $\zeta_2 = e^{4\pi/8}$, $\ldots$, $\zeta_7 = e^{14\pi/8}$. If we multiply all these by ζ_1, we get ζ_1, ζ_2, ζ_3, $\ldots$, ζ_7, ζ_0. The roots form a cyclic group. We also note that if $\zeta_i^N = 1$, then $(1/\zeta_i)^N = 1$. In this light we may write

$$\mathscr{F}_k(f(x)) = F_k = \frac{1}{N}\sum_{n=0}^{N-1} f_n \zeta_n^k. \tag{12}$$

We write out explicitly all the terms for $N = 8$, then $N = 4$, and then $N = 2$.

$$\begin{aligned}
F_0 &= f_0 + f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_8,\\
F_1 &= f_0 + \zeta_1 f_1 + \zeta_2 f_2 + \zeta_3 f_3 + \zeta_4 f_4 + \zeta_5 f_5 + \zeta_6 f_6 + \zeta_7 f_7,\\
F_2 &= f_0 + \zeta_2 f_1 + \zeta_4 f_2 + \zeta_6 f_3 + f_4 + \zeta_2 f_5 + \zeta_4 f_6 + \zeta_6 f_7,\\
F_3 &= f_0 + \zeta_3 f_1 + \zeta_6 f_2 + \zeta_1 f_3 + \zeta_4 f_4 + \zeta_7 f_5 + \zeta_2 f_6 + \zeta_5 f_7,\\
F_4 &= f_0 + \zeta_4 f_1 + f_2 + \zeta_4 f_3 + f_4 + \zeta_4 f_5 + f_6 + \zeta_4 f_7,\\
F_5 &= f_0 + \zeta_5 f_1 + \zeta_2 f_2 + \zeta_7 f_3 + \zeta_4 f_4 + \zeta_1 f_5 + \zeta_6 f_6 + \zeta_3 f_7,\\
F_6 &= f_0 + \zeta_6 f_1 + \zeta_4 f_2 + \zeta_2 f_3 + f_4 + \zeta_6 f_5 + \zeta_4 f_6 + \zeta_2 f_7,\\
F_7 &= f_0 + \zeta_7 f_1 + \zeta_6 f_2 + \zeta_5 f_3 + \zeta_4 f_4 + \zeta_3 f_5 + \zeta_2 f_6 + \zeta_1 f_7.
\end{aligned} \tag{13}$$

Here we have used the cyclic nature of the roots of unity. Clearly $\zeta_M = \zeta_{M mod(N)}$ and so all the powers or products of the ζ_i map onto the Nth roots of unity. For $N = 4$ we have

$$F_0 = f_0 + f_1 + f_2 + f_3,$$

$$F_1 = f_0 + \zeta_1 f_1 + \zeta_2 f_2 + \zeta_3 f3,$$

$$F_2 = f_0 + \zeta_2 f_1 + f_2 + \zeta_2 f_3,$$
$$F_3 = f_0 + \zeta_3 f_1 + \zeta_2 f_2 + \zeta_1 f_3. \tag{14}$$

Finally, for $N = 2$

$$F_0 = f_0 + f_1,$$
$$F_1 = f_0 - f_1, \tag{15}$$

since $\zeta_0 = 1$ and $\zeta_1 = -1$. Now, the roots for $N = 2, 4, 8$ are distributed on the input circle centred at zero in the complex plane as in Fig.1. Looking at eqn.14 in the light of eqn.15,

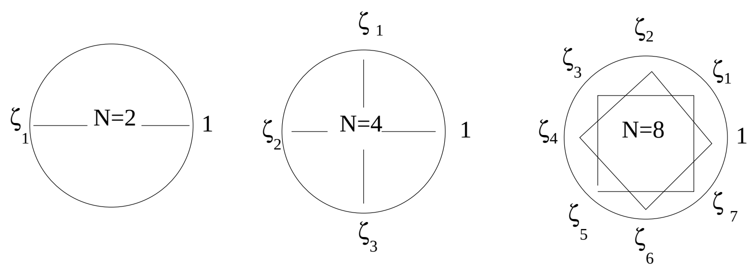

Figure 1: The Nth roots of unity for $N = 2$, $N = 4$, and $N = 8$.

we can write

$$F_0 = (f_0 + f_2) + (f_1 + f_3),$$
$$F_1 = (f_0 - f_2) + \zeta_1 (f_1 - f_3),$$
$$F_2 = (f_0 + f_2) - (f_1 + f_3),$$
$$F_3 = (f_0 - f_2) - \zeta_1 (f_1 - f_3). \tag{16}$$

That is to say, if we have the $N = 2$ DFTs on (f_0, f_2) and (f_1, f_3) we can construct the $N = 4$ transform as linear combinations of the $N = 2$ transforms. In Fig.1 we use the horizontal line for the $N = 2$ circle to represent the $N = 2$ DFT. For the $N = 4$ circle, we use the two lines to represent the fact that $N = 4$ DFT can be written in terms of two $N = 2$ DFTs. The reader may verify that the $N = 8$ DFT can be represented in terms of two $N = 4$ DFTs as represented by the two squares in Fig.1. The DFT was discovered by Gauss when he was devising interpolation methods for the orbit of the asteroid Ceres. In the same year (1805) Gauss also discovered the above method [1], which is the Fast Fourier Transform or FFT.

If we have 2^N data points, we need split this into transforms of order 2^{N-1} transforms, which turn can be split, and so on until we get down to order 2 DFTs. This means that instead of the order of N^2 operations required for the usual DFT, we have the number of operations reduced down to the order of $N \log_2 N$. Sometimes this approach is called a "divide and conquer" algorithm. This gives us the "Fast" in the FFT. Strangely, the FFT wasn't widely known until Cooley and Tukey re-invented it for engineering applications in 1965. However, there are now a large variety of FFT methods, and widespread and easily available computer programmes to perform them.

5 Some Theorems

5.1 Parseval's and Plancherel's theorem's

The reader may prove what is known as Plancherel's theorem. If F_i are the DFT components of f_i, and G_i are DFT components of g_i (of the same order and on the same interval) then

$$\sum_{n=0}^{N-1} f_n g_n^* = \frac{1}{N} \sum_{n=0}^{N-1} F_n G_n^*. \tag{17}$$

Parseval's theorem follows immediately, namely

$$\sum_{n=0}^{N-1} |f_n|^2 = \frac{1}{N} \sum_{n=0}^{N-1} |F_n|^2. \tag{18}$$

5.2 Periodicity and Shifting

Clearly, given the DFT of $f(x)$ we can extend the evaluation of $f(x)$ via the DFT coefficients outside the interval $(0, 2\pi)$. If we do so, it is obvious that $f(x + 2\pi) = f(x)$. That is, we have the extended $f(x)$ is periodic. Similarly, if F_k are the components of the DFT of $f(x)$, then $F_{k+N} = F_k$. We have already seen this via the cyclic nature of the roots of unity. That is the F_k are periodic, with period N.

If the function f is multiplied by a phase factor $e^{2\pi i p/N}$, then $F_k \to F_{(k+p) mod(N)}$. Obviously if $p = N$ we just see the inherent periodicity. This *circular* shift on multiplying f by a phase factor is known as the shifting theorem.

5.3 Convolution and Correlation

In the continuous Fourier transform, the Fourier transform of a convolution is the product of two Fourier transforms. That is if

$$\mathscr{F}\left(\int_{-\infty}^{\infty} f(s)g(x-s)ds\right) = \mathscr{F}(f)\mathscr{F}(g),$$

where

$$\mathscr{F}(f) = F(k) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ikx}dx. \tag{19}$$

So, what is the corresponding rule for the DFT? Suppose we have the F_k from DFT of a function f, and the G_k from another function g. Let's examine the inverse DFT of a function that has $F_k G_k$ as its DFT components. That is we have

$$C_n = \frac{1}{N}\sum_{k=0}^{N-1} F_k G_k e^{2\pi ikn/N}. \tag{20}$$

We can use the inverse transforms for the F_k and G_k, so

$$\begin{aligned} C_n &= \frac{1}{N}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\sum_{m=0}^{N-1} f_l g_m e^{2\pi ik(n-l-m)/N} \\ &= \sum_{l=0}^{N-1} f_l \sum_{m=0}^{N-1} g_m \left(\frac{1}{N}\sum_{k=0}^{N-1} e^{2\pi ik(n-l-m)/N}\right). \end{aligned} \tag{21}$$

Recall §1, eqns.5 and 6. By the same argument we see that the term in the parentheses is zero unless $m = n - l$, where it is one. In fact, if p is any integer, the term in the parentheses is unity if $m = n - l - pN$. For convenience, if we define any g_k to be zero if k is outside of $k = 0, 1, 2, \ldots, N-1$, then we can write this as

$$C_n = \mathscr{F}^{-1}(F_n G_n) = \sum_{l=0}^{N-1} f_l \sum_{p=-\infty}^{\infty} g_{n-l-pN}. \tag{22}$$

Eqn.22 defines the *circular convolution* of f with g. Let's have a look at this in practise with $n = 4$.

$$\begin{aligned} C_0 &= f_0 g_{0-0-0} + f_1 g_{0-1+4} + f_2 g_{0-2+4} + f_3 g_{0-3+4} = f_0 g_0 + f_1 g_3 + f_2 g_2 + f_3 g_1, \\ C_1 &= f_0 g_{1-0-0} + f_1 g_{1-1-0} + f_2 g_{1-2+4} + f_3 g_{1-3+4} = f_0 g_1 + f_1 g_0 + f_2 g_3 + f_3 g_2, \\ C_2 &= f_0 g_{2-0-0} + f_1 g_{2-1-0} + f_2 g_{2-2-0} + f_3 g_{2-3+4} = f_0 g_2 + f_1 g_1 + f_2 g_0 + f_3 g_3, \\ C_3 &= f_0 g_{3-0-0} + f_1 g_{3-1-0} + f_2 g_{3-2-0} + f_3 g_{3-3-0} = f_0 g_3 + f_1 g_2 + f_2 g_1 + f_3 g_0. \end{aligned} \tag{23}$$

That is, far from the sum with p varying from $-\infty$ to $+\infty$, we only require $p = 0$ or $p = -1$. So, the Fourier transform coefficients of a circular convolution of f and g are $F_k \times G_k$. Similarly the inverse transform of a circular convolution of the Fourier components are $f_k \times G_k$.

Note, that the number of operations required for a correlation is of order N^2, so it is much quicker to to use the FFT on the two functions, multiply the coefficients, and then use the inverse FFT. The reader may object that it will be quite rare that the order is an exact power of two, but we mentioned earlier that there are different FFT algorithms. In particular Rader's FFT works when the order is a prime, and Bluestein's FFT algorithm works on arbitrary sized data. (These are a bit slower than the divide and conquer FFT with N as a power of two). Alternatively, you can use just this divide and conquer method but *zero pad* the data.

Suppose, for instance, that we have a 1000 data points, and we use a 1024 point DFT. We would have hoped to have had data extended to 1024 points, but for some reason we just don't have it. What we do is set the last 24 points to have data values of zero. With this zero padding of the input, the DFT is a convolution of the DFT of our (missing) extended data set with a narrow sinc function. The larger the fraction of zero padding to the, the broader the sinc becomes, and the more blurred the DFT in frequency space. Of course, if our data is a section of a process that started long before we sample it, and long after, we have effectively zero padded. All that can be said is that the more data, the merrier!

A common data processing task is to look for correlations. We note that if we take the inverse transform of $(F_k x G_k^*)$ we get the *cross correlation* of f and g (which is the same as the convolution for real valued data). If $G_k = F_k$, the inverse transform is the *autocorrelation*.

6 The Sampling Theorem

We shall switch from the DFT to the continuous Fourier Transform

$$\mathscr{F}(f(x)) = F(k) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ikx}dx. \tag{24}$$

If we denote the convolution of two functions by an asterisk so that

$$h(x) = g(x) * f(x) = \int_{-\infty}^{\infty} f(s)g(x-s)ds. \tag{25}$$

Then we have the following convolution theorems. The first is

$$\mathscr{F}(g(x) * f(x)) = \mathscr{F}(g)\mathscr{F}(f) = G(k)F(k). \tag{26}$$

What we need for the sampling theorem is the second convolution theorem. Namely

$$\mathscr{F}(g(x)f(x)) = G(k) * F(k). \tag{27}$$

That is, the Fourier Transform of a product of two functions is the convolution of the Fourier Transforms of those two functions.

Now, an idealised sampling of a function can be modelled using the following well known property of the Dirac delta function: namely that if a is in the *any* interval (c, d), then

$$\int_c^d f(x)\delta(x-a)dx = f(a), \tag{28}$$

but the integral is zero if a is not in the interval (c, d). The delta function is said to *sift* the function f. Loosely speaking $\delta(x)$ is infinite at $x = 0$ and zero everywhere else. More strictly speaking, it isn't actually a function but the limit of a sequence of functions which have certain properties. One example would be that of the normal distribution with an appropriate sequence for $\sigma \to 0$.

We write

$$S(x, \Delta) = \sum_{k=-\infty}^{\infty} \delta(x - k\Delta). \tag{29}$$

Here, Δ is the sampling interval, and we have an infinite set of point samples of a function $f(x)$ represented by Sf. Using the second convolution theorem gives us

$$\mathscr{F}(S(x,\Delta)f(x)) = \int_{-\infty}^{\infty} F(f) \sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{\Delta}\right). \tag{30}$$

To avoid confusion with the spatial frequency f, we always use $f(x)$ rather than just f for the function of x. Also F is the Fourier transform of $f(x)$. The reader may verify that the Fourier transform of S (sometimes called a Dirac comb function) is a Dirac comb function in Fourier space with spacing $\Delta_f = 1/\Delta$.

Now, due to the sifting property, each delta function in eqn.30 produces an image of F centred at k/Δ. We suppose that the function $f(x)$ is band limited, or at least that $F(f)$ tends to zero rapidly beyond some frequency range. We depict the situation in Fig.2.

This is quite a remarkable result. We have drawn the Fourier transforms as being band limited. That is we suppose $F(f) \to 0$ if $|f| > B$. When we draw all the Fs and add them, we get something like the red line in Fig.2. We notice that the spacing is $1/\Delta$. That is, if we make a finer comb by decreasing Δ in real space (increasing the spatial frequency of the sampling), the images of F move further apart in frequency space. If the function is band limited, we can decrease the sampling space Δ until there is no longer any overlap of these images in frequency space as in Fig.3.

Now *if there is no overlap we may multiply by a "box function" (Fig.3) and remove all the images of F but one.* This box function is equal to one if $|f| < 1/\Delta$ and zero everywhere else. Now, the Fourier transform of a box function is a sinc function $\sin(\pi x)/(\pi x)$, and the Fourier transform of a sinc function is a box function. That means that multiplying by a box function to get $\mathscr{F}(f(x))$ instead of $\mathscr{F}(S(x,\Delta)f(x))$ means that when we take the inverse transform, we get a convolution with a sinc function in real space.

Now, to get from the pretty picture to something useful, we must observe that

$$\mathscr{F}(S(x,\Delta)f(x)) = \int_{-\infty}^{\infty} \sum_{k=-\infty}^{\infty} f(x)\delta(x-k\Delta)e^{-2\pi i f x}dx. = \sum_{k=-\infty}^{\infty} f(k\Delta)e^{-2\pi i f k\Delta}. \tag{31}$$

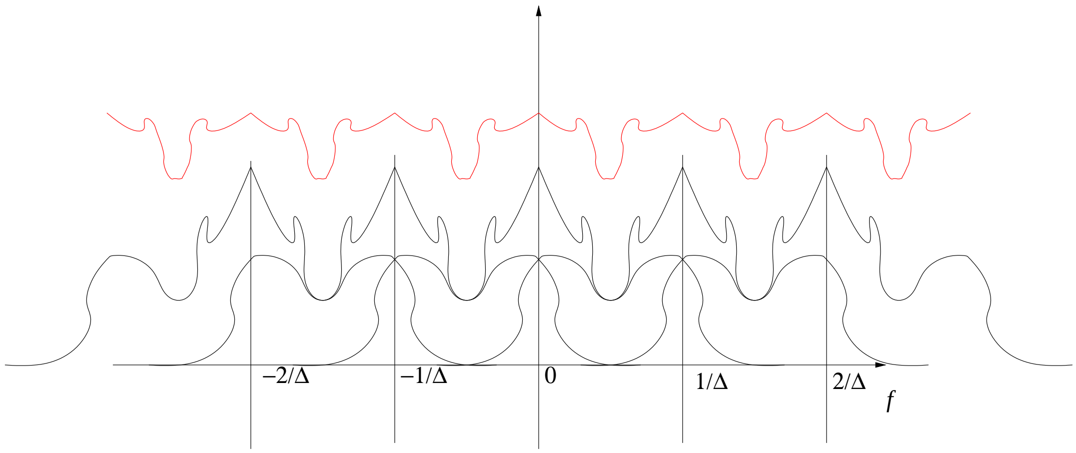

Figure 2: The Fourier transform of $f(x) \times S(x, \Delta)$ produces an infinity of images of $F = \mathscr{F}(f(x))$ at $f = k/T$, the red line is supposed to represent the sum of the images.

This is the representation of the repetition of the images of F in frequency space. The reason it now becomes useful is that it contains the actual data samples. It takes the form of a Fourier series with the function samples $f(k\Delta)$ as the coefficients. If $f \to f + 1/\Delta$ the function in frequency space is unchanged. Now it is *this* that we multiply by out box function, and the subsequent inverse transform is

$$f(x) = \sum_{k=-\infty}^{\infty} f(k\Delta) \mathrm{sinc}\left(\frac{x - k\Delta}{\Delta}\right). \tag{32}$$

(Note on recalling the factor π in the definition of the sinc function, we see that the sinc function is zero if $x = n\Delta$ for all integers $\pm n$ with $n \neq 0$.

So, if our function is band limited, and we sample at a rate of $2B$, the Fourier transform of the samples of $f(x)$ consist of non-overlapping copies of the Fourier transform $F(f)$ of $f(x)$. This means we can isolate F using the "box function", and this leads to eqn.32 being an exact reconstruction of $f(x)$. Of course, we can sample at a higher rate and it is still exact. The samples move closer together, the copies of F move further apart, the box becomes wider, and the sinc narrower.

If we sample at a lower rate, we have the Fourier transform of F plus overlapping copies of F in the box. This means we have aliasing, and can no longer reconstruct the function exactly. Of course $2B$ is the total bandwidth as used in §2, while here B is the one sided base bandwidth.

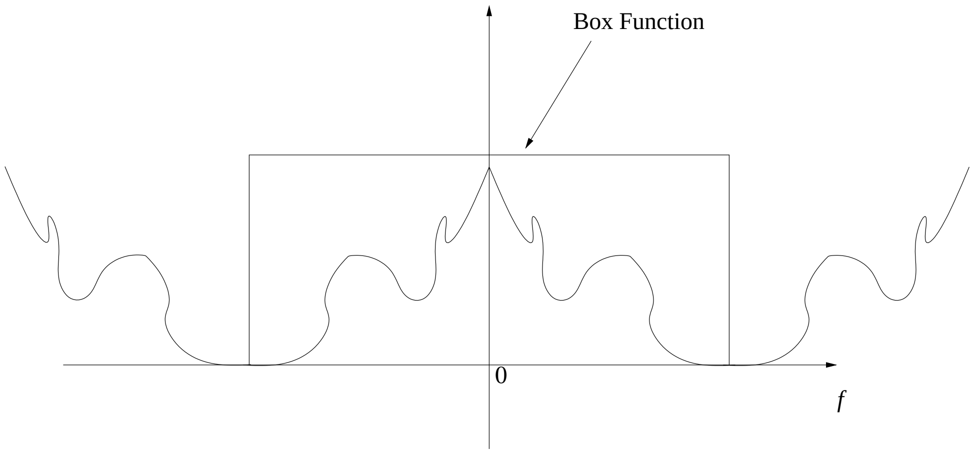

Figure 3: We have increased the spatial sampling frequency of the function, and so increased the spacing of the images until they no longer overlap.

7 Resampling in Two Dimensions

We consider that we are given data on a non-rectangular grid. Often we may wish to resample on to a uniform rectangular grid. If we are given irregular data as in Fig.4, we might try a linear interpolation. Given the irregular points we can generate a Delaunay triangulation. Any point in a triangle can be written as $\alpha\mathbf{a} + \beta\mathbf{b}$. Here $\mathbf{a}$ and $\mathbf{b}$ are two edge vectors, and α and β must be between 0 and 1, with the constraint $\alpha + \beta \leq 1$. These constraints ensure the point is inside the triangle. It is easy to show that the α is the fractional area of the triangle between the origin, the point in the triangle and $\mathbf{b}$ (since the (signed) area of a triangle is $1/2(\mathbf{a} \times \mathbf{b})$. Similarly, if we take the area of the triangle between the origin, the point and $\mathbf{a}$ and divide by the area of the total triangle, we have α. These coordinates are sometimes called *simplicial coordinates*, or *barycentric coordinates* or areal coordinates. The author prefers just local triangular coordinates. We note that the functions α, β, and $1 - \alpha - \beta$ are equal to one at the corners, and zero along the opposite edges. So the form a set of linear basis functions. So, $f(0)(1 - \alpha - \beta) + f(\mathbf{a})\alpha + f(\mathbf{b})\beta$ provides a linear interpolation of the data inside the triangle Given the (x, y) coordinates of the data, we just need to invert a 2×2 matrix to find α and β.

However, all this was an aside, we may often be faced with a different resampling problem as in Fig.5. In this case, if the underlying function is band limited, and the data is sufficiently sampled, we may resample the data to the regular rectangular grid with no

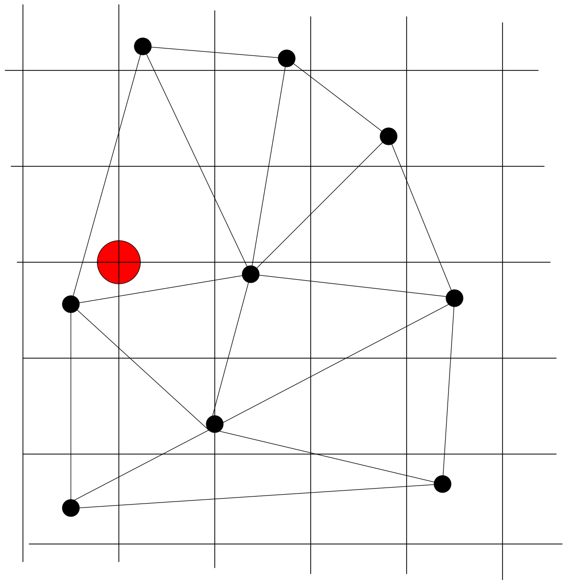

Figure 4: We have a triangulation of the data points, and want to resample to a point on the square rid (marked as a large red circle).

error.

We suppose we have two (non parallel and non zero) vectors $\mathbf{a}$ and $\mathbf{b}$. Then are samples will be at $\mathbf{r} = \mathbf{r_0} + \mathbf{ma} + \mathbf{nb}$, where m and n are integers. In the following, we shall assume that $\mathbf{r}_0 = 0$. That is to say we choose our origin to be one of the sampled points. We have the components of $\mathbf{a}$ and $\mathbf{b}$ in terms of the usual rectangular Cartesian coordinates. So $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j}$. Our idealised sampling can now be written down,

$$S(x, y, \mathbf{a}, \mathbf{b}) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(x - ma_1 - nb_1, y - ma_2 - nb_2). \tag{33}$$

The 2-D Fourier transform of S is then

$$\mathscr{S} = \mathscr{F}(S) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} e^{2\pi i[f_x(ma_1+nb_1)+f_y(ma_2+nb_2)]}$$

$$\mathscr{S} = \mathscr{F}(S) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} e^{2\pi i m \mathbf{f}\cdot\mathbf{a}} e^{2\pi i n \mathbf{f}\cdot\mathbf{b})}. \tag{34}$$

Here we have naturally put $\mathbf{f} = (f_x, f_y)$. The exponentials can be seen as standing plane waves in the directions $\mathbf{a}$ and $\mathbf{b}$ in frequency space. If we take frequency vector in the direction of $\mathbf{a}$, say $\mathbf{f} = \lambda\mathbf{a}$, then If $\lambda\mathbf{a} \to \lambda\mathbf{a} + \mathbf{a}/|\mathbf{a}|^2$, then $\lambda\mathbf{a}\cdot\mathbf{a} \to \lambda\mathbf{a}\cdot\mathbf{a} + 1/|\mathbf{a}|$. Then of course $e^{2\pi i m\lambda\mathbf{a}\cdot\mathbf{a}} \to e^{2\pi i m(\lambda\mathbf{a}\cdot\mathbf{a}+1)}$. So, the sampling grid has the same basic pattern, except with the displacement vectors $\mathbf{a}/|\mathbf{a}|^2$ and $\mathbf{b}/|\mathbf{b}|^2$ instead of $\mathbf{a}$ and $\mathbf{b}$. We call these new vectors $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$. Obviously, they are *not* unit vectors. Just as in the one dimensional case

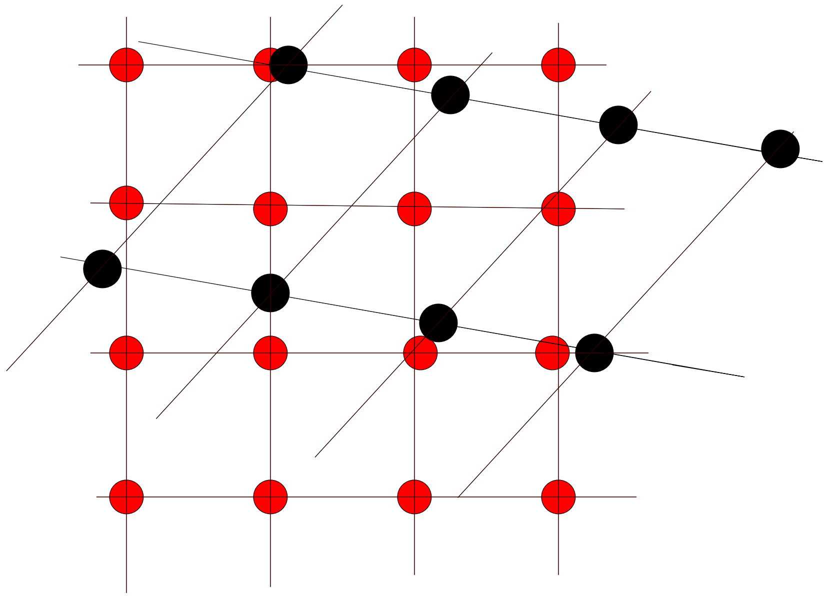

Figure 5: The data points are regular (black circles), but are tilted with respect to the desired resampling grid (red circles), are not rectangular, and have different spacings in different directions.

we have spacings in frequency space as the reciprocals of the spacings in real space. (In fact a crystallographer would call our grid in frequency space a reciprocal lattice.) Given the above, just as with eqn.30 in 1 dimension. we can write

$$\mathscr{S} = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(\mathbf{f} - m\hat{\mathbf{a}} - n\hat{\mathbf{b}}). \tag{35}$$

The Fourier transform of $f(x, y)S$ will just be $F(f_x, f_y)$ repeated at all the reciprocal lattice points. So far, there isn't any complication, but now we want our box function. In Fig.6 we see there is a slight problem. We *could* use the small square if the sampling is sufficiently fine (or the function has sufficiently small bandwidth, but mostly it will be far better to use the larger parallelepiped as box function. So, we need the inverse Fourier transform of this shape.

Now, if we put $\mathbf{f} = f_x\mathbf{i} + f_y\mathbf{j} = \alpha\hat{\mathbf{a}} + \beta\hat{\mathbf{b}}$, then we have $\hat{\mathbf{a}} = \hat{a_1}\mathbf{i} + \hat{a_2}\mathbf{j}$ and $\hat{\mathbf{b}} = \hat{b_1}\mathbf{i} + \hat{b_2}\mathbf{j}$. Or

$$\begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} \hat{a_1} & \hat{b_1} \\ \hat{a_2} & \hat{b_2} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{36}$$

In the x, y space, the corners of the parallelepiped are at $-\hat{\mathbf{a}}/2 - \hat{\mathbf{b}}/2$, $\hat{\mathbf{a}}/2 + \hat{\mathbf{b}}/2$, $\hat{\mathbf{a}}/2 - \hat{\mathbf{b}}/2$, and $-\hat{\mathbf{a}}/2 + \hat{\mathbf{b}}/2$. In the α, β space, they are at the corners of a square, parallel to the axes, centred at the origin, and of side length 1.

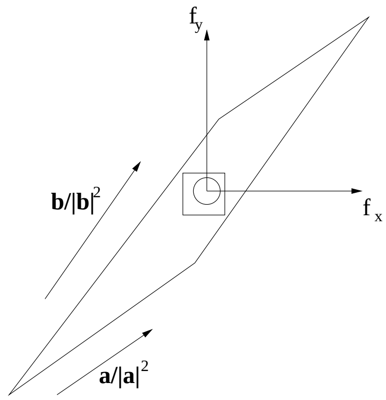

Figure 6: The Box function in frequency space.

Now

$$\int \int_R f(f_x(\alpha, \beta), f_y(\alpha, \beta)) df_x df_y$$

$$\int \int_{\mathscr{R}} f(\alpha, \beta) \left| \frac{\partial(f_x, f_y)}{\partial(\alpha, \beta)} \right| df_x df_y. \tag{37}$$

Here, the *Jacobian Determinant is given by*

$$\left| \frac{\partial(f_x, f_y)}{\partial(\alpha, \beta)} \right| = \begin{vmatrix} \frac{\partial f_x}{\partial \alpha} & \frac{\partial f_x}{\partial \beta} \\ \frac{\partial f_y}{\partial \alpha} & \frac{\partial f_y}{\partial \beta} \end{vmatrix} = (\hat{a_1}\hat{b_2} - \hat{a_2}\hat{b_1}). \tag{38}$$

In this linear transformation, the Jacobian is just a number (the area $\mathscr{A}$ of the paralleled in frequency space). We have put $R \to \mathcal{R}$ to represent the change in domain from the parallelepiped to the unit square. So, we need the real space convolution function $\mathscr{C}$ which is the inverse Fourier transform of the box function. Namely,

$$\mathscr{C} = \int \int_R e^{2\pi i x f_x} e^{2\pi i x f_x} df_x df_y \tag{39}$$

since the "box function" is unity inside R and zero outside R. From the above,

$$\mathscr{A} \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} e^{2\pi i x(\hat{a_1}\alpha + \hat{b_1}\beta)} e^{2\pi i y(\hat{a_2}\alpha + \hat{b_2}\beta)} d\alpha d\beta$$

$$= \mathscr{A} \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} e^{2\pi i(\mathbf{r}\cdot\hat{\mathbf{a}}\alpha)} e^{2\pi i(\mathbf{r}\cdot\hat{\mathbf{b}}\beta)} d\alpha d\beta \tag{40}$$

as the inverse transform of the box function. This is just

$$\frac{\mathscr{A}}{2\pi i\mathbf{r}\cdot\hat{\mathbf{a}} \times 2\pi i\mathbf{r}\cdot\hat{\mathbf{b}}} \left[e^{2\pi i\mathbf{r}\cdot\hat{\mathbf{a}}}\right]_{-1/2}^{1/2} \left[e^{2\pi i\mathbf{r}\cdot\hat{\mathbf{b}}}\right]_{-1/2}^{1/2}. \tag{41}$$

recalling

$$\sin x = \frac{1}{2i}\left(e^{ix} - e^{-ix}\right),$$

and

$$\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x} \tag{42}$$

we note that $f * g = g * f$, we want the convolution variables to be in the delta function, so we put

$$f(x,y) = \mathscr{A}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} f(s,t)\delta(s - ma_1 - nb_1)\delta(t - ma_2 - nb_2)$$

$$\mathrm{sinc}\left((x-s)\hat{a_1} + (y-t)\hat{a_2}\right)\mathrm{sinc}\left((x-s)\hat{b_1} + (y-t)\hat{b_2}\right) dsdt. \tag{43}$$

This gives us

$$f(x,y) = \mathscr{A}\sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} f(ma_1 + nb_1), ma_2 + n_b 2)$$

$$\mathrm{sinc}\left(x\hat{a_1} + y\hat{a_2}\right)\mathrm{sinc}\left(x\hat{b_1} + y\hat{b_2}\right).$$

$$= \mathscr{A}\sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} f(ma_1 + nb_1, ma_2 + n_b 2)$$

$$\mathrm{sinc}\left((x - ma_1 - nb_1)\hat{a_1} + (y - ma_1 - na_2)\hat{a_2}\right)$$

$$\mathrm{sinc}\left((x - ma_2 - nb_2)\hat{b_1} + (y - ma_2 - mb_2)\hat{b_2}\right). \tag{44}$$

If we put $\mathbf{r}_{mn} = m\mathbf{a} + n\mathbf{b}$ this becomes

$$f(x,y) = \mathscr{A}\sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} f(\mathbf{r}_{mn})\mathrm{sinc}\left((\mathbf{r} - \mathbf{r}_{mn})\cdot\frac{\mathbf{a}}{|\mathbf{a}|^2}\right)\mathrm{sinc}\left((\mathbf{r} - \mathbf{r}_{mn})\cdot\frac{\mathbf{b}}{|\mathbf{b}|^2}\right). \tag{45}$$

Here we have recalled $\mathbf{a}/|\mathbf{a}|^2 = \hat{\mathbf{a}}$ and $\mathbf{b}/|\mathbf{b}|^2 = \hat{\mathbf{b}}$. If $\mathbf{r} \to \mathbf{r} + \mathbf{a} + \mathbf{b}$, the arguments of the sinc functions both increase by one, so recalling the π in the definition of the sinc function, we see that if r is at any of the $\mathbf{r}_{mn}$, the sinc functions are unity there, but zero at all other $\mathbf{r}_{mn}$. The resampling of our non rectangular grid with neither basis vectors parallel to the x or y axes, is just as simple as resampling a regular rectangular grid parallel to the x, y axes.

We finish off by mentioning *deblurring*. A real world sample involves using some measuring device that cannot give a "point" value, but measures over some interval along x. That is, we are measuring point samples, not of $f(x)$, but of $f(x)$ convoluted with some device response function. In effect we are taking point samples of a blurred function. If we know this function, we know it's Fourier transform. So, under *some* circumstances it may be possible to enhance the reconstructed image so as to take out the instrument blur. (If course, this may just turn the image into nonsense given the wrong circumstances.) This is quite a subject in itself, so we just mention it in passing.

Often, the instrument function is known, for instance it may be Gaussian. In the Author's experience, some 2-D instrument functions may be split up as to be linear over triangular bases. Any triangle can be mapped onto the triangle (0,0), (1,0), (0,1) (which we did in taking the inverse transform of the box function as above.) We end with looking at this odd little transformation which maps this triangle onto a square! If we put $u = x$, $v = y/(1-x)$, or $x = u$, $y = (1-u)v$, then

$$\int_{x=0}^{1} \int_{y=0}^{(1-x)} f(x,y)dxdy = \int_0^1 \int_0^1 f(x(u,v), y(u,v))(1-u)dudv. \tag{46}$$

It's useless, but the author likes it!

References

[1] W. L. Briggs and V. E. Henson, *The DFT: An Owner's Manual for the Discrete Fourier Transform* (SIAM, Philadelphia, 1995).