The Fourier Transform and its Applications

In the previous lecture we studied three properties of Ш that make it so useful in many applications. They are:

- Periodizing
  - Convolving with Ш periodizes a function
- Sampling
  - Multiplying by Ш samples a function
- The Fourier transform of Ш is Ш.
  - Convolving and multiplying are themselves flip sides of the same coin via the convolution theorem for Fourier transforms.

We are now about to combine all of these ideas in a spectacular way to treat the problem of 'sampling and interpolation'. Let me state the problem this way:

- Given a signal $f(t)$ and a collection of *samples* of the signal, *i.e.* values of the signal at a set of points $f(t_0)$, $f(t_1)$, $f(t_2)$, $\ldots$, to what extent can one interpolate the values $f(t)$ at other points from the sample values?

This is an old question, and a broad one, and it would appear on the surface to have nothing to do with Ш's or Fourier transforms, or any of that. But we've already seen some clues, and the full solution is set to unfold.

Sampling sines, and band-limited signals

Why should we expect to be able to do interpolation at all? Imagine putting down a bunch of dots – maybe even infinitely many – and asking someone to pass a curve through them that *agrees everywhere exactly* with a predetermined mystery function passing through those dots. Ridiculous. But it's *not* ridiculous. If a relatively simple hypothesis is satisfied then interpolation can be done! Here's one way of getting some intuitive sense of the problem and what that hypothesis should be.

Suppose we *know* a signal is a single sinusoid. A sinusoid repeats, so if we have enough information to pin it down over one period, or cycle, then we know the whole thing. How many samples – how many values of the function – within one period do we need to know to know which sinusoid we have? We need three samples *strictly* within one cycle. You can think of the graph, or you can think of the equation: A general sinusoid is of the form $A\sin(2\pi\nu t+\phi)$. There are three unknowns, the amplitude A, the frequency ν and the phase ϕ. We would expect to need three equations to find the unknowns, hence we need values of the function at three points, three samples.

What if the signal is a sum of sinusoids, say

$$\sum_{n=1}^{N} A_n \sin(2\pi n\nu t+\phi_n).$$

Sample points for the sum are 'morally' sample points for the individual harmonics, though not explicitly. We need to take enough samples to get sufficient information to determine all of the unknowns for all of the harmonics. Now, in the time it takes for the combined signal to go through one cycle, the individual harmonics will have gone through several cycles, the lowest frequency harmonic through one cycle, the lower frequency harmonics through a few cycles, say, and the higher frequency harmonics through many. We have to take enough samples of the combined signal so that as the individual harmonics go rolling along we'll be sure to have at least three samples in *some* cycle of *every* harmonic.

To simplify and standardize we assume that we take evenly spaced samples (in t). Since we've phrased things in terms of cycles per second, to understand how many samples are enough it's then also better to think in terms of 'sampling rate', *i.e.* samples/sec instead of 'number of samples'. If we are to have at least three samples strictly within a cycle then the sample points must be strictly less than a half-cycle apart. A sinusoid of frequency ν goes through a half-cycle in $1/2\nu$ seconds so we want

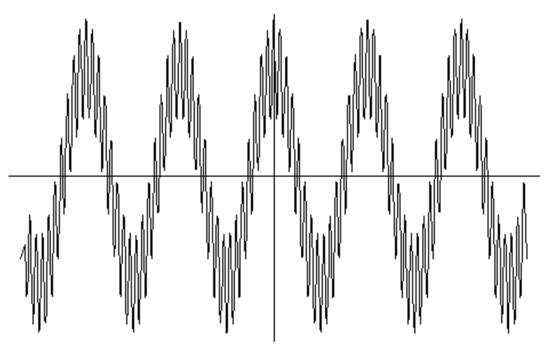
$$\text{spacing between samples} \;=\; \text{(number of seconds)/sample} < \frac{1}{2\nu}$$

The more usual way of putting this is

$$\text{Sampling rate} = \text{Samples/sec} > 2\nu.$$

This is the rate at which we should sample a given sinusoid of frequency ν to guarantee that a single cycle will contain at least three sample points. Furthermore, if we sample at this rate for a given frequency, we will certainly have more three sample points in some cycle of any harmonic at a *lower* frequency. Note that the sampling rate has units 1/seconds and, again, that the sample points are 1/(sampling rate) seconds apart.

For the combined signal – a sum of harmonics – the higher frequencies are driving up the sampling rate; specifically, the *highest* frequency is driving up the rate. To think of the interpolation problem geometrically, high frequencies cause more rapid oscillations, *i.e.* rapid changes in the function over small intervals, so to hope to interpolate such fluctuations accurately we'll need a lot of sample points and thus a high sampling rate. For example, here's a picture of the sum of two sinusoids one of low frequency and one of high frequency.



If we sample at too low rate we might miss the wiggles entirely. We might mistakenly think we had only the low frequency sinusoid, and, moreover, if all we had to go on were the samples we wouldn't even know we'd made a mistake! We'll come back to just this problem a little later.

If we sample at a rate greater than twice the highest frequency our sense is that we will be sampling often enough for all the lower harmonics as well, and we should be able to determine everything. The problem here is if the spectrum is *unbounded*. If, as for a square wave, we have a full Fourier series and not just a finite sum of sinusoids, then we have no hope of sampling frequently enough to determine the combined signal from the samples. For a square wave, for example, there is *no* 'highest frequency'. That's trouble. It's time to define ourselves out of this trouble.

Band-limited signals From the point of view of the preceding discussion, the problem for interpolation, is high frequencies, and the best thing a signal can be is a finite Fourier series. The latter is much too restrictive for applications, of course, so what's the 'next best' thing a signal can be? It's one for which there *is* a highest frequency. These are the *band-limited* signals – signals whose Fourier transform is identically zero outside of a finite interval. Such a signal has a bounded spectrum; there is a 'highest frequency'.

More formally:

- A signal $f(t)$ is band-limited if there is a $0 < p < \infty$ with $\mathcal{F}f(s) = 0$ for all $|s| \geq p/2$. The smallest number p for which this is true is called the *bandwidth* of $f(t)$.

There's a question about having $\mathcal{F}f$ be zero at the endpoints $\pm p/2$ as part of the definition. For the discussion on sampling and interpolation to follow it's easiest to assume this is the case, and treat separately some special cases when it isn't. For those want to know more, read the next paragraph.

Some technical remarks If $f(t)$ is an integrable function then $\mathcal{F}f(s)$ is continuous, so if $\mathcal{F}f(s) = 0$ for all $|s| > p/2$ then $\mathcal{F}f(\pm p/2) = 0$ as well. On the other hand, it's also common first to define the *support* of a function (integrable or not) as the complement of the largest open set on which the function is identically zero. (This definition can also be given for distributions.) This makes the support closed, being the complement of an open set. For example, if $\mathcal{F}f(s)$ is identically zero for $|s| > p/2$, and on no larger *open* set, then the support of $\mathcal{F}f$ is the *closed* interval $[-p/2, p/2]$. Thus, with this definition, even if $\mathcal{F}f(\pm p/2) = 0$ the endpoints $\pm p/2$ are included in the support of $\mathcal{F}f$.

One then says, as an alternate definition, that f is band-limited if the support of $\mathcal{F}f$ is closed and *bounded*. In mathematical terms, a closed, bounded set (in $\mathbf{R}^n$) is said to be *compact*, and so the shorthand definition of band-limited is that $\mathcal{F}f$ has compact support. A typical compact set is a closed interval, like $[-p/2, p/2]$, but we could also take finite unions of closed intervals. This definition is probably the one more often given, but it's a little more involved to set up, as you've just witnessed. Whichever definition of band-limited one adopts there are always questions about what happens at the endpoints anyway, as we'll see.

Sampling and interpolation for band-limited signals

We're about to solve the interpolation problem for band-limited signals. We'll show that interpolation is possible by finding an explicit formula that does the job. Before going through the solution, however, I want to make a general observation that's independent of the interpolation problem but is important to it.

It is unphysical to consider a signal as lasting forever in time. A physical signal $f(t)$ is naturally 'time-limited', meaning that $f(t)$ is identically zero on $|t| \geq q/2$ for some q – there just isn't any signal beyond a point. On the other hand, it is very physical to consider a band-limited signal, one with no frequencies beyond a certain point, or at least no frequencies that our instruments can register. Well, we can't have both, at least not in the ideal world of mathematics. Here is where mathematical description meets physical expectation – and they disagree. The fact is:

- A signal cannot be both time-limited and band-limited.

What this means in practice is that there must be inaccuracies in a mathematical model of a phenomenon that assumes a signal is both time-limited and band-limited. Such a model can be at best an approximation, and one has to be prepared to estimate the errors as they may affect measurements and conclusions.

Here's one argument why the statement is true; I'll give a more complete proof of a more general statement in Appendix 1. Suppose f is band-limited, say $\mathcal{F}f(s)$ is zero for $|s| \geq p/2$. Then

$$\mathcal{F}f = \Pi_p \cdot \mathcal{F}f.$$

Take the inverse Fourier transform of both sides to obtain

$$f(t) = p \operatorname{sinc} pt * f(t).$$

Now $\operatorname{sinc} pt$ 'goes on forever'; it decays but it has nonzero values all the way out to $\pm\infty$. Hence the convolution with f also goes on forever; it is not time-limited.

sinc as a 'convolution identity' There's an interesting observation that goes along with the argument we just gave. We're familiar with δ acting as an 'identity element' for convolution, meaning

$$f * \delta = f.$$

This important property of δ holds for *all* signals for which the convolution is defined. We've just seen for the more restricted class of band-limited functions, with spectrum from $-p/2$ to $p/2$, that the sinc function also has this property:

$$p \operatorname{sinc} pt * f(t) = f(t).$$

As an exercise you can show that sinc also has the sifting property for band-limited functions:

$$p \operatorname{sinc} p(t-a) * f(t) = f(t-a).$$

The Sampling Theorem Ready to solve the interpolation problem? It uses all the important properties of Ш, but it goes so fast that you might miss the fun entirely if you read too quickly.

Suppose $f(t)$ is band-limited with $\mathcal{F}f(s)$ identically zero for $|s| \geq p/2$. We periodize $\mathcal{F}f$ by $Ш_p$ and then cut-off to get $\mathcal{F}f$ back again:

$$\mathcal{F}f = \Pi_p(\mathcal{F}f * Ш_p).$$

This is the crucial equation.

Now take the inverse Fourier transform:

$$\begin{aligned}
f(t) = \mathcal{F}^{-1}\mathcal{F}f(t) &= \mathcal{F}^{-1}(\Pi_p(\mathcal{F}f * \text{III}_p))(t) \\
&= \mathcal{F}^{-1}\Pi_p(t) * \mathcal{F}^{-1}(\mathcal{F}f * \text{III}_p)(t) \\
&\qquad\qquad \text{(on the right, taking } \mathcal{F}^{-1} \text{ turns multiplication into convolution)} \\
&= \mathcal{F}^{-1}\Pi_p(t) * (\mathcal{F}^{-1}\mathcal{F}f(t) \cdot \mathcal{F}^{-1}\text{III}_p(t)) \\
&\qquad\qquad \text{(ditto, except it's convolution turning into multiplication)} \\
&= p \operatorname{sinc} pt * (f(t) \cdot \frac{1}{p}\text{III}_{1/p}(t)) \\
&= \operatorname{sinc} pt * \sum_{k=-\infty}^{\infty} f(\frac{k}{p})\delta(x - \frac{k}{p}) \quad \text{(the sampling property of III}_p\text{)} \\
&= \sum_{k=-\infty}^{\infty} f(\frac{k}{p}) \operatorname{sinc} pt * \delta(x - \frac{k}{p}) \\
&= \sum_{k=-\infty}^{\infty} f(\frac{k}{p}) \operatorname{sinc} p(t - \frac{k}{p}) \quad \text{(the sifting property of } \delta\text{)}
\end{aligned}$$

We've just established the classic 'Sampling Theorem', though it might be better to call it the interpolation theorem. Here it is as a single statement:

- If $f(t)$ is a signal with $\mathcal{F}f(s)$ identically zero for $|s| \geq p/2$ then

$$f(t) = \sum_{k=-\infty}^{\infty} f(\frac{k}{p}) \operatorname{sinc} p(t - \frac{k}{p}).$$

Some people write the formula as

$$f(t) = \sum_{k=-\infty}^{\infty} f(\frac{k}{p}) \operatorname{sinc}(pt - k),$$

but I generally prefer to emphasize the *sample points*

$$t_k = \frac{k}{p}$$

and then to write the formula as

$$f(t) = \sum_{k=-\infty}^{\infty} f(t_k) \operatorname{sinc} p(t - t_k).$$

What does the formula do, once again? It computes *any* value of f in terms of *sample values*. Here are a few general comments to keep in mind:

- The sample points are spaced $1/p$ apart, the *reciprocal* of the bandwidth.[1]

[1] That sort of reciprocal phenomenon is present again in higher dimensional versions of the sampling formula. This will be a later topic for us.

- The formula involves infinitely many sample points, all the points k/p for $k = 0, \pm 1, \pm 2, \cdots$.

  So don't think you're getting away too cheaply, and realize that any practical implementation can only involve a finite number of terms in the sum, so will necessarily be an approximation.

  - Since a band-limited signal cannot be time-limited we should expect to have to take samples all the way out to $\pm\infty$. However, sampling a band-limited *periodic* signal, *i.e.*, a finite Fourier series, requires only a finite number of samples. We'll cover this, below.

Put the outline of the argument for the sampling theorem into your head, it's important. Starting with a band-limited signal, there are three parts:

- Periodize the Fourier transform.
- Cut off this periodic function to get back where you started.
- Take the inverse Fourier transform.

Cutting off in the second step, a multiplication, exactly undoes periodizing in the first step, a convolution, *providing* one has $\mathcal{F}f = \Pi_p(\mathcal{F}f * Ш_p)$. But taking the inverse Fourier transform swaps multiplication with convolution and this is why something nontrivial happens. It's almost obscene the way this works.

Sampling rates and the Nyquist frequency The bandwidth determines the minimal sampling rate we can use to reconstruct the signal from its samples. I'd almost say that the bandwidth *is* the minimal sampling rate except for the slight ambiguity about where the spectrum starts being identically zero (the 'endpoint problem'). Here's the way the situation is usually expressed: If the (nonzero) spectrum runs from $-\nu_{\max}$ to $\nu_{\max}$ then we need

$$\text{Sampling rate} > 2\nu_{\max}$$

to reconstruct the signal from its samples.

The number $2\nu_{\max}$ is often called the *Nyquist frequency*, after Harry Nyquist, God of Sampling, who was the first engineer to consider these problems for the purpose of communications. There are other names associated with this circle of ideas, most notably E. Whittaker, a mathematician, and C. Shannon, an all around genius and founder of Information Theory. The formula as we've given it is often referred to as the Shannon Sampling Theorem.

The derivation of the formula gives us some one-sided freedom, or rather the opportunity to do more work than we have to. We cannot take p smaller than the length of the interval where $\mathcal{F}f$ is supported, twice the bandwidth, but we can take it larger. That is, if p is the bandwidth and $q > p$ we can periodize $\mathcal{F}f$ to have period q by convolving with $Ш_q$ and we still have the fundamental equation

$$\mathcal{F}f = \Pi_q(\mathcal{F}f * Ш_q).$$

(Draw a picture.) The derivation can then proceed exactly as above and we get

$$f(t) = \sum_{k=-\infty}^{\infty} f(\tau_k)\,\mathrm{sinc}\, q(t - \tau_k)$$

where the sample points are

$$\tau_k = \frac{k}{q}.$$

These sample points are spaced closer together than the sample points $t_k = k/p$. The sampling rate is higher than we need. We're doing more work than we have to.

Interpolation a little more generally

Effective approximation and interpolation of signals raises a lot of interesting and general questions. One approach that provides a good framework for many such questions is to bring in orthogonality. It's very much analogous to the way we looked at Fourier series.

Interpolation and orthogonality We begin with still another amazing property of sinc functions – they form an orthonormal collection. Specifically, the family of sinc functions $\{\mathrm{sinc}(t-n)\}$, $n = 0, \pm 1, \pm 2, \ldots$ is orthonormal with respect to the usual inner product on $L^2(\mathbf{R})$. Recall that the inner product is

$$(f, g) = \int_{-\infty}^{\infty} f(t)\overline{g(t)}\, dt.$$

The calculation to establish the orthonormality property of the sinc functions uses the general Parseval identity,

$$\int_{-\infty}^{\infty} f(t)\overline{g(t)}\, dt = \int_{-\infty}^{\infty} \mathcal{F}f(s)\overline{\mathcal{F}g(s)}\, ds.$$

We then have

$$\begin{aligned}
\int_{-\infty}^{\infty} \mathrm{sinc}(t-n)\ \mathrm{sinc}(t-m)\, dt &= \int_{-\infty}^{\infty} (e^{-2\pi i s n}\Pi(s))\,\overline{(e^{-2\pi i s m}\Pi(s))}\, ds \\
&= \int_{-\infty}^{\infty} e^{2\pi i s(m-n)}\Pi(s)\Pi(s)\, ds \\
&= \int_{-1/2}^{1/2} e^{2\pi i s(m-n)}\, ds
\end{aligned}$$

From here direct integration will give you that this is 1 when $n = m$ and 0 when $n \neq m$.

In case you're fretting over it, the sinc function *is* in $L^2(\mathbf{R})$ and the product of two sinc functions is integrable. Parseval's identity holds for functions in $L^2(\mathbf{R})$, though we did not establish this.

Now let's consider band-limited signals $g(t)$, and to be definite let's suppose the spectrum is contained in $-1/2 \leq s \leq 1/2$. Then the sampling rate is 1, *i.e.*, we sample at the integer points and the interpolation formula takes the form

$$g(t) = \sum_{n=-\infty}^{\infty} g(n)\,\mathrm{sinc}(t-n)$$

Coupled with the result on orthogonality, this formula suggest that the family of sinc functions form an orthonormal *basis* for the space of band-limited signals with spectrum in $[-1/2, 1/2]$, and that we're expressing $g(t)$ in terms of this basis. To see that this really is the case, we interpret the coefficients (the sample values $g(n)$) as the inner product of $g(t)$ with $\mathrm{sinc}(t-n)$. We have, again using Parseval,

$$\begin{aligned}
(g, \operatorname{sinc}(t-n)) &= \int_{-\infty}^{\infty} g(t)\ \operatorname{sinc}(t-n)\,dt \\
&= \int_{-\infty}^{\infty} \mathcal{F}g(s)\mathcal{F}(\operatorname{sinc}(t-n))\,ds \\
&= \int_{-\infty}^{\infty} \mathcal{F}g(s)\overline{(e^{-2\pi i s n}\Pi(s))}\,ds \\
&= \int_{-1/2}^{1/2} \mathcal{F}g(s)e^{2\pi i n s}\,ds \\
&= \int_{-\infty}^{\infty} \mathcal{F}g(s)e^{2\pi i n s}\,ds \quad \text{(because g is band-limited)} \\
&= g(n) \quad \text{(by Fourier inversion)}
\end{aligned}$$

It's perfect! The interpolation formula says that $g(t)$ is written in terms of an orthonormal basis, and the coefficient $g(n)$, the n'th sampled value of $g(t)$, is exactly the projection of $g(t)$ onto the n'th basis element:

$$g(t) = \sum_{n=-\infty}^{\infty} g(n)\operatorname{sinc}(t-n) = \sum_{n=-\infty}^{\infty} (g, \operatorname{sinc}(t-n))\operatorname{sinc}(t-n).$$

Lagrange interpolation Certainly for computational questions, going way back, it is desirable to find reasonably simple *approximations* of complicated functions, particularly those arising from solutions to differential equations.[2] The classic way to approximate is to interpolate. That is, to find a simple function that, at least, assumes the same values as the complicated function at a given finite set of points. Curve fitting, in other words. The classic way to do this is via polynomials. One method, presented here just for your general background and know-how, is due to Lagrange.

Suppose we have n points t_1, $t_2, \ldots$, t_n. We want a polynomial of degree $n-1$ that assumes given values at the n sample points. (Why degree $n-1$?)

For this, we start with an n'th degree polynomial that vanishes exactly at those points. This is given by

$$p(t) = (t-t_1)(t-t_2)\cdots(t-t_n).$$

Next put

$$p_k(t) = \frac{p(t)}{t-t_k}.$$

Then $p_k(t)$ is a polynomial of degree $n-1$; we divide out the factor $(t-t_k)$ and so $p_k(t)$ vanishes at the same points as $p(t)$ except at t_k. Next consider the quotient

$$\frac{p_k(t)}{p_k(t_k)}.$$

This is again a polynomial of degree $n-1$. The key property is that $p_k(t)/p_k(t_k)$ vanishes at the sample points t_j *except* at the point t_k where the value is 1:

$$\frac{p_k(t_j)}{p_k(t_k)} = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

[2]The sinc function may not really qualify as an 'easy approximation'. How is it computed, really?

the value is 1.

To interpolate a function by a polynomial (to fit a curve through a given set of points) we just scale and add. That is, suppose we have a function $g(t)$ and we want a *polynomial* that has values $g(t_1)$, $(g(t_2), \dots, g(t_n)$ at the points $t_1, t_2, \dots, t_n$. We get this by forming the sum

$$p(t) = \sum_{k=1}^{n} g(t_k) \frac{p_k(t)}{p_k(t_k)}.$$

This does the trick. It is known as the *Lagrange Interpolation Polynomial.* Remember, unlike the sampling formula we're not reconstructing *all* the values of $g(t)$ from a set of sample values. We're approximating $g(t)$ by a polynomial that has the same values as $g(t)$ at a prescribed set of points.

The sinc function is an analog of the $p_k(t)/p_k(t_k)$ for 'Fourier interpolation', if we can call it that. With

$$\operatorname{sinc} t = \frac{\sin \pi t}{\pi t}.$$

we recall some properties, analogous to the polynomials we built above:

- $\operatorname{sinc} t = 1$ when $t = 0$
- $\operatorname{sinc} t = 0$ at integer points $t = 0, \pm 1, \pm 2, \dots$.

Now shift this and consider

$$\operatorname{sinc}(t-k) = \frac{\sin \pi (t-k)}{\pi (t-k)}.$$

This has the value 1 at $t = k$ and is zero at the other integers.

Suppose we have our signal $g(t)$ and the sample points $\dots, g(-2),, g(-1), g(0), g(1), g(2), \dots$. So, again, we're sampling at evenly spaced points, and we've taken the sampling rate to be 1 just to simplify. To interpolate these values we would then form the sum

$$\sum_{n=-\infty}^{\infty} g(k) \operatorname{sinc}(t-k).$$

There it is again – the general interpolation formula. In the case that $g(t)$ is band-limited (bandwidth 1 in this example) we know we recover all values of $g(t)$ from the sample values.

Finite sampling for a band-limited periodic signal

We started this whole discussion of sampling and interpolation by arguing that one ought to be able to interpolate the values of a finite sum of sinusoids from knowledge of a finite number of samples. Let's see how this works out, but rather than starting from scratch let's use what we've learned about sampling for general band-limited signals.

As always, it's best to work with the complex form of a sum of sinusoids, so we consider a real signal given by

$$f(t) = \sum_{k=-N}^{N} c_k e^{2\pi i k t/q}, \quad c_{-k} = \overline{c_k}.$$

$f(t)$ is periodic of period q. Recall that $c_{-k} = \overline{c_k}$. Some of the coefficients may be zero, but we assume that $c_N \neq 0$.

There are $2N+1$ terms in the sum (don't forget $k=0$) and it should take $2N+1$ sampled values over one period to determine $f(t)$ completely. You might think it would take twice this many sampled values because the values of $f(t)$ are real and we have to determine *complex* coefficients. But remember that $c_{-k} = \overline{c_k}$, so if we know c_k we know c_{-k}. Think of the $2N+1$ sample values as enough information to determine the real number c_0 and the N complex numbers $c_1, c_2, \ldots, c_N$.

The Fourier transform of f is

$$\mathcal{F}f(s) = \sum_{k=-N}^{N} c_k \delta(s - \frac{k}{q})$$

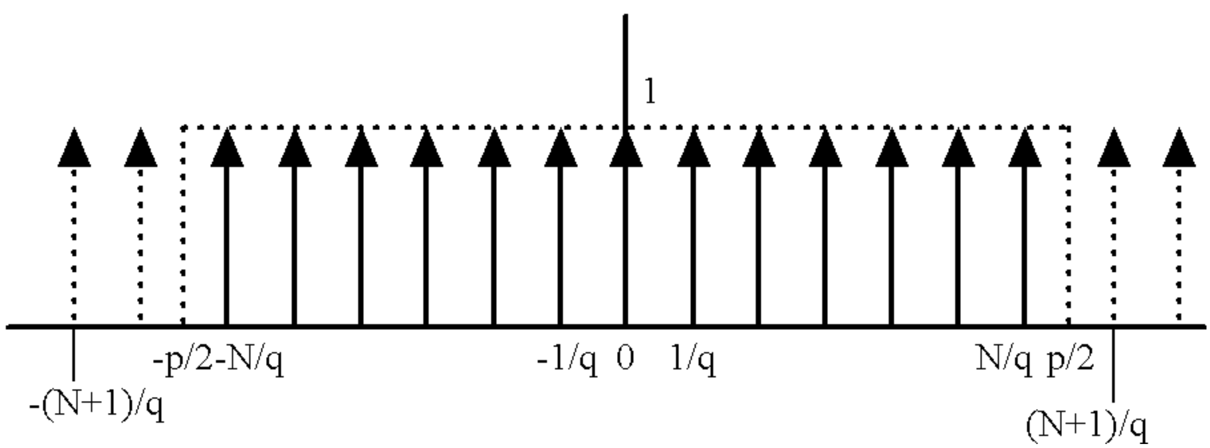
and the spectrum goes from $-N/q$ to N/q. The sampling formula applies to $f(t)$, and we can write an equation of the form

$$f(t) = \sum_{k=-\infty}^{\infty} f(t_k) \operatorname{sinc} p(t - t_k),$$

but it's a question of what to take for the sampling rate, and hence how to space the sample points.

We want to make use of the known periodicity of $f(t)$. If the sample points t_k are a fraction of a period apart, say q/M for an M to be determined, then the values $f(t_k)$ with $t_k = kq/M$, $k = 0, \pm 1, \pm 2, \ldots$ will repeat after M samples. We'll see how this collapses the interpolation formula.

To find the right sampling rate, p, think about the derivation of the sampling formula, the first step being: 'periodize $\mathcal{F}f$'. The Fourier transform $\mathcal{F}f$ is a bunch of δ's spaced $1/q$ apart (and scaled by the coefficients c_k). The natural periodization of $\mathcal{F}f$ is to keep the spacing $1/q$ in the periodized version, essentially making the periodized $\mathcal{F}f$ a scaled version of $Ш_{1/q}$. We do this by convolving $\mathcal{F}f$ with $Ш_p$ where $p/2$ is the midpoint between N/q, the last point in the spectrum of $\mathcal{F}f$, and the point $(N+1)/q$, which is the next point $1/q$ away. Here's a picture.



Thus we find p from

$$\frac{p}{2} = \frac{1}{2}\left(\frac{N}{q} + \frac{N+1}{q}\right) = \frac{(2N+1)}{2q}, \quad \text{or} \quad p = \frac{2N+1}{q}.$$

We periodize $\mathcal{F}f$ by $Ш_p$ (draw yourself a picture of this!), cut off by Π_p, then take the inverse Fourier transform. The sampling formula back in the time domain is

$$f(t) = \sum_{k=-\infty}^{\infty} f(t_k) \operatorname{sinc} p(t - t_k)$$

with

$$t_k = \frac{k}{p}.$$

With our particular choice of p let's now see how the q-periodicity of $f(t)$ comes into play. Write

$$M = 2N + 1,$$

so that then

$$t_k = \frac{k}{p} = \frac{kq}{M}.$$

Then, to repeat what we said earlier, the sample points are spaced a fraction of a period apart, q/M. and after $f(t_0)$, $f(t_1)$, $\ldots$, $f(t_{M-1})$ the sample values repeat, *e.g.* $f(t_M) = f(t_0)$, $f(t_{M+1}) = f(t_1)$ and so on. More succinctly,

$$t_{k+k'M} = t_k + k'q,$$

and so

$$f(t_{k+k'M}) = f(t_k + k'q) = f(t_k),$$

for any k and k'. Using this periodicity of the coefficients in the sampling formula, the single sampling sum splits into M sums as:

$$\begin{aligned}\sum_{k=-\infty}^{\infty} f(t_k)\,\mathrm{sinc}\,p(t-t_k) &= f(t_0)\sum_{m=-\infty}^{\infty}\mathrm{sinc}(pt-mM) + f(t_1)\sum_{m=-\infty}^{\infty}\mathrm{sinc}(pt-(1+mM)) \\ &\quad + f(t_2)\sum_{m=-\infty}^{\infty}\mathrm{sinc}(pt-(2+mM)) + \cdots + f(t_{M-1})\sum_{m=-\infty}^{\infty}\mathrm{sinc}(pt-(M-1+mM))\end{aligned}$$

Those sums of sincs on the right are periodizations of $\mathrm{sinc}\,pt$ and, remarkably, they have a simple closed form expression. The k'th sum is:

$$\begin{aligned}\sum_{m=-\infty}^{\infty}\mathrm{sinc}(pt-k-mM) &= \mathrm{sinc}(pt-k) * \text{III}_{M/p}(t) \\ &= \frac{\mathrm{sinc}(pt-k)}{\mathrm{sinc}(\frac{1}{M}(pt-k))} \\ &= \frac{\mathrm{sinc}(p(t-t_k))}{\mathrm{sinc}(\frac{1}{q}(t-t_k))}\end{aligned}$$

I'll give a derivation of this in Appendix 2. Using these identities, we find that the sampling formula, to interpolate

$$f(t) = \sum_{k=-N}^{N} c_k e^{2\pi i k t/q}$$

from $2N+1 = M$ sampled values, is then

$$f(t) = \sum_{k=0}^{2N} f(t_k)\frac{\mathrm{sinc}(p(t-t_k))}{\mathrm{sinc}(\frac{1}{q}(t-t_k))}, \quad p = \frac{2N+1}{q}, \quad t_k = \frac{k}{p} = \frac{kq}{2N+1}.$$

This is the 'finite sampling theorem' for periodic functions.

It might also be helpful to write the sampling formula in terms of frequencies. Thus, if the lowest frequency is $\nu_{\min} = 1/q$ and the highest frequency is $\nu_{\max} = N\nu_{\min}$ then

$$f(t) = \sum_{k=0}^{2N} f(t_k) \frac{\operatorname{sinc}((2\nu_{max} + \nu_{min})(t - t_k))}{\operatorname{sinc}(\nu_{\min}(t - t_k))}, \quad t_k = \frac{kq}{2N+1}.$$

The sampling rate is

$$\text{Sampling rate} = 2\nu_{\max} + \nu_{min}.$$

Compare this to

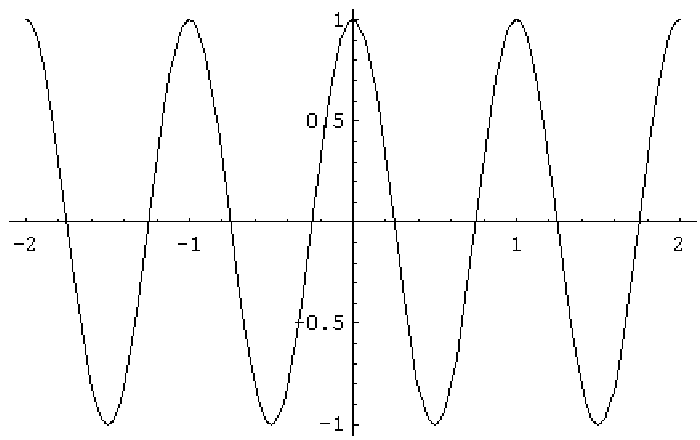
$$\text{Sampling rate} > 2\nu_{\max}$$

for a general band-limited function.

Here's a simple example of the formula. Take $f(t) = \cos 2\pi t$. There's only one frequency, and $\nu_{\min} = \nu_{\max} = 1$. Then $N = 1$, the sampling rate is 3 and the sample points are $t_0 = 0$, $t_1 = 1/3$ and $t_2 = 2/3$. The formula says

$$\cos 2\pi t = \frac{\operatorname{sinc} 3t}{\operatorname{sinc} t} + \cos\frac{2\pi}{3}\,\frac{\operatorname{sinc}(3(t - \frac{1}{3}))}{\operatorname{sinc}(t - \frac{1}{3})} + \cos\frac{4\pi}{3}\,\frac{\operatorname{sinc}(3(t - \frac{2}{3}))}{\operatorname{sinc}(t - \frac{2}{3})}$$

Does this really work? I'm certainly not going to plow through the trig identities needed to check it! However, here's a Mathematica plot of the right hand side.



Any questions? Ever thought you'd see such a complicated way of writing $\cos 2\pi t$?

Appendix 1: Time-limited *vs* band-limited signals

Here's a more careful treatment of the result that a band-limited signal cannot be time-limited, We'll actually prove a more general statement and perhaps I should have said that no *interesting* signal can be both time-limited and band-limited, because here's what we'll show precisely:

- Suppose $f(t)$ is a band-limited signal. If there is some interval $a \leq t \leq b$ on which $f(t)$ is identically zero then $f(t)$ is identically zero for all t.

This is a tricky argument. f is band-limited so $\mathcal{F}f(s)$ is zero, say, for $|s| \geq p/2$. The Fourier inversion formula says

$$f(t) = \int_{-\infty}^{\infty} \mathcal{F}f(s)e^{2\pi ist}\, ds = \int_{-p/2}^{p/2} \mathcal{F}f(s)e^{2\pi ist}\, ds.$$

(We assume the signal is such that Fourier inversion holds. You can take f to be a Schwartz function, but some more general signals will do.) Suppose $f(t)$ is zero for $a \leq t \leq b$. Then for t in this range,

$$\int_{-p/2}^{p/2} \mathcal{F}f(s)e^{2\pi ist}\, ds = 0.$$

Differentiate with respect to t under the integral. If we do this n-times we get

$$0 = \int_{-p/2}^{p/2} \mathcal{F}f(s)(2\pi is)^n e^{2\pi ist}\, ds = (2\pi i)^n \int_{-p/2}^{p/2} \mathcal{F}f(s)s^n e^{2\pi ist}\, ds,$$

so that

$$\int_{-p/2}^{p/2} \mathcal{F}f(s)s^n e^{2\pi ist}\, ds = 0.$$

Again, this holds for all t with $a \leq t \leq b$; pick one, say t_0. Then

$$\int_{-p/2}^{p/2} \mathcal{F}f(s)s^n e^{2\pi ist_0}\, ds = 0.$$

But now for *any* t (anywhere, not just between a and b) we can write

$$\begin{aligned}
f(t) &= \int_{-p/2}^{p/2} \mathcal{F}f(s)e^{2\pi ist}\, ds \\
&= \int_{-p/2}^{p/2} \mathcal{F}f(s)e^{2\pi is(t-t_0)}e^{2\pi ist_0}\, ds \\
&= \int_{-p/2}^{p/2} \sum_{n=0}^{\infty} \frac{(2\pi i(t-t_0))^n}{n!} s^n e^{2\pi ist_0} \mathcal{F}f(s)\, ds \\
&\qquad \text{(using the Taylor series expansion for } e^{2\pi is(t-t_0)}\text{)} \\
&= \sum_{n=0}^{\infty} \frac{(2\pi i(t-t_0))^n}{n!} \int_{-p/2}^{p/2} s^n e^{2\pi ist_0} \mathcal{F}f(s)\, ds = 0
\end{aligned}$$

Hence $f(t)$ is zero for all t.

The same argument *mutatis mutandis* will show:

- If $f(t)$ is time-limited and if $\mathcal{F}f(s)$ is identically zero on any interval $a \leq s \leq b$ then $\mathcal{F}f(s)$ is identically zero for all s.

Then $f(t)$ is identically zero, too, by Fourier inversions.

Remark 1, for eager seekers of knowledge This band-limited *vs* time-limited result is often proved by establishing a relationship between time-limited signals and analytic functions (of a complex variable), and then appealing to results from the theory of analytic functions. That connection opens up an important direction for applications of the Fourier transform, but we can't go there and the direct argument we just gave makes this approach unnecessary.

Remark 2, for overwrought math students and careful engineers Where in the preceding argument did we use that $p < \infty$? It's needed in switching integration and summation, in the line

$$\int_{-p/2}^{p/2} \sum_{n=0}^{\infty} \frac{(2\pi i(t-t_0))^n}{n!} s^n e^{2\pi i s t_0} \mathcal{F}f(s)\,ds = \sum_{n=0}^{\infty} \frac{(2\pi i(t-t_0))^n}{n!} \int_{-p/2}^{p/2} s^n e^{2\pi i s t_0} \mathcal{F}f(s)\,ds$$

The theorems that tell us 'the integral of the sum is the sum of the integral' require as an essential hypothesis that the series converges *uniformly*. 'Uniformly' means, loosely, that if we plug a particular value into the converging series we can estimate the rate at which the series converges *independent* of that particular value.[3] In the sum-and-integral expression, above, the variable s ranges over a finite interval, from $-p/2$ to $p/2$. Over such a finite interval the series for the exponential converges uniformly, essentially because the terms can only get so big – so they can be estimated uniformly – when s can only get so big. We can switch integration and summation in this case. If, however, we had to work with

$$\int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \frac{(2\pi i(t-t_0))^n}{n!} s^n e^{2\pi i s t_0} \mathcal{F}f(s)\,ds,$$

i.e. if we did not have the assumption of band-limitedness, then we could not make uniform estimates for the convergence of the series and switching integration and summation is not justified.

It's not only unjustified, it's really wrong. If we could drop the assumption that the signal is band-limited we'd be 'proving' the statement: If $f(t)$ is identically zero on an interval then it's identically zero. Think of the implications of such a dramatic statement. In a phone conversation if you paused for a few seconds to collect your thoughts your signal would be identically zero on that interval of time, and therefore you would have nothing to say at all, ever again. Be careful.[4]

Appendix 2: Periodizing sinc functions

In applying the general sampling theorem to the special case of a periodic signal we wound up with sums of sinc functions which we recognized (sharp eyed observers that we are) to be periodizations. Then, out of nowhere, came a closed form expression for such periodizations as a ratio of sinc functions. Here's where this comes from, and here's a fairly general result that covers it.

Lemma Let p, $q > 0$ and let N be the largest integer strictly less than $pq/2$. Then

$$\sum_{k=-\infty}^{\infty} \operatorname{sinc}(pt - kpq) = \operatorname{sinc}(pt) * Ш_q(t) = \frac{1}{pq} \frac{\sin((2N+1)\pi t/q)}{\sin(\pi t/q)}.$$

[3] We can make 'uniform' estimates, in other words. We saw this sort of thing in the notes on convergence of Fourier series.

[4] However, if $f(t)$ is a real analytic signal, that is if it is given by a convergent power series at each point in its domain, then the implication: '$f(t)$ identically zero on an interval $\implies f(t)$ identically zero everywhere' is true.

There's a version of this lemma with $N \le pq/2$, too, but that's not important for us. In terms of sinc functions the formula is

$$\operatorname{sinc}(pt) * Ш_q(t) = \frac{2N+1}{pq}\frac{\operatorname{sinc}((2N+1)t/q)}{\operatorname{sinc}(t/q)}.$$

It's then easy to extend the lemma slightly to include periodizing a shifted sinc function, $\operatorname{sinc}(pt+b)$, namely

$$\sum_{k=-\infty}^{\infty} \operatorname{sinc}(pt+b-kpq) = \operatorname{sinc}(pt+b) * Ш_q(t) = \frac{2N+1}{pq}\frac{\operatorname{sinc}(\frac{2N+1}{pq}(pt+b))}{\operatorname{sinc}(\frac{1}{pq}(pt+b))}$$

This is what is needed in the last part of the derivation of the finite sampling formula.

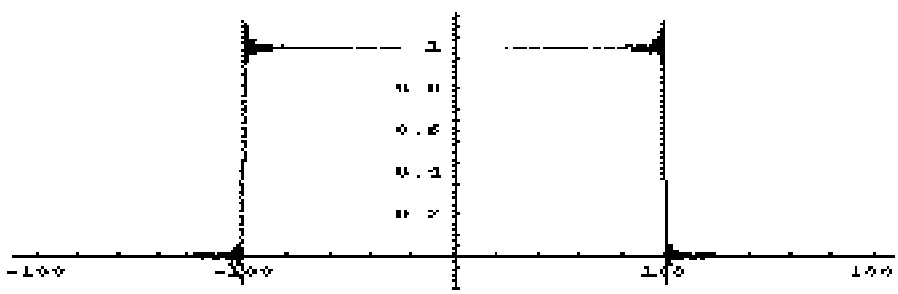
Having written this lemma down so grandly I now have to admit that it's really only a special case of the general sampling theorem as we've already developed it, though I think it's fair to say that this is only 'obvious' in retrospect. The fact is that the ratio of sine functions on the right hand side of the equation is a band-limited signal (we've seen it before, see below) and the sum for $\operatorname{sinc}(pt) * Ш_q(t)$ is just the sampling formula applied to that function. One usually thinks of the sampling theorem as going *from* the signal *to* the series of sampled values, but it can also go the other way. This admission notwithstanding, I still want to go through the derivation, from scratch

One more thing before we do that. If $p = q = 1$, so that $N = 0$, the formula in the lemma gives

$$\sum_{n=-\infty}^{\infty} \operatorname{sinc}(t-n) = \operatorname{sinc} t * Ш_1(t) = 1.$$

Striking. Still don't believe it? Here's a plot of

$$\sum_{n=-100}^{100} \operatorname{sinc}(t-n).$$



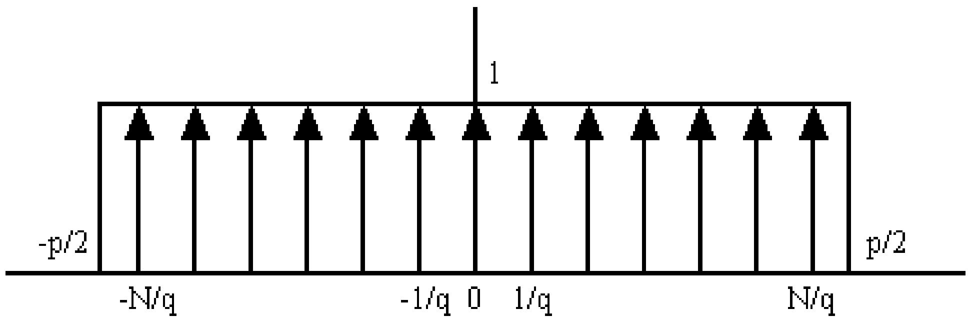
Note the Gibbs-like phenomena at the edges. This means there's some issue with what kind of convergence is involved, which is the last thing I want to worry about.

We proceed with the derivation of the formula

$$\operatorname{sinc}(pt) * Ш_q(t) = \frac{1}{pq}\frac{\sin((2N+1)\pi t/q)}{\sin(\pi t/q)}$$

This will look awfully familiar, indeed I'll really just be repeating the derivation of the general sampling formula for this special case. Take the Fourier transform of the convolution:

$$\begin{aligned}
\mathcal{F}(\operatorname{sinc}(pt) * Ш_q(t)) &= \mathcal{F}(\operatorname{sinc}(pt)) \cdot \mathcal{F}Ш_q(t) \\
&= \frac{1}{p}\Pi_p(s) \cdot \frac{1}{q}Ш_{1/q}(s) \\
&= \frac{1}{pq}\sum_{n=-N}^{N} \delta(s - \frac{n}{q}) \quad \text{(see the figure, below).}
\end{aligned}$$



And now take the inverse Fourier transform:

$$\begin{aligned}
\mathcal{F}^{-1}\left(\frac{1}{pq}\sum_{n=-N}^{N} \delta(s - \frac{n}{q})\right) &= \frac{1}{pq}\sum_{n=-N}^{N} e^{2\pi i n t/q} \\
&= \frac{1}{pq}\frac{\sin(\pi(2N+1)t/q))}{\sin(\pi t/q)}.
\end{aligned}$$

There it is. One reason I wanted to go through this is because it is another occurrence of the sum of exponentials and the identity

$$\sum_{n=-N}^{N} e^{2\pi i n t/q} = \frac{\sin(\pi(2N+1)t/q))}{\sin(\pi t/q)}$$

which we've now seen on at least two other occasions. Reading the equalities backwards we have

$$\mathcal{F}\left(\frac{\sin(\pi(2N+1)t/q))}{\sin(\pi t/q)}\right) = \mathcal{F}\left(\sum_{n=-N}^{N} e^{2\pi i n t/q}\right) = \sum_{n=-N}^{N} \delta(s - \frac{n}{q}).$$

This substantiates the earlier claim that the ratio of sines is band-limited, and hence we could have appealed to the sampling formula directly instead of going through the argument we just did. But who would have guessed it?