Lecture 18

The Sampling Theorem

Relevant section from Boggess and Narcowich: 2.7, pp. 117-120.

Electronic storage and transmission of signals and images has been of obvious importance in our civilization. From the telephone, to radio, and then to television, engineers and scientists have consistently been faced with the basic question of how to store and transmit information as efficiently as possible.

In the not-too-distant-past pre-digital age, the transmission and storage of audio and video (except for still images) was analog, i.e. continuous in time, in the form of reel-to-reel tapes and videotape. The advent of computers ushered in the digital age, where continuous signals were replaced by sequences of “bits”, i.e., 0’s and 1’s. This led to digital storage devices that mimicked the storage of information in computer memory: floppy disks of various sizes, followed by audio digital tape, compact discs (CDs) and, most recently, DVDs.

As mentioned earlier, there has always been, and perhaps always will be, the fundamental question of how to store and transmit information as efficiently as possible. Back in the pre-digital age of analog communication, Claude Shannon of Bell Labs (later to be AT &T Labs) provided a basic reference point for communication theory in a celebrated paper. (C. Shannon, “A mathematical theory of communication,” Bell System Technical Journal, vol. 27, pp. 379, 623 (1948).) Shannon’s classic paper gave birth to rapid advances in information and communication theory.

That being said, Shannon was actually not the first to come up with this fundamental result. There is a very interesting history behind the Sampling Theorem and so-called “cardinal series,” to be introduced below. A brief discussion is given in the introductory chapter of the book, Introduction to Shannon Sampling and Interpolation Theory, by R.J. Marks II (Springer-Verlag, NY, 1991). Marks writes that one historian (H.S. Black) credits the mathematician Cauchy for understanding the idea of sampling a function with finite support, citing a work of 1841. Another researcher (Higgins, 1985) disputes this claim and credits the mathematician E. Borel in an 1897 paper. The British mathematician E.T. Whittaker published a highly cited paper on the sampling theorem in 1915. (E.T. Whittaker, “On the functions which are represented by the expansions of the interpolation theory,” Proc. Roy. Soc. Edinburgh, vol. 35, pp. 181-194 (1915).) Whittaker’s formula was
later called the *cardinal series* by his son, J.M. Whittaker. V.A. Kotel’nikov reported the sampling theorem in a Soviet journal in 1933. As stated earlier, Shannon showed the importance of the sampling theorem to communication theory in his 1948 paper, in which he cited Whittaker’s 1915 paper. A number of other events in the development of the cardinal series are listed by Marks.

In any case, Shannon’s paper was fundamental in showing the application of the Sampling Theorem to communications, thereby attracting the attention of the communications research community.

The basic question asked by Shannon and others was as follows. Suppose that we have a continuous, or analog, signal $f(t)$ – for example, an audio signal – that is sampled to produce discrete data points, as we discussed in earlier lectures, i.e.,

$$f[n] = f(nT), \quad n \in \mathbb{Z}.$$  \hspace{1cm} (1)

Here, $T > 0$ is the sampling period. Can we reconstruct $f(t)$ perfectly for all $t \in \mathbb{R}$ from these samples?

Before we examine the Sampling Theorem of Shannon et al., let us step back and think a little about this problem. Suppose that you were given the data points $f[n]$. What could one do in an effort to construct $f(t)$, or at least approximations to it?

The simplest response would be to attempt various *interpolations* of the points $f[n]$. And the simplest interpolation would be:

**Piecewise constant interpolation:** We define the following approximation $g_0(t)$ to $f(t)$: For $n \in \mathbb{Z}$,

$$g_0(t) = f(nT), \quad nT \leq t < (n+1)T,$$  \hspace{1cm} (2)

sketched schematically below.

There is one obvious drawback to this approach: $g(t)$ is discontinuous at the sample points, which would probably be disastrous for audio signals. (In two-dimensions, it is not such a bad approximation for images. In fact, digital images are piecewise constant approximations to a “real” continuous photo or scene.)

There is another way of looking at this approximation which will be quite useful in our later
Piecewise constant approximation $g_0(t)$ to continuous signal $f(t)$.

discussions. Let us define the fundamental basis function $\phi(t)$ for $t \in \mathbb{R}$.

$$
\phi(t) = \begin{cases} 
1, & 0 \leq t < T, \\
0, & \text{otherwise}.
\end{cases}
$$

Then our piecewise constant function $g_0(t)$ may be written as

$$
g_0(t) = \sum_{n=-\infty}^{\infty} f(nT) \phi(t - nT).
$$

Each translate $\phi(t - nT)$ has value 1 over the interval $[nT, (n+1)T)$ and is zero outside this interval. This is what permits us to write Eq. (4). The set of all translates $\phi(t - nT)$, $n \in \mathbb{Z}$, serves as a basis for all functions on $\mathbb{R}$ that are piecewise constant on the intervals $[nT, (n+1)T)$. In fact, these basis functions are orthogonal to each other. This idea will be important in our study of wavelets.

**Piecewise linear interpolation:** Now define the approximation $g_1(t)$ to $f(t)$ as follows: For $n \in \mathbb{Z}$,

$$
g_1(t) = \frac{(n+1)T - t}{T} f(nT) + \frac{t - nT}{T} f((n+1)T), \quad nT \leq t < (n+1)T.
$$

By construction, $g_1(nT) = f(nT)$ for all $n$, and the graph of $g_1(t)$ from $f(nT)$ to $f((n+1)T)$ is a straight line, as sketched below.

We may also view the function $g_1(t)$ as a linear combination of basis functions which are translates of a fundamental basis function $h(t)$. To see this, consider the sketch below, where we have drawn triangular “hat” functions that have bases on the intervals $[(n-1)T, (n+1)T]$ and apices at points $nT$ with heights $f(nT)$. 

![Diagram of piecewise linear interpolation](image-url)
Piecewise linear approximation/interpolation \( g_1(t) \) to continuous signal \( f(t) \).

Each triangular function is a translation and vertically scaled version of the following function, which is sketched in the figure below.

\[
h(t) = \begin{cases} 
\frac{t}{T} + 1, & -T \leq t < 0, \\
1 - \frac{t}{T}, & 0 \leq t < T \\
0, & \text{otherwise.}
\end{cases}
\] (6)

The fact that \( h(0) = 1 \) dictates that the triangular function below the sample point at \( t = nT \) must be multiplied by the sample value \( f(nT) \). And the fact that \( h(-T) = h(T) = 0 \) produces the linear interpolation between adjacent sample values. As a result, the function \( g_1(t) \) may be written as

\[ g_1(t) = \sum_{n=-\infty}^{\infty} f(nT)h(t - nT). \] (7)

Notice the similarity in form between Eqs. (4) and (7). The translates \( h(t - nT), \ n \in \mathbb{Z} \), form a
Triangular hat function \( h(t) \) whose translates comprise a basis for piecewise linear functions.

(nonorthogonal) basis for piecewise linear functions over the intervals \([nT, (n+1)T]\) on \( \mathbb{R} \).

**Higher-order interpolations:** It is possible to construct \( k \)th degree polynomials that interpolate between the sample points \( f(nT) \) and \( f((n+1)T) \) using \( k+1 \) consecutive sample points that contain these points. These polynomial interpolation functions are called *splines* and will comprise the interpolation function \( g_k(t) \).

We now return to the Sampling Theorem. Shannon’s idea was to restrict attention to “bandlimited” functions: functions \( f(t) \) with Fourier transforms \( F(\omega) \) that were identically zero outside a finite interval, assumed to be the symmetric interval \([-\Omega, \Omega]\) for some \( \Omega > 0 \), i.e.,

\[
F(\omega) = 0 \quad \text{for} \quad |\omega| > \Omega.
\]  

\( \Omega \) is known as the *band limit* of \( f \) (or \( F \)).

Does this sound like an artificial constraint? Perhaps, but, in fact, it is practical, for the following reasons.

1. Sounds made by the human voice are contained well within the frequency range of an 88-key piano keyboard:

   “low A” at 27.5 Hz , (1 Hz = 1 cycle/second)

   “high C” at 4186 Hz.

Therefore, speech signals are essentially bandlimited, with \( \Omega = 4200 \times 2\pi = 8400\pi \).

**Note:** \( \Omega \) is the *angular velocity*, in units of radians/unit time. There are \( 2\pi \) radians/cycle. Equivalently, we have \( 1/(2\pi) \) cycle/radian.
2. The human ear can hear sounds in roughly the range 20-20,000 Hz. As such, audible sounds are bandlimited, with $\Omega \approx 20,000 \times 2\pi = 40,000\pi$.

**Proof of the “Whittaker-Shannon” Sampling Theorem**

Once again, we define *bandlimited* functions as follows:

A function $f(t)$, defined on the real line, i.e., $t \in \mathbb{R}$, is said to be bandlimited, or $\Omega$-bandlimited, if there exists an $\Omega > 0$ such that its Fourier transform behaves as follows,

$$F(\omega) = 0 \text{ for } |\omega| > \Omega. \quad (9)$$

In practice, one generally tries to find the smallest such frequency $\Omega$ for which (9) holds.

Associated with the angular frequency band limit $\Omega$ (radians/second) is the (cyclical) frequency

$$\nu = \frac{\Omega}{2\pi} \text{ Hz (cycles/second)}, \quad (10)$$

known as the *Nyquist frequency*. The *Nyquist rate* is given by

$$2\nu = \frac{\Omega}{\pi} \text{ Hz.} \quad (11)$$

Its importance will become clear after we study the Sampling Theorem.

**The Whittaker-Shannon Sampling Theorem:** Let $f(t)$ be an $\Omega$-bandlimited function, with Fourier transform $F(\omega)$ that satisfies Eq. (9) for some $\Omega > 0$. Furthermore, assume that $F(\omega)$ is piecewise continuous on $[-\Omega, \Omega]$. Then $f = \mathcal{F}^{-1}(F)$ is completely determined at any $t \in \mathbb{R}$ by its values at the points $t_k = \frac{k\pi}{\Omega}$, $k = 0, \pm 1, \pm 2, \cdots$, as follows,

$$f(t) = \sum_{k=-\infty}^{\infty} f\left(\frac{k\pi}{\Omega}\right) \frac{\sin(\Omega t - k\pi)}{\Omega t - k\pi} \quad (12)$$

$$= \sum_{k=-\infty}^{\infty} f\left(\frac{k\pi}{\Omega}\right) \text{sinc}\left(\frac{\Omega t}{\pi} - k\right). \quad (13)$$

Furthermore, the series on the right converges uniformly on closed subsets of $\mathbb{R}$. 

217
Note: Please note that we are now using the signal/image processing definition of the sinc function, i.e.,

\[
sinc(x) = \begin{cases} \frac{\sin \pi x}{\pi x}, & x \neq 0, \\ 1, & x = 0. \end{cases}
\]

This form of the sinc function includes the factor of \(\pi\) in the definition.

Just a couple of comments before we prove this theorem.

1. First of all, the series on the right, in whatever form, is known as the cardinal series. (As mentioned in the introduction to this lecture, the origin of this term is due to J.M. Whittaker.)

2. From Eqs. (10) and (11), we may express the cardinal series in the following alternate forms:

\[
f(t) = \sum_{k=-\infty}^{\infty} f \left( \frac{k}{2\nu} \right) \text{sinc} \left( 2\nu t - k \right) 
\]

\[
= \sum_{k=-\infty}^{\infty} f(kT) \text{sinc} \left( \frac{t}{T} - k \right),
\]

where

\[
T = \frac{1}{2\nu} = \frac{\pi}{\Omega}
\]

is the sampling period. Note that

\[
T = \frac{1}{2} \left( \frac{2\pi}{\Omega} \right).
\]

In other words, the period \(T\) is one-half the period associated with the bandwidth \(\Omega\). Another way to put this is as follows:

The sampling frequency is twice the bandwidth frequency \(\Omega\).

The above is in terms of angular frequency. In terms of cycles per unit time, this explains why the Nyquist rate of sampling is twice the Nyquist frequency (associated with the bandwidth).

We now prove the W-S Theorem.

Proof: We exploit the fact that the Fourier transform is supported on the finite interval \(\omega \in [-\Omega, \Omega]\) and expand \(F(\omega)\) in terms of a Fourier series. We’ll use the complex exponential functions \(u_k(\omega) = e^{i\pi k\omega/\Omega}\) that form an orthogonal set on \([-\Omega, \Omega]\). (You’ve seen these before, but for the independent variable \(t \in [-L, L]\) instead of \(\omega \in [-\Omega, \Omega]\).) The (unnormalized) Fourier series for \(F(\omega)\) has the form

\[
F(\omega) = \sum_{k=-\infty}^{\infty} c_k e^{i\pi k\omega/\Omega},
\]

218
where
\[
c_k = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\omega) e^{-i\pi k \omega/\Omega} d\omega.  \tag{20}
\]
Because \(F(\omega)\) is assumed to be piecewise continuous on \([-\Omega, \Omega]\), this series will converge uniformly over intervals that do not contain points of discontinuity of \(F\).

Since \(F(\omega) = 0\) for \(|\omega| > \Omega\), we may change the limits of integration to extend over the entire real line:

\[
c_k = \frac{1}{2\Omega} \int_{-\infty}^{\infty} F(\omega) e^{-i\pi k \omega/\Omega} d\omega
\]
\[
= \frac{\sqrt{2\pi}}{2\Omega} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega(-\frac{\pi k}{\Omega})} d\omega,  \tag{21}
\]
where we have rewritten the integrand slightly so that it has the form of an inverse Fourier transform (which also explains the introduction of the factor \(1/\sqrt{2\pi}\)). Indeed, we may write that

\[
c_k = \frac{\sqrt{2\pi}}{2\Omega} f\left(-\frac{k\pi}{\Omega}\right).  \tag{22}
\]

Now substitute this result into Eq. (19) to obtain

\[
F(\omega) = \frac{\sqrt{2\pi}}{2\Omega} \sum_{k=-\infty}^{\infty} f\left(-\frac{k\pi}{\Omega}\right) e^{i\pi k \omega/\Omega}.  \tag{23}
\]

For what is known as “cosmetic reasons,” we replace \(k\) with \(-k\) (or let \(l = -k\), then replace \(l\) with \(k\)) to give

\[
F(\omega) = \frac{\sqrt{2\pi}}{2\Omega} \sum_{k=-\infty}^{\infty} f\left(\frac{k\pi}{\Omega}\right) e^{-i\pi k \omega/\Omega}.  \tag{24}
\]

We'll now use this result for \(F(\omega)\) in the expression of \(f(t)\) as an inverse Fourier transform:

\[
f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega
\]
\[
= \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} F(\omega) e^{i\omega t} d\omega \quad \text{(since } F(\omega) = 0 \text{ for } |\omega| > \Omega) \]
\[
= \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} \sum_{k=-\infty}^{\infty} f\left(\frac{k\pi}{\Omega}\right) e^{-i\omega\left[\frac{\pi k}{\Omega} - t\right]} d\omega
\]
\[
= \frac{1}{2\Omega} \sum_{k=-\infty}^{\infty} f\left(\frac{k\pi}{\Omega}\right) \int_{-\Omega}^{\Omega} e^{-i\omega\left[\frac{\pi k}{\Omega} - t\right]} d\omega.  \tag{25}
\]

The uniform convergence of the Fourier series to \(F(\omega)\) (over intervals not containing discontinuities) permits the interchange of summation and integration that produced the final line.
We now evaluate the integrals:

\[
\int_{-\Omega}^{\Omega} e^{-i\omega \left( \frac{\pi}{\Omega} - t \right)} d\omega = \frac{1}{-i} \frac{1}{\frac{\pi k}{\Omega} - t} \left[ e^{-i\pi k - \Omega t} - e^{i\pi k - \Omega t} \right]
\]

\[
= \frac{2}{\frac{\pi k}{\Omega} - t} \sin(\pi k - \Omega t).
\]  

(26)

Substitution into Eq. (25) yields the desired result,

\[
f(t) = \sum_{k=\infty}^{\infty} f\left( \frac{k\pi}{\Omega} \right) \frac{\sin(\Omega t - k\pi)}{(\Omega t - k\pi)}
\]

\[
= \sum_{k=\infty}^{\infty} f\left( \frac{k\pi}{\Omega} \right) \text{sinc} \left( \frac{\Omega t}{\pi} - k \right).
\]  

(27)

This is the cardinal series for an \( \Omega \)-bandlimited function \( f(t) \) and the proof is complete.
Lecture 19

Sampling Theorem (cont’d)

Recall that if \( f(t) \) is \( \Omega \)-bandlimited, i.e., its Fourier transform \( F(\omega) \) is identically zero outside the interval \([−\Omega, \Omega]\) for some \( \Omega > 0 \), then from the Sampling Theorem, proved in the previous lecture, it may be reconstructed exactly for any \( t \in \mathbb{R} \) from samples taken at times \( t_k = \frac{k\pi}{\Omega}, \ k \in \mathbb{Z} \). The reconstruction is performed via the so-called cardinal series:

\[
    f(t) = \sum_{k=-\infty}^{\infty} f \left( \frac{k\pi}{\Omega} \right) \frac{\sin(\Omega t - k\pi)}{(\Omega t - k\pi)}
\]

\[
    = \sum_{k=-\infty}^{\infty} f \left( \frac{k\pi}{\Omega} \right) \text{sinc} \left( \frac{\Omega t}{\pi} - k \right). \tag{28}
\]

We now make some comments on this result:

1. For a fixed \( t \in \mathbb{R} \), the cardinal series can converge slowly, since the sinc function decays on the order of \( O(1/k) \) as \( k \to \infty \). (Actually, the function \( f(t) \) also decays to zero as \( t \to \pm \infty \) since it is assumed to be in \( L^2(\mathbb{R}) \), which can improve the convergence somewhat.) There are ways to increase the convergence rate, including “oversampling,” which may be addressed later.

2. Since the sampling period is given by \( T = \pi/\Omega \), functions \( f(t) \) with higher \( \Omega \) band limit values must be sampled more frequently. This makes sense: functions with higher \( \Omega \) values naturally have higher frequency components. It is necessary to sample the function more often in order to capture the greater variability produced by these higher frequencies.

3. From the previous comment, it is desirable to find the smallest \( \Omega \) for which the Fourier transform \( F(\omega) \) vanishes outside the interval \([−\Omega, \Omega]\). In this way, the function \( f(t) \) does not have to be sampled at too high a rate. That being said, a cardinal series employing a higher value of \( \Omega \) than necessary will still converge to \( f(t) \).

On the other hand, if we use an \( \Omega \) value that is too small, i.e., the cardinal series will not converge to \( f(t) \). This is because the Fourier transform \( F(\omega) \) is not being properly approximated by the Fourier series. For example, if \( \Omega = 1 \) and you use a value of \( 1/2 \) in the cardinal series, you are only approximating \( F(\omega) \) over \([-1/2, 1/2]\) and ignoring the portions of \( F(\omega) \) that lie in the
intervals $[1/2, 1]$ and $[-1, -1/2]$. This is the problem of “undersampling,” to which we shall return below.

4. The above cardinal series may be rewritten as follows,

$$f(t) = \sum_{k=-\infty}^{\infty} f\left(\frac{k\pi}{\Omega}\right) \text{sinc}\left[\frac{\Omega}{\pi}\left(t - \frac{k\pi}{\Omega}\right)\right]$$

$$= \sum_{k=-\infty}^{\infty} f(kT) \text{sinc}\left[\frac{1}{T}(t - kT)\right].$$

(29)

In this form, we see that it is a sum of shifted sinc functions that are multiplied by the sample values $f(kT)$, where $T = \pi/\Omega$. Moreover, the sinc functions are shifted by multiples of $T$ as well. In other words, the cardinal series has the same general form as the series for piecewise constant and piecewise linear interpolations of $f(t)$, shown in the previous lecture.

This indicates that the translated shifted sinc functions may be viewed as basis functions. Moreover, each sinc function is nonzero only at the point where it is multiplied by a sample point $f(kT)$ – at the other sample points it is zero. As such, the cardinal series provides a quite sophisticated interpolation of the sample points $f(kT)$. This property is sketched in the figure below.

![Plot of a generic function $f(t)$ along with the components $f(kT)\text{sinc}\left(\frac{1}{T}(t - kT)\right)$ of its cardinal series. Here, $T = 1$. The component $k = 4$ is labelled.](image)

**Example:** Consider the Fourier transform given by

$$F(\omega) = \begin{cases} \sqrt{2\pi}(1 - \omega^2), & -1 \leq \omega \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

(30)
the graph which is sketched below.

In this case, the band limit of $F(\omega)$ is $\Omega = 1$. We solve for the function $f = F^{-1} F$:

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} \, d\omega$$

$$= \int_{-1}^{1} (1 - \omega^2) e^{i\omega t} \, d\omega$$

$$= \frac{4}{t^3} [\sin t - t \cos t]. \quad (31)$$

(Details left as an exercise.) Note that despite the $t^3$ term in the denominator, $f(t)$ is continuous at $t = 0$: $f(t) \to \frac{4}{3}$ as $t \to 0$. (If $f(t)$ were discontinuous at $t = 0$, then the Fourier transform $F(\omega)$ would not be bandlimited.)

In this case, since $\Omega = 1$, the sampling period is given by $T = \frac{\pi}{\Omega} = \pi$. As such, the cardinal series for $f(t)$ has the form

$$f(t) = \sum_{k=-\infty}^{\infty} f(k\pi) \frac{\sin(t - k\pi)}{t - k\pi}. \quad (32)$$

The sample values $f(kT)$ are given by

$$f(0) = \frac{4}{3} \quad \text{and} \quad f(k\pi) = \frac{4 \cos(k\pi)}{(k\pi)^2} = \frac{(-1)^{k+1}}{(k\pi)^2}, \quad k \neq 0. \quad (33)$$

In the figure below are plotted several partial sums for this cardinal series having the form

$$S_N(t) = \sum_{k=-N}^{N} f(k\pi) \frac{\sin(t - k\pi)}{t - k\pi}, \quad (34)$$

on the interval $[-20, 20]$. To the accuracy of the plot, perfect reconstruction of $f(t)$ is achieved over this interval for $N = 10$ (21 terms) and even $N = 5$ (11 terms). For $N = 3$, errors are visible
only for \( t \geq 15 \). The case \( N = 0 \), i.e., using only the \( f(0) \) term, corresponds to a one-sinc-function approximation to \( f(t) \). In all figures, the true \( f(t) \) is shown as a dotted plot.

Approximations to \( f(t) = \frac{4}{t^3} \left[ \sin t - t \cos t \right] \) using partial sums \( S_N(t) \) of cardinal series given in main text. Where seen, the dotted plots correspond to the true values of \( f(t) \).

Now suppose that we did not know the exact bandlimiting value \( \Omega \) for a function \( f(t) \). We would then have to consider cardinal series that are constructed with estimates \( \Omega' \) of \( \Omega \), having the form

\[
g(t) = \sum_{k=-\infty}^{\infty} f \left( \frac{k\pi}{\Omega'} \right) \sin(\Omega't - k\pi) \frac{\sin(\Omega'\pi)}{\Omega'\pi}, \quad \Omega' < \Omega. \tag{35}\]

As mentioned earlier, if \( \Omega' \geq \Omega \), then the above cardinal series will converge to \( f(t) \). since it is still taking into consideration the interval \( [-\Omega, \Omega] \) on which the Fourier transform \( F(\omega) \) is supported. On the other hand, if \( \Omega' < \Omega \), then the above cardinal series will not converge to \( f(t) \) since the portions of the spectrum lying outside the interval \( [-\Omega', \Omega'] \) are not being approximated properly, as sketched in the figure below.
These parts of the spectrum $F(\omega)$ are ignored by undersampling.

The plots shown in the next figure demonstrate this problem. We have used a value of $\Omega' = 0.8$ in the cardinal series for the function $f(t)$ of the previous example, where the true band limit is $\Omega = 1$. The partial sums for $N = 10$ and $20$ are virtually identical, which indicates that convergence of the cardinal series to a limit has been obtained. This limit, however, is not $f(t)$.

Approximations to $f(t) = \frac{4}{t^3} \left[ \sin t - t \cos t \right]$ using partial sums of cardinal series, when the bandwidth $\Omega$ is underestimated. In the calculations above, we have used $\Omega' = 0.8$, which is below the true band limit $\Omega = 1$. Significant errors are incurred with $N = 10$. No improvement for $N = 20$. Plot of true $f(t)$ is dotted.

A generalized Sampling Theorem for samples obtained by convolutions

In practice, it is generally impossible to obtain perfect samples $f(kT)$ of a signal $f(t)$. Samples are generally obtained by some kind of convolution/local averaging process, which may be modelled as follows: At a point $t_k$, the sampling of $f(t)$ is given by

$$g(t_k) = \int_{-\Delta}^{\Delta} f(t_k - s) a_\Delta(s) \, ds$$
where \( a_\Delta(t) \) is a non-negative convolution kernel with finite support, i.e.,

1. \( a_\Delta(t) \geq 0 \) for \( t \in [-\Delta, \Delta] \),
2. \( a_\Delta(t) = 0 \) for \( |t| > \Delta \), and
3. \( \int_{-\Delta}^{\Delta} a_\Delta(t) \, dt = 1. \)

In practice, \( \Delta \) may be very small. In fact, it should be small compared to \( T \) the sampling period. We shall also assume that the following mathematical limit exists:

\[
a_\Delta(t) \to \delta(t), \quad \text{as} \quad \Delta \to 0,
\]

where \( \delta(t) \) denotes the Dirac delta function. In this limit \( g(t_k) \to f(t_k) \).

The goal is to reconstruct the function \( f(t) \) from the samples \( g(t_k) \) along the same lines as was done for the Whittaker-Shannon Theorem, where \( g(t_k) = f(t_k) \). As before, we assume that \( f(t) \) is bandlimited, i.e., its Fourier transform is identically zero outside the interval \([ -\Omega, \Omega ]\) for some \( \Omega > 0 \).

First, consider the function,

\[
g(t) = (f * a)(t),
\]

where \( a(t) \) is the sampling convolution kernel – we have dropped the \( \Delta \) subscript for notational convenience. Note that even though the sampling is performed at discrete sample points \( t_k = kT \), we shall, at least for the moment, assume that the convolution is performed at all \( t \in \mathbb{R} \) to produce \( g(t) \).

From the Convolution Theorem, the Fourier transform of \( g(t) \) is given by

\[
G(\omega) = \sqrt{2\pi} F(\omega) A(\omega),
\]

where \( A(\omega) = \mathcal{F}[a] \) is the Fourier transform of \( a(t) \). Note also that since

\[
F(\omega) = 0, \quad |\omega| > \Omega,
\]

it follows that

\[
G(\omega) = 0, \quad |\omega| > \Omega.
\]

In other words, \( g(t) \) is also \( \Omega \)-bandlimited.
We shall also assume that \( A(\omega) \neq 0 \) on \([-\Omega, \Omega]\) so that
\[
F(\omega) = \frac{1}{\sqrt{2\pi}} \frac{G(\omega)}{A(\omega)}.
\]
We’ll return to this expression later.

As was done for the Sampling Theorem, we expand \( G(\omega) \) in a Fourier series over its interval of support \([-\Omega, \Omega]\),
\[
G(\omega) = \sum_{k=-\infty}^{\infty} c_k e^{i\pi k\omega/\Omega}.
\]
In the same way as for the Sampling Theorem, we find that
\[
c_k = \frac{\sqrt{2\pi} g}{2\Omega} \left( -\frac{k\pi}{\Omega} \right).
\]
Substituting this result into Eq. (43), and replacing \( k \) with \(-k\), we obtain
\[
G(\omega) = \frac{\sqrt{2\pi}}{2\Omega} \sum_{k=-\infty}^{\infty} g \left( \frac{k\pi}{\Omega} \right) e^{-i\pi k\omega/\Omega}.
\]

We now return to \( f(t) \) and express it in terms of an inverse Fourier transform,
\[
f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega.
\]
However, the problem is that we do not know \( F(\omega) \). We’ll substitute the expression in (42) for \( F(\omega) \) above,
\[
f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{G(\omega)}{A(\omega)} e^{i\omega t} d\omega,
\]
and then substitute the Fourier series expansion for \( G(\omega) \):
\[
f(t) = \frac{1}{\sqrt{2\pi}} \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} \frac{1}{A(\omega)} \sum_{k=-\infty}^{\infty} g \left( \frac{k\pi}{\Omega} \right) e^{-i\omega \left[ \frac{\pi k}{\Omega} - t \right]} d\omega.
\]
Since the Fourier series to \( G(\omega) \) is assumed to converge uniformly, we may interchange the operations of integration and summation to give,
\[
f(t) = \sum_{k=-\infty}^{\infty} g \left( \frac{k\pi}{\Omega} \right) \frac{1}{\sqrt{2\pi}} \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} \frac{1}{A(\omega)} e^{i\omega \left[ t - \frac{\pi k}{\Omega} \right]} d\omega.
\]
The final portion of the RHS, starting with the \( 1/\sqrt{2\pi} \) factor, looks like some kind of inverse Fourier transform. Now consider the following change of variable,
\[
\nu = \frac{\omega}{\Omega}, \quad \omega = \Omega \nu, \quad d\omega = \Omega \, d\nu.
\]
227
We rewrite this final portion of the RHS as

\[
\frac{1}{\sqrt{2\pi}} \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} \frac{1}{A(\omega)} e^{i\omega[t-\frac{\pi k}{\Omega}]} d\omega = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} P(\nu) e^{i\nu[\Omega t - \pi k]} d\nu, \tag{51}
\]

where

\[
P(\nu) = \frac{1}{2A(\Omega \nu)}. \tag{52}
\]

The final integral is the inverse Fourier transform of \(P(\nu)\) and we may write

\[
f(t) = \sum_{k=-\infty}^{\infty} g(\frac{k\pi}{\Omega}) p(\Omega t - \pi k), \tag{53}
\]

which has the same form as the cardinal series for the classical Sampling Theorem. We have arrived at our generalized cardinal series which produces \(f(t)\) from its samples \(g(kT)\).

**Exercise:** Show that in the case \(a(t) = \delta(t)\), the Dirac delta function, for which \(A(\omega) = \frac{1}{\sqrt{2\pi}}\), the above generalized cardinal series reduces to the classical cardinal series involving shifted sinc functions.

**The Sampling Theorem - final remarks**

Finally, a few comments about sampling in everyday life. Recall our earlier comment that speech signals are essentially bandlimited at 4200 Hz. For this reason, a modest sampling rate of 8 Hz, typical of a computer sound card, is adequate for speech transmission, e.g., telephone, but certainly not for music!

The current standard for CD recording is 44,100 16-bit samples per second per channel, which is the Nyquist sampling rate for a bandlimit of 22,050 Hz, which is above the normal range of human hearing. This corresponds to a sample rate of 88.2 kB/sec (kilobytes/sec) per channel or

\[
\frac{44100 \ samples}{sec} \cdot \frac{3600 \ sec}{hour} \cdot \frac{2 \ bytes}{sample} = \frac{318 \ Mbytes}{hour}. \tag{54}
\]

Thus, a one-hour recording would require 636 Mb of storage, just under the capacity of a CD. This estimate, of course, neglects any savings in storage that would be made possible by compression, which is beyond the scope of the course, at least at this time.

We also mention that the above sampling for CDs is performed on quantized sample values, which corresponds to a finite and prescribed accuracy. 16 bits implies \(2^{16} = 65536\) possible values. Actually, if one bit is used for the sign of the sample, then only 65535 values are possible.
Lecture 20

The Uncertainty Principle

Relevant section from Boggess and Narcowich: 2.5, pp. 120-125.

Recall, from a couple of lectures ago, that the Fourier transform of a Gaussian function (in $t$-space) is a Gaussian function (in $\omega$-space). The particular “Fourier pair” of functions examined were $f_\sigma(t)$ and $F_\sigma(\omega) = \mathcal{F}[f]$, where

$$f_\sigma(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}, \quad F_\sigma(\omega) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}\omega^2}. \tag{55}$$

The variance of $f_\sigma(t)$ about its mean, zero, is $\sigma^2$ and the variance of $F_\sigma(\omega)$ about its mean, zero, is $1/\sigma^2$. For $\sigma$ very small, the graph of $f_\sigma(t)$ is concentrated about zero, and the graph of $F_\sigma(\omega)$ is spread-out. In the limit $\sigma \to 0$, $f_\sigma(t) \to \delta(t)$, the Dirac delta function, and $F_\sigma(\omega) \to 1/\sqrt{2\pi}$ a constant. This pair represents the ultimate in concentration vs. spreading.

Another example of this “concentration vs. spreading” between Fourier transforms is the “boxcar-sinc” pair of functions (details left as an exercise):

$$f_a(t) = \begin{cases} \frac{1}{2a}, & -a \leq t \leq a \\ 0, & \text{otherwise} \end{cases}, \quad F_a(\omega) = \frac{1}{\sqrt{2\pi}} \frac{\sin(a\omega)}{a\omega}. \tag{56}$$

As $a$ decreases to zero, the function $f_a(t)$ becomes more peaked around zero. The sinc function spreads out, since the distance between its nodes is $\pi/a$. And once again, as $a \to 0$, $f_a(t) \to \delta(t)$ and $F_a(\omega) \to 1/\sqrt{2\pi}$.

As we mentioned earlier, these are examples of the complementarity between the time and frequency spaces, which is described mathematically by the so-called “uncertainty theorem” or “uncertainty principle.” To develop this result, we need to have a measure of the “spread” or dispersion of a function $f(t)$. The dispersion is a generalization of the idea of the variance of a (positive) function about its mean value.

In what follows, we assume that $f(t) \in L(\mathbb{R})$. The dispersion of $f$ about a point $a \in \mathbb{R}$ is given by

$$\Delta_a f = \frac{\int_{-\infty}^{\infty} (t-a)^2 |f(t)|^2 \, dt}{\int_{-\infty}^{\infty} |f(t)|^2 \, dt} = \frac{1}{\|f\|^2} \int_{-\infty}^{\infty} (t-a)^2 |f(t)|^2 \, dt. \tag{57}$$
In this definition, the function \( f(t) \) is used to define a nonnegative distribution function \( g(t) = |f(t)|^2 \). In fact, the function

\[
\rho(t) = \frac{1}{\|f\|^2} |f(t)|^2, \quad t \in \mathbb{R},
\]

may be viewed as a probability distribution function, since

\[
\int_{-\infty}^{\infty} \rho(t) \, dt = 1.
\]

Aside: This is precisely the situation in quantum mechanics, where \( f(t) \) is the complex-valued \textit{wave-function} of a one-dimensional quantum mechanical “system,” e.g., particle. According to the Copenhagen interpretation of quantum mechanics, the probability of finding the particle in an interval of width \( dx \) and centered at \( x \) is \( \rho(x) \, dx = \|\psi(x)\|^2 \, dx \). The probability of finding the particle somewhere, i.e., over the entire real line \( \mathbb{R} \), is one.

If we view \( \rho(t) \) above as a probability distribution function over \( \mathbb{R} \), then the average or mean value of a function \( c(t) \) over \( \mathbb{R} \) is given by

\[
\bar{c} = \int_{-\infty}^{\infty} c(t) \rho(t) \, dt.
\]

(If \( t \) is considered to be a random real variable with probability distribution \( \rho(t) \), then the above mean value is the expectation value of \( c(t) \), i.e., \( E[c(t)] \).) In particular, the average value, or mean, of the variable \( t \) according to this distribution is

\[
\bar{t} = \int_{-\infty}^{\infty} t \rho(t) \, dt.
\]

Furthermore, recall that the variance of the function \( c(t) \) about its mean value \( \bar{c} \) is defined by

\[
\sigma^2(c) = \int_{-\infty}^{\infty} (c(t) - \bar{c})^2 \rho(t) \, dt.
\]

In particular, the variance of the variable \( t \) according to this distribution is

\[
\sigma^2(t) = \int_{-\infty}^{\infty} (t - \bar{t})^2 \rho(t) \, dt.
\]

The dispersion \( \Delta_a f \) in Eq. (57) is a generalization of the variance of the function \( |f(t)|^2 \) – it measures the variance with respect to the point \( a \in \mathbb{R} \). Graphically, it measures the local width of the graph of \( f(t)^2 \) about \( a \).
Example: Consider the Gaussian function examined earlier in this course,

\[ f_\sigma(t) = e^{-\frac{t^2}{2\sigma^2}}, \quad \sigma > 0. \] (64)

(We can ignore the multiplicative constant since it will be removed by the ratio in the formula for the dispersion.)

The dispersion \( \Delta_i f_\sigma \) in Eq. (57) is given by

\[
\Delta_i f_\sigma = \frac{\int_{-\infty}^{\infty} (t-a)^2 |f_\sigma(t)|^2 \, dt}{\int_{-\infty}^{\infty} |f_\sigma(t)|^2 \, dt} = \frac{\int_{-\infty}^{\infty} (t-a)^2 e^{-t^2/\sigma^2} \, dt}{\int_{-\infty}^{\infty} e^{-t^2/\sigma^2} \, dt} \]

\[ = 1 + \frac{1}{2} \sigma^2 + a^2. \] (65)

The details involving the computation of the integrals are left as an exercise. A few comments regarding this result:

1. For a given \( \sigma > 0 \), the dispersion \( \Delta_i f_\sigma \) increases as \( |a| \) increases. This is because the Gaussian function \( f_\sigma^2(t) \) flattens out as \( |a| \) increases. At \( a = 0 \), the Gaussian function is most “concentrated”.

2. For a given \( a \), in particular \( a = 0 \), the dispersion \( \Delta_i f_\sigma \) decreases as \( \sigma \) decreases, as expected.

3. Based on the earlier comments regarding the interpretation of the dispersion as a variance, one might have expected the dispersion at \( a = 0 \) to be the well known \( \sigma^2 \) associated with the Gaussian function of probability. Yes, indeed, \( \sigma^2 \) is the variance associated with the normalized distribution defined by \( f_\sigma(t) \). But, as mentioned above, dispersion is the variance associated with the distribution defined by \( |f(t)|^2 \) – in this case \( f_\sigma^2(t) \). As such, the exponential in the integral becomes \( e^{-t^2/\sigma^2} \), the variance of which is \( \frac{1}{2} \sigma^2 \).

We shall also be considering the dispersion of \( F(\omega) \), the Fourier transform of \( f(t) \), about a frequency value \( \alpha \in \mathbb{R} \):

\[
\Delta_\alpha F = \frac{1}{\|F\|^2} \int_{-\infty}^{\infty} (\omega - \alpha)^2 |F(\omega)|^2 \, d\omega. \] (66)

Note that even though \( f(t) \) is assumed to be real-valued, \( F(\omega) \) is, in general, complex-valued. Also note that by the Plancherel Theorem

\[ \|f\|^2 = \|F\|^2. \] (67)
Here is an interpretation of the dispersion $\Delta_{\alpha} F$ of the Fourier transform $F(\omega)$ of a function $f(t)$:

If the dispersion of $\Delta_{\omega} F$ is \( \begin{cases} \text{small} \\ \text{large} \end{cases} \), then the frequency range of $f(t)$ is \( \begin{cases} \text{concentrated} \\ \text{diffuse} \end{cases} \) near $\omega = \alpha$.

**Example:** Recall that the Fourier transform of the Gaussian function $f_{\sigma}(t)$ examined above is given by

$$F_{\sigma}(\omega) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \sigma^2 \omega^2}.$$  \hfill (68)

The dispersion of $F_{\sigma}(\omega)$ at any frequency value $\alpha \in \mathbb{R}$ is (Exercise)

$$\Delta_{\alpha} F_{\sigma} = \frac{1}{2\sigma^2} + \alpha^2.$$  \hfill (69)

Some comments:

1. Once again, since $F_{\sigma}(\omega)$ is a Gaussian centered at $\omega = 0$, its dispersion increases as $|\alpha|$ increases.

2. For a fixed $\alpha$, in particular $\alpha = 0$, the dispersion of $F_{\sigma}(\omega)$ increases as $\sigma$ decreases. This is a consequence of the increased concentration of $F_{\sigma}(\omega)$ – and broadening of $f_{\sigma}(t)$ – as $\sigma$ is increased.

We now state the Uncertainty Principle for a function $f$ and its Fourier transform $F$:

$$(\Delta_{a} f)(\Delta_{\alpha} F) \geq \frac{1}{4},$$

for any $a \in \mathbb{R}$ and $\alpha \in \mathbb{R}$.

**Sketch of Proof:** (This is a variation of a proof given by the mathematician/mathematical physicist H. Weyl in 1936.) The LHS of the above inequality is given by

$$(\Delta_{a} f)(\Delta_{\alpha} F) = \frac{1}{\|f\|^4} \left[ \int_{-\infty}^{\infty} |(t - a)f(t)|^2 \, dt \right] \left[ \int_{-\infty}^{\infty} |(\omega - \alpha)F(\omega)|^2 \, d\omega \right].$$  \hfill (71)

We shall work on the second integral on the RHS. First note that $i\omega F(\omega)$ is the Fourier transform of $f'(t)$ (derivative property of FTs). We’ll write this as follows,

$$\mathcal{F}[f'(t)] = i\omega F(\omega).$$  \hfill (72)
By linearity,
\[ \mathcal{F}[i\alpha f(t)] = i\alpha F(\omega). \] (73)

Subtracting the second result from the first yields,
\[ \mathcal{F}[f'(t) - i\alpha f(t)] = i(\omega - \alpha)F(\omega). \] (74)

We recall the Plancherel Theorem again, written as follows:
\[ \|g\|^2 = \|\mathcal{F}[g]\|^2, \text{ for any } g \in L^2(\mathbb{R}). \] (75)

From Eq. (74), we therefore have
\[ \|f' - i\alpha f\|^2 = \|\omega - \alpha\|^2. \] (76)

We now apply this result to the second integral in Eq. (71):
\[ (\Delta_a f)(\Delta_{\alpha} F) = \frac{1}{\|f\|^4} \left[ \int_{-\infty}^{\infty} |(t-a)f(t)|^2 \, dt \right] \left[ \int_{-\infty}^{\infty} |f'(t) - i\alpha f(t)|^2 \, dt \right]. \] (77)

Now recall the Cauchy-Schwarz inequality for the inner product in a Hilbert space \( H \):
\[ |\langle u, v \rangle| \leq \|u\| \|v\|, \quad u, v \in H. \] (78)

For \( H = L^2(\mathbb{R}) \), the Cauchy-Schwarz inequality becomes
\[ \left| \int_{-\infty}^{\infty} u(t)v(t) \, dt \right| \leq \left[ \int_{-\infty}^{\infty} |u(t)|^2 \, dt \right]^{1/2} \left[ \int_{-\infty}^{\infty} |v(t)|^2 \, dt \right]^{1/2}. \] (79)

We square both sides, let
\[ u(t) = (t-a)f(t), \quad v(t) = f'(t) - i\alpha f(t), \] (80)

and substitute into Eq. (77) to give
\[ (\Delta_a f)(\Delta_{\alpha} F) \geq \frac{1}{\|f\|^4} \left| \int_{-\infty}^{\infty} [(t-a)f(t)] \left[ f'(t) + i\alpha f(t) \right] \, dt \right|^2. \] (81)

Now rewrite the integral on the RHS as follows,
\[ \int_{-\infty}^{\infty} [(t-a)f(t)] \left[ f'(t) + i\alpha f(t) \right] \, dt = \int_{-\infty}^{\infty} (t-a)f(t)f'(t) \, dt + i\alpha \int_{-\infty}^{\infty} (t-a)f(t)^2 \, dt \]
\[ = A + iB, \] (82)
where \( A, B \in \mathbb{R} \), since we have assumed that \( f(t) \) is real-valued. Integral \( A \) may be written as follows,

\[
A = \int_{-\infty}^{\infty} (t-a) \frac{d}{dt} \left[ \frac{1}{2} f(t)^2 \right] dt
\]

\[
= \left[ \frac{1}{2} (t-a) f(t)^2 \right]_{-\infty}^{\infty} - \frac{1}{2} \int_{-\infty}^{\infty} f(t)^2 dt
\]

\[
= -\frac{1}{2} \|f\|^2,
\]

where we have used the fact that \( f \in L^2(\mathbb{R}) \), which implies that \( f(t)^2 \) must decay to zero at least as quickly as \( t^{-1-\epsilon} \) for \( \epsilon > 0 \). We now employ the following result,

\[
|A + iB|^2 = A^2 + B^2 \geq A^2 = \frac{1}{4} \|f\|^4,
\]

in Eq. (82) to arrive at the desired result,

\[
(\Delta_a f)(\Delta_\alpha F) \geq \frac{1}{\|f\|^2} \left( \frac{1}{4} \|f\|^4 \right) = \frac{1}{4}.
\]

(85)

Note that this result is valid for all times \( a \in \mathbb{R} \) and frequencies \( \alpha \in \mathbb{R} \), independent of each other. A low dispersion/high concentration at a value of \( a \), i.e., \( \Delta_a = \epsilon \) implies that the dispersion in the frequency domain is \( \Delta_\alpha \geq \frac{1}{4\epsilon} \) for all \( \alpha \). This indicates the global nature of the Fourier transform.

**Example:** We return to the Gaussian function \( f_\sigma(t) \) and its Fourier transform \( F_\sigma(\omega) \) examined earlier. It is not too difficult to establish that the Uncertainty Principle is satisfied here, i.e.,

\[
(\Delta_a f_\sigma)(\Delta_\alpha F_\sigma) = \left[ \frac{1}{2} \sigma^2 + a^2 \right] \left[ \frac{1}{2\sigma^2} + \alpha^2 \right] \geq \frac{1}{4}, \quad a, \alpha \in \mathbb{R}.
\]

(86)

Details are left as an exercise. Note that equality is achieved at \( a = \alpha = 0 \), i.e, at the peaks of the Gaussian functions.

**A function and its Fourier transform cannot both have finite support**

It is often stated that the uncertainty principle implies that a function \( f(t) \) and its Fourier transform \( F(\omega) \) cannot both be supported on finite intervals. (Recall that for a function \( f(t) \) to have finite support means that there exists a finite interval \([a,b]\) on which \( f(t) \) assumes nonzero values and \( f(t) = 0 \) for all \( t \notin [a,b] \).) The following theorem states this property more concretely.

**Theorem:** Suppose that \( f(t) \) has finite support, i.e., there exists an interval \([a,b]\), with \( a \) and \( b \) finite, such that \( f(t) \) is not identically zero on \([a,b]\) but \( f(t) = 0 \) for all \( t \notin [a,b] \). Then there exists no closed
interval \([c, d] \subset \mathbb{R}, c < d\), on which \(F(\omega)\), the Fourier transform of \(f(t)\), is identically zero. The other way also holds, i.e., if \(F(\omega)\) has finite support, then \(f(t)\) cannot be identically zero on a nontrivial closed interval.

**Proof:** We prove the second part since the first can be derived from the second by applying the Fourier transform. Without loss of generality, we assume that \(F(\omega)\) is nonzero only inside the interval \([-\Omega, \Omega]\), in which case we may write

\[
 f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} F(\omega) e^{i\omega t} d\omega \tag{87}
\]

from the definition of the inverse Fourier transform.

Now assume that \(f(t)\) is not identically zero on the real line \(\mathbb{R}\) but that \(f(t) = 0\) for all \(t \in [c, d]\) with \(c \neq d\). It follows that at any \(t_0 \in (c, d)\),

\[
 f^{(n)}(t_0) = 0, \quad n = 0, 1, 2, \ldots \tag{88}
\]

From Eq. (87), differentiating \(n\) times with respect to \(t\) and then setting \(t = t_0\), we obtain

\[
 f^{(n)}(t_0) = \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} F(\omega) (i\omega)^n e^{i\omega t_0} d\omega = 0, \quad n = 0, 1, 2, \ldots \tag{89}
\]

(The fact that the interval of integration \([-\Omega, \Omega]\) is finite ensures the existence of these integrals for all \(n \geq 0\), regardless of the nature of \(F(\omega)\).)

We return to the inverse FT in Eq. (87) and rewrite it slightly as follows,

\[
 f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} F(\omega) e^{i\omega(t-t_0)} e^{i\omega t_0} d\omega \tag{90}
\]

Now expand the function \(e^{i\omega(t-t_0)}\) in terms of its Taylor series at \(t = t_0\), i.e.,

\[
 e^{i\omega(t-t_0)} = \sum_{n=0}^{\infty} \frac{(i(t-t_0))^n}{n!}, \quad t \in \mathbb{R}, \tag{91}
\]

and substitute this result into Eq. (90) to obtain

\[
 f(t) = \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{(i(t-t_0))^n}{n!} \int_{-\Omega}^{\Omega} F(\omega) \omega^n e^{i\omega t_0} d\omega, \quad t \in \mathbb{R}. \tag{92}
\]

But from Eq. (89), all of the integrals in the above expression are zero. Therefore we have

\[
 f(t) = 0, \quad t \in \mathbb{R}, \tag{93}
\]

235
which contradicts the original assumption that \( f(t) \) is not identically zero on the real line \( \mathbb{R} \). This completes the proof of the second statement.

**Comment:** A consequence of this theorem is that if one of \( f \) or \( F \) has finite support, the other cannot have finite support. (If the “other” did, then it would necessarily be zero over an interval which, according to the above theorem, cannot happen.)