17
Fourier Analysis of Linear and Nonlinear Signals

In this chapter we examine Fourier series *and* Fourier integrals (or transforms). *This is the traditional tool for decomposing both periodic and nonperiodic motions, respectively, into an infinite number of harmonic functions. Because it represents a function as a series of sines and cosines of a time variable, a Fourier series always generates a periodic function. This clearly is desirable if the function we wish to approximate is periodic, in which case we construct the series to have the correct period. A nonperiodic function can frequently be well approximated by a Fourier series* over some limited time*, but the Fourier series will eventually show its periodicity for values outside this range. In this latter case of limited range, the Fourier integral is more appropriate.*

17.1
Harmonics of Nonlinear Oscillations (Problem 1)

Consider a particle oscillating in a nonharmonic potential. This could be the nonharmonic oscillator (15.4),

$$V(x) = \frac{1}{p}k|x|^p \tag{17.1}$$

for $p \neq 2$, the perturbed harmonic oscillator (15.2),

$$V(x) = \tfrac{1}{2}kx^2\left(1 - \tfrac{2}{3}\alpha x\right) \tag{17.2}$$

or the realistic pendulum of Section 19.1. While free oscillations in these potentials are always periodic, they are not truly sinusoidal. Your **problem** is to take the solution of one of these nonlinear oscillators and relate it to the solution

$$x(t) = A_0 \sin(\omega t + \phi_0) \tag{17.3}$$

of the linear, harmonic oscillator. (In your future study of chaos, you will want to extend this analysis to the response of a damped, oscillating system driven by an external force.) If your oscillator is sufficiently nonlinear to behave like

Computationyal Physics. Problem Solving with Computers (2nd edn).
Rubin H. Landau, Manuel José Páez, Cristian C. Bordeianu

ISBN: 978-3-527-40626-5

the sawtooth function (Fig. 17.1, left), then the Fourier spectrum you obtain should be similar to that shown on the right of Fig. 17.1.

In general, when we want to undertake such a spectral analysis, we want to analyze the steady-state behavior of a system. This means that the initial transient behavior has had a chance to die off. Just what is the initial transient is easy to identify for linear systems, but may be less so for nonlinear systems in which the "steady state" jumps among a number of configurations.

17.2 Fourier Analysis (Math)

Nonlinear oscillations are interesting in part because they hardly ever are studied in traditional courses. This is true even though the linear term is just a first approximation to a naturally oscillating system. If the force on a particle is always toward its equilibrium position (a restoring force), then the resulting motion will be *periodic*, but not necessarily *harmonic*. A good example is the motion in a highly anharmonic well $p \approx 10$, which produces an $x(t)$ looking like a series of pyramids; this is periodic but not harmonic.

On a computer, the distinction between a Fourier integral and a Fourier series is less clear because the integral is approximated as a finite series. We will illustrate both methods by analyzing anharmonic oscillations with the series and by analyzing the charge density of elementary systems with the integral.

In a sense, our approach is the inverse of the traditional one in which the *fundamental* oscillation is determined analytically, and the higher frequency *overtones* are determined in perturbation theory [31]. We start with the full (numerical) periodic solution and then decompose it into what may be called *harmonics.* When we speak of fundamentals, overtones, and harmonics, we speak of solutions to the *boundary-value problem*, for example, of waves on a plucked violin string. In this latter case, and when given the correct conditions (enough musical skill), it is possible to excite individual harmonics or sums of them within the series

$$y(t) = b_0 \sin \omega_0 t + b_1 \sin \frac{n\omega_0 t}{m} + \cdots \tag{17.4}$$

The anharmonic oscillator vibrates with a single frequency (which may change with changing amplitude) but not a sinusoidal waveform. Expanding the anharmonic vibration as a Fourier series does not imply that the individual harmonics can be "played."

You may recall from classical mechanics that the most general solution for some vibrating physical system can be expressed as the sum of the *normal modes* of that system. These expansions are possible because we have *linear*

operators and, subsequently, the *principle of superposition*: *If $x_1(t)$ and $x_2(t)$ are solutions of some linear equation, then $\alpha_1 x_1(t) + \alpha_2 x_2(t)$ is also a solution.*

The principle of linear superposition does not hold when we solve nonlinear problems. Nevertheless, it is always possible to expand a *periodic* solution of a *nonlinear* problem in terms of trigonometric functions that have frequencies that are integer multiples of the true frequency of the nonlinear oscillator. This is a consequence of *Fourier's theorem* being applicable to any single-valued, periodic function with only a finite number of discontinuities. We assume we know the period T, that is, that

$$y(t+T) = y(t) \tag{17.5}$$

This tells us the "true" frequency ω:[1]

$$\omega \equiv \omega_1 = \frac{2\pi}{T} \tag{17.6}$$

Any such periodic function can be expanded as a series of harmonic functions with frequencies that are multiples of the true frequency:

$$y(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos n\omega t + b_n \sin n\omega t) \tag{17.7}$$

This equation represents a signal as the simultaneous sum of pure tones of frequency $n\omega$. The coefficients a_n and b_n are a measure of the amount of $\cos n\omega t$ and $\sin n\omega t$ present in $y(t)$, specifically, the intensity or power at each frequency proportional to $a_n^2 + b_n^2$.

The Fourier series (17.7) is a "best fit" in the least-squares sense of Chap. 8 because it minimizes $\sum_i [y(t_i) - y_i]^2$. This means that the series converges to the average behavior of the function, but misses the function at discontinuities (at which points it converges to the mean) or at sharp corners (where it overshoots). A general function $y(t)$ may contain an infinite number of Fourier components, although a good approximation is usually possible with a small number of harmonics.

The coefficients a_n and b_n are determined by the standard techniques for function expansion. To find them, you multiply both sides of (17.7) by $\cos n\omega t$ or $\sin n\omega t$, integrate over one period, and project out a single a_n or b_n:

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \frac{2}{T} \int_0^T dt \begin{pmatrix} \cos n\omega t \\ \sin n\omega t \end{pmatrix} y(t) \qquad \omega \overset{\text{def}}{=} \frac{2\pi}{T} \tag{17.8}$$

[1] We remind the reader that every periodic system by definition has a period T and consequently a "true" frequency ω. Nonetheless, this does not imply that the system behaves like $\sin \omega t$. Only harmonic oscillators do that.

As seen in the b_n coefficients (Fig. 17.1, right), these coefficients usually decrease in magnitude as the frequency increases, and can occur with negative sign, the negative sign indicating opposite phase.

Awareness of the *symmetry* of the function $y(t)$ may eliminate the need to evaluate all the expansion coefficients. For example

- a_0 is twice the average value of y.

$$a_0 = 2\,\langle y(t)\rangle \tag{17.9}$$

- For an *odd function*, that is, one for which $y(-t) = -y(t)$, all the coefficients $a_n \equiv 0$ and only half the integration range is needed to determine b_n:

$$b_n = \frac{4}{T}\int_0^{T/2} dt\; y(t)\sin n\omega t \tag{17.10}$$

  However, if there is no input signal for $t < 0$, we do not have a truly odd function, and so small values of a_n may occur.

- For an *even function*, that is, one for which $y(-t) = y(t)$, the coefficient $b_n \equiv 0$ and only half the integration range is needed to determine a_n:

$$a_n = \frac{4}{T}\int_0^{T/2} dt\; y(t)\cos n\omega t \tag{17.11}$$

17.2.1 Example 1: Sawtooth Function

The sawtooth function (Fig. 17.1) is described mathematically as

$$y(t) = \begin{cases} \frac{t}{T/2} & \text{for } 0 \le t \le \frac{T}{2} \\ \frac{t-T}{T/2} & \text{for } \frac{T}{2} \le t \le T \end{cases} \tag{17.12}$$

It is clearly periodic, nonharmonic, and discontinuous. Yet it is also an odd function, and so can be represented more simply by shifting the period over to the left:

$$y(t) = \frac{t}{T/2} \qquad -\frac{T}{2} \le t \le \frac{T}{2} \tag{17.13}$$

While it is relatively easy to reproduce the general shape of this function with only a few terms of the Fourier series, many components are needed to reproduce the discontinuities at the sharp corners. Because the function is odd, the

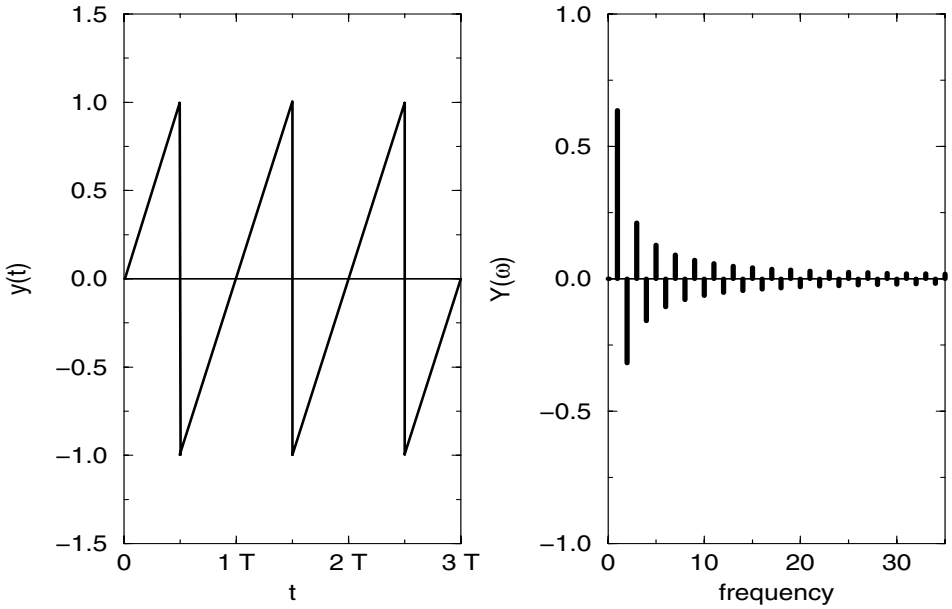

Fig. 17.1 *Left:* A sawtooth function in time. *Right:* The Fourier spectrum of frequencies in natural units contained in this sawtooth function.

Fourier series is a sine series, and (17.8) determines the values

$$b_n = \frac{2}{T}\int_{-T/2}^{+T/2} dt\ \sin n\omega t \frac{t}{(T/2)} = \frac{\omega}{\pi}\int_{-\pi/\omega}^{+\pi/\omega} dt\ \sin n\omega t \frac{\omega t}{\pi}$$

$$= \frac{2}{n\pi}(-1)^{n+1} \tag{17.14}$$

$$\Rightarrow \quad y(t) = \frac{2}{\pi}\left[\sin\omega t - \tfrac{1}{2}\sin 2\omega t + \tfrac{1}{3}\sin 3\omega t - \cdots\right] \tag{17.15}$$

Note that the terms in this series alternate in sign, which means that successive frequency components are out of phase with each other, yet that they all add together to give a straight line for $-T/2 \le t \le T/2$, with finite discontinuities as the interval repeats.

17.2.2
Example 2: Half-Wave Function

The half-wave function is

$$y(t) = \begin{cases} \sin\omega t & \text{for } 0 < t < T/2 \\ 0 & \text{for } T/2 < t < T \end{cases} \tag{17.16}$$

It is periodic, nonharmonic (the upper half of a sine wave), continuous, but with discontinuous derivatives. Because it lacks the sharp "corners" of the sawtooth, it is easier to reproduce with a finite Fourier series. Equations (17.8) determine

$$a_n = \begin{cases} \frac{-2}{\pi(n^2-1)}, & n \text{ even or } 0, \\ 0, & n \text{ odd}, \end{cases} \qquad b_n = \begin{cases} \frac{1}{2}, & n = 1, \\ 0, & n \neq 1, \end{cases}$$

$$\Rightarrow \quad y(t) = \tfrac{1}{2}\sin\omega t + \frac{1}{\pi} - \frac{2}{3\pi}\cos 2\omega t - \frac{2}{15\pi}\cos 4\omega t + \cdots . \tag{17.17}$$

17.3 Summation of Fourier Series(Exercise)

1. **Sawtooth function:** Sum the Fourier series for the *sawtooth function* up to order $n = 2, 4, 10, 20$, and plot the results over two periods.

   (a) Check that in each case the series gives the mean value of the function *at* the points of discontinuity.

   (b) Check that in each case the series *overshoots* by about 9% the value of the function on either side of the discontinuity (the *Gibbs phenomenon*).

2. **Half-wave function:** Sum the Fourier series for the *half-wave function* up to order $n = 2, 4, 10$, and plot the results over two periods. (The series converges quite well, doesn't it?)

17.4 Fourier Transforms (Theory)

Although a Fourier *series* is the right tool for approximating or analyzing periodic functions, the Fourier *transform* or *integral* is the right tool for nonperiodic functions. We transform the series formalism to the integral formalism by imagining a system described by a continuum of "fundamental" frequencies. We therefore deal with *wave packets* containing continuous rather than discrete frequencies.[2] While the difference between series and transform methods may appear clear mathematically, if we transform a function known only over a fi-

[2] We follow convention and consider time t as the function's variable and frequency ω as the transform's variable. Nonetheless, these can be reversed or other variables such as position x and wave vector k may also be used.

nite length of times, or approximate the Fourier integral as a finite sum, then the Fourier transform and the Fourier series become equivalent.

As an analogy to (17.7), we now imagine our function or signal $y(t)$ expressed in terms of a continuous series of harmonics

$$y(t) = \int_{-\infty}^{+\infty} d\omega Y(\omega) \frac{e^{i\omega t}}{\sqrt{2\pi}} \tag{17.18}$$

where for compactness we use a complex exponential function.[3] Here the expansion amplitude $Y(\omega)$ is analogous to (a_n, b_n) and is called the *Fourier transform* of $y(t)$.

A plot of the squared modulus $|Y(\omega)|^2$ versus ω is called the *power spectrum*. Actually, (17.18) is more properly called the *inverse transform* because it converts $Y(\omega)$ to $y(t)$. The *Fourier transform* converts $y(t)$ to $Y(\omega)$:

$$Y(\omega) = \int_{-\infty}^{+\infty} dt \frac{e^{-i\omega t}}{\sqrt{2\pi}} y(t) \tag{17.19}$$

You will note in (17.18) and (17.19) that only the relative sign in the exponential matters. In addition, you should note that we have chosen a symmetric $1/\sqrt{2\pi}$ normalization factor, which is common in quantum mechanics [30] but differs from that used in engineering.

If $y(t)$ represents the response of some system as a function of time, $Y(\omega)$ is a *spectral function* that measures the amount of frequency ω making up this response. While usually we think of measuring $y(t)$ in the laboratory and numerically transforming it to obtain $Y(\omega)$, some experiments may well measure $Y(\omega)$ directly [in which case a transform is needed to obtain $y(t)$]. Clearly, the mathematics is symmetric even if the real world is not.

If the Fourier transform and its inverse are consistent with each other, we should be able to substitute (17.18) into (17.19) and obtain an identity:

$$\begin{aligned} Y(\omega) &= \int_{-\infty}^{+\infty} dt \frac{e^{-i\omega t}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} d\omega' \frac{e^{i\omega' t}}{\sqrt{2\pi}} Y(\omega') \\ &= \int_{-\infty}^{+\infty} d\omega' \left\{ \int_{-\infty}^{+\infty} dt \frac{e^{i(\omega'-\omega)t}}{2\pi} \right\} Y(\omega') \end{aligned}$$

For this to be an identity the term in braces must be the *Dirac delta function*:

$$\int_{-\infty}^{+\infty} dt\, e^{i(\omega'-\omega)t} = 2\pi\delta(\omega' - \omega) \tag{17.20}$$

[3] Recall the principle of linear superposition and that $\exp(i\omega t) = \cos\omega t + i \sin\omega t$. This means that the real part of y gives the cosine series and the imaginary part the sine series.

While the delta function is one of the most common and useful functions in theoretical physics, it is not well behaved in a mathematical sense and is terribly misbehaved in a computational sense. While it is possible to create numerical approximations to $\delta(\omega' - \omega)$, they may well be borderline pathological. It is probably better for you to do the delta function part of an integration analytically and leave the nonsingular leftovers to the computer.

17.5 Discrete Fourier Transform Algorithm (DFT)

If $y(t)$ or $Y(\omega)$ is known analytically, the integral (17.18) or (17.19) can be evaluated analytically or numerically using the integration techniques studied earlier (particularly Gaussian quadrature). Likewise, if a table of N values for $y(t)$ is known, interpolations within the table can be used to evaluate the integral.

Here we will consider a technique for directly Fourier transforming functions that are known only for a finite number N of times t (for instance, as sampled in an experiment). This *discrete Fourier transform* (DFT) is an "approximate" procedure because the integrals are evaluated numerically.[4] By sampling a nonperiodic function at N times, we can determine N values of the Fourier transform of this function [N independent $y(t)$ values can produce N independent $Y(\omega)$ values]. We can then use those values of the transform to approximate the original function at any value of time. In this way the DFT can also be thought of as a technique for interpolating and extrapolating data.

Assume that the function $y(t)$ we wish to transform is measured or sampled at a discrete number $N + 1$ of times (N time intervals)

$$y_k \stackrel{\text{def}}{=} y(t_k) \qquad k = 0, 1, 2, \ldots, N \tag{17.21}$$

Assume that these times are evenly spaced with a time step h:

$$t_k = kh \qquad h = \Delta t \tag{17.22}$$

In other words, we measure the signal $y(t)$ once every h seconds during a total time interval of T and with a *sampling rate* s, with the two inversely related:

$$T \stackrel{\text{def}}{=} Nh \qquad s = \frac{N}{T} = \frac{1}{h} \tag{17.23}$$

Regardless of the true periodicity of the function being sampled, when we choose a total time T over which to sample the function, the mathematics produces a $y(t)$, which is periodic with period T. To make this self-consistent and

[4] More discussion can be found in the book [32] devoted to just this topic.

ensure that there are only N independent function values used in the transform, we require the first and the last y values to be the same:

$$y(t+T) = y(t) \qquad \Rightarrow \qquad y_0 = y_N \tag{17.24}$$

If we are analyzing a truly periodic function, then the first N points should all be within one period to guarantee their independence. Unless we make further assumptions, these N independent input data $y(t_k)$ can determine no more than N independent output Fourier components $Y(\omega_k)$.

The time interval T (which should be the period for periodic functions) is the largest time over which we consider variation of $y(t)$. Consequently, it determines the lowest frequency $\omega_1 = 2\pi/T$ contained in our Fourier representation of $y(t)$ (unless you want to be picky and argue that there may also be an $\omega = 0$ or "DC" component).

While we will be able to compute N independent values $Y(\omega_n)$ for $n = 1, N$, the values for the frequencies ω_n are determined by the number of samples taken and the total sampling time T. To make the connection with Fourier series, we choose the frequencies as

$$\omega_n = n\omega_1 = n\frac{2\pi}{Nh} \qquad n = 0, 1, \ldots, N \tag{17.25}$$

Here the $n = 0$ value corresponds to the zero frequency or DC component, $\omega_0 = 0$. We now see clearly that by limiting the time interval over which we sample the input function, we are making an approximation that limits the maximum frequency of the Fourier components we can compute.

Note that (17.25) indicates that the larger we make the time $T = Nh$ over which we sample the function, the smaller will be the frequency steps. Accordingly, if you want a smooth frequency spectrum, you need small steps in frequency, and, such being the case, large T. While the best approach would be to measure the input signal for longer times, in practise a measured signal $y(t)$ is often "padded" with zeros for $t > T$ in order to produce a smoother spectrum. Effectively, this is building into the analysis the experimentalist's belief that the signal does not repeat.

The discrete Fourier transform results from (1) evaluating the integral in (17.19), not from $-\infty$ to $+\infty$, but rather from the times 0 to T over which the

signal is measured, and from (2) using the trapezoid rule for the integration[5]

$$Y(\omega_n) \quad \stackrel{\text{def}}{=} \quad \int_{-\infty}^{+\infty} dt\, \frac{e^{-i\omega_n t}}{\sqrt{2\pi}} y(t) \;\simeq\; \int_0^T dt\, \frac{e^{-i\omega_n t}}{\sqrt{2\pi}} y(t) \tag{17.26}$$

$$\approx \sum_{k=1}^{N} h\, y(t_k) \frac{e^{-i\omega_n t_k}}{\sqrt{2\pi}} \;=\; h \sum_{k=1}^{N} y_k \frac{e^{-2\pi i k n/N}}{\sqrt{2\pi}} \tag{17.27}$$

To keep the final notation more symmetric, the step size h is factored from the transform Y and a discrete value Y_n is defined by

$$Y_n \quad \stackrel{\text{def}}{=} \quad \frac{1}{h} Y(\omega_n) \;=\; \sum_{k=1}^{N} y_k \frac{e^{-2\pi i k n/N}}{\sqrt{2\pi}} \tag{17.28}$$

With this same care in accounting, and $d\omega \to 2\pi/Nh$, we invert the Y_n's:

$$y(t) \quad \stackrel{\text{def}}{=} \quad \int_{-\infty}^{+\infty} d\omega\, \frac{e^{i\omega t}}{\sqrt{2\pi}} Y(\omega) \;\approx\; \sum_{n=1}^{N} \frac{2\pi}{Nh} \frac{e^{i\omega_n t}}{\sqrt{2\pi}} Y(\omega_n) \tag{17.29}$$

Once we know the N values of the transform, (17.29), which leaves off the DC component, is useful for evaluating $y(t)$ for any value of the time t.

There is nothing illegal about evaluating the DFT expressions for Y_n and y_k for arbitrarily large values of n and k, yet there is also nothing to be gained. Because the trigonometric functions are periodic, we just get the old answer back:

$$y(t_{k+N}) = y(t_k) \qquad Y(\omega_{n+N}) = Y(\omega_n) \tag{17.30}$$

Another way of stating this is to observe that none of the equations change if we replace $\omega_n t$ by $\omega_n t + 2\pi n$. There are still just N independent output numbers for N independent inputs.

While periodicity is expected for Fourier *series*, it is somewhat surprising for Fourier *integrals*, which have been touted as the right tool for nonperiodic functions. The periodicity arises from approximating the integral as a sum over a finite number of periodic functions. Clearly, if we input values of the signal for longer lengths of time, then the inherent period becomes longer, and if the repeat period is very long, it may be of little consequence for times short compared to the period.

If $y(t)$ is actually periodic with period Nh, then the integration formulas converge very rapidly and the DFT is an excellent way of obtaining Fourier

[5] The alert reader may be wondering what has happened to the $h/2$ with which the trapezoid rule weights the initial and final points. Actually, they are there, but because we have set $y_0 \equiv y_N$, two $h/2$'s have been added to produce one h.

series. If the input function is not periodic, then the DFT can be a bad approximation near the endpoints of the time interval (after which the function will repeat), or for the (low) frequencies corresponding to the entire function repeating.

The discrete Fourier transform and its inverse can be written in a concise and insightful way, and can be evaluated efficiently, by introducing a complex variable Z for the exponential, and then raising Z to various powers:

$$Y_n = \frac{1}{\sqrt{2\pi}} \sum_{k=1}^{N} Z^{nk} y_k \qquad n = 0, 1, \ldots, N \tag{17.31}$$

$$y_k = \frac{\sqrt{2\pi}}{N} \sum_{n=1}^{N} Z^{-nk} Y_n \qquad Z = e^{-2\pi i/N} \tag{17.32}$$

where $Z^{nk} \equiv [(Z)^n]^k$. Now the computer needs to compute only powers of Z. We give our DFT code in Listing 17.1 (which uses approximately half of its length on plotting).

If your preference is to avoid complex numbers, we can rewrite (17.31) in terms of separate real and imaginary parts by applying Euler's theorem:

$$Z = e^{-i\theta} \qquad \Rightarrow \qquad Z^{\pm nk} = e^{\mp ink\theta} = \cos nk\theta \mp i \sin nk\theta \tag{17.33}$$

where $\theta \stackrel{\text{def}}{=} 2\pi/N$. In terms of the explicit real and imaginary parts:

$$\begin{aligned} Y_n \;=\; & \frac{1}{\sqrt{2\pi}} \sum_{k=1}^{N} [\, (\cos(nk\theta)\, \mathrm{Re} y_k + \sin(nk\theta)\, \mathrm{Im} y_k \\ & \qquad + i\, (\cos(nk\theta)\, \mathrm{Im} y_k - \sin(nk\theta)\, \mathrm{Re} y_k)\,], \end{aligned} \tag{17.34}$$

$$\begin{aligned} y_k \;=\; & \frac{\sqrt{2\pi}}{N} \sum_{n=1}^{N} [\, (\cos(nk\theta)\, \mathrm{Re} Y_n - \sin(nk\theta)\, \mathrm{Im} Y_n \\ & \qquad + i\, (\cos(nk\theta)\, \mathrm{Im} Y_n + \sin(nk\theta)\, \mathrm{Re} Y_n)\,]. \end{aligned} \tag{17.35}$$

Equation (17.34) is interesting in that it shows that a real function produces a real Fourier transform only if all the $\sin nk\theta$ terms cancel out. This is to be expected if $y(t)$ is an even function of t and if we perform an exact transform with integration from $-\infty$ to $+\infty$. Yet because we have no signal for $t < 0$, and because we only approximate the transform integral, even real functions may end up with DFTs. However, the imaginary components should get smaller as we increase the number of integration (sampling) points.

Listing 17.1: `DFT.java` computes the discrete Fourier transform for the signal given in the method `f(signal[])`. You will have to add output and plotting to see the results. (The Instructor's version also does an inverse transform and plots the results with PtPlot.)

```
//DFT.java: Discrete Fourier Transform

import java.io.*;

public class DFT {
  static final int N = 1000, Np = N;                    // Global constants
  static double [] signal = new double[N + 1];
  static double twopi = 2.*Math.PI, sq2pi = 1./Math.sqrt(twopi);
  static double h = twopi/N;

  public static void main(String[] argv){

    double dftreal[] = new double[Np];
    double dftimag[] = new double[Np];

    f(signal);
    fourier(dftreal, dftimag);
  }

  public static void fourier(double dftreal[], double dftimag[]){
    double real, imag;                                  // Calc & plot Y(w)
    int n, k;
    for ( n = 0;  n < Np;  n++ ) {                      // Loop on frequency
      real = imag = 0. ;                                  // Clear variables
                                                             // Major loop
      for ( k = 0;  k < N;  k++ ) {
        real += signal[k] * Math.cos( twopi*k*n/N );
        imag += signal[k] * Math.sin( twopi*k*n/N );
      }
      dftreal[n] = real*sq2pi;
      dftimag[n] = -imag*sq2pi;
    }
  }

  public static void f(double [] signal) {              // Signal function
    int i;
    double step = twopi/N, x = 0.;
    for ( i=0;  i <= N;  i++ ){
      signal[i] = 5. + 10*Math.sin(x+2.);
      x += step;
    }
  }
}
```

The actual computation time for a discrete Fourier transform can be reduced even further by use of the *fast Fourier transform (FFT)* algorithm. An examination of (17.31) shows that the DFT is evaluated as a matrix multiplication of a vector of length N of Z values times a vector of length N of y value. The time for this DFT scales like N^2. With the FFT algorithm, the time would scale like $N \log_2 N$. While this may not seem like much at first, for $N = 10^{2-3}$, the dif-

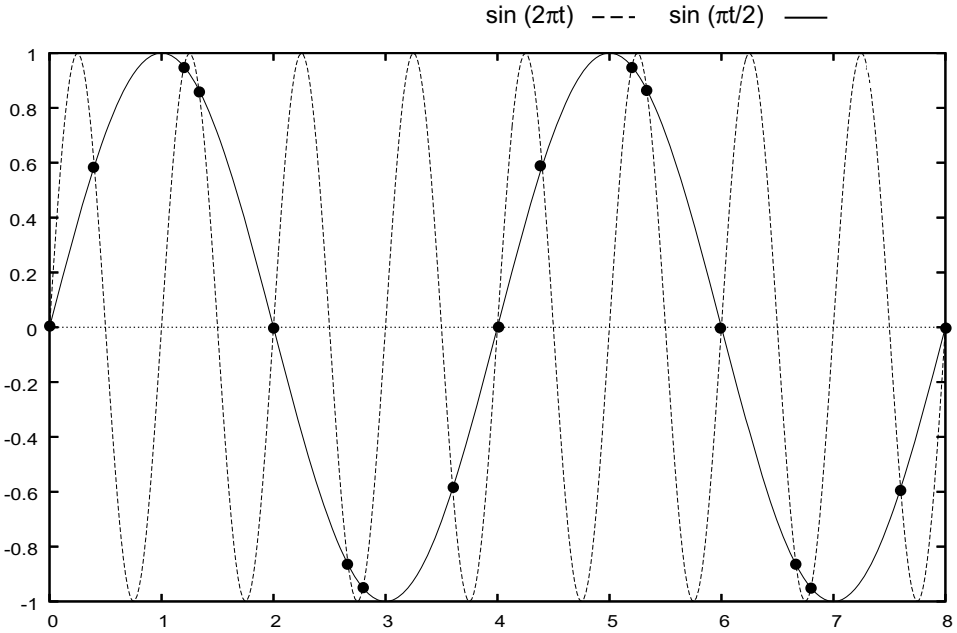

Fig. 17.2 A plot of the functions $\sin(\pi t/2)$ and $\sin(2\pi t)$. If the sampling rate is not high enough, these signals would appear indistinguishable. If both were present in a signal that was not sampled at a high enough rate, the deduced low-frequency component would be contaminated by the higher frequency component.

ference of a factor 10^{3-5} is the difference between a minute and a week. This is the reason FFT is often used for the online analysis of data.

17.6 Aliasing and Antialiasing (Assessment)⊙

A consequence of sampling a signal for only a finite number of times, as done with a DFT, is that this may lead to a poor deduction of the high-frequency components present in the signal. Clearly, obtaining good information about very high frequencies requires sampling the signal with very small time steps. While a poor deduction of the high-frequency components may be tolerable if all we care about are the low-frequency ones, those high-frequency components remain present in the signal and may contaminate the low-frequency components that we deduce. This effect is called *aliasing*, and is the cause of the Moire pattern distortion in digital images.

As an example, consider Fig. 17.2 showing the two functions $\sin(\frac{\pi}{2}t)$ and $\sin(2\pi t)$ for $0 \leq t \leq 8$, with their points of overlap emphasized. If we were unfortunate enough to sample a signal containing these functions at the times

$t = 0, 2, 4, 6, 8$, then we would measure 0 and assume that there was no signal at all. However, if we were unfortunate enough to measure the signal at the filled dots in Fig. 17.2 where $\sin(\frac{\pi}{2}t) = \sin(2\pi t)$, specifically, $t = 0, \frac{12}{10}, \frac{4}{3}, \ldots$, then our Fourier analysis would completely miss the high-frequency component. In DFT jargon, we would say that the high frequency has been *aliased* by the low-frequency component. In other cases, some high-frequency values may be included in our sampling of the signal, but our sampling rate may not be high enough to include enough of them to resolve the high-frequency component properly. In this case, some high-frequency signals get included spuriously as part of the low-frequency spectrum, and this leads to spurious low-frequency oscillations when the signal is synthesized from its Fourier components.

More precisely, aliasing occurs when a signal containing frequency f is sampled at a rate of $s = N/T$ measurements per unit time, with $s \leq f/2$. In this case, the frequencies f and $f - 2s$ yield the same DFT, and we would not be able to determine that there are two frequencies present. That being the case, to avoid aliasing we want no frequencies $f > s/2$ to be present in our input signal. This is known as the *Nyquist criterion*.

In practice we could avoid the effects of aliasing by filtering out the high frequencies from our signal and then analyzing the remaining low-frequency part. (The low-frequency *sinc filter* is often used for this.) Even though this approach eliminates some information about the high frequencies, it avoids the contamination, and consequent distortion, of the low-frequency components. This often leads to an improved reproduction of the signal.

If accurate values for the high frequencies are required, then we need to increase the sampling rate s by increasing the number N of samples taken within our fixed sampling time $T = Nh$. By keeping the sampling time constant and increasing the number of samples taken, we make the time step h smaller, and this picks up the higher frequencies. By increasing the number N of frequencies that you compute, you move the higher frequency components you are interested in closer to the middle of the spectrum and thus away from the error-prone ends.

While we are talking about these kind of practical matters, it is worthwhile to point out again the effect of varying the total time $T = Nh$ over which the signal is sampled, but not the sampling rate $s = N/T = 1/h$. Because the discrete frequencies of the DFT,

$$\omega_n = n\omega_1 = n\frac{2\pi}{T} \tag{17.36}$$

are measured in steps of ω_1, if we increase the total time T over which the signal is sampled, ω_1 gets smaller, and the corresponding frequency spectrum looks smoother. However, to keep the time step the same, we would also need

to increase the number of samples, N. And as we said, this is often done, after the fact, by *"padding"* the end of the data set with zeros.

17.7 DFT for Fourier Series (Algorithm)

For simplicity let us consider the Fourier cosine series:

$$y(t) = \sum_{n=0}^{\infty} a_n \cos(n\omega t) \qquad a_k = \frac{2}{T}\int_0^T dt \cos(k\omega t) y(t) \tag{17.37}$$

Here $T \overset{\text{def}}{=} 2\pi/\omega$ is the actual period of the system (not necessarily the period of the simple harmonic motion occurring for small amplitude). We assume that the function $y(t)$ is sampled for a discrete set of times

$$y(t = t_k) \equiv y_k \qquad k = 0, 1, \ldots, N \tag{17.38}$$

Because we are analyzing a periodic function, we will retain the conventions used in the DFT and require the function to repeat itself with period $T = Nh$, that is, we assume that the amplitude is the same at the first and last points:

$$y_0 = y_N \tag{17.39}$$

This means that there are only N independent values of y being used as input. For these N independent y_k values, we can determine uniquely only N expansion coefficients a_k. If we use the trapezoid rule to approximate the integration in (17.37), we determine the N independent Fourier components as

$$a_n \simeq \frac{2h}{T}\sum_{k=1}^{N} \cos(n\omega t_k)\, y(t_k) = \frac{2}{N}\sum_{k=1}^{N} \cos\left(\frac{2\pi nk}{N}\right) y_k, \quad n = 0, \ldots, N \tag{17.40}$$

Because for N independent $y(t)$ values we can determine only N Fourier components, our Fourier series for the function $y(t)$ must be in terms of only these components:

$$y(t) \simeq \sum_{n=0}^{N} a_n \cos(n\omega t) = \sum_{n=0}^{N} a_n \cos\left(\frac{2\pi nt}{Nh}\right) \tag{17.41}$$

In summary, we sample the function $y(t)$ at N times, t_1, ..., t_N. Because $y(t)$ is periodic, if we sample within one period, we are ensured of independent input data. You see that all N values of y sampled contribute to each a_k. Consequently, if we increase N in order to determine more coefficients, we must recompute all the a_n values. In the model-independent approach

discussed in Section 17.10, the theory is reformulated so that additional samplings determine higher Fourier components without affecting lower ones.

17.8 Assessments

- **Simple Analytic Input**

The simple checks here are generally good to do before examining more complex problems. If your system has some Fourier analysis packages (such as the graphing package *Ace/gr*), you may want to compare your results with those from the packages. Once you understand how the packages work, it makes sense to use them.

1. Sample the even signal,

   $$y(t) = 3\cos(\omega t) + 2\cos(3\omega t) + \cos(5\omega t)$$

   Decompose this into its components and check that they are essentially real and in the ratio 3:2:1 (or 9:4:1 for the power spectrum).

2. Experiment on the effect of picking different values of the step size h and, independently, of extending the measurement period $T = Nh$ to larger values.

3. Sample the odd signal,

   $$y(t) = \sin(\omega t) + 2\sin(3\omega t) + 3\sin(5\omega t)$$

   Decompose this into its components and check that they are imaginary and in the ratio 1:2:3 (or 1:4:9 if a power spectrum is plotted).

4. Sample the mixed-symmetry signal

   $$y(t) = 5\sin(\omega t) + 2\cos(3\omega t) + \sin(5\omega t)$$

   Decompose this into its components and see if there are three of them in the ratio 5:2:1 (or 25:4:1 if a power spectrum is plotted). Then check that your Y_n values can be resumed to reproduce this input.

5. In our discussion of aliasing, we examined Fig. 17.2, which shows the functions $\sin(\frac{\pi}{2}x)$ and $\sin(2\pi x)$. Sample the function

   $$y(t) = \sin(\tfrac{\pi}{2}x) + \sin(2\pi x)$$

   with a sampling rate that leads to aliasing, as well as a higher sampling rate at which there is no aliasing. Compare the resulting DFTs in each case, and check if your simulations agree with the Nyquist criterion.

- **Highly Nonlinear Oscillator**

Recall the numerical solution for oscillations of the spring with power $p = 11$, (Eq. (17.1)). Decompose the solution into a Fourier series and determine the number of higher harmonics that contribute at least 10%; for example, determine the n for which $|b_n/b_1| < 0.1$. Specifically, check that summing your series reproduces your original solution. (*Warning:* The ω you use in your series must correspond to the actual frequency of the system, not just that in the small oscillation limit.)

- **Nonlinearly Perturbed Oscillator**

Recall the harmonic oscillator with a nonlinear perturbation (15.2):

$$V(x) = \tfrac{1}{2}kx^2\left(1 - \tfrac{2}{3}\alpha x\right) \quad F(x) = -kx(1 - \alpha x) \tag{17.42}$$

For very small amplitudes of oscillation ($x \ll 1/\alpha$), the solution $x(t)$ will essentially be only the first term of a Fourier series.

1. We want to say that "we have approximately a 10% nonlinearity." Such being the case, fix your value of α so that $\alpha\, x_{\max} \simeq 10\%$, where $x_{\max}$ is the maximum amplitude of oscillation. For the rest of the problem, keep the value of α fixed.
2. Decompose your numerical solution into a discrete Fourier spectrum.
3. Plot a graph of the percentage importance of the first *two*, non-DC Fourier components as a function of the initial displacement for $0 < x_0 < 1/2\alpha$. You should find that higher harmonics are more important as the amplitude increases. Because there may be both even and odd components present, there may be both real and imaginary parts to Y_n. Consequently, one way to answer this question is to calculate $|Y_{n+1}|/|Y_n|$, where $|Y_n| = \sqrt{(\mathrm{Re}Y_n)^2 + (\mathrm{Im}Y_n)^2}$. Alternatively, you can also look at successive terms in the power spectrum, although then the 10% effect in amplitude becomes a 1% effect in power.
4. As always, make sure to check by resuming the series for $y(t)$ and seeing if the input is reproduced.

(*Warning:* The ω you use in your series must correspond to the *true* frequency of the system, not just ω in the small oscillation limit.)

17.9 DFT of Nonperiodic Functions (Exploration)

Consider a simple model of a "localized" electron that moves through space and time. We assume that the electron is described by a wave packet $\psi(x)$ that

is a function of the spatial coordinate x. A good model for an electron initially localized around $x = 5$ is a Gaussian multiplying a plane wave:

$$\psi(x, t = 0) = \exp\left[-\frac{1}{2}\left(\frac{x - 5.0}{\sigma_0}\right)^2\right] e^{ik_0 x} \tag{17.43}$$

This wave packet is not an eigenstate of the momentum operator[6] $p = id/dx$, and in fact contains a spread of momenta. The **problem** is to evaluate the Fourier transform,

$$\psi(p) = \int_{-\infty}^{+\infty} dx \, \frac{e^{ipx}}{\sqrt{2\pi}} \psi(x) \tag{17.44}$$

as a way of determining the momenta spectrum in (17.43).

17.10
Model Independent Data Analysis (Exploration)⊙

The scattering of electrons and x-rays from solids, atoms, molecules, and nuclei provides a means of determining the charge density $\rho(r)$ of the target. The experiments actually measure the *form factor* or *structure function* of the target,[7] that is, the Fourier transform of the charge density:

$$F(q) = \int d^3 r e^{i\mathbf{q}\cdot\mathbf{r}} \rho(r) = 4\pi \int_0^\infty r^2 dr \frac{\sin qr}{qr} \rho(r) \tag{17.45}$$

Here $\mathbf{q}$ is the momentum transferred during scattering and the 1D integral obtains for spherically symmetric $\rho(r)$. The **problem** is to determine $\rho(r)$ from experimental measurements of $F(q)$, that is, to *invert* the transform in (17.45) to obtain ρ as a function of r. The real problem is that a laboratory beam of particles has a finite momentum, which means that there is a finite limit to the largest q value $q_{\max}$ at which $F(q)$ can be measured, and that experiments must be finished in a finite period of time, which means that only a finite number of q values can be measured. The $q < q_{\max}$ limitation leads to uncertainties in those parts of $\rho(r)$ that oscillate with high frequencies. The discrete number of measurements means that there are uncertainties in lower frequency components as well.

[6] We use natural units in which $\hbar = 1$.

[7] While the form factor can be deduced directly from experiment only in first Born approximation, the method described here is more general.

The traditional solution to this problem has been to *assume* that some specific functional form for the density $\rho(r)$ containing a number of adjustable parameters, and then find a best fit to the data by varying the parameters. A shortcoming with this approach is that the values deduced for the density depend somewhat on the functional form assumed for $\rho(r)$.

A more general approach is called *model-independent analysis*. In it, $\rho(r)$ is expanded in a complete set of functions

$$\rho(r) = \sum_{n=1}^{\infty} \rho_n(r) \tag{17.46}$$

with, typically, one parameter in each $\rho_n(r)$. If the ρ_n values form a complete set and if the sum actually goes out to $n = \infty$, then this is an exact representation of any $\rho(r)$. In practice, the discrete and finite nature of the measurements limit the number of ρ_n values that can be determined to a maximum number N, and so we have an approximate representation:

$$\rho(r) \approx \sum_{n=1}^{N} \rho_n(r) \tag{17.47}$$

To make a clear separation of how the deduced $\rho_n(r)$ values are affected by the measurements of $F(q)$ for $q > q_{\max}$, we impose the constraint that the ρ_n values determined for measurement with $q < q_{\max}$ be orthogonal to those determined with $q > q_{\max}$. While the equations will appear similar to those in DFT, in DFT *all* N values of y_k are used to determine each Y_n, while here we require each ρ_n to be determined by only *one* measured $F(q)$. It is then manifest that each new experimental measurement determines one additional expansion coefficient.

Now for a specific example. The density $\rho(r)$ within a specific atomic nucleus is known to vanish rapidly beyond some radius $r \simeq R$ and to be approximately constant for $r \simeq 0$. We therefore model $\rho(r)$ as the sine series:

$$r\rho(r) \;=\; \begin{cases} \sum_{n=1}^{N} b_n \sin(q_n r) & \text{for } r \le R \\ 0 & \text{for } r > R \end{cases} \qquad q_n = \frac{n\pi}{R} \tag{17.48}$$

where we are a simple formula in place of measured momentum transfer. This makes it clear that each additional measurement is made at a higher q value. The coefficients b_n are determined from the inversion formula (17.8):

$$b_n \;=\; \frac{2}{R}\int_0^R \sin(q_n r) r\rho(r) dr \;=\; \frac{q_n F(q_n)}{2\pi R} \tag{17.49}$$

The corresponding expression for the form factor as determined by the measured values $F(q_n)$ is

$$F(q) = \frac{2\pi R}{q} \sum_{n=1}^{N} b_n \left\{ \frac{\sin[(q-q_n)R]}{(q-q_n)R} - \frac{\sin[(q+q_n)R]}{(q+q_n)R} \right\} \tag{17.50}$$

These relations show exactly how measurements at larger and larger momentum transfers, q_n values, determine the higher and higher Fourier components b_n values, which, in turn, are related to higher and higher frequency ripples in the charge density. (This is sometimes stated somewhat loosely as "large q are needed to measure small r.") Regardless of the model assumed for the density ρ, any experiment with a definite $q_{\max}$ has a limit on the components of the density it can deduce. As higher and higher q measurements are made, more components of ρ are determined.

17.11 Assessment

1. As a simple exercise, verify the sensitive relation between each measurement at q_n and the Fourier component b_n by evaluating the term in braces in (17.50) for $n = 10$ and plotting it as a function of qR. You should find a peaking that indicates the region of sensitivity, that is, the region to which a single experimental measurement is most sensitive.

2. Now try a computer experiment that simulates a model-independent data analysis. Assume at first that you are mother nature and so you know that the actual charge distribution $\rho(r)$ and form factor for some nucleus are

$$\rho(r) = \rho(0)\left[1 + \alpha(r/a)^2\right] e^{-(r/a)^2} \tag{17.51}$$

$$F(q) = \left[1 - \frac{\alpha(qa)^2}{2(2+3\alpha)}\right] e^{-(qa)^2/4} \tag{17.52}$$

   Use this analytic expression for $F(q)$ to determine the values of the experimental $F(q_m)$ as needed. In a real-world situation you would measure data and then fit them to determine the Fourier coefficients b_m.

   (a) Consider the nucleus ^{16}O which has $\alpha = \frac{4}{3}$. We will measure distances in fermis, fm $= 10^{-13}$ cm and momentum transfers q in inverse fermis. In these units, mother nature knows that $a \simeq 1.66$, and you as an outsider would have to try values of $R \simeq 5$ and $R \simeq 6$ as the radius beyond which the nuclear density vanishes.

   (b) Generate from (17.52) a table of $F(q_n)$ values for both values of R and for q values starting at zero and increasing to the point where $F(q) \approx 10^{-11}$. Plot up both (they should fall on the same curve).

(c) Use (17.48) to calculate and plot the contribution to $\rho(r)$ coming from progressively larger and larger values of q_n.

(d) For 10 terms, examine how the sum for $\rho(r)$ differs for the two R values used. This is a good measure of the *model dependence* in this "model independent" analysis.

(e) Again, for the sum of 10 terms, examine how the Fourier series for $\rho(r)$ differs from the actual functional form (17.51).

(f) Examine the series expansion for $F(q)$ and $\rho(r)$ and note any unphysical oscillations.

(g) Do a computer experiment in which you assume that a different form for the large q behavior of $F(q)$ [e.g., $1/q^4$, $\exp(-bq)$], and then see how this affects the deduced $\rho(r)$. You should find that the small r ripples are most sensitive to the assumed large q behavior.