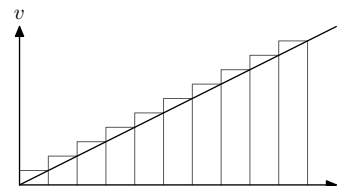# 3 Discretization

Discretization takes the fundamental idea of calculus and pushes it to the opposite extreme from what calculus uses. Calculus was invented to analyze changing processes such as orbits of planets or, as a one-dimensional illustration, how far a ball drops in time t. The usual computation

$$\text{distance} = \text{velocity} \times \text{time}$$

fails because the velocity changes (it increases linearly with time). However – and this next step is the fundamental idea of calculus – over a short time interval, its velocity is almost constant and the usual distance formula works for each short interval. Each short distance is the area of one rectangle, and the total distance fallen is approximately the combined area of the rectangles. To eliminate the error, calculus uses the extreme case of infinite rectangles, ever thinner (shorter intervals) until each shrinks to zero width. Then the approximation of constant speed becomes exact. Discretization uses the opposite extreme: one maybe two fat rectangles. This limitation means the error may not be zero, but it drastically simplifies any computations.

3.1 Exponential decay

The first example is this integral:

$$\int_0^\infty e^{-t}\, dt.$$

Since the derivative of e^t is e^t, the indefinite integral is easy to find exactly, and the limits make the computation even simpler. In an example where the exact answer is known, we can
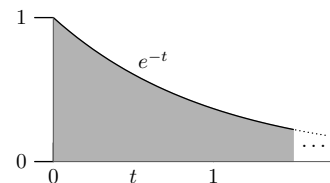
Discretization **32**

adjust the free parameters in the method of discretization until the method produces accurate values. So, replace the complicated, continuous, smooth exponential decay e^{-t} by a rectangle, and do the integral by finding the area of the rectangle.

With one rectangle, the approximate function remains constant until it abruptly falls to and remains zero. Finding the area of the rectangle requires choosing its height and width. A natural height is the maximum of e^{-t}, which is 1. A natural width is the time interval until $f(t) = e^{-t}$ changes significantly. A similar idea appeared in **Section 1.4** to approximate a derivative df/dx. Its numerator df was estimate as a typical value of $f(x)$. Its denominator dx became the x interval over which $f(x)$ changes significantly. For an exponential, a natural definition for significant change is to changes by a factor of e. When $f(t) = e^{-t}$, this change happens when t goes from t to $t + 1$. So the approximating rectangle, whose height we've chosen to be 1, also has unit width. It is a unit square with unit area. And this rectangle exactly estimates the integral since

$$\int_0^\infty e^{-t}\, dt = 1.$$
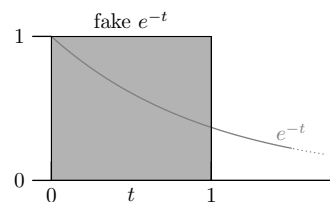
3.2   Circuit with exponential decay

In **Chapter 1** on dimensions, I insisted that declaring quantities prematurely dimensionless ties one hand behind your back. In the previous example I committed that sin by making the exponent be $-t$. Since an exponent is dimensionless, my choice made t dimensionless as well.

A more natural interpretation of t is as a time. So here is a similar example but where t has dimensions, which are useful in making and checking the approximations. Let's first investigate the initial conditions, just before the switch closes. No current is flowing since the circuit is not yet a closed loop. Furthermore, because the circuit has been waiting forever, the capacitor has had completely discharged. So capacitor has no charge on it. The charge determines the voltage across the capacitor by

$$Q = CV_{\mathrm{C}},$$

2008-03-06 13:24:47 / rev ebd336097912+
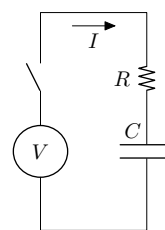
3.2 Circuit with exponential decay   **33**

where Q is the charge on the capacitor, C is its capacitance, and V_C is the resulting voltage. [See the classics on circuits [2] and electromagnetism [3] for more on capacitors.] So just before the switch closes, the capacitor has zero voltage on it ($V_C = 0$).

At time $t = 0$, I close the switch, which connects the resistor and capacitor to the source voltage V (which is constant). Since V_C starts at zero, the voltage drops in the resistor is the whole source voltage V:

$$V_R = V \quad \text{(initially)},$$

where V_R is the voltage across the resistor. This voltage drop is caused by a current I flowing through the resistor (which then flows through the capacitor). Ohm's law says that $V_R = IR$. Initially $V_R = V$ so the initial current is $I_0 = V/R$. This current charges the capacitor and increases V_C. As V_C increases, V_R decreases – which decreases the current I, which decreases how fast V_C increases, which ...Finding the current is a problem for calculus, in particular for a differential equation.

Instead, let's guess the current using dimensions, extreme cases, and the new technique of discretization. First apply extreme cases. At the $t = 0$ extreme, the current is $I_0 = V/R$. At the $t = \infty$ extreme, no current flows: The capacitor accumulates enough charge so that $V_C = V$, whereupon no voltage drops across the resistor. From Ohm's law again, a zero voltage drop is possible only if no current flows.
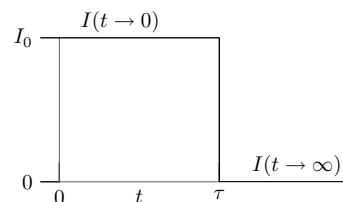

Between those extremes, we guess I using discretization. Pretend that I stays at its $t = 0$ value of I_0 for a time τ, then drops to its $t = \infty$ value of $I = 0$. So τ is the time for the current to change significantly. To determine τ, use dimensions. It can depend only on R and C. [It could depend on V, but because the system is linear, the time constants do not depend on amplitude.] The only way to combine R and C into a time is the product RC. A reasonable guess for τ is therefore $\tau = RC$. In this picture, the discretized current stays at V/R until $t = \tau$, then falls to 0 and remains zero forever.

Discretization

No physical current changes so abruptly. To guess the true current, use discretization in reverse. The exponential decay of **Section 3.1** produced the same rectangular shape after discretizing. So perhaps the true current here is also an exponential decay. In the other example, the function was e^{-t}, and the changeover from early- to late-time behavior happened

at $t = 1$ (in that example, t had no dimensions). By $t = 1$, the exponential decay e^{-t} had changed significantly (by a factor of e). For this circuit, the corresponding changeover time is τ. To change by a factor of e in time τ, the current should contain $e^{-t/\tau}$. The initial current is $I = I_0$, so the current should be
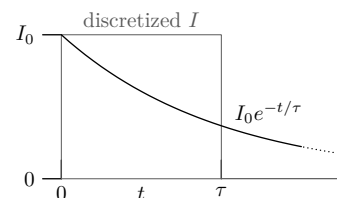
$$I = I_0 e^{-t/\tau} = \frac{V}{R} e^{-t/\tau}.$$

Having a solution, even a guess, turns the hard work of solving differential equations into the easier work of verifying a solution.

To test the guess for I, I derive the differential equation for the current. The source voltage V drops only in the resistor and capacitor, so their voltage drops must add to V:

$$V = V_R + V_C.$$

The capacitor voltage is $V_C = Q/C$. The resistor voltage is $V_R = IR$, so

$$V = IR + \frac{Q}{C}.$$

It seems that there are too many variables: V and C are constants, but I and Q are unknown. Fortunately current I and charge Q are connected: charge is the time integral of current and $I = dQ/dt$. Differentiating each term with respect to time simplifies the equation:

$$0 = R\frac{dI}{dt} + \frac{1}{C}\underbrace{\left(\frac{dQ}{dt}\right)}_{I} = R\frac{dI}{dt} + \frac{I}{C}.$$

Move the R to be near its companion C (divide by R):

$$0 = \frac{dI}{dt} + \underbrace{\frac{I}{RC}}_{\tau} = \frac{dI}{dt} + \frac{I}{\tau}.$$

Dimensions, extreme cases, and reverse discretization produced this current:

2008-03-06 13:24:47 / rev ebd336097912+

$$I = I_0 e^{-t/\tau}.$$

Amazing! It satisfies the differential equation:

$$\frac{d}{dt}\left(I_0 e^{-t/\tau}\right) + \frac{I_0 e^{-t/\tau}}{\tau} = 0$$

because the time derivative brings down a factor of $-1/\tau$, making the first and second terms equal except for a minus sign.

3.3   Population

Not all problems are exponential decays. In the next example, the true functions are unknown and exact answers are not available. The problem is to estimate the number of babies in the United States. To specify the problem, define babies as children less than two years old. One estimate comes from census data, which is accurate within the limits of statistical sampling. You integrate the population curve over the range $t = 0$ to 2 years. But that method relies on the massive statistical efforts of the US census bureau and would not work on a desert island. If only the population were constant (didn't depend on age), then the integrals are easy! The desert-island, back-of-the-envelope method is to replace the complicated population curve by a single rectangle.

How high is the rectangle and how wide is it? The width τ, which is a time, has a reasonable estimate as the average life expectancy. So $\tau \sim$ 70 years. How high is the rectangle? The height does not have such an obvious direct answer as the width. In the exponential-decay examples, the height was the the initial value, from which we found the area. Here, the procedure reverses. You know the area – the population of the United States – from which you find the height. So, with the area being $3 \cdot 10^8$, the height is

$$\text{height} \sim \frac{\text{area}}{\text{width}} \sim \frac{3 \cdot 10^8}{75\,\text{years}},$$
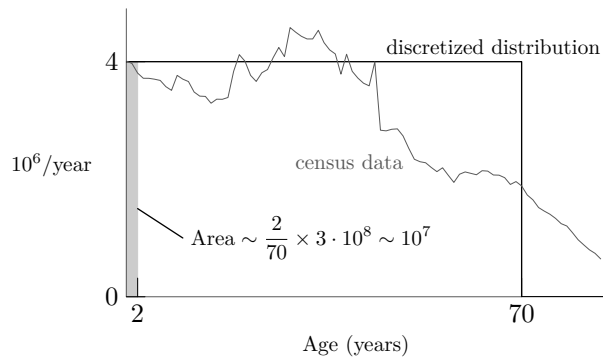
Discretization **36**

since the width is the life expectancy, for which we used 70 years. How did it become 75 years? The answer is by a useful fudge. The new number 75 divides into 3 (or 300) more easily than 70 does. So change the life expectancy to ease the mental calculations. The inaccuracies caused by that fudge are no worse than in replacing the complex population curve by a rectangle. So

$$\text{height} \sim 4 \cdot 10^6 \, \text{year}^{-1}.$$

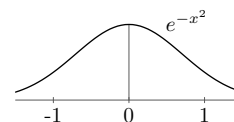Integrating a rectangle of that height over the infancy duration of 2 years gives

$$N_{\text{babies}} \sim \underbrace{4 \cdot 10^6 \, \text{years}^{-1}}_{\text{height}} \times \underbrace{2 \, \text{years}}_{\text{infancy}} = 8 \cdot 10^6.$$

Thus roughly 8 million babies live in the United States. From this figure, you can estimate the landfill volume used each year by disposable diapers (nappies).
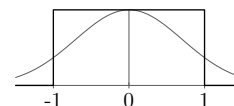
3.4   Full width at half maximum

The Gaussian integral

$$\int_{-\infty}^{\infty} e^{-x^2} \, dx$$

has appeared in several examples, and you've seen the trick (in **Section 2.2**) of squaring it to show that its value is $\sqrt{\pi}$. The exponential in the integrand is a difficult, continuous function. Except over the infinite integration range, the integral has no closed form, which is why statistics tables enumerate the related error function numerically. I introduce that evidence to show you how difficult the integral is without infinite limits, and it is not easy even with infinite limits.

Pretend therefore that you forget the trick. You can approximate the integral using discretization by replacing the integrand with a rectangle. How high and how wide should the rectangle be? The recipe is to take the height as the maximum height of the function and the width as the distance until the function falls significantly. In the exponential-decay examples, significant meant changing by a factor of e. The maximum of e^{-x^2} is at $x = 0$ when $e^{-x^2} = 1$, so the approximating rectangle has unit height. It falls to $1/e$ when $x = \pm 1$, so the approximating rectangle has width 2 and therefore area 2. This estimate is

3.4 Full width at half maximum $\qquad$ **37**

half decent. The true value is $\sqrt{\pi} = 1.77\ldots$, so the error is about 13%: a reasonable trade for one line of work.

Another recipe, also worth knowing because it is sometimes more accurate, arose in the bygone days of spectroscopy. Spectroscopes measure the wavelengths (or frequencies) where a molecule absorbed radiation and the corresponding absorption strengths. These data provided an early probe into the structure of atoms and molecules, and was essential to the development of quantum theory [4]. An analogous investigation occurs in today's particle accelerators – colloquially, atom smashers – such as SLAC in California and CERN and in Geneva: particles, perhaps protons and neutrons, collide at high energies, showering fragments that carry information about the structure of the original particles. Or, to understand how a finely engineered wristwatch works, hammer it and see what the flying shards and springs reveal.

The spectroscope was a milder tool. A chart recorder plotted the absorption as the spectroscope swept through the wavelength range. The area of the peaks was an important datum, and whole books like [5] are filled with these measurements. Over half a century before digital chart recorders and computerized numerical integration , how did one compute these areas? The recipe was the FWHM.

> FWHM = full width at half maximum

Unpack the acronym in slow motion:

1. **M**. Find the maximum value (the peak value).

2. **HM**. Find one-half of the maximum value, which is the half maximum.

3. **FWHM**. Find the two wavelengths – above and below the peak – where the function has fallen to one-half of the maximum value. The full width is the difference between the above and below wavelengths.

The FWHM approximation recipe replaces the peak by a rectangle with height equal to the peak height and width equal to the the width estimated by the preceding three-step procedure.

Try this recipe on the Gaussian integral and compare the estimate with the estimate from the old recipe of finding where the function changed by a factor of e. The Gaussian has maximum height 1 at $x = 0$. The half maximum is 1/2, which
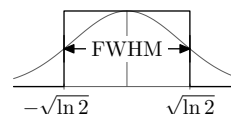
Discretization **38**

happens when $x = \pm\sqrt{\ln 2}$. The full width is then $2\sqrt{\ln 2}$, and the area of the rectangle – which estimates the original integral – is $2\sqrt{\ln 2}$. Here, side by side, are the estimate and the exact integral:

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \begin{cases} \sqrt{\pi} = 1.7724\ldots & \text{(exact)}, \\ 2\sqrt{\ln 2} = 1.6651\ldots & \text{(estimate)}. \end{cases}$$

The FWHM estimate is accurate to 6%, twice as accurate as the previous recipe. It's far better than one has a right to expect for doing only two lines of algebra!

3.5 Stirling's formula

The FWHM result accurately estimates one of the most useful quantities in applied mathematics:

$$n! \equiv n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1.$$

We meet this quantity again as a picture proof in **Section 4.6**. Here we estimate $n!$ by discretizing an integral representing $n!$:

$$\int_0^{\infty} t^n e^{-t}\,dt = n!$$

You may not yet know that this integral is $n!$; you can show it either with integration by parts or see **??** on generalization to learn differentiation under the integral sign. For now accept the integral representation on faith, with a promise to redeem the trust in a later chapter.

To approximate the integral, imagine what the integrand $t^n e^{-t}$ looks like. It is the product of the increasing function t^n and the decreasing function e^{-t}. Such a product usually peaks. A familiar example of this principle is the product of the increasing function x and the decreasing function $1 - x$. over the range $x \in [0, 1]$ where both functions are positive. The product rises from and then falls back to zero, with a peak at $x = 1/2$.
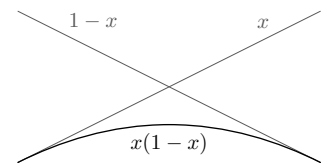

You can check that the product $t^n e^{-t}$ has a peak by looking at its behavior in two extreme cases: in the short run $t \to 0$ and in the long run $t \to \infty$. When $t \to 0$, the exponential is 1, but the polynomial factor t^n wipes it out by multiplying by zero. When $t \to \infty$, the polynomial factor t^n pushes the product to infinity while the exponential factor e^{-t} pushes it to zero.

An exponential beats any polynomial. To see why and avoid the negative exponent $-t$ muddying this issue, compare instead e^t with t^n as $t \to \infty$. The Taylor series for e^t contains all powers of t, so it is like an infinite-degree polynomial. So e^t/t^n goes to infinity once t gets large enough. Similarly, its reciprocal $t^n e^{-t}$ goes to zero as $t \to \infty$. Being zero at also $t = 0$, the product is zero at both extremes and positive elsewhere. Therefore it peaks in between. Maybe it has more than one peak, but it should have at least one peak. Furthermore, as n increases, the t^n polynomial factor strengthens, so the e^{-t} requires a larger t to beat down the t^n. Therefore, as n increases the peak moves right.

With $t^n e^{-t}$ having a peak, the FWHM recipe can approximate its area. The recipe requires finding the height (the maximum of the function) and the width (the FWHM) of the approximating rectangle. To find these parameters, slurp the t^n into the exponent:
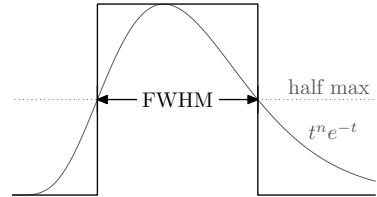
$$t^n e^{-t} = e^{n \ln t} e^{-t} = e^{n \ln t - t}.$$

The exponent $f(t) \equiv n \ln t - t$ is interesting. As $t \to 0$, the $\ln t$ takes $f(t)$ to $-\infty$. As $t \to \infty$, the $-t$ takes $f(t)$ again to $-\infty$. Between these limits, it peaks. To find the maximum, set $f'(t) = 0$:

$$f'(t) = \frac{n}{t} - 1 = 0,$$

or $t_{\text{peak}} = n$. As we predicted, the peak moves right as n increases. The height of the peak is one item needed to estimate the rectangle's area. At the peak, $f(t)$ is $f(n) = n \ln n - n$, so the original integrand, which is $e^{f(t)}$, is

$$e^{f(t_{\text{peak}})} = e^{f(n)} = e^{n \ln n - n} = \frac{n^n}{e^n} = \left(\frac{n}{e}\right)^n.$$

To find the width, look closely at how $f(t)$ behaves near the peak $t = n$ by writing it as a Taylor series around the peak:

$$f(t) = f(n) + f'(n)(t - n) + \frac{1}{2} f''(n)(t - n)^2 + \cdots.$$

The first derivative is zero because the expansion point, $t = n$, is a maximum and there $f'(n) = 0$. So the second term in the Taylor series vanishes. To evaluate the third term, compute the second derivative of f at $t = n$:

$$f''(n) = -\frac{n}{t^2} = -\frac{1}{n}.$$

So

$$f(t) = \underbrace{n \ln n - n}_{f(n)} + \frac{1}{2} \times \underbrace{\left(-\frac{1}{n}\right)}_{f''(n)} (t - n)^2 + \cdots.$$

The first term gives the height of the peak that we already computed. The second term says how the height falls as t moves away from n. The result is an approximation for the integrand:

$$e^{f(t)} = \left(\frac{n}{e}\right)^n e^{-(t-n)^2/2n}.$$

The first factor is a constant, the peak height. The second factor is the familiar Gaussian. This one is centered at $t = n$ and contains $1/2n$ in the exponent but otherwise it's the usual Gaussian with a quadratic exponent. It falls by a factor of 2 when $(t - n)^2/2n = \ln 2$, which is when

$$t_{\pm} = n \pm \sqrt{2n \ln 2}.$$
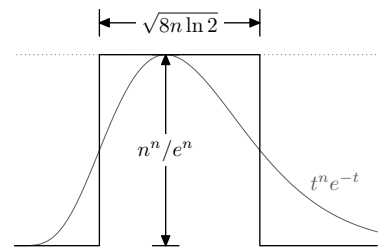
The FWHM is $t_+ - t_-$, which is $\sqrt{8n \ln 2}$. The estimated area under $e^{f(t)}$ is then

$$\left(\frac{n}{e}\right)^n \times \sqrt{8n \ln 2}.$$

As an estimate for $n!$, each piece is correct except for the constant factor. The more accurate answer has $\sqrt{2\pi}$ instead of $\sqrt{8 \ln 2}$. However, 2π is roughly $8 \ln 2$ so the approximate is, like the estimate the vanilla Gaussian integral (coincidence?), accurate to 6%.

3.6 Pendulum period

The period of a pendulum is by now a familiar topic in this book. Its differential equation becomes tractable with a bit of discretization. The differential equation that describes pendulum motion is

$$\frac{d^2\theta}{dt^2} + \frac{g}{l} \sin \theta = 0$$

This nonlinear equation has no solution in terms of the usual functions – to put it more precisely, in terms of elementary functions. But you can

3.6 Pendulum period   **41**

understand a lot about how it behaves by discretizing. If only the equation
were

$$\frac{d^2\theta}{dt^2} + \frac{g}{l}\theta = 0.$$

This equation is linear, and therefore possible to solve without too much
misery – I hesitate to say that any differential equation is 'easy' – and its
solution is an oscillation with angular frequency $\omega = \sqrt{g/l}$:

$$\theta(t) = \theta_0 \cos\left(\frac{g}{l}t\right).$$

Its period is $2\pi\sqrt{g/l}$, which is independent of amplitude θ_0.

The complexity of the unapproximated pendulum equation
arises because it has the torque-producing factor $\sin\theta$ instead
of its approximation θ. The two functions match perfectly as
$\theta \to 0$. But as θ grows – which happens with large amplitudes
– the equivalence becomes less accurate. One way to compare
them is to look at their ratio $(\sin\theta)/\theta$. As expected, when $\theta =$
0, the ratio is 1. As θ grows, the ratio falls, making the simple-
harmonic approximation less accurate. We can discretize to
find a more accurate approximation than the usual simple-
harmonic one, yet still produce a linear differential equation. The upcoming
figures illustrate making and refining that approximation.
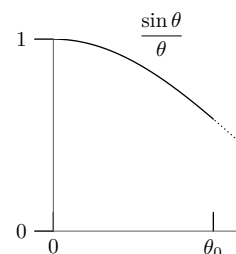
We need a discrete approximation to the difficult function
$\sin\theta$ in the range $[0, \theta_0]$. Look at the two functions θ and $\sin\theta$
after dividing by θ; we are taking out the common big part, the
topic of **Chapter 5**. The difficult function becomes $(\sin\theta)/\theta$.
The other function, a straight line, is the simple harmonic ap-
proximation, and is a useful zeroth approximation. But it does
not produce any change in period as a function of amplitude
(since the height of the replacement line is independent of θ_0).

2008–03–06 13:24:47 / rev ebd336097912+
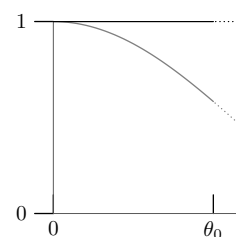
Discretization   **42**

The next approximation does fixes that problem. Use a flat line with height $(\sin\theta_0)/\theta_0$. This line is the minimum height of $(\sin\theta)/\theta$. Why is that choice an improvement on the first approximation, using the maximum height of 1? Because in this choice, the height varies with amplitude, so the period varies with amplitude: This choice explains a physical effect that the first choice approximated into oblivion. In this second approximation, the torque term $(g/l)\sin\theta$ becomes
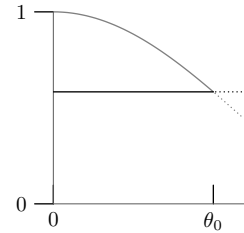
$$\frac{g}{l}\theta\frac{\sin\theta_0}{\theta_0}.$$

Starting from the simple-harmonic approximation, this choice is equivalent to replacing gravity by a slightly weaker gravity:

$$g \rightarrow g \times \frac{\sin\theta_0}{\theta_0}.$$

The Taylor series for sin gives

$$\frac{\sin\theta_0}{\theta_0} \approx 1 - \frac{\theta_0^2}{6}.$$

The fake g is then

$$g_{\text{fake}} = g\left(1 - \frac{\theta_0^2}{6}\right).$$

Using this fake g, the period becomes

$$T \approx 2\pi\sqrt{\frac{l}{g_{\text{fake}}}}.$$

To compute $g_{\text{fake}}^{-1/2}$ requires another Taylor series:

$$(1+x)^{-1/2} \approx 1 - \frac{x}{2}.$$

Then

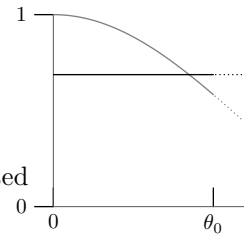$$T \approx 2\pi\sqrt{\frac{l}{g}}\left(1 + \frac{\theta_0^2}{12}\right).$$

This period is an overestimate because it assumed the weakest torque adjustment factor: scaling the torque by the value of $(\sin\theta)/\theta$ at the endpoints of the swing when $\theta = \pm\theta_0$. The next approximation comes from using an intermediate height for the replacement line. Equivalently, say that the pendulum spends half its flight acting like a spring, where the torque contains just θ; and half its flight where the torque has the term $\theta(\sin\theta_0)/\theta_0$. Then the period is an average of the simple-harmonic period $T = 2\pi\sqrt{l/g}$ with the preceding underestimate:
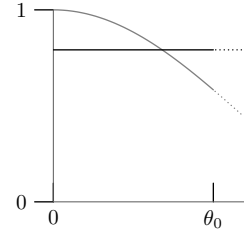
$$T = 2\pi\sqrt{\frac{l}{g}}\left(1 + \frac{\theta_0^2}{24}\right).$$

The next step – and here I am pushing this method perhaps farther than is justified – is to notice that the pendulum spends most of its time where it moves the slowest. So it spends most of time near the endpoints of the swings, where the simple-harmonic approximation is the least accurate. So the endpoint-only underestimate estimate for T should be weighted slightly more than the simple-harmonic overestimate. The most recent estimate weighted these pieces equally. To improve it, count the endpoint estimate, say, twice and the center estimate once. This recipe has a further justification in that there are two endpoints and only one center! Then the period becomes

$$T = 2\pi\sqrt{\frac{l}{g}}\left(1 + \frac{\theta_0^2}{18}\right)$$

The true coefficient, which comes from doing a power-series solution, is $1/16$ so this final weighted estimate is very accurate!

3.7 What have you learnt

Discretization makes hard problems simple. The recipe is to replace a complicated function by a rectangle. The art is in choosing the height and width of the rectangle, and you saw two recipes. In both, the height is the maximum of the original function. In the first, easier recipe, the width is the range over which the function changes by a factor of e; this recipe is useful for linear exponential decays. The second recipe, the FWHM, is useful for messy functions like spectroscope absorption peaks and Gaussians. In that

Discretization **44**

recipe, the width is the width over which the function goes from one-half the maximum and then returns to that value.