Chapter 1

Infinite and Power series

In ancient times, there was a sense of mystery, that an infinite sequence of numbers could sum to be a finite number. Consider uniform motion that proceeds along the x-axis from zero to one with unit velocity. The time it takes is $t = 1$, right? But the ancient Zeno said: The cat in Figure 1.1 goes

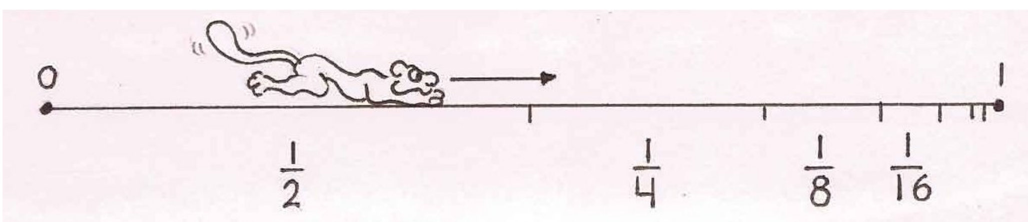

Figure 1.1

half the distance, from $x = 0$ to $x = \frac{1}{2}$, then half again the remaining distance, from $x = \frac{1}{2}$ to $x = \frac{3}{4}$, and so on. Shouldn't these individual intervals, an infinity of them, take forever? The sequence of time increments $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$ are indicated in Figure 1.1, and visual common sense suggests that they add up to one,

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 1.$$

It seems that the verbal, non-visual argument leads us astray.

From your previous elementary calculus courses, you are aware that mathematics has long reconciled itself to infinite series. Not only reconciled, but mathematics, and sciences it serves, can scarcely do without. This is because of the versatility of infinite series in representing the functions encountered in everyday applications. In particular, power series are refined instruments to discern the structures of functions locally (that is, in small intervals) by simple, easily manipulated polynomial approximations. Here are some "preview" examples to demonstrate how useful these local approximations are in practice.

Figure 1.2 shows the cross sections of a spherical mirror of radius R, and a

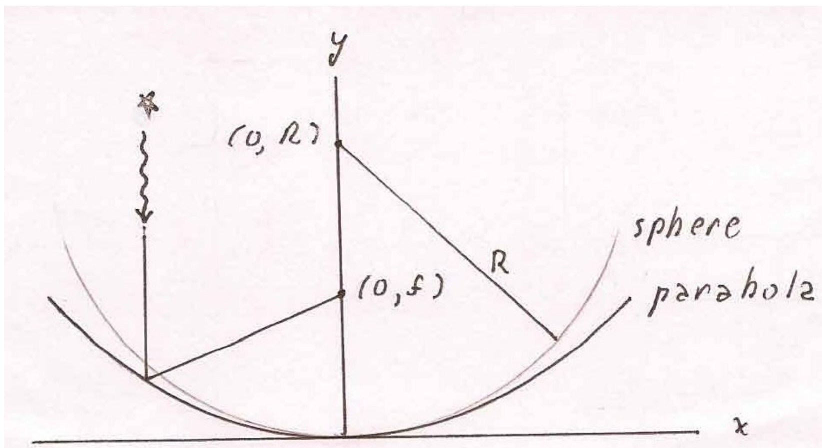

Figure 1.2

parabolic mirror of focal length f. The parabolic mirror focuses rays of light parallel to the y axis to the point $(0, f)$. This is how you want your telescope mirror to work. But the process of rubbing two glass disks over each other with abrasives in between naturally leads to a sphere. In particular, the optician grinds the initial sphere, and then "preferential polishing" refines the sphere into the parabola. This is practical only if the final "polishing to a parabola" involves the removal of very little material. First question: What is the radius R of the sphere so the sphere is a "good first approximation" to the parabola with focal length f? Next, what is the difference between the sphere and the parabola? The spherical cross section in Figure 1.2 is represented by

$$(y - R)^2 + x^2 = R^2,$$

or

$$y = R - \sqrt{R^2 - x^2} = R\left\{1 - \sqrt{1 - \left(\frac{x}{R}\right)^2}\right\}. \tag{1.1}$$

The final step of factoring out R is "good housekeeping": The dimensionless ratio $\frac{x}{R}$ appears and for most telescope mirrors, we are interested in $\frac{|x|}{R} \ll 1$. For $\frac{|x|}{R} \ll 1$, the function $\sqrt{1 - \left(\frac{x}{R}\right)^2}$ is well approximated by the first few terms of a power series (in powers of $\frac{x}{R}$),

$$\sqrt{1 - \left(\frac{x}{R}\right)^2} = 1 - \frac{1}{2}\left(\frac{x}{R}\right)^2 - \frac{1}{8}\left(\frac{x}{R}\right)^4 - \dots . \tag{1.2}$$

We'll review the construction of power series such as (1.2), but for now, accept, and see what happens: Substituting (1.2) into (1.1), we have the approximation of the sphere,

$$y = \frac{x^2}{2R} + \frac{x^4}{8R^3} + \dots . \tag{1.3}$$

The parabola in Figure 1.2 is represented by

$$y = \frac{x^2}{4f} . \tag{1.4}$$

The first term in the right-hand side of (1.3) matches the parabola if

$$R = 2f. \tag{1.5}$$

If $\frac{|x|}{R} \ll 1$, this "parabola term" of (1.3) is much greater than the second term. In fact, its much greater than the sum of all the remaining terms. Setting $R = 2f$ in (1.3) and subtracting (1.4) from (1.3) gives the difference between sphere and parabola,

$$\delta y = \frac{1}{64}\frac{x^4}{f^3} + \dots . \tag{1.6}$$

Again, the sum of the neglected terms is much less than the first term $\frac{1}{64}\frac{x^4}{f^3}$ if $\frac{|x|}{R} \ll 1$. The Mount Palomar mirror has radius $x = 100''$ and a focal length $f = 660''$. Hence, the approximate δy at the edge of the mirror amounts to

$$\delta y \simeq \frac{1}{64}\frac{(100)^4}{(660)^3} = \frac{1}{64}\left(\frac{1}{6.6}\right)^3 100'' \simeq .00543'',$$

or "about 5 mills". That's the thickness of a housepainter's plastic drop-cloth.

The second example is "Russian roulette with an n-chambered revolver". The probability that you'll get the bullet with your first squeeze of the trigger is $\frac{1}{n}$. The probability that you'll live to squeeze it again is $1 - \frac{1}{n}$. The probability that you are alive after N squeezes is $\left(1 - \frac{1}{N}\right)^N$ (provided you spin the cylinder after each squeeze!). The probability that you get killed before you completed N squeezes is

$$p = 1 - \left(1 - \frac{1}{n}\right)^N . \tag{1.7}$$

The question is: What is N so that the chances you are dead before completing N squeezes is close to 50/50? In (1.7) set $p = \frac{1}{2}$ and solve for N to find

$$N = \frac{-\log 2}{\log\left(1 - \frac{1}{n}\right)} . \tag{1.8}$$

For $n \gg 1$ $\left(\frac{1}{n} \ll 1\right)$ we evoke the power series (in powers of $\frac{1}{n}$),

$$\log\left(1 - \frac{1}{n}\right) = \left(-\frac{1}{n}\right) - \frac{1}{2}\left(-\frac{1}{n}\right)^2 + \frac{1}{3}\left(-\frac{1}{n}\right)^3 - \dots .$$

For $n \gg 1$, the first term on the right-hand side suffices, and we obtain

$$N \simeq n \log 2 \simeq .6931n.$$

What follows now is a review of fundamentals. The content is subtly different from your first exposure in calculus class: The point is to become a skilled craftsman in constructing and applying infinite series in everyday scientific calculations.

Basic definitions and examples

A *sequence* is a list of numbers in a particular order,

$$a_1, a_2, a_3, \dots , \tag{1.9}$$

and its *n-th partial sum* is

$$s_n := a_1 + a_2 + \dots a_n = \sum_1^n a_k.$$

So: starting from a sequence (1.9), we can generate the sequence

$$s_1, s_2, s_3, \dots \tag{1.10}$$

of its partial sums. The *infinite series*, traditionally written as

$$a_1 + a_2 + \cdots = \sum_1^\infty a_k \tag{1.11}$$

really stands for the sequence (1.10) of partial sums. If the sequence of partial sums converges, so

$$s := \lim_{n\to\infty} s_n \tag{1.12}$$

exists, the infinite series (1.11) is called *convergent*. Otherwise, its *divergent*.

For instance, consider the *geometric series* with $a_k = ar^k, a, r$ given numbers. The number r is called the *ratio* of the geometric series. Its partial sums are computed explicitly by a "telescoping sum trick": We have

$$\begin{aligned} s_n &= a + ar + \dots ar^{n-1} \\ rs_n &= ar + \dots ar^{n-1} + ar^n \end{aligned}$$

and subtracting the second line from the first gives

$$(1-r)s_n = a(1-r^n).$$

Hence, we have

$$s_n = \left[\begin{array}{ll} na, & r = 1, \\ a\dfrac{1-r^n}{1-r}, & r \neq 1. \end{array} \right. \tag{1.13}$$

For instance, on behalf of the cat in Figure 1.1, we can take $a = \frac{1}{2}$ and $r = \frac{1}{2}$, so $s_n = \frac{1}{2}\frac{1-\left(\frac{1}{2}\right)^n}{1-\frac{1}{2}} = 1 - \frac{1}{2^n} \to 1$ as $n \to \infty$. Indeed, the cat in Figure 1.1 runs from $x = 0$ to $x = 1$ in time 1. In general, the geometric series converges,

$$a + ar + ar^2 + \cdots = \frac{a}{1-r}, \tag{1.14}$$

for ratios r less than one in absolute value, $|r| < 1$.

As an example of a divergent geometric series, consider

$$1 + 2 + 4 + \dots. \tag{1.15}$$

Its n-th partial sum is

$$s_n = \frac{2^n - 1}{2 - 1} = 2^n - 1, \tag{1.16}$$

which clearly diverges to $+\infty$ as $n \to \infty$. Suppose we do the "telescoping sum trick" but under the delusion that (1.15) converges to some s. We have

$$s = 1 + 2 + 4 + \dots$$

and

$$2s = 2 + 4 + \dots$$

and formal subtraction of first line from the second gives the nonsense, $s = -1$.

An application: The bookkeeping of a Ponzi scheme is the telescoping sum trick. Here is the version with ratio $r = 2$. Bernie gives investors \$1, and they give Bernie a return of \$2. Bernie takes that \$2 and gives it to investors (a presumably different and bigger group) and Bernie gets a return of \$4. After n cycles of investment and return, Bernie gets

$$\begin{aligned}
&(-1+2) + (-2+4) + \dots (-2^{n-1} + 2^n) \\
&\quad = -1 + (2-2) + (4-4) + \dots (2^{n-1} - 2^{n-1}) + 2^n \\
&\quad = 2^n - 1
\end{aligned}$$

dollars. Notice that the first line is $1+2+\dots 2^{n-1} = s_n$, so this "bookkeeping" contains a rederivation of (1.16). The "gain" of the investors after n cycles is

$$\begin{aligned}
&(1-2) + (2-4) + \dots (2^{n-1} - 2^n) \\
&\quad = 1 + (-2+2) + (-4+4) + \dots (-2^{n-1} + 2^{n-1}) - 2^n \\
&\quad = 1 - 2^n
\end{aligned}$$

dollars. Big negative number.

Convergence tests

We review the standard tests for convergence/divergence of given infinite series. You've seen most of these before. The point for you now is skillful recognition of which ones are relevant for a given series, and then to administer one of them quickly and mercifully. We'd also like to see how these "simple" techniques illuminate seemingly "difficult and mysterious" examples.

The *preliminary test* establishes the divergence of certain series immediately, so no further effort is wasted on them. It is the simple observation, that if $\lim_{k\to\infty} a_k$ is non-zero or does not exist, then the series $a_1 + a_2 + \dots$ is divergent. The "common sense" argument: Suppose the series is convergent, with partial sums s_n converging to some s as $n \to \infty$. Then $a_k = s_k - s_{k-1} \to s - s = 0$ as $k \to \infty$. Hence, if a_k does anything other than converge to zero as $k \to \infty$, then the infinite series $a_1 + a_2 + \dots$ diverges.

Here is an example in which the partial sums don't run off to $+\infty$ or $-\infty$. They "just loiter around in some finite interval, but never settle down" so the infinite series is still divergent. The series is

$$\cos\theta + \cos 3\theta + \cos 5\theta + \dots ,$$

where θ is any real number. Referring to Figure 1.3, we discern that the

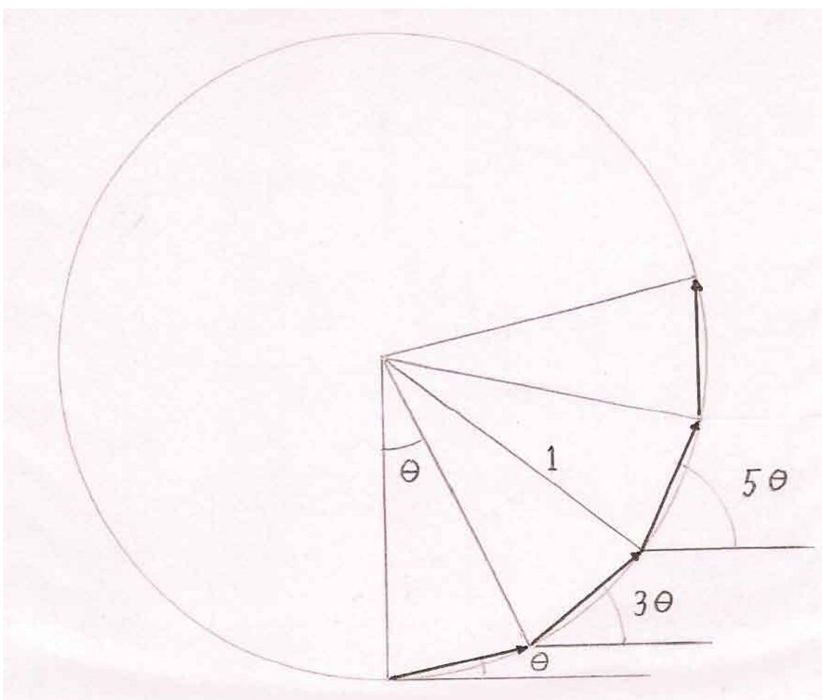

Figure 1.3

partial sums are

$$\cos\theta + \cos 3\theta + \dots \cos(2n-1)\theta = \frac{\sin 2n\theta}{2\sin\theta} . \qquad (1.17)$$

This looks complicated, but all it really amounts to is the x-component of a vector addition. The vectors in question are the directed chords inscribed in a circle of radius one. The calculation eading to (1.17) is presented as an exercise. As n increases, the partial sums (1.17) "oscillate" between $-\frac{1}{2\sin\theta}$ and $+\frac{1}{2\sin\theta}$.

Many elementary convergence/divergence tests rely on comparing a given "test" series with another series whose convergence/divergence is known. The *basic comparison test* has "convergence" and "divergence" parts. First, the convergence part: Suppose we have a convergent "comparison" series of positive terms $m_1 + m_2 + \dots$ and the test series $a_1 + a_2 + \dots$ has

$$|a_k| < m_k \tag{1.18}$$

for k *sufficiently large*. Then the test series is *absolutely convergent*, meaning that $|a_1| + |a_2| + \dots$ converges. In an exercise, it is shown that absolute convergence implies convergence. The sufficiency of the bound (1.18) for "k sufficiently large" is common sense: If for a positive integer K, the "tail" of the series, $a_{K+1} + a_{K+2} + \dots$ converges, we need only add to the sum of tail terms the first K terms, to sum the whole series. In this sense, "only the tail matters for convergence". Its a piece of common sense not to be overlooked. Very commonly, the identification of "nice simple m_k" is easiest for k large. But that's all you need.

The "divergence" part of the basic comparison test is what you think it is: This time, we assume that $|a_k| > d_k$ for k sufficiently large, where the series of non-negative d's, $d_1 + d_2 + \dots$ diverges. Then $|a_1| + |a_2| + \dots$ diverges. But now, be careful: $|a_1| + |a_2| + \dots$ divergent doesn't necessarily imply $a_1 + a_2 + \dots$ divergent.

There are two most useful corollaries of the basic comparison test which allow us to dispatch almost all "everyday business". These are the ratio and integral tests.

Ratio test

Very often you'll encounter series so that $\rho := \lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right|$ exists. This condition invites comparison with a geometric series. The rough idea is that $|a_k|$ behave like a constant times ρ^k for k large. Knowing that the geometric series $\sum_1^\infty \rho^k$ converges (diverges) for $\rho < 1$ ($\rho > 1$) we surmise that $\rho < 1$ indicates convergence of $a_1 + a_2 + \dots$, and $\rho > 1$, divergence.

Lets be a little more rigorous, so as to exercise our understanding of the basic comparison test. First, assume $\rho < 1$. Then for k greater than sufficiently large K, we'll have

$$\left|\frac{a_{k+1}}{a_k}\right| < \rho + \frac{1-\rho}{2} = \frac{1}{2} + \frac{\rho}{2} =: r < 1. \tag{1.19}$$

Figure 1.4 visualizes this inequality. From (1.19), deduce

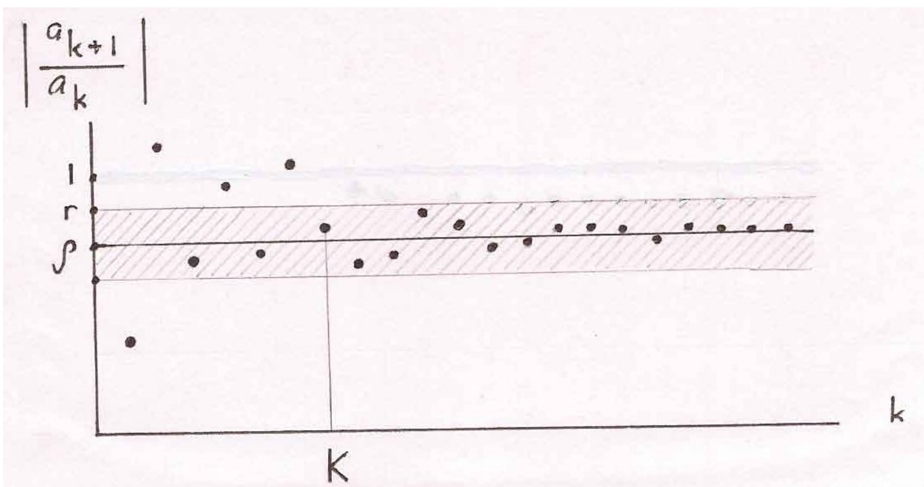

Figure 1.4

$$|a_{K+1}| < |a_K|r, |a_{K+2}| < |a_{K+1}|r, \ldots$$

and from these,

$$|a_{K+j}| < |a_K|r^j$$

for $j \geq 1$. Hence the appropriate comparison series in the "basic comparison test" has $m_k = Mr^k$, where $M := |a_K|r^{-K}$. The comparison series is geometric with ratio r, $0 < r < 1$, hence convergent. Hence, $a_1 + a_2 + \ldots$ converges. The proof that the test series diverges if $\rho > 1$ is an exercise.

Here is a juicy example of the ratio test at work: The series is

$$\sum_1^{\infty} k! e^{-(k^{1+\sigma})}, \tag{1.20}$$

with σ a positive contest. We have

$$a_k = k!e^{-(k^{1+\sigma})} = (\text{Godzilla} := k!) \div (\text{Mothra} := e^{(k^{1+\sigma})}).$$

Who wins? Godzilla or Mothra? We calculate

$$\frac{a_{k+1}}{a_k} = (k+1)e^{-\{(k+1)^{1+\sigma} - k^{1+\sigma}\}}. \tag{1.21}$$

Since $\sigma > 0$, we have a premonition that "Mothra wins" and $\frac{a_{k+1}}{a_k} \to 0$ as $n \to \infty$. We bound the exponent in (1.21) by applying the mean value theorem to $f(x) = x^{1+\sigma}$. Specifically,

$$(k+1)^\sigma - k^\sigma = f(k+1) - f(k) = f'(\zeta) = (1+\sigma)\zeta^\sigma,$$

where ζ is between k and $k+1$. Hence, $(k+1)^\sigma - k^\sigma \geq (1+\sigma)k^\sigma > k^\sigma$, and from (1.21) we deduce the inequality

$$\frac{a_{k+1}}{a_k} \leq \frac{k+1}{e^{(k^\sigma)}}\,.$$

In an exercise, repeated application of L'Hopital's rule shows that the right-hand side vanishes as $k \to \infty$. Hence, the series (1.20) converges for $\sigma > 0$. Here is a challenge for you: What is the fate (convergence or divergence) of the series

$$\sum_1^\infty k!e^{-(k^{1-\sigma})}$$

for $0 < \sigma < 1$. Mothra $:= e^{(k^{1-\sigma})}$ still diverges to $+\infty$ as $k \to \infty$, but does it diverge fast enough to overpower Godzilla $:= k!$.

If $\rho = 1$, the ratio test is inconclusive. This happens a lot. Here are two of the "usual suspects":

$$\begin{aligned} &\textit{Harmonic series} \qquad \sum_1^\infty \frac{1}{k}\,, \\ &\textit{"p" series} \qquad \sum_1^\infty \frac{1}{k^p}\,. \end{aligned} \tag{1.22}$$

These, and many others can be dispatched by the

Integral test

Often, you can identify a function $f(x)$ defined for sufficiently large, real x so that the terms a_k of the test series for k sufficiently large are given by

$$a_k = f(k).$$

Let's assume $f(x)$ is positive and non-increasing for sufficiently large x. A simple picture spells out the essence of the integral test at a glance: We see

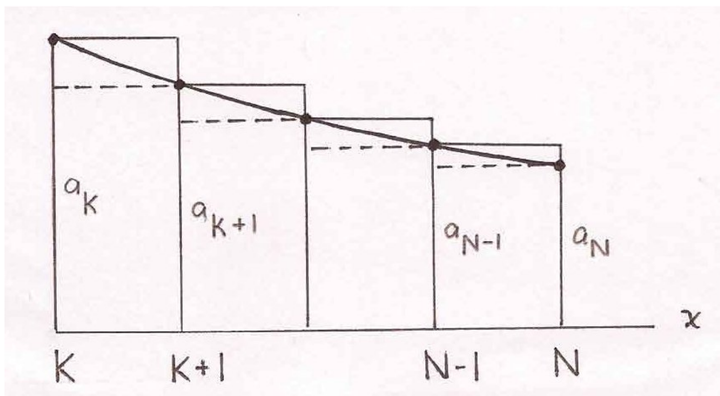

Figure 1.5

that

$$a_{K+1} + \dots a_N \leq \int_K^N f(x)dx \leq a_K + \dots a_{N-1}, \tag{1.23}$$

and its evident that the series $a_1 + a_2 + \dots$ converges (diverges) if the integral $\int_K^\infty f(x)dx$ converges (diverges).

The standard application is to the p-series (1.22). Here, $f(x) = x^{-p}$, and we calculate

$$\int_1^N x^{-p}dx = \left[\begin{array}{ll} \frac{1}{p-1}(1 - N^{1-p}), & p \neq 1, \\ & \\ \log N, & p = 1. \end{array} \right.$$

We see that the p series converges for $p > 1$, and in the special case of the harmonic series, $p = 1$, diverges. Here is a concrete example from electrostatics. We examine the electric potential and electric field at the origin due to charges Q at positions along the x-axis, $x = r, 2r, 3r, \dots$.

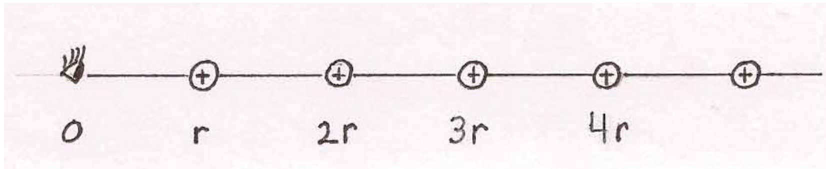

Figure 1.6

The electric potential at the origin is formally given by the series

$$U = \sum_{1}^{\infty} \frac{Q}{kr} = \frac{Q}{r} \sum_{1}^{\infty} \frac{1}{k},$$

which diverges. This means that the work required to move a charge $-Q$ from the origin to spatial infinity away from the positive x-axis is infinite. Nevertheless, the *electric field* at the origin is a well-defined convergent series,

$$\mathbf{E} = -\frac{Q}{r^2} \left(\sum_{1}^{\infty} \frac{1}{k^2} \right) \hat{\mathbf{x}}.$$

Finally, two examples of "borderline brinksmanship": For $\sum_2^\infty \frac{1}{k \log k}$, we look at

$$\int_2^\infty \frac{dx}{x \log x} = \int_{\log 2}^\infty \frac{du}{u} = +\infty,$$

and we conclude "divergence". But for $\sum_2^\infty \frac{1}{k(\log k)^2}$, we have

$$\int_2^\infty \frac{dx}{x(\log x)^2} = \int_{\log 2}^\infty \frac{du}{u^2} = \frac{1}{\log 2},$$

and now, "convergence".

Mystery and magic in the vineyard

You stand at the origin in a vineyard that covers the whole infinite $\mathbb{R}^2$. There are grape stakes at integer lattice points $\mathbf{k} = (m, n) \neq \mathbf{0}$. You look in various directions, θ, as depicted in Figure 1.7a. We say that the direction θ is *rational* if your line of sight is blocked by a stake at some "prime" lattice point $\mathbf{k}_* = (m_*, n_*)$ as depicted in Figure 1.7b. Here "prime" means that the line segment from 0 to $\mathbf{k}_*$ does not cross any other lattice point.

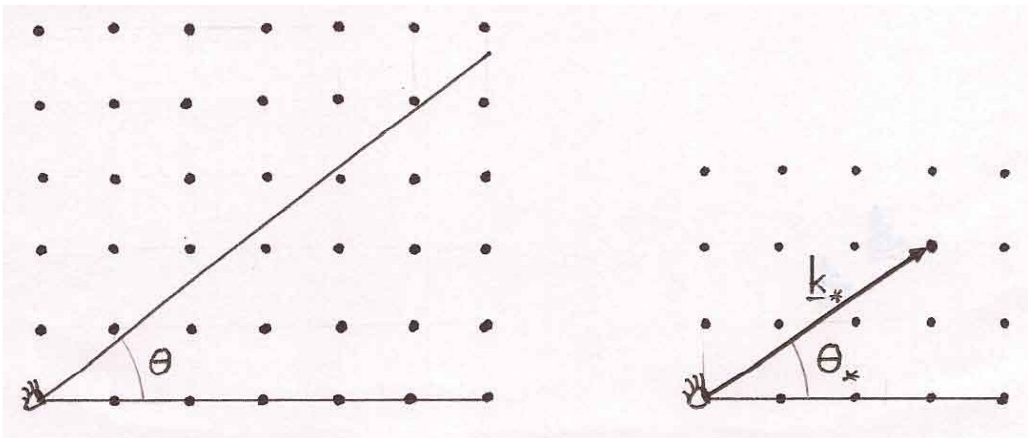

Figure 1.7

In arithmetic talk, "m_* and n_* have no common factor". We denote the direction angle of $\mathbf{k}_*$ by θ_*. If a direction θ is not rational, its called *irrational.* Your line of sight in an irrational direction continues with no obstruction to spatial infinity. The rational directions are dense on $[0, 2\pi]$, like the rational numbers. Nevertheless the union of all the rational directions is in a certain sense "negligible" relative to the union of all irrational directions. Here is what is meant: Let I_* denote the interval of angles about a given θ_*, given by

$$I_* : |\theta - \theta_*| \leq K|\mathbf{k}_*|^{-2-\sigma}, \tag{1.24}$$

where K and σ are given positive numbers. Think of the grape stake centered about $\mathbf{k}_*$ as now having a finite thickness so its angular radius as seen from the origin is $K|\mathbf{k}_*|^{-2-\sigma}$ as in (1.24). The I_* represent "blocked out directions". Consider the union of all the I_* whose $\mathbf{k}_*$ lie within distance R from the origin. What fraction of all the directions in $[0, 2\pi]$ do they block out? Due to possible intersections of different I_*'s, this fraction f has the upper bound,

$$f \leq \frac{K}{2\pi} \sum_{|\mathbf{k}_*|<R} |\mathbf{k}_*|^{-2-\sigma}.$$

Extending the summation on the right-hand side from prime lattice points $\mathbf{k}_*$ to all $\mathbf{k} \neq 0$ in $|\mathbf{k}| < R$ gives the bigger upper bound

$$f < \frac{K}{2\pi} \sum_{|\mathbf{k}|<R} |\mathbf{k}|^{-2-\sigma}. \tag{1.25}$$

Finally we bound *this* sum from above by a double integral, like in the integral test. We get

$$f < \frac{K}{2\pi} \int_1^R r^{-2-\sigma} 2\pi r dr = \frac{K}{\sigma}(1 - R^{-\sigma}) < \frac{K}{\sigma}\,, \tag{1.26}$$

for *all* $R > 1$. Now the magic and mystery begins: For $K < \sigma$, the fraction of blotted out directions must be less than one. Even though the θ_* are dense, and around each an interval of "blocked out" directions, there are still "gaps" through which lines of sight "escape to ∞". And even more than this: As $K \to 0$, the fraction of obstructed directions goes to zero, and almost all lines of sight reach ∞.

Alternating series

are characterized by any two successive terms having opposite signs. For instance, let's replace the positive charges at $x = 2r, 4r, 6r, \ldots$ in Figure 1.6 by negative changes. The electric potential at the origin is formally an alter-

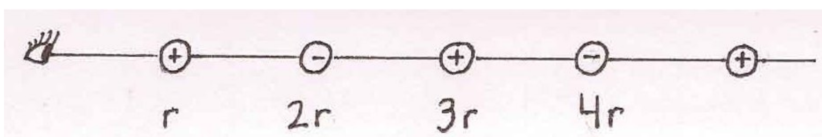

Figure 1.8

nating series,

$$U = \frac{Q}{r}\left\{1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots\right\}. \tag{1.27}$$

In general, an alternating series converges if the $|a_k|$ are monotone decreasing for k sufficiently large, and $\lim_{k\to\infty} a_k = 0$. The proof is an exercise. Here, we discuss the subtle business of rearranging series. Rearrangement means putting the numbers a_k of the sequence (1.9) in a different order, and then computing the partial sums of the reordered series. If the series is absolutely convergent, rearrangement doesn't change the sum. Not so for some alternating series. For instance, let's do partial sums of the series (1.27) like this:

$$s_n := \left(1 + \frac{1}{3} + \ldots \frac{1}{2N-1}\right) - \left(\frac{1}{2} + \frac{1}{4} + \ldots \frac{1}{2n}\right).$$

The first parentheses contain the first $N = N(n)$ positive terms, where $N(n)$ is an increasing function of n so $N(n) \to \infty$ as $n \to \infty$. The second parentheses contain the first n negative terms. In the limit $n \to \infty$, we cover all the terms in the series (1.27). We have

$$1 + \frac{1}{3} + \dots \frac{1}{2N-1} > \frac{1}{2} + \frac{1}{4} + \dots \frac{1}{2N} = \sum_{k=1}^{N} \frac{1}{2k},$$

so

$$s_n > \sum_{k=1}^{N} \frac{1}{2k} - \sum_{k=1}^{n} \frac{1}{2k} = \sum_{n+1}^{N} \frac{1}{2k} > \int_{n+1}^{N+1} \frac{dx}{2x}$$
$$= \frac{1}{2} \log \frac{N+1}{n+1}.$$

We have $s_n \to \infty$ as $n \to \infty$ if $\frac{N(n)}{n} \to +\infty$ as $n \to \infty$, and $s_n \to -\infty$ if $\frac{N(n)}{n} \to 0$ as $n \to \infty$. More generally, we can "cherry pick" the order of terms to get partial sums which converge to any number you want. This is true as mathematics, but *not* relevant to the physical example of alternating charges along a line. In particular, you expect that the "monofilament crystal" in Figure 1.8 ends after something on the order of 10^{23} charges. Hence, it is the standard alternating series with *no* rearrangement that matters.

Power series
take the form

$$\sum_{0}^{\infty} a_n x^n \tag{1.28}$$

where $a_0, a_1, a_2, \dots$ is a given sequence of constants, and x is a real variable. If the power series converges for x in some interval, then its sum represents a function $f(x)$ of x in that interval. The first order of business is to establish its interval of convergence. Garden variety power series from everyday applications are often characterized by the existence of

$$R := \lim_{k \to \infty} \left| \frac{a_k}{a_{k+1}} \right|. \tag{1.29}$$

Applying the ratio test, with $a_k x^k$ replacing a_k, we calculate

$$\rho := \lim_{k \to \infty} \left| \frac{a_{k+1} x^{k+1}}{a_k x^k} \right| = \frac{|x|}{R},$$

so the power series converges if $|x| < R$, and diverges if $|x| > R$. R is called the *radius of convergence*. The endpoints $x = -R, +R$ with $\rho = 1$ have to be checked individually by some other methods. For instance, consider

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots = \sum_1^\infty \frac{(-1)^{n-1}}{n} x^n. \tag{1.30}$$

We have $a_n = \frac{(-1)^{n-1}}{n}$, and $R = 1$, so there is convergence in $|x| < 1$ and divergence in $|x| > 1$. At the endpoints $x = +1$, we have

$$1 - \frac{1}{2} + \frac{1}{3} - \ldots,$$

a convergent alternating series. At $x = -1$, we have

$$-1 - \frac{1}{2} - \frac{1}{3} - \ldots,$$

a divergent harmonic series. Figure 1.9 visualizes the interval of convergence.

Figure 1.9

If you replace x in the series (1.30) by some function $y(x)$, you have to figure out the domain of x-values so the range of y values is in the “half dumbbell” in Figure 1.9. For instance, if $y = 3x + 2$, we have convergence for x so $-1 < 3x + 2 \leq 1$ or $\frac{1}{3} < x \leq 1$.

Elementary manipulations of power series

You can construct power series of a great many functions by manipulating easily remembered power series of a few elementary functions. The main ones are e^x, $\cos x$, $\sin x$, the geometric series for $\frac{1}{1+x}$, and more generally, the *binomial* $(1 + x)^p$, $p =$ constant.

The most elementary manipulations are algebraic: In their common interval of convergence, we can add and subtract two series termwise. Multiplication is like “multiplying endless polynomials”. We can express the quotient of two series as a series.

We present some techniques for multiplication and division of series. Suppose you want to compute a series that represents the product of two series $\sum_0^\infty a_n x^n$ and $\sum_0^\infty b_n x^n$. First, you convert the product into a "double sum":

$$\left(\sum_{m=0}^{\infty} a_m x^m\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} a_m b_n x^{m+n}. \tag{1.31}$$

In the left-hand side, notice that we used different letters m and n for indices of summation, so you don't confuse the *two* summation processes. In the right-hand side we rearrange the double sum: For each non-negative integer N we first sum over m and n so $m+n=N$, and then sum over N. For instance, in Figure 1.9, we've highlighted the (m,n) with $m+n=3$, and we see that the coefficient of x^3 in the product series is $a_0a_3+a_1a_2+a_2a_1+a_3a_0 = 2a_0a_3+2a_1a_2$. The rearrangement converts the right-hand side into

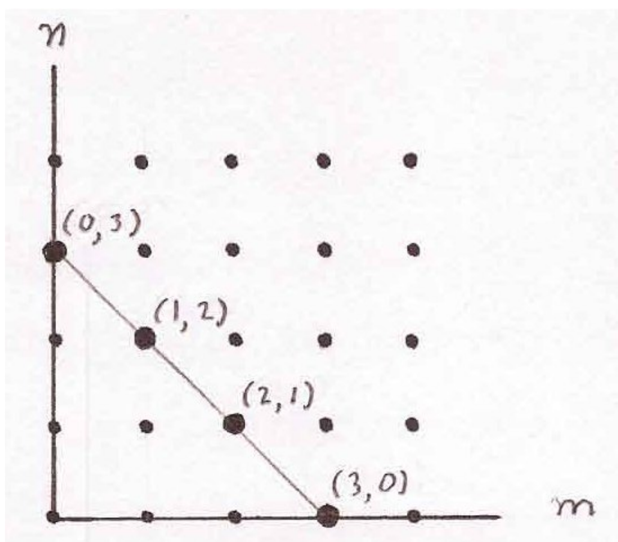

Figure 1.10

$$\sum_{N=0}^{\infty}\left(\sum_{m=0}^{N} a_m b_{N-m}\right) x^N. \tag{1.32}$$

Here is an example: In the "telescope mirror" example that begins on page 1, we used the first three terms of the series for $\sqrt{1+x}$,

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \dots\,. \tag{1.33}$$

We use the "rearrangement" technique to deduce the coefficients $\frac{1}{2}$ and $-\frac{1}{8}$ in (1.33). Let a_m represent the coefficients of the series for $\sqrt{1+x}$. In (1.31), (1.32) take $b_m = a_m$, and we have

$$1 + x = \left(\sum_{0}^{\infty} a_m x^m \right)^2 = \sum_{N=0}^{\infty} \left(\sum_{m=0}^{\infty} a_m a_{N-m} \right) x^N.$$

Equating coefficients of the far right-hand side and left-hand side, we have

$$\begin{aligned} 1 &= a_0^2, \\ 1 &= a_0 a_1 + a_1 a_0 = 2a_0 a_1, \\ 0 &= a_0 a_2 + a_1 a_1 + a_2 a_0 = 2a_0 a_2 + a_1^2. \end{aligned}$$

The recursive solution of these equations give:

$$\begin{aligned} a_0 &= +1 \text{ (for \textit{positive} square root)} \\ 1 = 2a_1 &\Rightarrow a_1 = \frac{1}{2}, \\ 2a_2 + \left(\frac{1}{2}\right)^2 = 0 &\Rightarrow a_2 = -\frac{1}{8}. \end{aligned}$$

Division of series can be formulated like the "long division" you do in elementary school. Here is the construction of the geometric series for $\frac{1}{1-x}$ by "long division":

$$\begin{array}{rl} & \underline{\;1 + x + x^2 + \ldots\qquad\qquad\qquad} \\ 1 - x & |\,1 + 0x + 0x^2 + 0x^2 + \ldots \\ & \underline{1 - x} \\ & \quad x + 0x^2 \\ & \quad \underline{x - x^2} \\ & \qquad x^2 + 0x^3 \\ & \qquad \underline{x^2 - x^3} \\ & \qquad\quad x^3 + 0x^4 \\ & \qquad\quad \vdots \end{array} \tag{1.34}$$

A second example computes the first four terms in the reciprocal of the "full" series

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots . \tag{1.35}$$

We have

$$
\begin{array}{r|rrrrrrrrr}
 & 1 & - & x & + & \frac{x^2}{2} & - & \frac{x^3}{0} & + & \dots \\
\hline
1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \dots & 1 & + & 0x & + & 0x^2 & + & 0x^3 & + & \dots \\
 & 1 & + & x & + & \frac{x^2}{2} & + & \frac{x^3}{6} & + & \dots \\
\cline{2-10}
 & & & -x & - & \frac{x^2}{2} & - & \frac{x^3}{6} & - & \dots \\
 & & & -x & - & x^2 & - & \frac{x^3}{2} & - & \dots \\
\cline{4-10}
 & & & & & \frac{x^2}{2} & + & \frac{x^3}{3} & - & \dots \\
 & & & & & \frac{x^2}{2} & + & \frac{x^3}{2} & + & \dots \\
\cline{6-10}
 & & & & & & - & \frac{x^3}{6} & + & \dots \\
 & & & & & & & \vdots & &
\end{array}
\tag{1.36}
$$

It looks like the "reciprocal" series is obtained by replacing x in the original series (1.35) by $-x$. In fact this is true. The series (1.35) *is* special, and you will be reminded of its meaning soon.

Substitution is replacing x in a power series by a function of x, most commonly $-x$, x^2 or $\frac{1}{x}$. For instance replacing x by x^2 in the geometric series $1 + x + x^2 + \dots$ for $\frac{1}{1-x}$, we see right away that

$$\frac{1}{1-x^2} = 1 + (x^2) + (x^2)^2 + \dots = 1 + x^2 + x^4 + \dots .$$

You could have done partial fractions, so

$$\frac{1}{1-x^2} = \frac{1}{2}\frac{1}{1-x} + \frac{1}{2}\frac{1}{1+x}$$

and evoke the geometric series for $\frac{1}{1-x}$ and $\frac{1}{1+x}$. Not too hard but clearly less efficient. Here is another quintessential example: It is easy to see that $\frac{1}{1+x} \simeq \frac{1}{x}$ for large $|x|$. But how do we generate successive refined approximations? Since we are examining $|x| > 1$ we can't do the standard geometric series in powers of x. Instead, write

$$\frac{1}{1+x} = \frac{1}{x}\frac{1}{1+\frac{1}{x}} = \frac{1}{x}\left\{1 - \frac{1}{x} + \frac{1}{x^2} + \dots\right\} = \frac{1}{x} - \frac{1}{x^2} + \frac{1}{x^3} - \dots .$$

For $|x| > 1$, the right-hand side is a nice convergent power series in powers of $\frac{1}{x}$.

Finally, there is "inversion": We start with a function $y(x)$ whose power series has $a_0 = 0$, $a_1 \neq 0$. For instance, take

$$y = x - \frac{x^2}{2}\,. \tag{1.37}$$

To find the inverse function in a neighborhood of $x = 0$, interchange x and y and "solve" for y: We have

$$x = y - \frac{y^2}{2} \tag{1.38}$$

and the solution of this quadratic equation for y given x with $y = 0$ at $x = 0$ is

$$y = 1 - \sqrt{1 - 2x}, \tag{1.39}$$

apparently for $x < \frac{1}{2}$. Figure 1.11 shows what's going on: The darkened curve

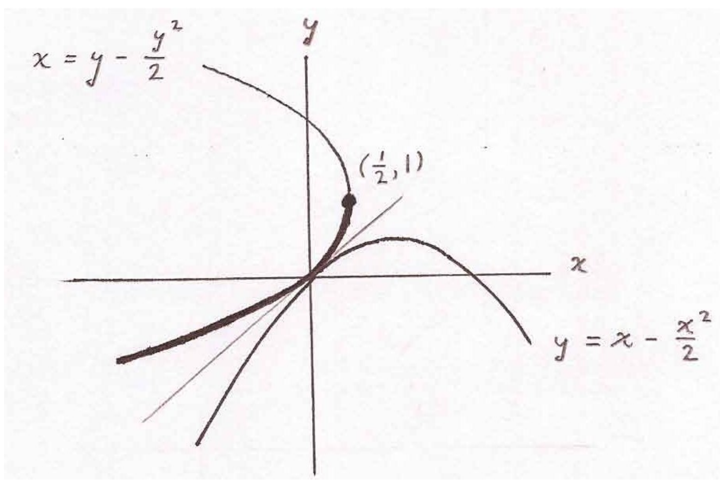

Figure 1.11

segment is the graph of the inverse function whose series we seek. We're going to ignore the explicit formula (1.39) since in real life we an rarely find explicit formulas for the inverse function. Instead substitute into (1.38) the series for y in powers of x,

$$y = b_1 x + b_2 x^2 + b_3 x^3 + \dots\,.$$

We obtain

$$\begin{aligned} x &= (b_1 x + b_2 x^2 + b_3 x^3 + \dots) - \frac{1}{2}(b_1 x + b_2 x^2 + \dots)^2 \\ &= (b_1 x + b_2 x^2 + b_3 x^3 + \dots) - \frac{1}{2}x^2(b_1 + b_2 x + + \dots)^2 \\ &= b_1 x + \left(b_2 - \frac{1}{2}b_1^2\right) x^2 + (b_3 - b_1 b_2) x^3 + \dots \end{aligned}$$

and equating coefficients of powers of x, we have

$$\begin{aligned} 1 &= b_1 \\ 0 &= b_2 - \frac{1}{2}b_1^2 \Rightarrow b_2 = \frac{1}{2} \\ 0 &= b_3 - b_1 b_2 \Rightarrow b_3 = \frac{1}{2} \end{aligned}$$

and the first three terms in the series of the inverse function are

$$y = x + \frac{x^2}{2} + \frac{x^3}{3} + \dots .$$

You might have noticed that these techniques for multiplication, division and inversion of series are often practical, only for computing the first few terms. That's just the way it is. But for practical scientists, "the first few terms" are almost always sufficient.

The two main "calculus manipulations" are termwise differentiation and integration, valid in the interval of convergence. For instance, termwise differentiation of the geometric series

$$\frac{1}{1+x} = 1 - x + x^2 - \dots \tag{1.40}$$

in $|x| < 1$ gives

$$\frac{1}{(1+x)^2} = 1 - 2x + 3x^2 - \dots$$

also in $|x| < 1$. You could "square the sum" in (1.40), and grind away, but this is much more tedious. Integration leads to a result that no *algebraic* manipulation can achieve:

$$\log(1+x) = \int_0^x \frac{dt}{1+t} = \int_0^x (1 - t + t^2 - \dots) dt = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \tag{1.41}$$

in $|x| < 1$. We recognize the series on the right-hand side from (1.30). In particular, the series converges at the endpoint $x = 1$ and setting $x = 1$ in (1.41) formally leads to

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \dots,$$

so the electric potential in (1.27) is $\frac{Q \log 2}{r}$. Turns out this is right.

Analytic functions

Suppose the coefficients a_k of the power series (1.28) satisfy (1.29) for some $R > 0$, so its sum is a function $f(x)$ in $|x| < R$,

$$f(x) = \sum_0^\infty a_n x^n. \tag{1.42}$$

Term by term differentiation gives

$$f'(x) = \sum_1^\infty k a_k x^{k-1} = \sum_0^\infty (k+1) a_{k+1} x^k$$

and the ratio test easily establishes the convergence of the $f'(x)$ series in $|x| < R$. Successive termwise differentiations produce power series for all derivatives, all converging in $|x| < R$. Evidently $f(x)$ has derivatives of all orders in $|x| < R$, and for this reason is called *analytic* in $|x| < R$. Taking n derivatives of $f(x)$ in (1.42) and setting $x = 0$, we identify the coefficients a_n in terms of derivatives of f evaluated at $x = 0$: First observe

$$(x^k)^{(n)}(0) = \begin{bmatrix} 0, & k \neq n, \\ \\ n!, & k = n \end{bmatrix}$$

so we deduce from (1.42) that

$$f^{(n)}(0) = \sum_0^\infty a_k (x^k)^{(n)}(0) = a_n n!,$$

and then

$$a_n = \frac{f^{(n)}(0)}{n!}.$$

Hence, $f(x)$ has the *Taylor series* in $|x| < R$,

$$f(x) = \sum_{0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k. \tag{1.43}$$

If successive derivatives of $f(x)$ are easy to compute explicitly, then (1.43) is an excellent tool to generate the power series of $f(x)$. Such is the case for e^x, $\cos x$, and $\sin x$. Their Taylor series are embedded in your brain at the molecular level, or you'll review them and as Captain Picard says, you'll "make it so".

Here, we review in detail the binomial series for $(1+x)^p$, $p =$ real number. It has an interesting historical narrative: By pure algebra, its been long known that for $p = n =$ integer, that

$$(1+x)^n = \sum_{0}^{n} \binom{n}{k} x^k, \tag{1.44}$$

where

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} \tag{1.45}$$

are *binomial coefficients*. For n not too large, the fastest way to generate the polynomial on the right-hand side of (1.44) is the *Pascal triangle* in Figure 1.12:

row											
0						1					
1					1		1				
2				1		2		1			
3			1		3		3		1		
4		1		4		6		4		1	
5	1		5		10		10		5		1

Figure 1.12

The fifth row lists the binomial coefficients $\binom{5}{k}$, $k = 0, 1, 2, 3, 4, 5$, and we have

$$(1+x)^5 = 1 + 5x + 10x^2 + 10x^3 + 5x^4 + x^5.$$

Lets rewrite the binomial coefficient in (1.44) as

$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!} \,. \tag{1.46}$$

Notice that for $k > n$, the numerator always has the factor $n - n$, so $\binom{n}{k} \equiv 0$ for $k > n$.[1] Now Mr. Newton says: "In (1.46), let's replace the integer n by $p =$ any non-integer real number, and define for $k =$ integer,

$$\binom{p}{k} := \frac{p(p-1)\dots(p-k+1)}{k!} \,.$$

It now appears that the binomial expansion in (1.44) becomes

$$(1+x)^p = \sum_1^\infty \binom{p}{k} x^k = 1+px+\frac{p(p-1)}{2!}x^2+\frac{p(p-1)(p-2)}{3!}x^3+\dots \,, \tag{1.47}$$

and the series on the right-hand side "goes on forever". Nowadays you'd just recognize that $\binom{p}{k}$ is just the k-th derivative of $(1+x)^p$ evaluated at $x = 0$, and divided by $k!$, and that the right-hand side of (1.47) is the Taylor series of $(1+x)^p$.

Often we require local approximations to an analytic function in a small neighborhood of some $x = a \neq 0$. In this case, its natural to replace powers of x by powers of $x - a$, so we obtain a representation of $f(x)$,

$$f(x) = \sum_0^\infty \frac{f^{(k)}(a)}{k!}(x-a)^k \tag{1.48}$$

for $|x - a| < R =$ positive constant. (1.48) is called "*the Taylor series of $f(x)$ about $x = a$*". Here's an example: Define

$$f(x) := \sum_0^\infty \frac{x^n}{n!} \,.$$

You easily check that the power series on the right-hand side converges for all x. Pretend like you don't know that $f(x) = e^x$, and you're going to investigate it. Differentiating term by term, you find

$$f'(x) = \sum_1^\infty \frac{x^{n-1}}{(n-1)!} = \sum_0^\infty \frac{x^n}{n!} = f(x), \tag{1.49}$$

[1] A neat gag for algebra teachers: $(x-a)(x-b)(x-c)\dots(x-z) = ?$

for all x. Now evaluate the Taylor series about $x = a$, with the help of (1.49):

$$\begin{aligned} f(x) &= \sum_0^\infty \frac{f^{(k)}(a)}{k!}(x-a)^k \\ &= f(a)\sum_0^\infty \frac{(x-a)^k}{k!} = f(a)f(x-a). \end{aligned}$$

Now set $b = x - a$, and we have

$$f(a+b) = f(a)f(b). \tag{1.50}$$

Next, you recall the exponentiation rule, $c^{a+b} = c^a c^b$, for c a positive real number, and a, b integers. Then you argue it for a, b rational, and due to sequences of rationals converging to my real number, you believe it for a, b real. Finally, you sense it in the Force that $f(x)$ in (1.50) is "some positive constant to the x-th power".

Skillful construction of Power series

The Taylor series formula (1.43) can be a terrible temptation: "Its all there in a nutshell. Just take the given $f(x)$ and generate the values of $f(0), f'(0), f''(0), \ldots$ like sausages, and stuff them into the Taylor series, and serve it up." Most of the time: Not so! For instance, look at

$$\begin{aligned} f(x) &= \operatorname{arctanh} x, \\ f'(x) &= \frac{1}{1-x^2}\,, \\ f''(x) &= \frac{2x}{(1-x^2)^2}\,, \\ f'''(x) &= \frac{2}{(1-x^2)^2} - \frac{8x^2}{(1-x^2)^3}\,. \end{aligned}$$

Had enough? It would be much better if you stopped at the formula for $f'(x)$, did the geometric series, followed by integration. This leads to

$$\operatorname{arctanh} x = \int_0^x (1 + t^2 + t^4 + \ldots)dt = x + \frac{x^3}{3} + \frac{x^5}{5} + \ldots$$

in $|x| < 1$. Let's try another,

$$f(x) := \log\sqrt{\frac{1+x}{1-x}} \tag{1.51}$$

in $|x| < 1$. Again, please don't do $f^{(k)}(0)$ by hand. Instead, in $|x| < 1$,

$$\begin{aligned} f(x) &= \frac{1}{2}\log(1+x) - \frac{1}{2}\log(1-x) \\ &= \frac{1}{2}\int_0^x \left(\frac{1}{1+t} + \frac{1}{1-t}\right) dt \\ &= \int_0^x \frac{dt}{1-t^2} \\ &= \int_0^x (1 + t^2 + t^4 + \dots) dt \\ &= x + \frac{x^3}{3} + \frac{x^5}{5} + \dots \\ &= \operatorname{arctanh} x. \end{aligned}$$

Actually we could have recognized $\operatorname{arctanh} x$ by the third line. Notice that we can also get the log formula (1.51) by inversion. That is, solve $x = \tanh y = \frac{e^y - e^{-y}}{e^y + e^{-y}}$ for y in terms of x. There is a quadratic equation for e^y and then use $y = \log e^y$.

The asymptotic character of Taylor series

Don't be mislead by the pretty "blackboard examples" in which complete Taylor series are quickly whipped up by some (usually short) sequence of clever techniques. The bad news: "Not so in real life." The good: "You usually don't need to." Now why is that?

Let's investigate the *truncation error* due to replacing the whole Taylor series by the K-th degree *Taylor polynomial*

$$P_K(x) := f(0) + f'(0)x + \dots \frac{f^{(K)}(0)}{K!}x^K.$$

The elegant form of truncation error comes from the *generalized mean value theorem*, which says: For $|x| < R =$ radius of convergence,

$$f(x) - P_K(x) = \frac{f^{(K+1)}(\zeta)}{(K+1)!}x^{K+1} \tag{1.52}$$

where ζ is between 0 and x. The analysis behind (1.52) is more intricate than anything we've been doing so far. Nevertheless its so useful we'll just grab it and run. In particular, it explains why we can neglect the $k > K$ tail of Taylor series for $\frac{|x|}{R}$ sufficiently small.

Restrict x to a subinterval of $|x| < R$, say $|x| < \frac{R}{2}$. Since $f^{(K+1)}(x)$ is analytic in $|x| < R$, $\frac{|f^{(K+1)}(x)|}{K!}$ has some positive upper bound M in $|x| < \frac{R}{2}$ and it follows from (1.52) that

$$|f(x) - P_K(x)| < M|x|^{K+1}. \tag{1.53}$$

In words: "The truncation error in approximating $f(x)$ by its K-th degree Taylor polynomial has an upper bound that scales with $|x|$ like $|x|^{K+1}$, just like the "next term" $\frac{f^{(K+1)}(0)}{(K+1)!}x^{K+1}$. Also, notice that as $|x| \to 0$, the upper bound $M|x|^{K+1}$ on error goes to zero *faster* than the "*smallest non-zero term*" of $P_K(x)$, which is $\frac{f^{(L)}(0)}{L!}x^L$, for some L, $0 \leq L \leq K$. For this reason, the Taylor polynomials $P_K(x)$ are called *asymptotic expansions* of $f(x)$ as $|x| \to 0$.

Asymptotic expansions represent a sense of approximation different from convergence. Let's contrast the two verbally:

> *Convergence*: "For *fixed* x in $|x| < R$, take K *big enough* so $P_K(x)$ is 'close enough' to $f(x)$."

> *Asymptotic expansion*: For *fixed* K, $P_K(x)$ approximates $f(x)$ as $|x| \to 0$, with an error that vanishes as $|x| \to 0$ faster than the smallest term in $P_K(x)$.

Asymptotic approximations tend to be "practical" because very often, "only a few terms suffice". In a sense this is "nothing new". For a long time you've been doing approximations like $e^x \simeq 1 + x$, $\sin x \simeq x$, etc. What's changed is an *explicit consciousness* of what it really means.

"O" notation

In practical approximations, we don't want to be bogged down by long derivations of rigorous error bounds. We want quick order of magnitude estimates. In particular, the most important aspect of the error bound (1.53) is the power of x. This is expressed by a kind of shorthand,

$$f(x) = P_K(x) + O(x^{K+1}), \tag{1.54}$$

which means: There is a positive constant M so that

$$|f(x) - P_K(x)| < M|x|^{K+1}$$

for $|x|$ sufficiently small. More generally, we say that

$$f(x) = O(g(x))$$

as $x \to 0$ if for $|x|$ sufficiently small, there is a positive constant M so

$$|f(x)| < M|g(x)|.$$

Sometimes you *do* want a decent idea of how big M is. For Taylor polynomials, we see from the generalized mean value theorem (1.52) that "M is close to $\frac{|f^{(K+1)}(0)|}{(K+1)!}$" for $|x|$ sufficiently small. Here is an example that shows the typical use of Taylor polynomials and the "O" notation.

Difference equation approximations to ODE

Computer analysis of ODE often proceeds by replacing the original ODE by a *difference equation*. For instance consider the initial value problem

$$\begin{aligned} y'(x) &= y(x) \text{ in } x > 0, \\ y(0) &= 1, \end{aligned} \tag{1.55}$$

whose solution is $y(x) = e^x$. The simplest difference equation approximation to (1.55) goes like this: y_k denotes the approximation to $y(x = kh)$, where h is a given "stepsize" and k assumes integer values $k = 0, 1, 2, \dots$. The derivative $y'(x = kh)$ is approximated by the *difference quotient* $\frac{y_{k+1} - y_k}{h}$ and the ODE at $x = kh$ is approximated by the *difference equation*

$$\frac{y_{k+1} - y_k}{h} = y_k$$

or

$$y_{k+1} = (1 + h)y_k, \ k = 0, 1, 2, \dots . \tag{1.56}$$

The solution for the y_k subject to $y_0 = 1$ is

$$y_k = (1 + h)^k. \tag{1.57}$$

Lets suppose we want to construct an approximation to the exact solution $y(x) = e^x$ at a particular $x > 0$. Starting from $x = 0$, we'll reach the given

$x > 0$ in N iterations if $h = \frac{x}{N}$, and the approximation to e^x is going to be (1.57) with $h = \frac{x}{N}$ and $k = N$:

$$y_N = \left(1 + \frac{x}{N}\right)^N. \tag{1.58}$$

In general, replacing ODE by difference equations introduces inevitable error. For our example here, the error is

$$\delta y = \left(1 + \frac{x}{N}\right)^N - e^x. \tag{1.59}$$

Since we examine the limit $N \to \infty$ with $x > 0$ fixed, we are seeking a Taylor polynomial approximation to δy *in powers of* $\frac{1}{N}$. First observe that

$$\left(1 + \frac{x}{N}\right)^N = e^{N \log\left(1 + \frac{x}{N}\right)}.$$

The exponent has Taylor series (in powers of $\frac{1}{N}$)

$$\begin{aligned} N \log\left(1 + \frac{x}{N}\right) &= N\left\{\frac{x}{N} - \frac{1}{2}\left(\frac{x}{N}\right)^2 + \frac{1}{3}\left(\frac{x}{N}\right)^3 - \dots\right\} \\ &= x - \frac{x^2}{2N} + \frac{x^3}{3N^2} - \dots . \end{aligned}$$

Hence,

$$\left(1 + \frac{x}{N}\right)^N = e^x e^{-\frac{x^2}{2N} + \frac{x^3}{3n^2} - \dots}. \tag{1.60}$$

Notice that the second exponential converges to one as $N \to \infty$, so we already see that $\left(1 + \frac{x}{N}\right)^N$ converges to e^x as $N \to \infty$. Next, we evoke the Taylor series for the exponential,

$$e^h = 1 + h + \frac{h^2}{2} + \dots$$

with $h := -\frac{x^2}{2N} + \frac{x^3}{3N^2} - \dots$, and (1.60) becomes

$$\left(1 + \frac{x}{N}\right)^N = e^x\left\{1 + \left(-\frac{x^2}{2N} + \frac{x^3}{3N^2} - \dots\right) + \frac{1}{2}\left(-\frac{x^2}{2N} + \dots\right)^2 + \dots\right\}.$$

The Taylor series in parenthesis looks like a mess but in order to resolve the $\frac{1}{N}$ component of δy, *we need only the first two terms*, so we write

$$\left(1+\frac{x}{N}\right)^N = e^x\left\{1-\frac{x^2}{2N}+O\left(\frac{1}{N^2}\right)\right\}. \tag{1.61}$$

Inserting (1.61) into (1.59), we deduce

$$\delta y = -\frac{x^2 e^x}{2N}+O\left(\frac{1}{N^2}\right). \tag{1.62}$$