

Asymptotics

EXACT ANSWERS are great when we can find them; there's something very satisfying about complete knowledge. But there's also a time when approximations are in order. If we run into a sum or a recurrence whose solution doesn't have a closed form (as far as we can tell), we still would like to know something about the answer; we don't have to insist on all or nothing. And even if we do have a closed form, our knowledge might be imperfect, since we might not know how to compare it with other closed forms.

For example, there is (apparently) no closed form for the sum

$$S_n = \sum_{k=0}^n \binom{3n}{k}.$$

But it is nice to know that

$$S_n \sim 2 \binom{3n}{n}, \quad \text{as } n \rightarrow \infty;$$

Uh oh ... here comes that A-word.

we say that the sum is “asymptotic to” $2 \binom{3n}{n}$. It's even nicer to have more detailed information, like

$$S_n = \binom{3n}{n} \left(2 - \frac{4}{n} + O\left(\frac{1}{n^2}\right) \right), \quad (9.1)$$

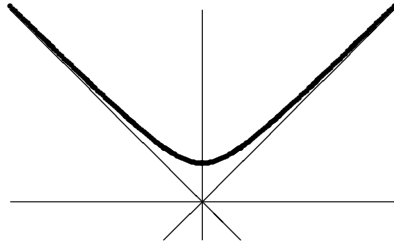
which gives us a “relative error of order $1/n^2$.” But even this isn't enough to tell us how big S_n is, compared with other quantities. Which is larger, S_n or the Fibonacci number F_{4n} ? Answer: We have $S_2 = 22 > F_8 = 21$ when $n = 2$; but F_{4n} is eventually larger, because $F_{4n} \sim \phi^{4n}/\sqrt{5}$ and $\phi^4 \approx 6.8541$, while

$$S_n = \sqrt{\frac{3}{\pi n}} (6.75)^n \left(1 - \frac{151}{72n} + O\left(\frac{1}{n^2}\right) \right). \quad (9.2)$$

Our goal in this chapter is to learn how to understand and to derive results like this without great pain.

The word *asymptotic* stems from a Greek root meaning “not falling together.” When ancient Greek mathematicians studied conic sections, they considered hyperbolas like the graph of $y = \sqrt{1+x^2}$,

Other words like ‘symptom’ and ‘ptomaine’ also come from this root.



which has the lines $y = x$ and $y = -x$ as “asymptotes.” The curve approaches but never quite touches these asymptotes, when $x \rightarrow \infty$. Nowadays we use “asymptotic” in a broader sense to mean any approximate value that gets closer and closer to the truth, when some parameter approaches a limiting value. For us, asymptotics means “almost falling together.”

Some asymptotic formulas are very difficult to derive, well beyond the scope of this book. We will content ourselves with an introduction to the subject; we hope to acquire a suitable foundation on which further techniques can be built. We will be particularly interested in understanding the definitions of ‘ \sim ’ and ‘ O ’ and similar symbols, and we’ll study basic ways to manipulate asymptotic quantities.

9.1 A HIERARCHY

Functions of n that occur in practice usually have different “asymptotic growth ratios”; one of them will approach infinity faster than another. We formalize this by saying that

$$f(n) \prec g(n) \iff \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0. \quad (9.3)$$

This relation is transitive: If $f(n) \prec g(n)$ and $g(n) \prec h(n)$ then $f(n) \prec h(n)$. We also may write $g(n) \succ f(n)$ if $f(n) \prec g(n)$. This notation was introduced in 1871 by Paul du Bois-Reymond [85].

All functions great and small.

For example, $n \prec n^2$; informally we say that n grows more slowly than n^2 . In fact,

$$n^\alpha \prec n^\beta \iff \alpha < \beta, \quad (9.4)$$

when α and β are arbitrary real numbers.

There are, of course, many functions of n besides powers of n . We can use the \prec relation to rank lots of functions into an asymptotic pecking order

that includes entries like this:

$$1 \prec \log \log n \prec \log n \prec n^\epsilon \prec n^c \prec n^{\log n} \prec c^n \prec n^n \prec c^{c^n}.$$

(Here ϵ and c are arbitrary constants with $0 < \epsilon < 1 < c$.)

All functions listed here, except 1, go to infinity as n goes to infinity. Thus when we try to place a new function in this hierarchy, we're not trying to determine *whether* it becomes infinite but rather *how fast*.

It helps to cultivate an expansive attitude when we're doing asymptotic analysis: We should THINK BIG, when imagining a variable that approaches infinity. For example, the hierarchy says that $\log n \prec n^{0.0001}$; this might seem wrong if we limit our horizons to teeny-tiny numbers like one googol, $n = 10^{100}$. For in that case, $\log n = 100$, while $n^{0.0001}$ is only $10^{0.01} \approx 1.0233$. But if we go up to a googolplex, $n = 10^{10^{100}}$, then $\log n = 10^{100}$ pales in comparison with $n^{0.0001} = 10^{10^{96}}$.

Even if ϵ is extremely small (smaller than, say, $1/10^{10^{100}}$), the value of $\log n$ will be much smaller than the value of n^ϵ , if n is large enough. For if we set $n = 10^{10^{2k}}$, where k is so large that $\epsilon \geq 10^{-k}$, we have $\log n = 10^{2k}$ but $n^\epsilon \geq 10^{10^k}$. The ratio $(\log n)/n^\epsilon$ therefore approaches zero as $n \rightarrow \infty$.

The hierarchy shown above deals with functions that go to infinity. Often, however, we're interested in functions that go to zero, so it's useful to have a similar hierarchy for those functions. We get one by taking reciprocals, because when $f(n)$ and $g(n)$ are never zero we have

$$f(n) \prec g(n) \iff \frac{1}{g(n)} \prec \frac{1}{f(n)}. \quad (9.5)$$

Thus, for example, the following functions (except 1) all go to zero:

$$\frac{1}{c^{c^n}} \prec \frac{1}{n^n} \prec \frac{1}{c^n} \prec \frac{1}{n^{\log n}} \prec \frac{1}{n^c} \prec \frac{1}{n^\epsilon} \prec \frac{1}{\log n} \prec \frac{1}{\log \log n} \prec 1.$$

Let's look at a few other functions to see where they fit in. The number $\pi(n)$ of primes less than or equal to n is known to be approximately $n/\ln n$. Since $1/n^\epsilon \prec 1/\ln n \prec 1$, multiplying by n tells us that

$$n^{1-\epsilon} \prec \pi(n) \prec n.$$

We can in fact generalize (9.4) by noticing, for example, that

$$\begin{aligned} n^{\alpha_1} (\log n)^{\alpha_2} (\log \log n)^{\alpha_3} &\prec n^{\beta_1} (\log n)^{\beta_2} (\log \log n)^{\beta_3} \\ \iff (\alpha_1, \alpha_2, \alpha_3) &< (\beta_1, \beta_2, \beta_3). \end{aligned} \quad (9.6)$$

Here ' $(\alpha_1, \alpha_2, \alpha_3) < (\beta_1, \beta_2, \beta_3)$ ' means lexicographic order (dictionary order); in other words, either $\alpha_1 < \beta_1$, or $\alpha_1 = \beta_1$ and $\alpha_2 < \beta_2$, or $\alpha_1 = \beta_1$ and $\alpha_2 = \beta_2$ and $\alpha_3 < \beta_3$.

A loerarchy?

How about the function $e^{\sqrt{\log n}}$; where does it live in the hierarchy? We can answer questions like this by using the rule

$$e^{f(n)} \prec e^{g(n)} \iff \lim_{n \rightarrow \infty} (f(n) - g(n)) = -\infty, \quad (9.7)$$

which follows in two steps from definition (9.3) by taking logarithms. Consequently

$$1 \prec f(n) \prec g(n) \implies e^{|f(n)|} \prec e^{|g(n)|}.$$

And since $1 \prec \log \log n \prec \sqrt{\log n} \prec \epsilon \log n$, we have $\log n \prec e^{\sqrt{\log n}} \prec n^\epsilon$.

When two functions $f(n)$ and $g(n)$ have the *same* rate of growth, we write ' $f(n) \asymp g(n)$ '. The official definition is:

$$f(n) \asymp g(n) \iff |f(n)| \leq C|g(n)| \text{ and } |g(n)| \leq C|f(n)|, \\ \text{for some } C \text{ and for all sufficiently large } n. \quad (9.8)$$

This holds, for example, if $f(n)$ is constant and $g(n) = \cos n + \arctan n$. We will prove shortly that it holds whenever $f(n)$ and $g(n)$ are polynomials of the same degree. There's also a stronger relation, defined by the rule

$$f(n) \sim g(n) \iff \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 1. \quad (9.9)$$

In this case we say that " $f(n)$ is asymptotic to $g(n)$."

G. H. Hardy [179] introduced an interesting and important concept called the class of *logarithmico-exponential functions*, defined recursively as the smallest family \mathfrak{L} of functions satisfying the following properties:

- The constant function $f(n) = \alpha$ is in \mathfrak{L} , for all real α .
- The identity function $f(n) = n$ is in \mathfrak{L} .
- If $f(n)$ and $g(n)$ are in \mathfrak{L} , so is $f(n) - g(n)$.
- If $f(n)$ is in \mathfrak{L} , so is $e^{f(n)}$.
- If $f(n)$ is in \mathfrak{L} and is "eventually positive," then $\ln f(n)$ is in \mathfrak{L} .

A function $f(n)$ is called "eventually positive" if there is an integer n_0 such that $f(n) > 0$ whenever $n \geq n_0$.

We can use these rules to show, for example, that $f(n) + g(n)$ is in \mathfrak{L} whenever $f(n)$ and $g(n)$ are, because $f(n) + g(n) = f(n) - (0 - g(n))$. If $f(n)$ and $g(n)$ are eventually positive members of \mathfrak{L} , their product $f(n)g(n) = e^{\ln f(n) + \ln g(n)}$ and quotient $f(n)/g(n) = e^{\ln f(n) - \ln g(n)}$ are in \mathfrak{L} ; so are functions like $\sqrt{f(n)} = e^{\frac{1}{2} \ln f(n)}$, etc. Hardy proved that every logarithmico-exponential function is eventually positive, eventually negative, or identically zero. Therefore the product and quotient of any two \mathfrak{L} -functions is in \mathfrak{L} , except that we cannot divide by a function that's identically zero.

Hardy's main theorem about logarithmico-exponential functions is that they form an asymptotic hierarchy: If $f(n)$ and $g(n)$ are any functions in \mathfrak{L} , then either $f(n) \prec g(n)$, or $f(n) \succ g(n)$, or $f(n) \asymp g(n)$. In the last case there is, in fact, a constant α such that

$$f(n) \sim \alpha g(n).$$

The proof of Hardy's theorem is beyond the scope of this book; but it's nice to know that the theorem exists, because almost every function we ever need to deal with is in \mathfrak{L} . In practice, we can generally fit a given function into a given hierarchy without great difficulty.

9.2 O NOTATION

A wonderful notational convention for asymptotic analysis was introduced by Paul Bachmann in 1894 and popularized in subsequent years by Edmund Landau and others. We have seen it in formulas like

$$H_n = \ln n + \gamma + O(1/n), \quad (9.10)$$

which tells us that the n th harmonic number is equal to the natural logarithm of n plus Euler's constant, plus a quantity that is "Big Oh of 1 over n ." This last quantity isn't specified exactly; but whatever it is, the notation claims that its absolute value is no more than a constant times $1/n$.

The beauty of O -notation is that it suppresses unimportant detail and lets us concentrate on salient features: The quantity $O(1/n)$ is negligibly small, if constant multiples of $1/n$ are unimportant.

Furthermore we get to use O right in the middle of a formula. If we want to express (9.10) in terms of the notations in Section 9.1, we must transpose ' $\ln n + \gamma$ ' to the left side and specify a weaker result like

$$H_n - \ln n - \gamma \prec \frac{\log \log n}{n}$$

or a stronger result like

$$H_n - \ln n - \gamma \asymp \frac{1}{n}.$$

The Big Oh notation allows us to specify an appropriate amount of detail in place, without transposition.

The idea of imprecisely specified quantities can be made clearer if we consider some additional examples. We occasionally use the notation ' ± 1 ' to stand for something that is either $+1$ or -1 ; we don't know (or perhaps we don't care) which it is, yet we can manipulate it in formulas.

"... wir durch das Zeichen $O(n)$ eine Größe ausdrücken, deren Ordnung in Bezug auf n die Ordnung von n nicht überschreitet; ob sie wirklich Glieder von der Ordnung n in sich enthält, bleibt bei dem bisherigen Schlußverfahren dahingestellt."
—P. Bachmann [17]

N. G. de Bruijn begins his book *Asymptotic Methods in Analysis* [74] by considering a Big Ell notation that helps us understand Big Oh. If we write $L(5)$ for a number whose absolute value is less than 5 (but we don't say what the number is), then we can perform certain calculations without knowing the full truth. For example, we can deduce formulas such as $1 + L(5) = L(6)$; $L(2) + L(3) = L(5)$; $L(2)L(3) = L(6)$; $e^{L(5)} = L(e^5)$; and so on. But we cannot conclude that $L(5) - L(3) = L(2)$, since the left side might be $4 - 0$. In fact, the most we can say is $L(5) - L(3) = L(8)$.

Bachmann's O-notation is similar to L-notation but it's even less precise: $O(\alpha)$ stands for a number whose absolute value is at most a constant times $|\alpha|$. We don't say what the number is and we don't even say what the constant is. Of course the notion of a "constant" is nonsense if there is nothing variable in the picture, so we use O-notation only in contexts when there's at least one quantity (say n) whose value is varying. The formula

$$f(n) = O(g(n)) \quad \text{for all } n \quad (9.11)$$

means in this context that there is a constant C such that

$$|f(n)| \leq C|g(n)| \quad \text{for all } n; \quad (9.12)$$

and when $O(g(n))$ stands in the middle of a formula it represents a function $f(n)$ that satisfies (9.12). The values of $f(n)$ are unknown, but we do know that they aren't too large. Similarly, de Bruijn's ' $L(n)$ ' represents an unspecified function $f(n)$ whose values satisfy $|f(n)| < |n|$. The main difference between L and O is that O-notation involves an unspecified constant C ; each appearance of O might involve a different C , but each C is independent of n .

For example, we know that the sum of the first n squares is

$$\square_n = \frac{1}{3}n(n + \frac{1}{2})(n + 1) = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n.$$

We can write

$$\square_n = O(n^3)$$

because $|\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n| \leq \frac{1}{3}|n|^3 + \frac{1}{2}|n|^2 + \frac{1}{6}|n| \leq \frac{1}{3}|n^3| + \frac{1}{2}|n^3| + \frac{1}{6}|n^3| = |n^3|$ for all integers n . Similarly, we have the more specific formula

$$\square_n = \frac{1}{3}n^3 + O(n^2);$$

we can also be sloppy and throw away information, saying that

$$\square_n = O(n^{10}).$$

Nothing in the definition of O requires us to give a best possible bound.

*It's not nonsense,
but it is pointless.*

*I've got a little
list — I've got a
little list,
Of annoying terms
and details that
might well be under
ground,
And that never
would be missed —
that never would be
missed.*

But wait a minute. What if the variable n isn't an integer? What if we have a formula like $S(x) = \frac{1}{3}x^3 + \frac{1}{2}x^2 + \frac{1}{6}x$, where x is a real number? Then we cannot say that $S(x) = O(x^3)$, because the ratio $S(x)/x^3 = \frac{1}{3} + \frac{1}{2}x^{-1} + \frac{1}{6}x^{-2}$ becomes unbounded when $x \rightarrow 0$. And we cannot say that $S(x) = O(x)$, because the ratio $S(x)/x = \frac{1}{3}x^2 + \frac{1}{2}x + \frac{1}{6}$ becomes unbounded when $x \rightarrow \infty$. So we apparently can't use O -notation with $S(x)$.

The answer to this dilemma is that variables used with O are generally subject to side conditions. For example, if we stipulate that $|x| \geq 1$, or that $x \geq \epsilon$ where ϵ is any positive constant, or that x is an integer, then we can write $S(x) = O(x^3)$. If we stipulate that $|x| \leq 1$, or that $|x| \leq c$ where c is any positive constant, then we can write $S(x) = O(x)$. The O -notation is governed by its environment, by constraints on the variables involved.

These constraints are often specified by a limiting relation. For example, we might say that

$$f(n) = O(g(n)) \quad \text{as } n \rightarrow \infty. \quad (9.13)$$

This means that the O -condition is supposed to hold when n is "near" ∞ ; we don't care what happens unless n is quite large. Moreover, we don't even specify exactly what "near" means; in such cases each appearance of O implicitly asserts the existence of *two* constants C and n_0 , such that

$$|f(n)| \leq C|g(n)| \quad \text{whenever } n \geq n_0. \quad (9.14)$$

The values of C and n_0 might be different for each O , but they do not depend on n . Similarly, the notation

$$f(x) = O(g(x)) \quad \text{as } x \rightarrow 0$$

means that there exist two constants C and ϵ such that

$$|f(x)| \leq C|g(x)| \quad \text{whenever } |x| \leq \epsilon. \quad (9.15)$$

The limiting value does not have to be ∞ or 0 ; we can write

$$\ln z = z - 1 + O((z - 1)^2) \quad \text{as } z \rightarrow 1,$$

because it can be proved that $|\ln z - z + 1| \leq |z - 1|^2$ when $|z - 1| \leq \frac{1}{2}$.

Our definition of O has gradually developed, over a few pages, from something that seemed pretty obvious to something that seems rather complex; we now have O representing an undefined function and either one or two unspecified constants, depending on the environment. This may seem complicated enough for any reasonable notation, but it's still not the whole story! Another

*You are the fairest
of your sex,
Let me be your
hero;
I love you as
one over x ,
As x approaches
zero.
Positively.*

subtle consideration lurks in the background. Namely, we need to realize that it's fine to write

$$\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n = O(n^3),$$

but we should *never* write this equality with the sides reversed. Otherwise we could deduce ridiculous things like $n = n^2$ from the identities $n = O(n^2)$ and $n^2 = O(n^2)$. When we work with O -notation and any other formulas that involve imprecisely specified quantities, we are dealing with *one-way equalities*. The right side of an equation does not give more information than the left side, and it may give less; the right is a “crudification” of the left.

From a strictly formal point of view, the notation $O(g(n))$ does not stand for a single function $f(n)$, but for the *set* of all functions $f(n)$ such that $|f(n)| \leq C|g(n)|$ for some constant C . An ordinary formula $g(n)$ that doesn't involve O -notation stands for the set containing a single function $f(n) = g(n)$. If S and T are sets of functions of n , the notation $S + T$ stands for the set of all functions of the form $f(n) + g(n)$, where $f(n) \in S$ and $g(n) \in T$; other notations like $S - T$, ST , S/T , \sqrt{S} , e^S , $\ln S$ are defined similarly. Then an “equation” between such sets of functions is, strictly speaking, a *set inclusion*; the ‘=’ sign really means ‘ \subseteq ’. These formal definitions put all of our O manipulations on firm logical ground.

For example, the “equation”

$$\frac{1}{3}n^3 + O(n^2) = O(n^3)$$

means that $S_1 \subseteq S_2$, where S_1 is the set of all functions of the form $\frac{1}{3}n^3 + f_1(n)$ such that there exists a constant C_1 with $|f_1(n)| \leq C_1|n^2|$, and where S_2 is the set of all functions $f_2(n)$ such that there exists a constant C_2 with $|f_2(n)| \leq C_2|n^3|$. We can formally prove this “equation” by taking an arbitrary element of the left-hand side and showing that it belongs to the right-hand side: Given $\frac{1}{3}n^3 + f_1(n)$ such that $|f_1(n)| \leq C_1|n^2|$, we must prove that there's a constant C_2 such that $|\frac{1}{3}n^3 + f_1(n)| \leq C_2|n^3|$. The constant $C_2 = \frac{1}{3} + C_1$ does the trick, since $n^2 \leq |n^3|$ for all integers n .

If ‘=’ really means ‘ \subseteq ’, why don't we use ‘ \subseteq ’ instead of abusing the equals sign? There are four reasons.

First, tradition. Number theorists started using the equals sign with O -notation and the practice stuck. It's sufficiently well established by now that we cannot hope to get the mathematical community to change.

Second, tradition. Computer people are quite used to seeing equals signs abused—for years FORTRAN and BASIC programmers have been writing assignment statements like ‘ $N = N + 1$ ’. One more abuse isn't much.

Third, tradition. We often read ‘=’ as the word ‘is’. For instance we verbalize the formula $H_n = O(\log n)$ by saying “ H sub n is Big Oh of $\log n$.”

“And to auoide the tedious repetition of these woordes: is equalle to: I will sette as I doe often in woorke use, a paire of paralleles, or Gemowe lines of one lengthe, thus: \equiv , bicause noe .2. thynges, can be moare equalle.”
—R. Recorde [305]

"It is obvious that the sign = is really the wrong sign for such relations, because it suggests symmetry, and there is no such symmetry. . . . Once this warning has been given, there is, however, not much harm in using the sign =, and we shall maintain it, for no other reason than that it is customary."
—N. G. de Bruijn [74]

And in English, this 'is' is one-way. We say that a bird is an animal, but we don't say that an animal is a bird; "animal" is a crudification of "bird."

Fourth, for our purposes it's natural. If we limited our use of O-notation to situations where it occupies the whole right side of a formula—as in the harmonic number approximation $H_n = O(\log n)$, or as in the description of a sorting algorithm's running time $T(n) = O(n \log n)$ —it wouldn't matter whether we used '=' or something else. But when we use O-notation in the middle of an expression, as we usually do in asymptotic calculations, our intuition is well satisfied if we think of the equals sign as an equality, and if we think of something like $O(1/n)$ as a very small quantity.

So we'll continue to use '=', and we'll continue to regard $O(g(n))$ as an incompletely specified function, knowing that we can always fall back on the set-theoretic definition if we must.

But we ought to mention one more technicality while we're picking nits about definitions: If there are several variables in the environment, O-notation formally represents sets of functions of two or more variables, not just one. The domain of each function is every variable that is currently "free" to vary.

This concept can be a bit subtle, because a variable might be defined only in parts of an expression, when it's controlled by a \sum or something similar. For example, let's look closely at the equation

$$\sum_{k=0}^n (k^2 + O(k)) = \frac{1}{3}n^3 + O(n^2), \quad \text{integer } n \geq 0. \quad (9.16)$$

The expression $k^2 + O(k)$ on the left stands for the set of all two-variable functions of the form $k^2 + f(k, n)$ such that there exists a constant C with $|f(k, n)| \leq Ck$ for $0 \leq k \leq n$. The sum of this set of functions, for $0 \leq k \leq n$, is the set of all functions $g(n)$ of the form

$$\sum_{k=0}^n (k^2 + f(k, n)) = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n + f(0, n) + f(1, n) + \cdots + f(n, n),$$

where f has the stated property. Since we have

$$\begin{aligned} & \left| \frac{1}{2}n^2 + \frac{1}{6}n + f(0, n) + f(1, n) + \cdots + f(n, n) \right| \\ & \leq \frac{1}{2}n^2 + \frac{1}{6}n^2 + C \cdot 0 + C \cdot 1 + \cdots + C \cdot n \\ & < n^2 + C(n^2 + n)/2 < (C+1)n^2, \end{aligned}$$

(Now is a good time to do warmup exercises 3 and 4.)

all such functions $g(n)$ belong to the right-hand side of (9.16); therefore (9.16) is true.

People sometimes abuse O-notation by assuming that it gives an exact order of growth; they use it as if it specifies a lower bound as well as an upper bound. For example, an algorithm to sort n numbers might be called

inefficient “because its running time is $O(n^2)$.” But a running time of $O(n^2)$ does not imply that the running time is not also $O(n)$. There’s another notation, Big Omega, for lower bounds:

$$f(n) = \Omega(g(n)) \iff |f(n)| \geq C|g(n)| \quad \text{for some } C > 0. \quad (9.17)$$

We have $f(n) = \Omega(g(n))$ if and only if $g(n) = O(f(n))$. A sorting algorithm whose running time is $\Omega(n^2)$ is inefficient compared with one whose running time is $O(n \log n)$, if n is large enough.

Finally there’s Big Theta, which specifies an exact order of growth:

$$f(n) = \Theta(g(n)) \iff \begin{array}{l} f(n) = O(g(n)) \\ \text{and } f(n) = \Omega(g(n)). \end{array} \quad (9.18)$$

We have $f(n) = \Theta(g(n))$ if and only if $f(n) \asymp g(n)$ in the notation we saw previously, equation (9.8).

Edmund Landau [238] invented a “little oh” notation,

$$f(n) = o(g(n)) \iff |f(n)| \leq \epsilon |g(n)| \quad \begin{array}{l} \text{for all } n \geq n_0(\epsilon) \text{ and} \\ \text{for all constants } \epsilon > 0. \end{array} \quad (9.19)$$

This is essentially the relation $f(n) \prec g(n)$ of (9.3). We also have

$$f(n) \sim g(n) \iff f(n) = g(n) + o(g(n)). \quad (9.20)$$

Many authors use ‘o’ in asymptotic formulas, but a more explicit ‘O’ expression is almost always preferable. For example, the average running time of a computer method called “bubblesort” depends on the asymptotic value of the sum $P(n) = \sum_{k=0}^n k^{n-k} k! / n!$. Elementary asymptotic methods suffice to prove the formula $P(n) \sim \sqrt{\pi n / 2}$, which means that the ratio $P(n) / \sqrt{\pi n / 2}$ approaches 1 as $n \rightarrow \infty$. However, the true behavior of $P(n)$ is best understood by considering the *difference*, $P(n) - \sqrt{\pi n / 2}$, not the ratio:

n	$P(n) / \sqrt{\pi n / 2}$	$P(n) - \sqrt{\pi n / 2}$
1	0.798	−0.253
10	0.878	−0.484
20	0.904	−0.538
30	0.918	−0.561
40	0.927	−0.575
50	0.934	−0.585

The numerical evidence in the middle column is not very compelling; it certainly is far from a dramatic proof that $P(n) / \sqrt{\pi n / 2}$ approaches 1 rapidly,

Since Ω and Θ are uppercase Greek letters, the O in O-notation must be a capital Greek Omicron. After all, Greeks invented asymptotics.

if at all. But the right-hand column shows that $P(n)$ is very close indeed to $\sqrt{\pi n/2}$. Thus we can characterize the behavior of $P(n)$ much better if we can derive formulas of the form

$$P(n) = \sqrt{\pi n/2} + O(1),$$

or even sharper estimates like

$$P(n) = \sqrt{\pi n/2} - \frac{2}{3} + O(1/\sqrt{n}).$$

Stronger methods of asymptotic analysis are needed to prove O-results, but the additional effort required to learn these stronger methods is amply compensated by the improved understanding that comes with O-bounds.

Moreover, many sorting algorithms have running times of the form

$$T(n) = A n \lg n + B n + O(\log n)$$

for some constants A and B . Analyses that stop at $T(n) \sim A n \lg n$ don't tell the whole story, and it turns out to be a bad strategy to choose a sorting algorithm based just on its A value. Algorithms with a good ' A ' often achieve this at the expense of a bad ' B '. Since $n \lg n$ grows only slightly faster than n , the algorithm that's faster asymptotically (the one with a slightly smaller A value) might be faster only for values of n that never actually arise in practice. Thus, asymptotic methods that allow us to go past the first term and evaluate B are necessary if we are to make the right choice of method.

Also \lg , the Duraflame logarithm.

Before we go on to study O , let's talk about one more small aspect of mathematical style. Three different notations for logarithms have been used in this chapter: \lg , \ln , and \log . We often use ' \lg ' in connection with computer methods, because binary logarithms are often relevant in such cases; and we often use ' \ln ' in purely mathematical calculations, since the formulas for natural logarithms are nice and simple. But what about ' \log '? Isn't this the "common" base-10 logarithm that students learn in high school—the "common" logarithm that turns out to be very uncommon in mathematics and computer science? Yes; and many mathematicians confuse the issue by using ' \log ' to stand for natural logarithms or binary logarithms. There is no universal agreement here. But we can usually breathe a sigh of relief when a logarithm appears inside O -notation, because O ignores multiplicative constants. There is no difference between $O(\lg n)$, $O(\ln n)$, and $O(\log n)$, as $n \rightarrow \infty$; similarly, there is no difference between $O(\lg \lg n)$, $O(\ln \ln n)$, and $O(\log \log n)$. We get to choose whichever we please; and the one with ' \log ' seems friendlier because it is more pronounceable. Therefore we generally use ' \log ' in all contexts where it improves readability without introducing ambiguity.

Notice that $\log \log \log n$ is undefined when $n \leq 10$.

9.3 O MANIPULATION

Like any mathematical formalism, the O-notation has rules of manipulation that free us from the grungy details of its definition. Once we prove that the rules are correct, using the definition, we can henceforth work on a higher plane and forget about actually verifying that one set of functions is contained in another. We don't even need to calculate the constants C that are implied by each O, as long as we follow rules that guarantee the existence of such constants.

For example, we can prove once and for all that

$$n^m = O(n^{m'}), \quad \text{when } m \leq m'; \quad (9.21)$$

$$O(f(n)) + O(g(n)) = O(|f(n)| + |g(n)|). \quad (9.22)$$

Then we can say immediately that $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n = O(n^3) + O(n^3) + O(n^3) = O(n^3)$, without the laborious calculations in the previous section.

Here are some more rules that follow easily from the definition:

$$f(n) = O(f(n)); \quad (9.23)$$

$$c \cdot O(f(n)) = O(f(n)), \quad \text{if } c \text{ is constant}; \quad (9.24)$$

$$O(O(f(n))) = O(f(n)); \quad (9.25)$$

$$O(f(n))O(g(n)) = O(f(n)g(n)); \quad (9.26)$$

$$O(f(n)g(n)) = f(n)O(g(n)). \quad (9.27)$$

Exercise 9 proves (9.22), and the proofs of the others are similar. We can always replace something of the form on the left by what's on the right, regardless of the side conditions on the variable n .

Equations (9.27) and (9.23) allow us to derive the identity $O(f(n)^2) = O(f(n))^2$. This sometimes helps avoid parentheses, since we can write

$$O(\log n)^2 \quad \text{instead of} \quad O((\log n)^2).$$

Both of these are preferable to ' $O(\log^2 n)$ ', which is ambiguous because some authors use it to mean ' $O(\log \log n)$ '.

Can we also write

$$O(\log n)^{-1} \quad \text{instead of} \quad O((\log n)^{-1})?$$

No! This is an abuse of notation, since the set of functions $1/O(\log n)$ is neither a subset nor a superset of $O(1/\log n)$. We could legitimately substitute $\Omega(\log n)^{-1}$ for $O((\log n)^{-1})$, but this would be awkward. So we'll restrict our use of "exponents outside the O" to constant, positive integer exponents.

*The secret of being
a bore is to tell
everything.*

—Voltaire

(Note: The formula $O(f(n))^2$ does not denote the set of all functions $g(n)^2$ where $g(n)$ is in $O(f(n))$; such functions $g(n)^2$ cannot be negative, but the set $O(f(n))^2$ includes negative functions. In general, when S is a set, the notation S^2 stands for the set of all products $s_1 s_2$ with s_1 and s_2 in S , not for the set of all squares s^2 with $s \in S$.)

Power series give us some of the most useful operations of all. If the sum

$$S(z) = \sum_{n \geq 0} a_n z^n$$

converges absolutely for some complex number $z = z_0$, then

$$S(z) = O(1), \quad \text{for all } |z| \leq |z_0|.$$

This is obvious, because

$$|S(z)| \leq \sum_{n \geq 0} |a_n| |z|^n \leq \sum_{n \geq 0} |a_n| |z_0|^n = C < \infty.$$

In particular, $S(z) = O(1)$ as $z \rightarrow 0$, and $S(1/n) = O(1)$ as $n \rightarrow \infty$, provided only that $S(z)$ converges for at least one nonzero value of z . We can use this principle to truncate a power series at any convenient point and estimate the remainder with O . For example, not only is $S(z) = O(1)$, but

$$\begin{aligned} S(z) &= a_0 + O(z), \\ S(z) &= a_0 + a_1 z + O(z^2), \end{aligned}$$

and so on, because

$$S(z) = \sum_{0 \leq k < m} a_k z^k + z^m \sum_{n \geq m} a_n z^{n-m}$$

and the latter sum, like $S(z)$ itself, converges absolutely for $z = z_0$ and is $O(1)$. Table 452 lists some of the most useful asymptotic formulas, half of which are simply based on truncation of power series according to this rule.

Dirichlet series, which are sums of the form $\sum_{k \geq 1} a_k/k^z$, can be truncated in a similar way: If a Dirichlet series converges absolutely when $z = z_0$, we can truncate it at any term and get the approximation

$$\sum_{1 \leq k < m} a_k/k^z + O(m^{-z}),$$

Remember that \Re stands for "real part."

valid for $\Re z \geq \Re z_0$. The asymptotic formula for Bernoulli numbers B_n in Table 452 illustrates this principle.

On the other hand, the asymptotic formulas for H_n , $n!$, and $\pi(n)$ in Table 452 are not truncations of convergent series; if we extended them indefinitely they would diverge for all values of n . This is particularly easy to see in the case of $\pi(n)$, since we have already observed in Section 7.3, Example 5, that the power series $\sum_{k \geq 0} k!/(n \ln n)^k$ is everywhere divergent. Yet these truncations of divergent series turn out to be useful approximations.

Table 452 Asymptotic approximations, valid as $n \rightarrow \infty$ and $z \rightarrow 0$.

$$H_n = \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} + O\left(\frac{1}{n^6}\right). \quad (9.28)$$

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3} + O\left(\frac{1}{n^4}\right)\right). \quad (9.29)$$

$$B_n = 2[n \text{ even}](-1)^{n/2-1} \frac{n!}{(2\pi)^n} (1 + 2^{-n} + 3^{-n} + O(4^{-n})). \quad (9.30)$$

$$\pi(n) = \frac{n}{\ln n} + \frac{n}{(\ln n)^2} + \frac{2!n}{(\ln n)^3} + \frac{3!n}{(\ln n)^4} + O\left(\frac{n}{(\log n)^5}\right). \quad (9.31)$$

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + O(z^5). \quad (9.32)$$

$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + O(z^5). \quad (9.33)$$

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + z^4 + O(z^5). \quad (9.34)$$

$$(1+z)^\alpha = 1 + \alpha z + \binom{\alpha}{2} z^2 + \binom{\alpha}{3} z^3 + \binom{\alpha}{4} z^4 + O(z^5). \quad (9.35)$$

An asymptotic approximation is said to have *absolute error* $O(g(n))$ if it has the form $f(n) + O(g(n))$ where $f(n)$ doesn't involve O . The approximation has *relative error* $O(g(n))$ if it has the form $f(n)(1 + O(g(n)))$ where $f(n)$ doesn't involve O . For example, the approximation for H_n in Table 452 has absolute error $O(n^{-6})$; the approximation for $n!$ has relative error $O(n^{-4})$. (The right-hand side of (9.29) doesn't actually have the required form $f(n)(1 + O(n^{-4}))$, but we could rewrite it

$$\sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3}\right) (1 + O(n^{-4}))$$

if we wanted to; a similar calculation is the subject of exercise 12.) The absolute error of this approximation is $O(n^{-3.5}e^{-n})$. Absolute error is related to the number of correct decimal digits to the right of the decimal point if the O term is ignored; relative error corresponds to the number of correct "significant figures."

(Relative error is nice for taking reciprocals, because $1/(1 + O(\epsilon)) = 1 + O(\epsilon)$.)

We can use truncation of power series to prove the general laws

$$\ln(1 + O(f(n))) = O(f(n)), \quad \text{if } f(n) \prec 1; \quad (9.36)$$

$$e^{O(f(n))} = 1 + O(f(n)), \quad \text{if } f(n) = O(1). \quad (9.37)$$

(Here we assume that $n \rightarrow \infty$; similar formulas hold for $\ln(1 + O(f(x)))$ and $e^{O(f(x))}$ as $x \rightarrow 0$.) For example, let $\ln(1 + g(n))$ be any function belonging to the left side of (9.36). Then there are constants C , n_0 , and c such that

$$|g(n)| \leq C|f(n)| \leq c < 1, \quad \text{for all } n \geq n_0.$$

It follows that the infinite sum

$$\ln(1 + g(n)) = g(n) \cdot (1 - \frac{1}{2}g(n) + \frac{1}{3}g(n)^2 - \dots)$$

converges for all $n \geq n_0$, and the parenthesized series is bounded by the constant $1 + \frac{1}{2}c + \frac{1}{3}c^2 + \dots$. This proves (9.36), and the proof of (9.37) is similar. Equations (9.36) and (9.37) combine to give the useful formula

$$(1 + O(f(n)))^{O(g(n))} = 1 + O(f(n)g(n)), \quad \begin{array}{l} \text{if } f(n) \prec 1 \text{ and} \\ f(n)g(n) = O(1). \end{array} \quad (9.38)$$

Problem 1: Return to the Wheel of Fortune.

Let's try our luck now at a few asymptotic problems. In Chapter 3 we derived equation (3.13) for the number of winning positions in a certain game:

$$W = \lfloor N/K \rfloor + \frac{1}{2}K^2 + \frac{5}{2}K - 3, \quad K = \lfloor \sqrt[3]{N} \rfloor.$$

And we promised that an asymptotic version of W would be derived in Chapter 9. Well, here we are in Chapter 9; let's try to estimate W , as $N \rightarrow \infty$.

The main idea here is to remove the floor brackets, replacing K by $N^{1/3} + O(1)$. Then we can go further and write

$$K = N^{1/3}(1 + O(N^{-1/3}));$$

this is called "pulling out the large part." (We will be using this trick a lot.) Now we have

$$\begin{aligned} K^2 &= N^{2/3}(1 + O(N^{-1/3}))^2 \\ &= N^{2/3}(1 + O(N^{-1/3})) = N^{2/3} + O(N^{1/3}) \end{aligned}$$

by (9.38) and (9.26). Similarly

$$\begin{aligned} \lfloor N/K \rfloor &= N^{1-1/3}(1 + O(N^{-1/3}))^{-1} + O(1) \\ &= N^{2/3}(1 + O(N^{-1/3})) + O(1) = N^{2/3} + O(N^{1/3}). \end{aligned}$$

It follows that the number of winning positions is

$$\begin{aligned} W &= N^{2/3} + O(N^{1/3}) + \frac{1}{2}(N^{2/3} + O(N^{1/3})) + O(N^{1/3}) + O(1) \\ &= \frac{3}{2}N^{2/3} + O(N^{1/3}). \end{aligned} \quad (9.39)$$

Notice how the O terms absorb one another until only one remains; this is typical, and it illustrates why O -notation is useful in the middle of a formula.

Problem 2: Perturbation of Stirling's formula.

Stirling's approximation for $n!$ is undoubtedly the most famous asymptotic formula of all. We will prove it later in this chapter; for now, let's just try to get better acquainted with its properties. We can write one version of the approximation in the form

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \frac{a}{n} + \frac{b}{n^2} + O(n^{-3})\right), \quad \text{as } n \rightarrow \infty, \quad (9.40)$$

for certain constants a and b . Since this holds for all large n , it must also be asymptotically true when n is replaced by $n-1$:

$$\begin{aligned} (n-1)! &= \sqrt{2\pi(n-1)} \left(\frac{n-1}{e}\right)^{n-1} \\ &\times \left(1 + \frac{a}{n-1} + \frac{b}{(n-1)^2} + O((n-1)^{-3})\right). \end{aligned} \quad (9.41)$$

We know, of course, that $(n-1)! = n!/n$; hence the right-hand side of this formula must simplify to the right-hand side of (9.40), divided by n .

Let us therefore try to simplify (9.41). The first factor becomes tractable if we pull out the large part:

$$\begin{aligned} \sqrt{2\pi(n-1)} &= \sqrt{2\pi n} (1 - n^{-1})^{1/2} \\ &= \sqrt{2\pi n} \left(1 - \frac{1}{2n} - \frac{1}{8n^2} + O(n^{-3})\right). \end{aligned}$$

Equation (9.35) has been used here.

Similarly we have

$$\begin{aligned} \frac{a}{n-1} &= \frac{a}{n} (1 - n^{-1})^{-1} = \frac{a}{n} + \frac{a}{n^2} + O(n^{-3}); \\ \frac{b}{(n-1)^2} &= \frac{b}{n^2} (1 - n^{-1})^{-2} = \frac{b}{n^2} + O(n^{-3}); \\ O((n-1)^{-3}) &= O(n^{-3} (1 - n^{-1})^{-3}) = O(n^{-3}). \end{aligned}$$

The only thing in (9.41) that's slightly tricky to deal with is the factor $(n-1)^{n-1}$, which equals

$$n^{n-1} (1 - n^{-1})^{n-1} = n^{n-1} (1 - n^{-1})^n (1 + n^{-1} + n^{-2} + O(n^{-3})).$$

(We are expanding everything out until we get a relative error of $O(n^{-3})$, because the relative error of a product is the sum of the relative errors of the individual factors. All of the $O(n^{-3})$ terms will coalesce.)

In order to expand $(1 - n^{-1})^n$, we first compute $\ln(1 - n^{-1})$ and then form the exponential, $e^{n \ln(1 - n^{-1})}$:

$$\begin{aligned}
 (1 - n^{-1})^n &= \exp(n \ln(1 - n^{-1})) \\
 &= \exp(n(-n^{-1} - \tfrac{1}{2}n^{-2} - \tfrac{1}{3}n^{-3} + O(n^{-4}))) \\
 &= \exp(-1 - \tfrac{1}{2}n^{-1} - \tfrac{1}{3}n^{-2} + O(n^{-3})) \\
 &= \exp(-1) \cdot \exp(-\tfrac{1}{2}n^{-1}) \cdot \exp(-\tfrac{1}{3}n^{-2}) \cdot \exp(O(n^{-3})) \\
 &= \exp(-1) \cdot (1 - \tfrac{1}{2}n^{-1} + \tfrac{1}{8}n^{-2} + O(n^{-3})) \\
 &\quad \cdot (1 - \tfrac{1}{3}n^{-2} + O(n^{-4})) \cdot (1 + O(n^{-3})) \\
 &= e^{-1}(1 - \tfrac{1}{2}n^{-1} - \tfrac{5}{24}n^{-2} + O(n^{-3})).
 \end{aligned}$$

Here we use the notation $\exp z$ instead of e^z , since it allows us to work with a complicated exponent on the main line of the formula instead of in the superscript position. We must expand $\ln(1 - n^{-1})$ with absolute error $O(n^{-4})$ in order to end with a relative error of $O(n^{-3})$, because the logarithm is being multiplied by n .

The right-hand side of (9.41) has now been reduced to $\sqrt{2\pi n}$ times n^{n-1}/e^n times a product of several factors:

$$\begin{aligned}
 &(1 - \tfrac{1}{2}n^{-1} - \tfrac{1}{8}n^{-2} + O(n^{-3})) \\
 &\quad \cdot (1 + n^{-1} + n^{-2} + O(n^{-3})) \\
 &\quad \cdot (1 - \tfrac{1}{2}n^{-1} - \tfrac{5}{24}n^{-2} + O(n^{-3})) \\
 &\quad \cdot (1 + an^{-1} + (a+b)n^{-2} + O(n^{-3})).
 \end{aligned}$$

Multiplying these out and absorbing all asymptotic terms into one $O(n^{-3})$ yields

$$1 + an^{-1} + (a + b - \tfrac{1}{12})n^{-2} + O(n^{-3}).$$

Hmmm; we were hoping to get $1 + an^{-1} + bn^{-2} + O(n^{-3})$, since that's what we need to match the right-hand side of (9.40). Has something gone awry? No, everything is fine, provided that $a + b - \frac{1}{12} = b$.

This perturbation argument doesn't prove the validity of Stirling's approximation, but it does prove something: It proves that formula (9.40) cannot be valid unless $a = \frac{1}{12}$. If we had replaced the $O(n^{-3})$ in (9.40) by $cn^{-3} + O(n^{-4})$ and carried out our calculations to a relative error of $O(n^{-4})$, we could have deduced that b must be $\frac{1}{288}$, as claimed in Table 452. (This is not the easiest way to determine the values of a and b , but it works.)

Problem 3: The n th prime number.

Equation (9.31) is an asymptotic formula for $\pi(n)$, the number of primes that do not exceed n . If we replace n by $p = P_n$, the n th prime number, we have $\pi(p) = n$; hence

$$n = \frac{p}{\ln p} + O\left(\frac{p}{(\log p)^2}\right) \quad (9.42)$$

as $n \rightarrow \infty$. Let us try to “solve” this equation for p ; then we will know the approximate size of the n th prime.

The first step is to simplify the O term. If we divide both sides by $p/\ln p$, we find that $n \ln p/p \rightarrow 1$; hence $p/\ln p = O(n)$ and

$$O\left(\frac{p}{(\log p)^2}\right) = O\left(\frac{n}{\log p}\right) = O\left(\frac{n}{\log n}\right).$$

(We have $(\log p)^{-1} \leq (\log n)^{-1}$ because $p \geq n$.)

The second step is to transpose the two sides of (9.42), except for the O term. This is legal because of the general rule

$$a_n = b_n + O(f(n)) \iff b_n = a_n + O(f(n)). \quad (9.43)$$

(Each of these equations follows from the other if we multiply both sides by -1 and then add $a_n + b_n$ to both sides.) Hence

$$\frac{p}{\ln p} = n + O\left(\frac{n}{\log n}\right) = n(1 + O(1/\log n)),$$

and we have

$$p = n \ln p (1 + O(1/\log n)). \quad (9.44)$$

This is an “approximate recurrence” for $p = P_n$ in terms of itself. Our goal is to change it into an “approximate closed form,” and we can do this by unfolding the recurrence asymptotically. So let’s try to unfold (9.44).

By taking logarithms of both sides we deduce that

$$\ln p = \ln n + \ln \ln p + O(1/\log n). \quad (9.45)$$

This value can be substituted for $\ln p$ in (9.44), but we would like to get rid of all p ’s on the right before making the substitution. Somewhere along the line, that last p must disappear; we can’t get rid of it in the normal way for recurrences, because (9.44) doesn’t specify initial conditions for small p .

One way to do the job is to start by proving the weaker result $p = O(n^2)$. This follows if we square (9.44) and divide by pn^2 ,

$$\frac{p}{n^2} = \frac{(\ln p)^2}{p} (1 + O(1/\log n)),$$

since the right side approaches zero as $n \rightarrow \infty$. OK, we know that $p = O(n^2)$; therefore $\log p = O(\log n)$ and $\log \log p = O(\log \log n)$. We can now conclude from (9.45) that

$$\ln p = \ln n + O(\log \log n);$$

in fact, with this new estimate in hand we can conclude that $\ln \ln p = \ln \ln n + O(\log \log n / \log n)$, and (9.45) now yields

$$\ln p = \ln n + \ln \ln n + O(\log \log n / \log n).$$

And we can plug this into the right-hand side of (9.44), obtaining

$$p = n \ln n + n \ln \ln n + O(n).$$

This is the approximate size of the n th prime.

We can refine this estimate by using a better approximation of $\pi(n)$ in place of (9.42). The next term of (9.31) tells us that

$$n = \frac{p}{\ln p} + \frac{p}{(\ln p)^2} + O\left(\frac{p}{(\log p)^3}\right); \quad (9.46)$$

Get out the scratch paper again, gang.

proceeding as before, we obtain the recurrence

Boo, Hiss.

$$p = n \ln p (1 + (\ln p)^{-1})^{-1} (1 + O(1/\log n)^2), \quad (9.47)$$

which has a relative error of $O(1/\log n)^2$ instead of $O(1/\log n)$. Taking logarithms and retaining proper accuracy (but not too much) now yields

$$\begin{aligned} \ln p &= \ln n + \ln \ln p + O(1/\log n) \\ &= \ln n \left(1 + \frac{\ln \ln p}{\ln n} + O(1/\log n)^2\right); \\ \ln \ln p &= \ln \ln n + \frac{\ln \ln n}{\ln n} + O\left(\frac{\log \log n}{\log n}\right)^2. \end{aligned}$$

Finally we substitute these results into (9.47) and our answer finds its way out:

$$P_n = n \ln n + n \ln \ln n - n + n \frac{\ln \ln n}{\ln n} + O\left(\frac{n}{\log n}\right). \quad (9.48)$$

For example, when $n = 10^6$ this estimate comes to $15631363.8 + O(n/\log n)$; the millionth prime is actually 15485863. Exercise 21 shows that a still more accurate approximation to P_n results if we begin with a still more accurate approximation to $\pi(n)$ in place of (9.46).

Problem 4: A sum from an old final exam.

When Concrete Mathematics was first taught at Stanford University during the 1970–1971 term, students were asked for the asymptotic value of the sum

$$S_n = \frac{1}{n^2+1} + \frac{1}{n^2+2} + \cdots + \frac{1}{n^2+n}, \quad (9.49)$$

with an absolute error of $O(n^{-7})$. Let's imagine that we've just been given this problem on a (take-home) final; what is our first instinctive reaction?

No, we don't panic. Our first reaction is to THINK BIG. If we set $n = 10^{100}$, say, and look at the sum, we see that it consists of n terms, each of which is slightly less than $1/n^2$; hence the sum is slightly less than $1/n$. In general, we can usually get a decent start on an asymptotic problem by taking stock of the situation and getting a ballpark estimate of the answer.

Let's try to improve the rough estimate by pulling out the largest part of each term. We have

$$\frac{1}{n^2+k} = \frac{1}{n^2(1+k/n^2)} = \frac{1}{n^2} \left(1 - \frac{k}{n^2} + \frac{k^2}{n^4} - \frac{k^3}{n^6} + O\left(\frac{k^4}{n^8}\right) \right),$$

and so it's natural to try summing all these approximations:

$$\begin{aligned} \frac{1}{n^2+1} &= \frac{1}{n^2} - \frac{1}{n^4} + \frac{1^2}{n^6} - \frac{1^3}{n^8} + O\left(\frac{1^4}{n^{10}}\right) \\ \frac{1}{n^2+2} &= \frac{1}{n^2} - \frac{2}{n^4} + \frac{2^2}{n^6} - \frac{2^3}{n^8} + O\left(\frac{2^4}{n^{10}}\right) \\ &\vdots \\ \frac{1}{n^2+n} &= \frac{1}{n^2} - \frac{n}{n^4} + \frac{n^2}{n^6} - \frac{n^3}{n^8} + O\left(\frac{n^4}{n^{10}}\right) \\ \hline S_n &= \frac{n}{n^2} - \frac{n(n+1)}{2n^4} + \cdots. \end{aligned}$$

It looks as if we're getting $S_n = n^{-1} - \frac{1}{2}n^{-2} + O(n^{-3})$, based on the sums of the first two columns; but the calculations are getting hairy.

If we persevere in this approach, we will ultimately reach the goal; but we won't bother to sum the other columns, for two reasons: First, the last column is going to give us terms that are $O(n^{-6})$, when $n/2 \leq k \leq n$, so we will have an error of $O(n^{-5})$; that's too big, and we will have to include yet another column in the expansion. Could the exam-giver have been so sadistic? We suspect that there must be a better way. Second, there is indeed a much better way, staring us right in the face.

Do pajamas have buttons?

Namely, we know a closed form for S_n : It's just $H_{n^2+n} - H_{n^2}$. And we know a good approximation for harmonic numbers, so we just apply it twice:

$$\begin{aligned} H_{n^2+n} &= \ln(n^2+n) + \gamma + \frac{1}{2(n^2+n)} - \frac{1}{12(n^2+n)^2} + O\left(\frac{1}{n^8}\right); \\ H_{n^2} &= \ln n^2 + \gamma + \frac{1}{2n^2} - \frac{1}{12n^4} + O\left(\frac{1}{n^8}\right). \end{aligned}$$

Now we can pull out large terms and simplify, as we did when looking at Stirling's approximation. We have

$$\begin{aligned} \ln(n^2+n) &= \ln n^2 + \ln\left(1 + \frac{1}{n}\right) = \ln n^2 + \frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} - \cdots; \\ \frac{1}{n^2+n} &= \frac{1}{n^2} - \frac{1}{n^3} + \frac{1}{n^4} - \cdots; \\ \frac{1}{(n^2+n)^2} &= \frac{1}{n^4} - \frac{2}{n^5} + \frac{3}{n^6} - \cdots. \end{aligned}$$

So there's lots of helpful cancellation, and we find

$$\begin{aligned} S_n &= n^{-1} - \frac{1}{2}n^{-2} + \frac{1}{3}n^{-3} - \frac{1}{4}n^{-4} + \frac{1}{5}n^{-5} - \frac{1}{6}n^{-6} \\ &\quad - \frac{1}{2}n^{-3} + \frac{1}{2}n^{-4} - \frac{1}{2}n^{-5} + \frac{1}{2}n^{-6} \\ &\quad + \frac{1}{6}n^{-5} - \frac{1}{4}n^{-6} \end{aligned}$$

plus terms that are $O(n^{-7})$. A bit of arithmetic and we're home free:

$$S_n = n^{-1} - \frac{1}{2}n^{-2} - \frac{1}{6}n^{-3} + \frac{1}{4}n^{-4} - \frac{2}{15}n^{-5} + \frac{1}{12}n^{-6} + O(n^{-7}). \quad (9.50)$$

It would be nice if we could check this answer numerically, as we did when we derived exact results in earlier chapters. Asymptotic formulas are harder to verify; an arbitrarily large constant may be hiding in a O term, so any numerical test is inconclusive. But in practice, we have no reason to believe that an adversary is trying to trap us, so we can assume that the unknown O -constants are reasonably small. With a pocket calculator we find that $S_4 = \frac{1}{17} + \frac{1}{18} + \frac{1}{19} + \frac{1}{20} = 0.2170107$; and our asymptotic estimate when $n = 4$ comes to

$$\frac{1}{4}\left(1 + \frac{1}{4}\left(-\frac{1}{2} + \frac{1}{4}\left(-\frac{1}{6} + \frac{1}{4}\left(\frac{1}{4} + \frac{1}{4}\left(-\frac{2}{15} + \frac{1}{4} \cdot \frac{1}{12}\right)\right)\right)\right)\right) = 0.2170125.$$

If we had made an error of, say, $\frac{1}{12}$ in the term for n^{-6} , a difference of $\frac{1}{12} \cdot \frac{1}{4096}$ would have shown up in the fifth decimal place; so our asymptotic answer is probably correct.

Problem 5: An infinite sum.

We turn now to an asymptotic question posed by Solomon Golomb [152]: What is the approximate value of

$$S_n = \sum_{k \geq 1} \frac{1}{k N_n(k)^2}, \quad (9.51)$$

where $N_n(k)$ is the number of digits required to write k in radix n notation?

First let's try again for a ballpark estimate. The number of digits, $N_n(k)$, is approximately $\log_n k = \log k / \log n$; so the terms of this sum are roughly $(\log n)^2 / k (\log k)^2$. Summing on k gives $\approx (\log n)^2 \sum_{k \geq 2} 1/k (\log k)^2$, and this sum converges to a constant value because it can be compared to the integral

$$\int_2^\infty \frac{dx}{x (\ln x)^2} = -\frac{1}{\ln x} \Big|_2^\infty = \frac{1}{\ln 2}.$$

Therefore we expect S_n to be about $C(\log n)^2$, for some constant C .

Hand-wavy analyses like this are useful for orientation, but we need better estimates to solve the problem. One idea is to express $N_n(k)$ exactly:

$$N_n(k) = \lfloor \log_n k \rfloor + 1. \quad (9.52)$$

Thus, for example, k has three radix n digits when $n^2 \leq k < n^3$, and this happens precisely when $\lfloor \log_n k \rfloor = 2$. It follows that $N_n(k) > \log_n k$, hence $S_n = \sum_{k \geq 1} 1/k N_n(k)^2 < 1 + (\log n)^2 \sum_{k \geq 2} 1/k (\log k)^2$.

Proceeding as in Problem 1, we can try to write $N_n(k) = \log_n k + O(1)$ and substitute this into the formula for S_n . The term represented here by $O(1)$ is always between 0 and 1, and it is about $\frac{1}{2}$ on the average, so it seems rather well-behaved. But still, this isn't a good enough approximation to tell us about S_n ; it gives us zero significant figures (that is, high relative error) when k is small, and these are the terms that contribute the most to the sum. We need a different idea.

The key (as in Problem 4) is to use our manipulative skills to put the sum into a more tractable form, before we resort to asymptotic estimates. We can introduce a new variable of summation, $m = N_n(k)$:

$$\begin{aligned} S_n &= \sum_{k, m \geq 1} \frac{[m = N_n(k)]}{k m^2} \\ &= \sum_{k, m \geq 1} \frac{[n^{m-1} \leq k < n^m]}{k m^2} \\ &= \sum_{m \geq 1} \frac{1}{m^2} (H_{n^m-1} - H_{n^{m-1}-1}). \end{aligned}$$

This may look worse than the sum we began with, but it's actually a step forward, because we have very good approximations for the harmonic numbers.

Still, we hold back and try to simplify some more. No need to rush into asymptotics. Summation by parts allows us to group the terms for each value of $H_{n^{k-1}}$ that we need to approximate:

$$S_n = \sum_{k \geq 1} H_{n^{k-1}} \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right).$$

For example, $H_{n^{2-1}}$ is multiplied by $1/2^2$ and then by $-1/3^2$. (We have used the fact that $H_{n^{0-1}} = H_0 = 0$.)

Now we're ready to expand the harmonic numbers. Our experience with estimating $(n-1)!$ has taught us that it will be easier to estimate H_{n^k} than $H_{n^{k-1}}$, since the $(n^k - 1)$'s will be messy; therefore we write

$$\begin{aligned} H_{n^{k-1}} &= H_{n^k} - \frac{1}{n^k} = \ln n^k + \gamma + \frac{1}{2n^k} + O\left(\frac{1}{n^{2k}}\right) - \frac{1}{n^k} \\ &= k \ln n + \gamma - \frac{1}{2n^k} + O\left(\frac{1}{n^{2k}}\right). \end{aligned}$$

Our sum now reduces to

$$\begin{aligned} S_n &= \sum_{k \geq 1} \left(k \ln n + \gamma - \frac{1}{2n^k} + O\left(\frac{1}{n^{2k}}\right) \right) \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right) \\ &= (\ln n) \Sigma_1 + \gamma \Sigma_2 - \frac{1}{2} \Sigma_3(n) + O(\Sigma_3(n^2)). \end{aligned} \quad (9.53)$$

There are four easy pieces left: Σ_1 , Σ_2 , $\Sigma_3(n)$, and $\Sigma_3(n^2)$.

Let's do the Σ_3 's first, since $\Sigma_3(n^2)$ is the O term; then we'll see what sort of error we're getting. (There's no sense carrying out other calculations with perfect accuracy if they will be absorbed into a O anyway.) This sum is simply a power series,

Into a Big Oh.

$$\Sigma_3(x) = \sum_{k \geq 1} \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right) x^{-k},$$

and the series converges when $x \geq 1$ so we can truncate it at any desired point. If we stop $\Sigma_3(n^2)$ at the term for $k = 1$, we get $\Sigma_3(n^2) = O(n^{-2})$; hence (9.53) has an absolute error of $O(n^{-2})$. (To decrease this absolute error, we could use a better approximation to H_{n^k} ; but $O(n^{-2})$ is good enough for now.) If we truncate $\Sigma_3(n)$ at the term for $k = 2$, we get

$$\Sigma_3(n) = \frac{3}{4} n^{-1} + O(n^{-2});$$

this is all the accuracy we need.

We might as well do Σ_2 now, since it is so easy:

$$\Sigma_2 = \sum_{k \geq 1} \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right).$$

This is the telescoping series $(1 - \frac{1}{4}) + (\frac{1}{4} - \frac{1}{9}) + (\frac{1}{9} - \frac{1}{16}) + \cdots = 1$.

Finally, Σ_1 gives us the leading term of S_n , the coefficient of $\ln n$ in (9.53):

$$\Sigma_1 = \sum_{k \geq 1} k \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right).$$

This is $(1 - \frac{1}{4}) + (\frac{2}{4} - \frac{2}{9}) + (\frac{3}{9} - \frac{3}{16}) + \cdots = \frac{1}{1} + \frac{1}{4} + \frac{1}{9} + \cdots = H_\infty^{(2)} = \pi^2/6$. (If we hadn't applied summation by parts earlier, we would have seen directly that $S_n \sim \sum_{k \geq 1} (\ln n)/k^2$, because $H_{n^{k-1}} - H_{n^{k-1}-1} \sim \ln n$; so summation by parts didn't help us to evaluate the leading term, although it did make some of our other work easier.)

Now we have evaluated each of the Σ 's in (9.53), so we can put everything together and get the answer to Golomb's problem:

$$S_n = \frac{\pi^2}{6} \ln n + \gamma - \frac{3}{8n} + O\left(\frac{1}{n^2}\right). \quad (9.54)$$

Notice that this grows more slowly than our original hand-wavy estimate of $C(\log n)^2$. Sometimes a discrete sum fails to obey a continuous intuition.

Problem 6: Big Phi.

Near the end of Chapter 4, we observed that the number of fractions in the Farey series \mathcal{F}_n is $1 + \Phi(n)$, where

$$\Phi(n) = \varphi(1) + \varphi(2) + \cdots + \varphi(n);$$

and we showed in (4.62) that

$$\Phi(n) = \frac{1}{2} \sum_{k \geq 1} \mu(k) \lfloor n/k \rfloor \lfloor 1 + n/k \rfloor. \quad (9.55)$$

Let us now try to estimate $\Phi(n)$ when n is large. (It was sums like this that led Bachmann to invent O -notation in the first place.)

Thinking BIG tells us that $\Phi(n)$ will probably be proportional to n^2 . For if the final factor were just $\lfloor n/k \rfloor$ instead of $\lfloor 1 + n/k \rfloor$, we would have $|\Phi(n)| \leq \frac{1}{2} \sum_{k \geq 1} \lfloor n/k \rfloor^2 \leq \frac{1}{2} \sum_{k \geq 1} (n/k)^2 = \frac{\pi^2}{12} n^2$, because the Möbius function $\mu(k)$ is either -1 , 0 , or $+1$. The additional ' $1 +$ ' in that final factor adds $\sum_{k \geq 1} \mu(k) \lfloor n/k \rfloor$; but this is zero for $k > n$, so it cannot be more than $nH_n = O(n \log n)$ in absolute value.

This preliminary analysis indicates that we'll find it advantageous to write

$$\begin{aligned}\Phi(n) &= \frac{1}{2} \sum_{k=1}^n \mu(k) \left(\left(\frac{n}{k} \right) + O(1) \right)^2 = \frac{1}{2} \sum_{k=1}^n \mu(k) \left(\left(\frac{n}{k} \right)^2 + O\left(\frac{n}{k} \right) \right) \\ &= \frac{1}{2} \sum_{k=1}^n \mu(k) \left(\frac{n}{k} \right)^2 + \sum_{k=1}^n O\left(\frac{n}{k} \right) \\ &= \frac{1}{2} \sum_{k=1}^n \mu(k) \left(\frac{n}{k} \right)^2 + O(n \log n).\end{aligned}$$

This removes the floors; the remaining problem is to evaluate the unfloored sum $\frac{1}{2} \sum_{k=1}^n \mu(k) n^2/k^2$ with an accuracy of $O(n \log n)$; in other words, we want to evaluate $\sum_{k=1}^n \mu(k)/k^2$ with an accuracy of $O(n^{-1} \log n)$. But that's easy; we can simply run the sum all the way up to $k = \infty$, because the newly added terms are

$$\begin{aligned}\sum_{k>n} \frac{\mu(k)}{k^2} &= O\left(\sum_{k>n} \frac{1}{k^2} \right) = O\left(\sum_{k>n} \frac{1}{k(k-1)} \right) \\ &= O\left(\sum_{k>n} \left(\frac{1}{k-1} - \frac{1}{k} \right) \right) = O\left(\frac{1}{n} \right).\end{aligned}$$

We proved in (7.89) that $\sum_{k \geq 1} \mu(k)/k^z = 1/\zeta(z)$. Hence $\sum_{k \geq 1} \mu(k)/k^2 = 1/(\sum_{k \geq 1} 1/k^2) = 6/\pi^2$, and we have our answer:

$$\Phi(n) = \frac{3}{\pi^2} n^2 + O(n \log n). \quad (9.56)$$

(The error term was shown to be at most $O(n(\log n)^{2/3} \times (\log \log n)^{1+\epsilon})$ by Saltykov in 1960 [316]. On the other hand, it is not as small as $o(n(\log \log n)^{1/2})$, according to Montgomery [275].)

9.4 TWO ASYMPTOTIC TRICKS

Now that we have some facility with O manipulations, let's look at what we've done from a slightly higher perspective. Then we'll have some important weapons in our asymptotic arsenal, when we need to do battle with tougher problems.

Trick 1: Bootstrapping.

When we estimated the n th prime P_n in Problem 3 of Section 9.3, we solved an asymptotic recurrence of the form

$$P_n = n \ln P_n (1 + O(1/\log n)).$$

We proved that $P_n = n \ln n + O(n)$ by first using the recurrence to show the weaker result $O(n^2)$. This is a special case of a general method called *bootstrapping*, in which we solve a recurrence asymptotically by starting with

a rough estimate and plugging it into the recurrence; in this way we can often derive better and better estimates, “pulling ourselves up by our bootstraps.”

Here’s another problem that illustrates bootstrapping nicely: What is the asymptotic value of the coefficient $g_n = [z^n] G(z)$ in the generating function

$$G(z) = \exp\left(\sum_{k \geq 1} \frac{z^k}{k^2}\right), \quad (9.57)$$

as $n \rightarrow \infty$? If we differentiate this equation with respect to z , we find

$$G'(z) = \sum_{n=0}^{\infty} n g_n z^{n-1} = \left(\sum_{k \geq 1} \frac{z^{k-1}}{k}\right) G(z);$$

equating coefficients of z^{n-1} on both sides gives the recurrence

$$n g_n = \sum_{0 \leq k < n} \frac{g_k}{n-k}. \quad (9.58)$$

Our problem is equivalent to finding an asymptotic formula for the solution to (9.58), with the initial condition $g_0 = 1$. The first few values

n	0	1	2	3	4	5	6
g_n	1	1	$\frac{3}{4}$	$\frac{19}{36}$	$\frac{107}{288}$	$\frac{641}{2400}$	$\frac{51103}{259200}$

don’t reveal much of a pattern, and the integer sequence $\langle n!^2 g_n \rangle$ doesn’t appear in Sloane’s *Handbook* [330]; therefore a closed form for g_n seems out of the question, and asymptotic information is probably the best we can hope to derive.

Our first handle on this problem is the observation that $0 < g_n \leq 1$ for all $n \geq 0$; this is easy to prove by induction. So we have a start:

$$g_n = O(1).$$

This equation can, in fact, be used to “prime the pump” for a bootstrapping operation: Plugging it in on the right of (9.58) yields

$$n g_n = \sum_{0 \leq k < n} \frac{O(1)}{n-k} = H_n O(1) = O(\log n);$$

hence we have

$$g_n = O\left(\frac{\log n}{n}\right), \quad \text{for } n > 1.$$

And we can bootstrap yet again:

$$\begin{aligned}
 ng_n &= \frac{1}{n} + \sum_{0 < k < n} \frac{O((1 + \log k)/k)}{n - k} \\
 &= \frac{1}{n} + \sum_{0 < k < n} \frac{O(\log n)}{k(n - k)} \\
 &= \frac{1}{n} + \sum_{0 < k < n} \left(\frac{1}{k} + \frac{1}{n - k} \right) \frac{O(\log n)}{n} \\
 &= \frac{1}{n} + \frac{2}{n} H_{n-1} O(\log n) = \frac{1}{n} O(\log n)^2,
 \end{aligned}$$

obtaining

$$g_n = O\left(\frac{\log n}{n}\right)^2. \quad (9.59)$$

Will this go on forever? Perhaps we'll have $g_n = O(n^{-1} \log n)^m$ for all m .

Actually no; we have just reached a point of diminishing returns. The next attempt at bootstrapping involves the sum

$$\begin{aligned}
 \sum_{0 < k < n} \frac{1}{k^2(n - k)} &= \sum_{0 < k < n} \left(\frac{1}{nk^2} + \frac{1}{n^2k} + \frac{1}{n^2(n - k)} \right) \\
 &= \frac{1}{n} H_{n-1}^{(2)} + \frac{2}{n^2} H_{n-1},
 \end{aligned}$$

which is $\Omega(n^{-1})$; so we cannot get an estimate for g_n that falls below $\Omega(n^{-2})$.

In fact, we now know enough about g_n to apply our old trick of pulling out the largest part:

$$\begin{aligned}
 ng_n &= \sum_{0 \leq k < n} \frac{g_k}{n} + \sum_{0 \leq k < n} g_k \left(\frac{1}{n - k} - \frac{1}{n} \right) \\
 &= \frac{1}{n} \sum_{k \geq 0} g_k - \frac{1}{n} \sum_{k \geq n} g_k + \frac{1}{n} \sum_{0 \leq k < n} \frac{kg_k}{n - k}.
 \end{aligned} \quad (9.60)$$

The first sum here is $G(1) = \exp(\frac{1}{1} + \frac{1}{4} + \frac{1}{9} + \dots) = e^{\pi^2/6}$, because $G(z)$ converges for all $|z| \leq 1$. The second sum is the tail of the first; we can get an upper bound by using (9.59):

$$\sum_{k \geq n} g_k = O\left(\sum_{k \geq n} \frac{(\log k)^2}{k^2}\right) = O\left(\frac{(\log n)^2}{n}\right).$$

This last estimate follows because, for example,

$$\sum_{k > n} \frac{(\log k)^2}{k^2} < \sum_{m \geq 1} \sum_{n^m < k \leq n^{m+1}} \frac{(\log n^{m+1})^2}{k(k-1)} < \sum_{m \geq 1} \frac{(m+1)^2 (\log n)^2}{n^m}.$$

(Exercise 54 discusses a more general way to estimate such tails.)

The third sum in (9.60) is

$$O\left(\sum_{0 \leq k < n} \frac{(\log n)^2}{k(n-k)}\right) = O\left(\frac{(\log n)^3}{n}\right),$$

by an argument that's already familiar. So (9.60) proves that

$$g_n = \frac{e^{\pi^2/6}}{n^2} + O(\log n/n^3). \quad (9.61)$$

Finally, we can feed this formula back into the recurrence, bootstrapping once more; the result is

$$g_n = \frac{e^{\pi^2/6}}{n^2} + O(\log n/n^3). \quad (9.62)$$

(Exercise 23 peeks inside the remaining O term.)

Trick 2: Trading tails.

We derived (9.62) in somewhat the same way we derived the asymptotic value (9.56) of $\Phi(n)$: In both cases we started with a finite sum but got an asymptotic value by considering an infinite sum. We couldn't simply get the infinite sum by introducing O into the summand; we had to be careful to use one approach when k was small and another when k was large.

Those derivations were special cases of an important three-step asymptotic summation method we will now discuss in greater generality. Whenever we want to estimate the value of $\sum_k a_k(n)$, we can try the following approach:

(This important method was pioneered by Laplace [240].)

- 1 First break the sum into two disjoint ranges, D_n and T_n . The summation over D_n should be the "dominant" part, in the sense that it includes enough terms to determine the significant digits of the sum, when n is large. The summation over the other range T_n should be just the "tail" end, which contributes little to the overall total.
- 2 Find an asymptotic estimate

$$a_k(n) = b_k(n) + O(c_k(n))$$

that is valid when $k \in D_n$. The O bound need not hold when $k \in T_n$.

3 Now prove that each of the following three sums is small:

$$\begin{aligned}\Sigma_a(n) &= \sum_{k \in T_n} a_k(n); & \Sigma_b(n) &= \sum_{k \in T_n} b_k(n); \\ \Sigma_c(n) &= \sum_{k \in D_n} |c_k(n)|.\end{aligned}\tag{9.63}$$

If all three steps can be completed successfully, we have a good estimate:

$$\sum_{k \in D_n \cup T_n} a_k(n) = \sum_{k \in D_n \cup T_n} b_k(n) + O(\Sigma_a(n)) + O(\Sigma_b(n)) + O(\Sigma_c(n)).$$

Here's why. We can "chop off" the tail of the given sum, getting a good estimate in the range D_n where a good estimate is necessary:

$$\sum_{k \in D_n} a_k(n) = \sum_{k \in D_n} (b_k(n) + O(c_k(n))) = \sum_{k \in D_n} b_k(n) + O(\Sigma_c(n)).$$

And we can replace the tail with another one, even though the new tail might be a terrible approximation to the old, because the tails don't really matter:

$$\begin{aligned}\sum_{k \in T_n} a_k(n) &= \sum_{k \in T_n} (b_k(n) - b_k(n) + a_k(n)) \\ &= \sum_{k \in T_n} b_k(n) + O(\Sigma_b(n)) + O(\Sigma_a(n)).\end{aligned}$$

*Asymptotics is
the art of knowing
where to be sloppy
and where to be
precise.*

When we evaluated the sum in (9.60), for example, we had

$$\begin{aligned}a_k(n) &= [0 \leq k < n] g_k / (n - k), \\ b_k(n) &= g_k / n, \\ c_k(n) &= k g_k / n(n - k);\end{aligned}$$

the ranges of summation were

$$D_n = \{0, 1, \dots, n-1\}, \quad T_n = \{n, n+1, \dots\};$$

and we found that

$$\Sigma_a(n) = 0, \quad \Sigma_b(n) = O((\log n)^2/n^2), \quad \Sigma_c(n) = O((\log n)^3/n^2).$$

This led to (9.61).

Similarly, when we estimated $\Phi(n)$ in (9.55) we had

$$\begin{aligned}a_k(n) &= \mu(k) \lfloor n/k \rfloor \lfloor 1+n/k \rfloor, & b_k(n) &= \mu(k) n^2/k^2, & c_k(n) &= n/k; \\ D_n &= \{1, 2, \dots, n\}, & T_n &= \{n+1, n+2, \dots\}.\end{aligned}$$

We derived (9.56) by observing that $\Sigma_a(n) = 0$, $\Sigma_b(n) = O(n)$, and $\Sigma_c(n) = O(n \log n)$.

Here's another example where tail switching is effective. (Unlike our previous examples, this one illustrates the trick in its full generality, with $\Sigma_a(n) \neq 0$.) We seek the asymptotic value of

$$L_n = \sum_{k \geq 0} \frac{\ln(n + 2^k)}{k!}.$$

The big contributions to this sum occur when k is small, because of the $k!$ in the denominator. In this range we have

$$\ln(n + 2^k) = \ln n + \frac{2^k}{n} - \frac{2^{2k}}{2n^2} + O\left(\frac{2^{3k}}{n^3}\right). \quad (9.64)$$

We can prove that this estimate holds for $0 \leq k < \lfloor \lg n \rfloor$, since the original terms that have been truncated with O are bounded by the convergent series

$$\sum_{m \geq 3} \frac{2^{km}}{mn^m} \leq \frac{2^{3k}}{n^3} \sum_{m \geq 3} \frac{2^{k(m-3)}}{n^{m-3}} \leq \frac{2^{3k}}{n^3} \left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) = \frac{2^{3k}}{n^3} \cdot 2.$$

(In this range, $2^k/n \leq 2^{\lfloor \lg n \rfloor - 1}/n \leq \frac{1}{2}$.)

Therefore we can apply the three-step method just described, with

$$\begin{aligned} a_k(n) &= \ln(n + 2^k)/k!, \\ b_k(n) &= (\ln n + 2^k/n - 4^k/2n^2)/k!, \\ c_k(n) &= 8^k/n^3 k!; \\ D_n &= \{0, 1, \dots, \lfloor \lg n \rfloor - 1\}, \\ T_n &= \{\lfloor \lg n \rfloor, \lfloor \lg n \rfloor + 1, \dots\}. \end{aligned}$$

All we have to do is find good bounds on the three Σ 's in (9.63), and we'll know that $\sum_{k \geq 0} a_k(n) \approx \sum_{k \geq 0} b_k(n)$.

The error we have committed in the dominant part of the sum, $\Sigma_c(n) = \sum_{k \in D_n} 8^k/n^3 k!$, is obviously bounded by $\sum_{k \geq 0} 8^k/n^3 k! = e^8/n^3$, so it can be replaced by $O(n^{-3})$. The new tail error is

$$\begin{aligned} |\Sigma_b(n)| &= \left| \sum_{k \geq \lfloor \lg n \rfloor} b_k(n) \right| \\ &< \sum_{k \geq \lfloor \lg n \rfloor} \frac{\ln n + 2^k + 4^k}{k!} \\ &< \frac{\ln n + 2^{\lfloor \lg n \rfloor} + 4^{\lfloor \lg n \rfloor}}{\lfloor \lg n \rfloor!} \sum_{k \geq 0} \frac{4^k}{k!} = O\left(\frac{n^2}{\lfloor \lg n \rfloor!}\right). \end{aligned}$$

Also, horses switch their tails when feeding time approaches.

*"We may not be big,
but we're small."*

Since $[\lg n]!$ grows faster than any power of n , this minuscule error is overwhelmed by $\Sigma_c(n) = O(n^{-3})$. The error that comes from the original tail,

$$\Sigma_a(n) = \sum_{k \geq [\lg n]} a_k(n) < \sum_{k \geq [\lg n]} \frac{k + \ln n}{k!},$$

is smaller yet.

Finally, it's easy to sum $\sum_{k \geq 0} b_k(n)$ in closed form, and we have obtained the desired asymptotic formula:

$$\sum_{k \geq 0} \frac{\ln(n + 2^k)}{k!} = e \ln n + \frac{e^2}{n} - \frac{e^4}{2n^2} + O\left(\frac{1}{n^3}\right). \quad (9.65)$$

The method we've used makes it clear that, in fact,

$$\sum_{k \geq 0} \frac{\ln(n + 2^k)}{k!} = e \ln n + \sum_{k=1}^{m-1} (-1)^{k+1} \frac{e^{2^k}}{kn^k} + O\left(\frac{1}{n^m}\right), \quad (9.66)$$

for any fixed $m > 0$. (This is a truncation of a series that diverges for all fixed n if we let $m \rightarrow \infty$.)

There's only one flaw in our solution: We were too cautious. We derived (9.64) on the assumption that $k < [\lg n]$, but exercise 53 proves that the stated estimate is actually valid for all values of k . If we had known the stronger general result, we wouldn't have had to use the two-tail trick; we could have gone directly to the final formula! But later we'll encounter problems where exchange of tails is the only decent approach available.

9.5 EULER'S SUMMATION FORMULA

And now for our next trick—which is, in fact, the last important technique that will be discussed in this book—we turn to a general method of approximating sums that was first published by Leonhard Euler [101] in 1732. (The idea is sometimes also associated with the name of Colin Maclaurin, a professor of mathematics at Edinburgh who discovered it independently a short time later [263, page 305].)

Here's the formula:

$$\sum_{a \leq k < b} f(k) = \int_a^b f(x) dx + \sum_{k=1}^m \frac{B_k}{k!} f^{(k-1)}(x) \Big|_a^b + R_m, \quad (9.67)$$

$$\text{where } R_m = (-1)^{m+1} \int_a^b \frac{B_m(\{x\})}{m!} f^{(m)}(x) dx, \quad \begin{array}{l} \text{integers } a \leq b; \\ \text{integer } m \geq 1. \end{array} \quad (9.68)$$

On the left is a typical sum that we might want to evaluate. On the right is another expression for that sum, involving integrals and derivatives. If $f(x)$ is a sufficiently “smooth” function, it will have m derivatives $f'(x), \dots, f^{(m)}(x)$, and this formula turns out to be an identity. The right-hand side is often an excellent approximation to the sum on the left, in the sense that the remainder R_m is often small. For example, we’ll see that Stirling’s approximation for $n!$ is a consequence of Euler’s summation formula; so is our asymptotic approximation for the harmonic number H_n .

The numbers B_k in (9.67) are the Bernoulli numbers that we met in Chapter 6; the function $B_m(\{x\})$ in (9.68) is the Bernoulli polynomial that we met in Chapter 7. The notation $\{x\}$ stands for the fractional part $x - \lfloor x \rfloor$, as in Chapter 3. Euler’s summation formula sort of brings everything together.

Let’s recall the values of small Bernoulli numbers, since it’s always handy to have them listed near Euler’s general formula:

$$\begin{aligned} B_0 &= 1, & B_1 &= -\frac{1}{2}, & B_2 &= \frac{1}{6}, & B_4 &= -\frac{1}{30}, & B_6 &= \frac{1}{42}, & B_8 &= -\frac{1}{30}; \\ B_3 &= B_5 = B_7 = B_9 = B_{11} = \dots = 0. \end{aligned}$$

Jakob Bernoulli discovered these numbers when studying the sums of powers of integers, and Euler’s formula explains why: If we set $f(x) = x^{m-1}$, we have $f^{(m)}(x) = 0$; hence $R_m = 0$, and (9.67) reduces to

$$\begin{aligned} \sum_{a \leq k < b} k^{m-1} &= \left. \frac{x^m}{m} \right|_a^b + \sum_{k=1}^m \frac{B_k}{k!} (m-1)^{\overline{k-1}} x^{m-k} \Big|_a^b \\ &= \frac{1}{m} \sum_{k=0}^m \binom{m}{k} B_k \cdot (b^{m-k} - a^{m-k}). \end{aligned}$$

For example, when $m = 3$ we have our favorite example of summation:

$$\sum_{0 \leq k < n} k^2 = \frac{1}{3} \left(\binom{3}{0} B_0 n^3 + \binom{3}{1} B_1 n^2 + \binom{3}{2} B_2 n \right) = \frac{n^3}{3} - \frac{n^2}{2} + \frac{n}{6}.$$

(This is the last time we shall derive this famous formula in this book.)

Before we prove Euler’s formula, let’s look at a high-level reason (due to Lagrange [234]) why such a formula ought to exist. Chapter 2 defines the difference operator Δ and explains that \sum is the inverse of Δ , just as \int is the inverse of the derivative operator D . We can express Δ in terms of D using Taylor’s formula as follows:

$$f(x + \epsilon) = f(x) + \frac{f'(x)}{1!} \epsilon + \frac{f''(x)}{2!} \epsilon^2 + \dots.$$

*All good things
must come to
an end.*

Setting $\epsilon = 1$ tells us that

$$\begin{aligned}\Delta f(x) &= f(x+1) - f(x) \\ &= f'(x)/1! + f''(x)/2! + f'''(x)/3! + \cdots \\ &= (D/1! + D^2/2! + D^3/3! + \cdots) f(x) = (e^D - 1) f(x).\end{aligned}\quad (9.69)$$

Here e^D stands for the differential operation $1 + D/1! + D^2/2! + D^3/3! + \cdots$. Since $\Delta = e^D - 1$, the inverse operator $\Sigma = 1/\Delta$ should be $1/(e^D - 1)$; and we know from Table 351 that $z/(e^z - 1) = \sum_{k \geq 0} B_k z^k/k!$ is a power series involving Bernoulli numbers. Thus

$$\Sigma = \frac{B_0}{D} + \frac{B_1}{1!} + \frac{B_2}{2!} D + \frac{B_3}{3!} D^2 + \cdots = \int + \sum_{k \geq 1} \frac{B_k}{k!} D^{k-1}.\quad (9.70)$$

Applying this operator equation to $f(x)$ and attaching limits yields

$$\sum_a^b f(x) \delta x = \int_a^b f(x) dx + \sum_{k \geq 1} \frac{B_k}{k!} f^{(k-1)}(x) \Big|_a^b,\quad (9.71)$$

which is exactly Euler's summation formula (9.67) without the remainder term. (Euler did not, in fact, consider the remainder, nor did anybody else until S. D. Poisson [295] published an important memoir about approximate summation in 1823. The remainder term is important, because the infinite sum $\sum_{k \geq 1} (B_k/k!) f^{(k-1)}(x) \Big|_a^b$ often diverges. Our derivation of (9.71) has been purely formal, without regard to convergence.)

Now let's prove (9.67), with the remainder included. It suffices to prove the case $a = 0$ and $b = 1$, namely

$$f(0) = \int_0^1 f(x) dx + \sum_{k=1}^m \frac{B_k}{k!} f^{(k-1)}(x) \Big|_0^1 - (-1)^m \int_0^1 \frac{B_m(x)}{m!} f^{(m)}(x) dx,$$

because we can then replace $f(x)$ by $f(x+l)$ for any integer l , getting

$$f(l) = \int_l^{l+1} f(x) dx + \sum_{k=1}^m \frac{B_k}{k!} f^{(k-1)}(x) \Big|_l^{l+1} - (-1)^m \int_l^{l+1} \frac{B_m(\{x\})}{m!} f^{(m)}(x) dx.$$

The general formula (9.67) is just the sum of this identity over the range $a \leq l < b$, because intermediate terms telescope nicely.

The proof when $a = 0$ and $b = 1$ is by induction on m , starting with $m = 1$:

$$f(0) = \int_0^1 f(x) dx - \frac{1}{2}(f(1) - f(0)) + \int_0^1 (x - \frac{1}{2}) f'(x) dx.$$

(The Bernoulli polynomial $B_m(x)$ is defined by the equation

$$B_m(x) = \binom{m}{0} B_0 x^m + \binom{m}{1} B_1 x^{m-1} + \cdots + \binom{m}{m} B_m x^0 \quad (9.72)$$

in general, hence $B_1(x) = x - \frac{1}{2}$ in particular.) In other words, we want to prove that

$$\frac{f(0) + f(1)}{2} = \int_0^1 f(x) dx + \int_0^1 (x - \frac{1}{2}) f'(x) dx.$$

But this is just a special case of the formula

$$u(x)v(x)|_0^1 = \int_0^1 u(x) dv(x) + \int_0^1 v(x) du(x) \quad (9.73)$$

for integration by parts, with $u(x) = f(x)$ and $v(x) = x - \frac{1}{2}$. Hence the case $m = 1$ is easy.

To pass from $m - 1$ to m and complete the induction when $m > 1$, we need to show that $R_{m-1} = (B_m/m!)f^{(m-1)}(x)|_0^1 + R_m$, namely that

$$\begin{aligned} (-1)^m \int_0^1 \frac{B_{m-1}(x)}{(m-1)!} f^{(m-1)}(x) dx \\ = \frac{B_m}{m!} f^{(m-1)}(x) \Big|_0^1 - (-1)^m \int_0^1 \frac{B_m(x)}{m!} f^{(m)}(x) dx. \end{aligned}$$

This reduces to the equation

$$(-1)^m B_m f^{(m-1)}(x) \Big|_0^1 = m \int_0^1 B_{m-1}(x) f^{(m-1)}(x) dx + \int_0^1 B_m(x) f^{(m)}(x) dx.$$

Once again (9.73) applies to these two integrals, with $u(x) = f^{(m-1)}(x)$ and $v(x) = B_m(x)$, because the derivative of the Bernoulli polynomial (9.72) is *Will the authors never get serious?*

$$\begin{aligned} \frac{d}{dx} \sum_k \binom{m}{k} B_k x^{m-k} &= \sum_k \binom{m}{k} (m-k) B_k x^{m-k-1} \\ &= m \sum_k \binom{m-1}{k} B_k x^{m-1-k} = m B_{m-1}(x). \end{aligned} \quad (9.74)$$

(The absorption identity (5.7) was useful here.) Therefore the required formula will hold if and only if

$$(-1)^m B_m f^{(m-1)}(x) \Big|_0^1 = B_m(x) f^{(m-1)}(x) \Big|_0^1.$$

In other words, we need to have

$$(-1)^m B_m = B_m(1) = B_m(0), \quad \text{for } m > 1. \quad (9.75)$$

This is a bit embarrassing, because $B_m(0)$ is obviously equal to B_m , not to $(-1)^m B_m$. But there's no problem really, because $m > 1$; we know that B_m is zero when m is odd. (Still, that was a close call.)

To complete the proof of Euler's summation formula we need to show that $B_m(1) = B_m(0)$, which is the same as saying that

$$\sum_k \binom{m}{k} B_k = B_m, \quad \text{for } m > 1.$$

But this is just the definition of Bernoulli numbers, (6.79), so we're done.

The identity $B'_m(x) = mB_{m-1}(x)$ implies that

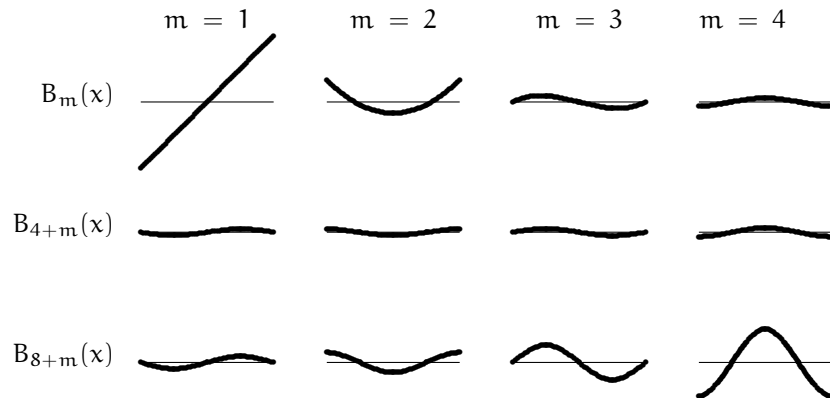
$$\int_0^1 B_m(x) dx = \frac{B_{m+1}(1) - B_{m+1}(0)}{m+1},$$

and we know now that this integral is zero when $m \geq 1$. Hence the remainder term in Euler's formula,

$$R_m = \frac{(-1)^{m+1}}{m!} \int_a^b B_m(\{x\}) f^{(m)}(x) dx,$$

multiplies $f^{(m)}(x)$ by a function $B_m(\{x\})$ whose average value is zero. This means that R_m has a reasonable chance of being small.

Let's look more closely at $B_m(x)$ for $0 \leq x \leq 1$, since $B_m(x)$ governs the behavior of R_m . Here are the graphs for $B_m(x)$ for the first twelve values of m :



Although $B_3(x)$ through $B_9(x)$ are quite small, the Bernoulli polynomials and numbers ultimately get quite large. Fortunately R_m has a compensating factor $1/m!$, which helps to calm things down.

The graph of $B_m(x)$ begins to look very much like a sine wave when $m \geq 3$; exercise 58 proves that $B_m(x)$ can in fact be well approximated by a negative multiple of $\cos(2\pi x - \frac{1}{2}\pi m)$, with relative error $1/2^m$.

In general, $B_{4k+1}(x)$ is negative for $0 < x < \frac{1}{2}$ and positive for $\frac{1}{2} < x < 1$. Therefore its integral, $B_{4k+2}(x)/(4k+2)$, decreases for $0 < x < \frac{1}{2}$ and increases for $\frac{1}{2} < x < 1$. Moreover, we have

$$B_{4k+1}(1-x) = -B_{4k+1}(x), \quad \text{for } 0 \leq x \leq 1,$$

and it follows that

$$B_{4k+2}(1-x) = B_{4k+2}(x), \quad \text{for } 0 \leq x \leq 1.$$

The constant term B_{4k+2} causes the integral $\int_0^1 B_{4k+2}(x) dx$ to be zero; hence $B_{4k+2} > 0$. The integral of $B_{4k+2}(x)$ is $B_{4k+3}(x)/(4k+3)$, which must therefore be positive when $0 < x < \frac{1}{2}$ and negative when $\frac{1}{2} < x < 1$; furthermore $B_{4k+3}(1-x) = -B_{4k+3}(x)$, so $B_{4k+3}(x)$ has the properties stated for $B_{4k+1}(x)$, but negated. Therefore $B_{4k+4}(x)$ has the properties stated for $B_{4k+2}(x)$, but negated. Therefore $B_{4k+5}(x)$ has the properties stated for $B_{4k+1}(x)$; we have completed a cycle that establishes the stated properties inductively for all k .

According to this analysis, the maximum value of $B_{2m}(x)$ must occur either at $x = 0$ or at $x = \frac{1}{2}$. Exercise 17 proves that

$$B_{2m}(\frac{1}{2}) = (2^{1-2m} - 1)B_{2m}; \quad (9.76)$$

hence we have

$$|B_{2m}(\{x\})| \leq |B_{2m}|. \quad (9.77)$$

This can be used to establish a useful upper bound on the remainder in Euler's summation formula, because we know from (6.89) that

$$\frac{|B_{2m}|}{(2m)!} = \frac{2}{(2\pi)^{2m}} \sum_{k \geq 1} \frac{1}{k^{2m}} = O((2\pi)^{-2m}), \quad \text{when } m > 0.$$

Therefore we can rewrite Euler's formula (9.67) as follows:

$$\begin{aligned} \sum_{a \leq k < b} f(k) &= \int_a^b f(x) dx - \frac{1}{2}f(x)\Big|_a^b + \sum_{k=1}^m \frac{B_{2k}}{(2k)!} f^{(2k-1)}(x)\Big|_a^b \\ &\quad + O((2\pi)^{-2m}) \int_a^b |f^{(2m)}(x)| dx. \end{aligned} \quad (9.78)$$

For example, if $f(x) = e^x$, all derivatives are the same and this formula tells us that $\sum_{a \leq k < b} e^k = (e^b - e^a)(1 - \frac{1}{2} + B_2/2! + B_4/4! + \cdots + B_{2m}/(2m)!) +$

$O((2\pi)^{-2m})$. Of course, we know that this sum is actually a geometric series, equal to $(e^b - e^a)/(e - 1) = (e^b - e^a) \sum_{k \geq 0} B_k/k!$.

If $f^{(2m)}(x) \geq 0$ for $a \leq x \leq b$, the integral $\int_a^b |f^{(2m)}(x)| dx$ is just $f^{(2m-1)}(x)|_a^b$, so we have

$$|R_{2m}| \leq \left| \frac{B_{2m}}{(2m)!} f^{(2m-1)}(x) \right|_a^b;$$

in other words, the remainder is bounded by the magnitude of the *final term* (the term just before the remainder), in this case. We can give an even better estimate if we know that

$$f^{(2m+2)}(x) \geq 0 \quad \text{and} \quad f^{(2m+4)}(x) \geq 0, \quad \text{for } a \leq x \leq b. \quad (9.79)$$

For it turns out that this implies the relation

$$R_{2m} = \theta_m \frac{B_{2m+2}}{(2m+2)!} f^{(2m+1)}(x) \Big|_a^b, \quad \text{for some } 0 < \theta_m < 1; \quad (9.80)$$

in other words, the remainder will then lie between 0 and the *first discarded term* in (9.78)—the term that would follow the final term if we increased m .

Here's the proof: Euler's summation formula is valid for all m , and $B_{2m+1} = 0$ when $m > 0$; hence $R_{2m} = R_{2m+1}$, and the first discarded term must be

$$R_{2m} - R_{2m+2}.$$

We therefore want to show that R_{2m} lies between 0 and $R_{2m} - R_{2m+2}$; and this is true if and only if R_{2m} and R_{2m+2} have opposite signs. We claim that

$$f^{(2m+2)}(x) \geq 0 \quad \text{for } a \leq x \leq b \quad \text{implies} \quad (-1)^m R_{2m} \geq 0. \quad (9.81)$$

This, together with (9.79), will prove that R_{2m} and R_{2m+2} have opposite signs, so the proof of (9.80) will be complete.

It's not difficult to prove (9.81) if we recall the definition of R_{2m+1} and the facts we proved about the graph of $B_{2m+1}(x)$. Namely, we have

$$R_{2m} = R_{2m+1} = \int_a^b \frac{B_{2m+1}(\{x\})}{(2m+1)!} f^{(2m+1)}(x) dx,$$

and $f^{(2m+1)}(x)$ is increasing because its derivative $f^{(2m+2)}(x)$ is positive. (More precisely, $f^{(2m+1)}(x)$ is nondecreasing because its derivative is non-negative.) The graph of $B_{2m+1}(\{x\})$ looks like $(-1)^{m+1}$ times a sine wave, so it is geometrically obvious that the second half of each sine wave is more influential than the first half when it is multiplied by an increasing function. This makes $(-1)^m R_{2m+1} \geq 0$, as desired. Exercise 16 proves the result formally.

9.6 FINAL SUMMATIONS

Now comes the summing up, as we prepare to conclude this book. We will apply Euler's summation formula to some interesting and important examples.

Summation 1: This one is too easy.

But first we will consider an interesting *unimportant* example, namely a sum that we already know how to do. Let's see what Euler's summation formula tells us if we apply it to the telescoping sum

$$S_n = \sum_{1 \leq k \leq n} \frac{1}{k(k+1)} = \sum_{1 \leq k \leq n} \left(\frac{1}{k} - \frac{1}{k+1} \right) = 1 - \frac{1}{n}.$$

It can't hurt to embark on our first serious application of Euler's formula with the asymptotic equivalent of training wheels.

We might as well start by writing the function $f(x) = 1/x(x+1)$ in partial fraction form,

$$f(x) = \frac{1}{x} - \frac{1}{x+1},$$

since this makes it easier to integrate and differentiate. Indeed, we have $f'(x) = -1/x^2 + 1/(x+1)^2$ and $f''(x) = 2/x^3 - 2/(x+1)^3$; in general

$$f^{(k)}(x) = (-1)^k k! \left(\frac{1}{x^{k+1}} - \frac{1}{(x+1)^{k+1}} \right), \quad \text{for } k \geq 0.$$

Furthermore

$$\int_1^n f(x) dx = \ln x - \ln(x+1) \Big|_1^n = \ln \frac{2n}{n+1}.$$

Plugging this into the summation formula (9.67) gives

$$S_n = \ln \frac{2n}{n+1} - \sum_{k=1}^m (-1)^k \frac{B_k}{k} \left(\frac{1}{n^k} - \frac{1}{(n+1)^k} - 1 + \frac{1}{2^k} \right) + R_m(n),$$

where $R_m(n) = - \int_1^n B_m(\{x\}) \left(\frac{1}{x^{m+1}} - \frac{1}{(x+1)^{m+1}} \right) dx$.

For example, the right-hand side when $m = 4$ is

$$\begin{aligned} \ln \frac{2n}{n+1} - \frac{1}{2} \left(\frac{1}{n} - \frac{1}{n+1} - \frac{1}{2} \right) - \frac{1}{12} \left(\frac{1}{n^2} - \frac{1}{(n+1)^2} - \frac{3}{4} \right) \\ + \frac{1}{120} \left(\frac{1}{n^4} - \frac{1}{(n+1)^4} - \frac{15}{16} \right) + R_4(n). \end{aligned}$$

This is kind of a mess; it certainly doesn't look like the real answer $1 - n^{-1}$. But let's keep going anyway, to see what we've got. We know how to expand the right-hand terms in negative powers of n up to, say, $O(n^{-5})$:

$$\begin{aligned}\ln \frac{n}{n+1} &= -n^{-1} + \frac{1}{2}n^{-2} - \frac{1}{3}n^{-3} + \frac{1}{4}n^{-4} + O(n^{-5}); \\ \frac{1}{n+1} &= -n^{-1} - n^{-2} + n^{-3} - n^{-4} + O(n^{-5}); \\ \frac{1}{(n+1)^2} &= n^{-2} - 2n^{-3} + 3n^{-4} + O(n^{-5}); \\ \frac{1}{(n+1)^4} &= n^{-4} + O(n^{-5}).\end{aligned}$$

Therefore the terms on the right of our approximation add up to

$$\begin{aligned}\ln 2 + \frac{1}{4} + \frac{1}{16} - \frac{1}{128} + \left(-1 - \frac{1}{2} + \frac{1}{2}\right)n^{-1} + \left(\frac{1}{2} - \frac{1}{2} - \frac{1}{12} + \frac{1}{12}\right)n^{-2} \\ + \left(-\frac{1}{3} + \frac{1}{2} - \frac{2}{12}\right)n^{-3} + \left(\frac{1}{4} - \frac{1}{2} + \frac{3}{12} + \frac{1}{120} - \frac{1}{120}\right)n^{-4} + R_4(n) \\ = \ln 2 + \frac{39}{128} - n^{-1} + R_4(n) + O(n^{-5}).\end{aligned}$$

The coefficients of n^{-2} , n^{-3} , and n^{-4} cancel nicely, as they should.

If all were well with the world, we would be able to show that $R_4(n)$ is asymptotically small, maybe $O(n^{-5})$, and we would have an approximation to the sum. But we can't possibly show this, because we happen to know that the correct constant term is 1, not $\ln 2 + \frac{39}{128}$ (which is approximately 0.9978). So $R_4(n)$ is actually equal to $\frac{89}{128} - \ln 2 + O(n^{-4})$, but Euler's summation formula doesn't tell us this.

In other words, we lose.

One way to try fixing things is to notice that the constant terms in the approximation form a pattern, if we let m get larger and larger:

$$\ln 2 - \frac{1}{2}B_1 + \frac{1}{2} \cdot \frac{3}{4}B_2 - \frac{1}{3} \cdot \frac{7}{8}B_3 + \frac{1}{4} \cdot \frac{15}{16}B_4 - \frac{1}{5} \cdot \frac{31}{32}B_5 + \dots$$

Perhaps we can show that this series approaches 1 as the number of terms becomes infinite? But no; the Bernoulli numbers get very large. For example, $B_{22} = \frac{854513}{138} > 6192$; therefore $|R_{22}(n)|$ will be much larger than $|R_4(n)|$. We lose totally.

There is a way out, however, and this escape route will turn out to be important in other applications of Euler's formula. The key is to notice that $R_4(n)$ approaches a definite limit as $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} R_4(n) = - \int_1^\infty B_4(\{x\}) \left(\frac{1}{x^5} - \frac{1}{(x+1)^5} \right) dx = R_4(\infty).$$

The integral $\int_1^\infty B_m(\{x\}) f^{(m)}(x) dx$ will exist whenever $f^{(m)}(x) = O(x^{-2})$ as $x \rightarrow \infty$, and in this case $f^{(4)}(x)$ surely qualifies. Moreover, we have

$$\begin{aligned} R_4(n) &= R_4(\infty) + \int_n^\infty B_4(\{x\}) \left(\frac{1}{x^5} - \frac{1}{(x+1)^5} \right) dx \\ &= R_4(\infty) + O\left(\int_n^\infty x^{-6} dx\right) = R_4(\infty) + O(n^{-5}). \end{aligned}$$

Thus we have used Euler's summation formula to prove that

$$\begin{aligned} \sum_{1 \leq k < n} \frac{1}{k(k+1)} &= \ln 2 + \frac{39}{128} - n^{-1} + R_4(\infty) + O(n^{-5}) \\ &= C - n^{-1} + O(n^{-5}) \end{aligned}$$

for some constant C . We do not know what the constant is—some other method must be used to establish it—but Euler's summation formula is able to let us deduce that the constant exists.

Suppose we had chosen a much larger value of m . Then the same reasoning would tell us that

$$R_m(n) = R_m(\infty) + O(n^{-m-1}),$$

and we would have the formula

$$\sum_{1 \leq k < n} \frac{1}{k(k+1)} = C - n^{-1} + c_2 n^{-2} + c_3 n^{-3} + \cdots + c_m n^{-m} + O(n^{-m-1})$$

for certain constants c_2, c_3, \dots . We know that the c 's happen to be zero in this case; but let's prove it, just to restore some of our confidence (in Euler's formula if not in ourselves). The term $\ln \frac{n}{n+1}$ contributes $(-1)^m/m$ to c_m ; the term $(-1)^{m+1}(B_m/m)n^{-m}$ contributes $(-1)^{m+1}B_m/m$; and the term $(-1)^k(B_k/k)(n+1)^{-k}$ contributes $(-1)^m \binom{m-1}{k-1} B_k/k$. Therefore

$$\begin{aligned} (-1)^m c_m &= \frac{1}{m} - \frac{B_m}{m} + \sum_{k=1}^m \binom{m-1}{k-1} \frac{B_k}{k} \\ &= \frac{1}{m} - \frac{B_m}{m} + \frac{1}{m} \sum_{k=1}^m \binom{m}{k} B_k = \frac{1}{m} (1 - B_m + B_m(1) - 1). \end{aligned}$$

Sure enough, it's zero, when $m > 1$. We have proved that

$$\sum_{1 \leq k < n} \frac{1}{k(k+1)} = C - n^{-1} + O(n^{-m-1}), \quad \text{for all } m \geq 1. \quad (9.82)$$

This is not enough to prove that the sum is exactly equal to $C - n^{-1}$; the actual value might be $C - n^{-1} + 2^{-n}$ or something. But Euler's summation

formula does give us the error bound $O(n^{-m-1})$ for arbitrarily large m , even though we haven't evaluated any remainders explicitly.

Summation 1, again: Recapitulation and generalization.

Before we leave our training wheels, let's review what we just did from a somewhat higher perspective. We began with a sum

$$S_n = \sum_{1 \leq k \leq n} f(k)$$

and we used Euler's summation formula to write

$$S_n = F(n) - F(1) + \sum_{k=1}^m (T_k(n) - T_k(1)) + R_m(n), \quad (9.83)$$

where $F(x)$ was $\int f(x) dx$ and where $T_k(x)$ was a certain term involving B_k and $f^{(k-1)}(x)$. We also noticed that there was a constant c such that

$$f^{(m)}(x) = O(x^{c-m}) \quad \text{as } x \rightarrow \infty, \quad \text{for all large } m.$$

(Namely, $f(k)$ was $1/k(k+1)$; $F(x)$ was $\ln(x/(x+1))$; c was -2 ; and $T_k(x)$ was $(-1)^{k+1} (B_k/k) (x^{-k} - (x+1)^{-k})$.) For all large enough values of m , this implied that the remainders had a small tail,

$$\begin{aligned} R'_m(n) &= R_m(\infty) - R_m(n) \\ &= (-1)^{m+1} \int_n^\infty \frac{B_m(\{x\})}{m!} f^{(m)}(x) dx = O(n^{c+1-m}). \end{aligned} \quad (9.84)$$

Therefore we were able to conclude that there exists a constant C such that

$$S_n = F(n) + C + \sum_{k=1}^m T_k(n) - R'_m(n). \quad (9.85)$$

(Notice that C nicely absorbed the $T_k(1)$ terms, which were a nuisance.)

We can save ourselves unnecessary work in future problems by simply asserting the existence of C whenever $R_m(\infty)$ exists.

Now let's suppose that $f^{(2m+2)}(x) \geq 0$ and $f^{(2m+4)}(x) \geq 0$ for $1 \leq x \leq n$. We have proved that this implies a simple bound (9.80) on the remainder,

$$R_{2m}(n) = \theta_{m,n} (T_{2m+2}(n) - T_{2m+2}(1)),$$

where $\theta_{m,n}$ lies somewhere between 0 and 1. But we don't really want bounds that involve $R_{2m}(n)$ and $T_{2m+2}(1)$; after all, we got rid of $T_k(1)$ when we introduced the constant C . What we really want is a bound like

$$-R'_{2m}(n) = \phi_{m,n} T_{2m+2}(n),$$

where $0 < \phi_{m,n} < 1$; this will allow us to conclude from (9.85) that

$$S_n = F(n) + C + T_1(n) + \sum_{k=1}^m T_{2k}(n) + \phi_{m,n} T_{2m+2}(n), \quad (9.86)$$

hence the remainder will truly be between zero and the first discarded term.

A slight modification of our previous argument will patch things up perfectly. Let us assume that

$$f^{(2m+2)}(x) \geq 0 \quad \text{and} \quad f^{(2m+4)}(x) \geq 0, \quad \text{as } x \rightarrow \infty. \quad (9.87)$$

The right-hand side of (9.85) is just like the negative of the right-hand side of Euler's summation formula (9.67) with $a = n$ and $b = \infty$, as far as remainder terms are concerned, and successive remainders are generated by induction on m . Therefore our previous argument can be applied.

Summation 2: Harmonic numbers harmonized.

Now that we've learned so much from a trivial (but safe) example, we can readily do a nontrivial one. Let us use Euler's summation formula to derive the approximation for H_n that we have been claiming for some time.

In this case, $f(x) = 1/x$. We already know about the integral and derivatives of f , because of Summation 1; also $f^{(m)}(x) = O(x^{-m-1})$ as $x \rightarrow \infty$. Therefore we can immediately plug into formula (9.85):

$$\sum_{1 \leq k < n} \frac{1}{k} = \ln n + C + B_1 n^{-1} - \sum_{k=1}^m \frac{B_{2k}}{2k n^{2k}} - R'_{2m}(n),$$

for some constant C . The sum on the left is H_{n-1} , not H_n ; but it's more convenient to work with H_{n-1} and to add $1/n$ later, than to mess around with $(n+1)$'s on the right-hand side. The $B_1 n^{-1}$ will then become $(B_1 + 1)n^{-1} = 1/(2n)$. Let us call the constant γ instead of C , since Euler's constant γ is, in fact, defined to be $\lim_{n \rightarrow \infty} (H_n - \ln n)$.

The remainder term can be estimated nicely by the theory we developed a minute ago, because $f^{(2m)}(x) = (2m)!/x^{2m+1} \geq 0$ for all $x > 0$. Therefore (9.86) tells us that

$$H_n = \ln n + \gamma + \frac{1}{2n} - \sum_{k=1}^m \frac{B_{2k}}{2k n^{2k}} + \theta_{m,n} \frac{B_{2m+2}}{(2m+2)n^{2m+2}}, \quad (9.88)$$

where $\theta_{m,n}$ is some fraction between 0 and 1. This is the general formula whose first few terms are listed in Table 452. For example, when $m = 2$ we get

$$H_n = \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} - \frac{\theta_{2,n}}{252n^6}. \quad (9.89)$$

This equation, incidentally, gives us a good approximation to γ even when $n = 2$:

$$\gamma = H_2 - \ln 2 - \frac{1}{4} + \frac{1}{48} - \frac{1}{1920} + \epsilon = 0.577165\dots + \epsilon,$$

where ϵ is between zero and $\frac{1}{16128}$. If we take $n = 10^4$ and $m = 250$, we get the value of γ correct to 1271 decimal places, beginning thus [205]:

$$\gamma = 0.57721\,56649\,01532\,86060\,65120\,90082\,40243\dots \quad (9.90)$$

But Euler's constant appears also in other formulas that allow it to be evaluated even more efficiently [345].

Summation 3: Stirling's approximation.

If $f(x) = \ln x$, we have $f'(x) = 1/x$, so we can evaluate the sum of logarithms using almost the same calculations as we did when summing reciprocals. Euler's summation formula yields

$$\begin{aligned} \sum_{1 \leq k < n} \ln k &= n \ln n - n + \sigma - \frac{\ln n}{2} \\ &\quad + \sum_{k=1}^m \frac{B_{2k}}{2k(2k-1)n^{2k-1}} + \varphi_{m,n} \frac{B_{2m+2}}{(2m+2)(2m+1)n^{2m+1}} \end{aligned}$$

where σ is a certain constant, "Stirling's constant," and $0 < \varphi_{m,n} < 1$. (In this case $f^{(2m)}(x)$ is negative, not positive; but we can still say that the remainder is governed by the first discarded term, because we could have started with $f(x) = -\ln x$ instead of $f(x) = \ln x$.) Adding $\ln n$ to both sides gives

$$\ln n! = n \ln n - n + \frac{\ln n}{2} + \sigma + \frac{1}{12n} - \frac{1}{360n^3} + \frac{\varphi_{2,n}}{1260n^5} \quad (9.91)$$

when $m = 2$. And we can get the approximation in Table 452 by taking 'exp' of both sides. (The value of e^σ turns out to be $\sqrt{2\pi}$, but we aren't quite ready to derive that formula. In fact, Stirling didn't discover the closed form for σ until several years after de Moivre [76] had proved that the constant exists.)

If m is fixed and $n \rightarrow \infty$, the general formula gives a better and better approximation to $\ln n!$ in the sense of absolute error, hence it gives a better and better approximation to $n!$ in the sense of relative error. But if n is fixed and m increases, the error bound $|B_{2m+2}|/(2m+2)(2m+1)n^{2m+1}$ decreases to a certain point and then begins to increase. Therefore the approximation reaches a point beyond which a sort of uncertainty principle limits the amount by which $n!$ can be approximated.

*Heisenberg may
have been here.*

In Chapter 5, equation (5.83), we generalized factorials to arbitrary real α by using a definition

$$\frac{1}{\alpha!} = \lim_{n \rightarrow \infty} \binom{n+\alpha}{n} n^{-\alpha}$$

suggested by Euler. Suppose α is a large number; then

$$\ln \alpha! = \lim_{n \rightarrow \infty} \left(\alpha \ln n + \ln n! - \sum_{k=1}^n \ln(\alpha + k) \right),$$

and Euler's summation formula can be used with $f(x) = \ln(x + \alpha)$ to estimate this sum:

$$\begin{aligned} \sum_{k=1}^n \ln(k + \alpha) &= F_m(\alpha, n) - F_m(\alpha, 0) + R_{2m}(\alpha, n), \\ F_m(\alpha, x) &= (x + \alpha) \ln(x + \alpha) - x + \frac{\ln(x + \alpha)}{2} \\ &\quad + \sum_{k=1}^m \frac{B_{2k}}{2k(2k-1)(x + \alpha)^{2k-1}}, \\ R_{2m}(\alpha, n) &= \int_0^n \frac{B_{2m}(\{x\})}{2m} \frac{dx}{(x + \alpha)^{2m}}. \end{aligned}$$

(Here we have used (9.67) with $a = 0$ and $b = n$, then added $\ln(n + \alpha) - \ln \alpha$ to both sides.) If we subtract this approximation for $\sum_{k=1}^n \ln(k + \alpha)$ from Stirling's approximation for $\ln n!$, then add $\alpha \ln n$ and take the limit as $n \rightarrow \infty$, we get

$$\begin{aligned} \ln \alpha! &= \alpha \ln \alpha - \alpha + \frac{\ln \alpha}{2} + \sigma \\ &\quad + \sum_{k=1}^m \frac{B_{2k}}{(2k)(2k-1)\alpha^{2k-1}} - \int_0^\infty \frac{B_{2m}(\{x\})}{2m} \frac{dx}{(x + \alpha)^{2m}}, \end{aligned}$$

because $\alpha \ln n + n \ln n - n + \frac{1}{2} \ln n - (n + \alpha) \ln(n + \alpha) + n - \frac{1}{2} \ln(n + \alpha) \rightarrow -\alpha$ and the other terms not shown here tend to zero. Thus Stirling's approximation behaves for generalized factorials (and for the Gamma function $\Gamma(\alpha + 1) = \alpha!$) exactly as for ordinary factorials.

Summation 4: A bell-shaped summand.

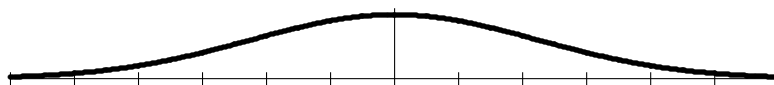
Let's turn now to a sum that has quite a different flavor:

$$\begin{aligned} \Theta_n &= \sum_k e^{-k^2/n} \\ &= \dots + e^{-9/n} + e^{-4/n} + e^{-1/n} + 1 + e^{-1/n} + e^{-4/n} + e^{-9/n} + \dots \end{aligned} \quad (9.92)$$

This is a doubly infinite sum, whose terms reach their maximum value $e^0 = 1$ when $k = 0$. We call it Θ_n because it is a power series involving the quantity $e^{-1/n}$ raised to the $p(k)$ th power, where $p(k)$ is a polynomial of degree 2; such power series are traditionally called “theta functions.” If $n = 10^{100}$, we have

$$e^{-k^2/n} = \begin{cases} e^{-.01} \approx 0.99005, & \text{when } k = 10^{49}; \\ e^{-1} \approx 0.36788, & \text{when } k = 10^{50}; \\ e^{-100} < 10^{-43}, & \text{when } k = 10^{51}. \end{cases}$$

So the summand stays very near 1 until k gets up to about \sqrt{n} , when it drops off and stays very near zero. We can guess that Θ_n will be proportional to \sqrt{n} . Here is a graph of $e^{-k^2/n}$ when $n = 10$:



Larger values of n just stretch the graph horizontally by a factor of \sqrt{n} .

We can estimate Θ_n by letting $f(x) = e^{-x^2/n}$ and taking $a = -\infty$, $b = +\infty$ in Euler's summation formula. (If infinities seem too scary, let $a = -A$ and $b = +B$, then take limits as $A, B \rightarrow \infty$.) The integral of $f(x)$ is

$$\int_{-\infty}^{+\infty} e^{-x^2/n} dx = \sqrt{n} \int_{-\infty}^{+\infty} e^{-u^2} du = \sqrt{n} C,$$

if we replace x by $u\sqrt{n}$. The value of $\int_{-\infty}^{+\infty} e^{-u^2} du$ is well known, but we'll call it C for now and come back to it after we have finished plugging into Euler's summation formula.

The next thing we need to know is the sequence of derivatives $f'(x)$, $f''(x)$, \dots , and for this purpose it's convenient to set

$$f(x) = g(x/\sqrt{n}), \quad g(x) = e^{-x^2}.$$

Then the chain rule of calculus says that

$$\frac{df(x)}{dx} = \frac{dg(y)}{dy} \frac{dy}{dx}, \quad y = \frac{x}{\sqrt{n}};$$

and this is the same as saying that

$$f'(x) = \frac{1}{\sqrt{n}} g'(x/\sqrt{n}).$$

By induction we have

$$f^{(k)}(x) = n^{-k/2} g^{(k)}(x/\sqrt{n}).$$

For example, we have $g'(x) = -2xe^{-x^2}$ and $g''(x) = (4x^2 - 2)e^{-x^2}$; hence

$$f'(x) = \frac{1}{\sqrt{n}} \left(-2 \frac{x}{\sqrt{n}} \right) e^{-x^2/n}, \quad f''(x) = \frac{1}{n} \left(4 \left(\frac{x}{\sqrt{n}} \right)^2 - 2 \right) e^{-x^2/n}.$$

It's easier to see what's going on if we work with the simpler function $g(x)$.

We don't have to evaluate the derivatives of $g(x)$ exactly, because we're only going to be concerned about the limiting values when $x = \pm\infty$. And for this purpose it suffices to notice that every derivative of $g(x)$ is e^{-x^2} times a polynomial in x :

$$g^{(k)}(x) = P_k(x)e^{-x^2}, \quad \text{where } P_k \text{ is a polynomial of degree } k.$$

This follows by induction.

The negative exponential e^{-x^2} goes to zero much faster than $P_k(x)$ goes to infinity, when $x \rightarrow \pm\infty$, so we have

$$f^{(k)}(+\infty) = f^{(k)}(-\infty) = 0$$

for all $k \geq 0$. Therefore all of the terms

$$\sum_{k=1}^m \frac{B_k}{k!} f^{(k-1)}(x) \Big|_{-\infty}^{+\infty}$$

vanish, and we are left with the term from $\int f(x) dx$ and the remainder:

$$\begin{aligned} \Theta_n &= C\sqrt{n} + (-1)^{m+1} \int_{-\infty}^{+\infty} \frac{B_m(\{x\})}{m!} f^{(m)}(x) dx \\ &= C\sqrt{n} + \frac{(-1)^{m+1}}{n^{m/2}} \int_{-\infty}^{+\infty} \frac{B_m(\{x\})}{m!} g^{(m)}\left(\frac{x}{\sqrt{n}}\right) dx \\ &= C\sqrt{n} + \frac{(-1)^{m+1}}{n^{(m-1)/2}} \int_{-\infty}^{+\infty} \frac{B_m(\{u\sqrt{n}\})}{m!} P_m(u) e^{-u^2} du \quad (x = u\sqrt{n}) \\ &= C\sqrt{n} + O(n^{(1-m)/2}). \end{aligned}$$

The O estimate here follows since $|B_m(\{u\sqrt{n}\})|$ is bounded and the integral $\int_{-\infty}^{+\infty} |P(u)| e^{-u^2} du$ exists whenever P is a polynomial. (The constant implied by this O depends on m .)

We have proved that $\Theta_n = C\sqrt{n} + O(n^{-M})$, for arbitrarily large M ; the difference between Θ_n and $C\sqrt{n}$ is "exponentially small." Let us therefore determine the constant C that plays such a big role in the value of Θ_n .

One way to determine C is to look the integral up in a table; but we prefer to know how the value can be derived, so that we can do integrals even

when they haven't been tabulated. Elementary calculus suffices to evaluate C if we are clever enough to look at the double integral

$$C^2 = \int_{-\infty}^{+\infty} e^{-x^2} dx \int_{-\infty}^{+\infty} e^{-y^2} dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dx dy.$$

Converting to polar coordinates gives

$$\begin{aligned} C^2 &= \int_0^{2\pi} \int_0^\infty e^{-r^2} r dr d\theta \\ &= \frac{1}{2} \int_0^{2\pi} d\theta \int_0^\infty e^{-u} du \quad (u = r^2) \\ &= \frac{1}{2} \int_0^{2\pi} d\theta = \pi. \end{aligned}$$

So $C = \sqrt{\pi}$. The fact that $x^2 + y^2 = r^2$ is the equation of a circle whose circumference is $2\pi r$ somehow explains why π gets into the act.

Another way to evaluate C is to replace x by \sqrt{t} and dx by $\frac{1}{2}t^{-1/2} dt$:

$$C = \int_{-\infty}^{+\infty} e^{-x^2} dx = 2 \int_0^\infty e^{-x^2} dx = \int_0^\infty t^{-1/2} e^{-t} dt.$$

This integral equals $\Gamma(\frac{1}{2})$, since $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ according to (5.84). Therefore we have demonstrated that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Our final formula, then, is

$$\Theta_n = \sum_k e^{-k^2/n} = \sqrt{\pi n} + O(n^{-M}), \quad \text{for all fixed } M. \quad (9.93)$$

The constant in the O depends on M ; that's why we say that M is "fixed."

When $n = 2$, for example, the infinite sum Θ_2 is approximately equal to 2.506628288; this is already very close to $\sqrt{2\pi} \approx 2.506628275$, even though n is quite small. The value of Θ_{100} agrees with $10\sqrt{\pi}$ to 427 decimal places! Exercise 59 uses advanced methods to derive a rapidly convergent series for Θ_n ; it turns out that

$$\Theta_n/\sqrt{\pi n} = 1 + 2e^{-n\pi^2} + O(e^{-4n\pi^2}). \quad (9.94)$$

Summation 5: The clincher.

Now we will do one last sum, which will turn out to tell us the value of Stirling's constant σ . This last sum also illustrates many of the other techniques of this last chapter (and of this whole book), so it will be a fitting way for us to conclude our explorations of Concrete Mathematics.

The final task seems almost absurdly easy: We will try to find the asymptotic value of

$$A_n = \sum_k \binom{2n}{k}$$

by using Euler's summation formula.

This is another case where we already know the answer (right?); but it's always interesting to try new methods on old problems, so that we can compare facts and maybe discover something new.

So we THINK BIG and realize that the main contribution to A_n comes from the middle terms, near $k = n$. It's almost always a good idea to choose notation so that the biggest contribution to a sum occurs near $k = 0$, because we can then use the tail-exchange trick to get rid of terms that have large $|k|$. Therefore we replace k by $n + k$:

$$A_n = \sum_k \binom{2n}{n+k} = \sum_k \frac{(2n)!}{(n+k)!(n-k)!}.$$

Things are looking reasonably good, since we know how to approximate $(n \pm k)!$ when n is large and k is small.

Now we want to carry out the three-step procedure associated with the tail-exchange trick. Namely, we want to write

$$\frac{(2n)!}{(n+k)!(n-k)!} = a_k(n) = b_k(n) + O(c_k(n)), \quad \text{for } k \in D_n,$$

so that we can obtain the estimate

$$A_n = \sum_k b_k(n) + O\left(\sum_{k \notin D_n} a_k(n)\right) + O\left(\sum_{k \notin D_n} b_k(n)\right) + \sum_{k \in D_n} O(c_k(n)).$$

Let us therefore try to estimate $\binom{2n}{n+k}$ in the region where $|k|$ is small. We could use Stirling's approximation as it appears in Table 452, but it's easier to work with the logarithmic equivalent in (9.91):

$$\begin{aligned} \ln a_k(n) &= \ln(2n)! - \ln(n+k)! - \ln(n-k)! \\ &= 2n \ln 2n - 2n + \frac{1}{2} \ln 2n + \sigma + O(n^{-1}) \\ &\quad - (n+k) \ln(n+k) + n+k - \frac{1}{2} \ln(n+k) - \sigma + O((n+k)^{-1}) \\ &\quad - (n-k) \ln(n-k) + n-k - \frac{1}{2} \ln(n-k) - \sigma + O((n-k)^{-1}). \end{aligned} \tag{9.95}$$

We want to convert this to a nice, simple O estimate.

The tail-exchange method allows us to work with estimates that are valid only when k is in the "dominant" set D_n . But how should we define D_n ?

Actually I'm not into dominance.

We have to make D_n small enough that we can make a good estimate; for example, we had better not let k get near n , or the term $O((n-k)^{-1})$ in (9.95) will blow up. Yet D_n must be large enough that the tail terms (the terms with $k \notin D_n$) are negligibly small compared with the overall sum. Trial and error is usually necessary to find an appropriate set D_n ; in this problem the calculations we are about to make will show that it's wise to define things as follows:

$$k \in D_n \iff |k| \leq n^{1/2+\epsilon}. \quad (9.96)$$

Here ϵ is a small positive constant that we can choose later, after we get to know the territory. (Our O estimates will depend on the value of ϵ .) Equation (9.95) now reduces to

$$\begin{aligned} \ln a_k(n) = & (2n + \tfrac{1}{2}) \ln 2 - \sigma - \tfrac{1}{2} \ln n + O(n^{-1}) \\ & - (n+k+\tfrac{1}{2}) \ln(1+k/n) - (n-k+\tfrac{1}{2}) \ln(1-k/n). \end{aligned} \quad (9.97)$$

(We have pulled out the large parts of the logarithms, writing

$$\ln(n \pm k) = \ln n + \ln(1 \pm k/n),$$

and this has made a lot of $\ln n$ terms cancel out.)

Now we need to expand the terms $\ln(1 \pm k/n)$ asymptotically, until we have an error term that approaches zero as $n \rightarrow \infty$. We are multiplying $\ln(1 \pm k/n)$ by $(n \pm k + \frac{1}{2})$, so we should expand the logarithm until we reach $o(n^{-1})$, using the assumption that $|k| \leq n^{1/2+\epsilon}$:

$$\ln\left(1 \pm \frac{k}{n}\right) = \pm \frac{k}{n} - \frac{k^2}{2n^2} + O(n^{-3/2+3\epsilon}).$$

Multiplication by $n \pm k + \frac{1}{2}$ yields

$$\pm k - \frac{k^2}{2n} + \frac{k^2}{n} + O(n^{-1/2+3\epsilon}),$$

plus other terms that are absorbed in the $O(n^{-1/2+3\epsilon})$. So (9.97) becomes

$$\ln a_k(n) = (2n + \tfrac{1}{2}) \ln 2 - \sigma - \tfrac{1}{2} \ln n - k^2/n + O(n^{-1/2+3\epsilon}).$$

Taking exponentials, we have

$$a_k(n) = \frac{2^{2n+1/2}}{e^\sigma \sqrt{n}} e^{-k^2/n} (1 + O(n^{-1/2+3\epsilon})). \quad (9.98)$$

This is our approximation, with

$$b_k(n) = \frac{2^{2n+1/2}}{e^\sigma \sqrt{n}} e^{-k^2/n}, \quad c_k(n) = 2^{2n} n^{-1+3\epsilon} e^{-k^2/n}.$$

Notice that k enters $b_k(n)$ and $c_k(n)$ in a very simple way. We're in luck, because we will be summing over k .

The tail-exchange trick tells us that $\sum_k a_k(n)$ will be approximately $\sum_k b_k(n)$ if we have done a good job of estimation. Let us therefore evaluate

$$\begin{aligned} \sum_k b_k(n) &= \frac{2^{2n+1/2}}{e^\sigma \sqrt{n}} \sum_k e^{-k^2/n} \\ &= \frac{2^{2n+1/2}}{e^\sigma \sqrt{n}} \Theta_n = \frac{2^{2n} \sqrt{2\pi}}{e^\sigma} (1 + O(n^{-M})). \end{aligned}$$

(Another stroke of luck: We get to use the sum Θ_n from the previous example.) This is encouraging, because we know that the original sum is actually

$$A_n = \sum_k \binom{2n}{k} = (1+1)^{2n} = 2^{2n}.$$

Therefore it looks as if we will have $e^\sigma = \sqrt{2\pi}$, as advertised.

But there's a catch: We still need to prove that our estimates are good enough. So let's look first at the error contributed by $c_k(n)$:

$$\sum_{|k| \leq n^{1/2+\epsilon}} 2^{2n} n^{-1+3\epsilon} e^{-k^2/n} \leq 2^{2n} n^{-1+3\epsilon} \Theta_n = O(2^{2n} n^{-\frac{1}{2}+3\epsilon}).$$

Good; this is asymptotically smaller than the previous sum, if $3\epsilon < \frac{1}{2}$.

Next we must check the tails. We have

$$\begin{aligned} \sum_{k > n^{1/2+\epsilon}} e^{-k^2/n} &< \exp(-\lfloor n^{1/2+\epsilon} \rfloor^2/n) (1 + e^{-1/n} + e^{-2/n} + \dots) \\ &= O(e^{-n^{2\epsilon}}) \cdot O(n), \end{aligned}$$

which is $O(n^{-M})$ for all M ; so $\sum_{k \notin D_n} b_k(n)$ is asymptotically negligible. (We chose the cutoff at $n^{1/2+\epsilon}$ just so that $e^{-k^2/n}$ would be exponentially small outside of D_n . Other choices like $n^{1/2} \log n$ would have been good enough too, and the resulting estimates would have been slightly sharper, but the formulas would have come out more complicated. We need not make the strongest possible estimates, since our main goal is to establish the value of the constant σ .) Similarly, the other tail

$$\sum_{k > n^{1/2+\epsilon}} \binom{2n}{n+k}$$

What an amazing coincidence.

I'm tired of getting to the end of long, hard books and not even getting a word of good wishes from the author. It would be nice to read a "thanks for reading this, hope it comes in handy," instead of just running into a hard, cold, cardboard cover at the end of a long, dry proof. You know?

is bounded by $2n$ times its largest term, which occurs at the cutoff point $k \approx n^{1/2+\epsilon}$. This term is known to be approximately $b_k(n)$, which is exponentially small compared with A_n ; and an exponentially small multiplier wipes out the factor of $2n$.

Thus we have successfully applied the tail-exchange trick to prove the estimate

$$2^{2n} = \sum_k \binom{2n}{k} = \frac{\sqrt{2\pi}}{e^\sigma} 2^{2n} + O(2^{2n} n^{-\frac{1}{2}+3\epsilon}), \quad \text{if } 0 < \epsilon < \frac{1}{6}. \quad (9.99)$$

*Thanks for reading
this, hope it comes
in handy.*

—The authors

We may choose $\epsilon = \frac{1}{8}$ and conclude that

$$\sigma = \frac{1}{2} \ln 2\pi.$$

QED.

Exercises

Warmups

- 1 Prove or disprove: If $f_1(n) \prec g_1(n)$ and $f_2(n) \prec g_2(n)$, then we have $f_1(n) + f_2(n) \prec g_1(n) + g_2(n)$.
- 2 Which function grows faster:
 - a $n^{(\ln n)}$ or $(\ln n)^n$?
 - b $n^{(\ln \ln \ln n)}$ or $(\ln n)!$?
 - c $(n!)!$ or $((n-1)!)! (n-1)!^{n!}$?
 - d $F_{\lfloor H_n \rfloor}^2$ or H_{F_n} ?
- 3 What's wrong with the following argument? "Since $n = O(n)$ and $2n = O(n)$ and so on, we have $\sum_{k=1}^n kn = \sum_{k=1}^n O(n) = O(n^2)$."
- 4 Give an example of a valid equation that has O -notation on the left but not on the right. (Do not use the trick of multiplying by zero; that's too easy.) *Hint:* Consider taking limits.
- 5 Prove or disprove: $O(f(n) + g(n)) = f(n) + O(g(n))$, if $f(n)$ and $g(n)$ are positive for all n . (Compare with (9.27).)
- 6 Multiply $(\ln n + \gamma + O(1/n))$ by $(n + O(\sqrt{n}))$, and express your answer in O -notation.
- 7 Estimate $\sum_{k \geq 0} e^{-k/n}$ with absolute error $O(n^{-1})$.

Basics

- 8 Give an example of functions $f(n)$ and $g(n)$ such that none of the three relations $f(n) \prec g(n)$, $f(n) \succ g(n)$, $f(n) \asymp g(n)$ is valid, although $f(n)$ and $g(n)$ both increase monotonically to ∞ .

- 9 Prove (9.22) rigorously by showing that the left side is a subset of the right side, according to the set-of-functions definition of O .
- 10 Prove or disprove: $\cos O(x) = 1 + O(x^2)$ for all real x .
- 11 Prove or disprove: $O(x + y)^2 = O(x^2) + O(y^2)$.

12 Prove that

$$1 + \frac{2}{n} + O(n^{-2}) = \left(1 + \frac{2}{n}\right)(1 + O(n^{-2})),$$

as $n \rightarrow \infty$.

- 13 Evaluate $(n + 2 + O(n^{-1}))^n$ with relative error $O(n^{-1})$.
- 14 Prove that $(n + \alpha)^{n+\beta} = n^{n+\beta} e^\alpha (1 + \alpha(\beta - \frac{1}{2}\alpha)n^{-1} + O(n^{-2}))$.
- 15 Give an asymptotic formula for the “middle” trinomial coefficient $\binom{3n}{n, n, n}$, correct to relative error $O(n^{-3})$.
- 16 Show that if $B(1 - x) = -B(x) \geq 0$ for $0 < x < \frac{1}{2}$, we have

$$\int_a^b B(\{x\}) f(x) dx \geq 0$$

if we assume also that $f'(x) \geq 0$ for $a \leq x \leq b$.

- 17 Use generating functions to show that $B_m(\frac{1}{2}) = (2^{1-m} - 1)B_m$, for all $m \geq 0$.
- 18 Find $\sum_k \binom{2n}{k}^\alpha$ with relative error $O(n^{-1/4})$, when $\alpha > 0$.

Homework exercises

- 19 Use a computer to compare the left and right sides of the approximations in Table 452, when $n = 10$, $z = \alpha = 0.1$, and $O(f(n)) = O(f(z)) = 0$.
- 20 Prove or disprove the following estimates, as $n \rightarrow \infty$:
- a $O\left(\left(\frac{n^2}{\log \log n}\right)^{1/2}\right) = O([\sqrt{n}]^2)$.
- b $e^{(1+O(1/n))^2} = e + O(1/n)$.
- c $n! = O(((1 - 1/n)^n n)^n)$.
- 21 Equation (9.48) gives the n th prime with relative error $O(\log n)^{-2}$. Improve the relative error to $O(\log n)^{-3}$ by starting with another term of (9.31) in (9.46).
- 22 Improve (9.54) to $O(n^{-3})$.
- 23 Push the approximation (9.62) further, getting absolute error $O(n^{-3})$.
Hint: Let $g_n = c/(n+1)(n+2) + h_n$; what recurrence does h_n satisfy?

- 24 Suppose $a_n = O(f(n))$ and $b_n = O(f(n))$. Prove or disprove that the convolution $\sum_{k=0}^n a_k b_{n-k}$ is also $O(f(n))$, in the following cases:
- $f(n) = n^{-\alpha}$, $\alpha > 1$.
 - $f(n) = \alpha^{-n}$, $\alpha > 1$.
- 25 Prove (9.1) and (9.2), with which we opened this chapter.
- 26 Equation (9.91) shows how to evaluate $\ln 10!$ with an absolute error $< \frac{1}{126000000}$. Therefore if we take exponentials, we get $10!$ with a relative error that is less than $e^{1/126000000} - 1 < 10^{-8}$. (In fact, the approximation gives 3628799.9714.) If we now round to the nearest integer, knowing that $10!$ is an integer, we get an exact result.
- Is it always possible to calculate $n!$ in a similar way, if enough terms of Stirling's approximation are computed? Estimate the value of m that gives the best approximation to $\ln n!$, when n is a fixed (large) integer. Compare the absolute error in this approximation with $n!$ itself.
- 27 Use Euler's summation formula to find the asymptotic value of $H_n^{(-\alpha)} = \sum_{k=1}^n k^{-\alpha}$, where α is any fixed real number. (Your answer may involve a constant that you do not know in closed form.)
- 28 Exercise 5.13 defines the hyperfactorial function $Q_n = 1^1 2^2 \dots n^n$. Find the asymptotic value of Q_n with relative error $O(n^{-1})$. (Your answer may involve a constant that you do not know in closed form.)
- 29 Estimate the function $1^{1/1} 2^{1/2} \dots n^{1/n}$ as in the previous exercise.
- 30 Find the asymptotic value of $\sum_{k \geq 0} k^l e^{-k^2/n}$ with absolute error $O(n^{-3})$, when l is a fixed nonnegative integer.
- 31 Evaluate $\sum_{k \geq 0} 1/(c^k + c^m)$ with absolute error $O(c^{-3m})$, when $c > 1$ and m is a positive integer.

Exam problems

- 32 Evaluate $e^{H_n + H_n^{(2)}}$ with absolute error $O(n^{-1})$.
- 33 Evaluate $\sum_{k \geq 0} \binom{n}{k} / n^{\bar{k}}$ with absolute error $O(n^{-3})$.
- 34 Determine values A through F such that $(1 + 1/n)^{nH_n}$ is

$$An + B(\ln n)^2 + C \ln n + D + \frac{E(\ln n)^2}{n} + \frac{F \ln n}{n} + O(n^{-1}).$$

- 35 Evaluate $\sum_{k=1}^n 1/kH_k$ with absolute error $O(1)$.
- 36 Evaluate $S_n = \sum_{k=1}^n 1/(n^2 + k^2)$ with absolute error $O(n^{-5})$.
- 37 Evaluate $\sum_{k=1}^n (n \bmod k)$ with absolute error $O(n \log n)$.
- 38 Evaluate $\sum_{k \geq 0} k^k \binom{n}{k}$ with relative error $O(n^{-1})$.

- 39 Evaluate $\sum_{0 \leq k < n} \ln(n-k)(\ln n)^k/k!$ with absolute error $O(n^{-1})$. *Hint:* Show that the terms for $k \geq 10 \ln n$ are negligible.
- 40 Let m be a (fixed) positive integer. Evaluate $\sum_{k=1}^n (-1)^k H_k^m$ with absolute error $O(1)$.
- 41 Evaluate the “Fibonacci factorial” $\prod_{k=1}^n F_k$ with relative error $O(n^{-1})$ or better. Your answer may involve a constant whose value you do not know in closed form.
- 42 Let α be a constant in the range $0 < \alpha < \frac{1}{2}$. We’ve seen in previous chapters that there is no general closed form for the sum $\sum_{k \leq \alpha n} \binom{n}{k}$. Show that there is, however, an asymptotic formula

$$\sum_{k \leq \alpha n} \binom{n}{k} = 2^{nH(\alpha) - \frac{1}{2} \lg n + O(1)},$$

where $H(\alpha) = \alpha \lg \frac{1}{\alpha} + (1-\alpha) \lg \frac{1}{1-\alpha}$. *Hint:* Show that $\binom{n}{k-1} < \frac{\alpha}{1-\alpha} \binom{n}{k}$ for $0 < k \leq \alpha n$.

- 43 Show that C_n , the number of ways to change n cents (as considered in Chapter 7) is asymptotically $cn^4 + O(n^3)$ for some constant c . What is that constant?
- 44 Prove that

$$x^{1/2} = x^{1/2} \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} - x^{-1/2} \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix} + x^{-3/2} \begin{bmatrix} 1/2 \\ -3/2 \end{bmatrix} + O(x^{-5/2})$$

as $x \rightarrow \infty$. (Recall the definition $x^{1/2} = x!/(x - \frac{1}{2})!$ in (5.88), and the definition of generalized Stirling numbers in Table 272.)

- 45 Let α be an irrational number between 0 and 1. Chapter 3 discusses the quantity $D(\alpha, n)$, which measures the maximum discrepancy by which the fractional parts $\{k\alpha\}$ for $0 \leq k < n$ deviate from a uniform distribution. The recurrence

$$D(\alpha, n) \leq D(\{\alpha^{-1}\}, \lfloor \alpha n \rfloor) + \alpha^{-1} + 2$$

was proved in (3.31); we also have the obvious bounds

$$0 \leq D(\alpha, n) \leq n.$$

Prove that $\lim_{n \rightarrow \infty} D(\alpha, n)/n = 0$. *Hint:* Chapter 6 discusses continued fractions.

- 46 Show that the Bell number $\omega_n = e^{-1} \sum_{k \geq 0} k^n/k!$ of exercise 7.15 is asymptotically equal to

$$m(n)^n e^{m(n)-n-1/2} / \sqrt{\ln n},$$

where $m(n) \ln m(n) = n - \frac{1}{2}$, and estimate the relative error in this approximation.

- 47 Let m be an integer ≥ 2 . Analyze the two sums

$$\sum_{k=1}^n \lfloor \log_m k \rfloor \quad \text{and} \quad \sum_{k=1}^n \lceil \log_m k \rceil;$$

which is asymptotically closer to $\log_m n!$?

- 48 Consider a table of the harmonic numbers H_k for $1 \leq k \leq n$ in decimal notation. The k th entry \hat{H}_k has been correctly rounded to d_k significant digits, where d_k is just large enough to distinguish this value from the values of H_{k-1} and H_{k+1} . For example, here is an extract from the table, showing five entries where H_k passes 10:

k	H_k	\hat{H}_k	d_k
12364	9.99980041–	9.9998	5
12365	9.99988128+	9.9999	5
12366	9.99996215–	9.99996	6
12367	10.00004301–	10.0000	6
12368	10.00012386+	10.0001	6

Estimate the total number of digits in the table, $\sum_{k=1}^n d_k$, with an absolute error of $O(n)$.

- 49 In Chapter 6 we considered the tale of a worm that reaches the end of a stretching band after n seconds, where $H_{n-1} < 100 \leq H_n$. Prove that if n is a positive integer such that $H_{n-1} \leq \alpha \leq H_n$, then

$$\lfloor e^{\alpha-\gamma} \rfloor \leq n \leq \lceil e^{\alpha-\gamma} \rceil.$$

- 50 Venture capitalists in Silicon Valley are being offered a deal giving them a chance for an exponential payoff on their investments: For an n million dollar investment, where $n \geq 2$, the GKP consortium promises to pay up to N million dollars after one year, where $N = 10^n$. Of course there's some risk; the actual deal is that GKP pays k million dollars with probability $1/(k^2 H_N^{(2)})$, for each integer k in the range $1 \leq k \leq N$. (All payments are in megabucks, that is, in exact multiples of \$1,000,000; the payoff is determined by a truly random process.) Notice that an investor always gets at least a million dollars back.

- a What is the asymptotic expected return after one year, if n million dollars are invested? (In other words, what is the mean value of the payment?) Your answer should be correct within an absolute error of $O(10^{-n})$ dollars.
- b What is the asymptotic probability that you make a profit, if you invest n million? (In other words, what is the chance that you get back more than you put in?) Your answer here should be correct within an absolute error of $O(n^{-3})$.

*I once earned
 $O(10^{-n})$ dollars.*

Bonus problems

- 51 Prove or disprove: $\int_n^\infty O(x^{-2}) dx = O(n^{-1})$ as $n \rightarrow \infty$.
- 52 Show that there exists a power series $A(z) = \sum_{k \geq 0} a_k z^k$, convergent for all complex z , such that

$$A(n) \asymp n^{n^{n^{\dots^n}}}.$$

- 53 Prove that if $f(x)$ is a function whose derivatives satisfy

$$f'(x) \leq 0, \quad -f''(x) \leq 0, \quad f'''(x) \leq 0, \quad \dots, \quad (-1)^m f^{(m+1)}(x) \leq 0$$

for all $x \geq 0$, then we have

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \dots + \frac{f^{(m-1)}(0)}{(m-1)!}x^{m-1} + O(x^m), \quad \text{for } x \geq 0.$$

In particular, the case $f(x) = -\ln(1+x)$ proves (9.64) for all $k, n > 0$.

- 54 Let $f(x)$ be a positive, differentiable function such that $xf'(x) \prec f(x)$ as $x \rightarrow \infty$. Prove that

$$\sum_{k \geq n} \frac{f(k)}{k^{1+\alpha}} = O\left(\frac{f(n)}{n^\alpha}\right), \quad \text{if } \alpha > 0.$$

Hint: Consider the quantity $f(k - \frac{1}{2})/(k - \frac{1}{2})^\alpha - f(k + \frac{1}{2})/(k + \frac{1}{2})^\alpha$.

- 55 Improve (9.99) to relative error $O(n^{-3/2+5\epsilon})$.
- 56 The quantity $Q(n) = 1 + \frac{n-1}{n} + \frac{n-1}{n} \frac{n-2}{n} + \dots = \sum_{k \geq 1} n^k/n^k$ occurs in the analysis of many algorithms. Find its asymptotic value, with absolute error $o(1)$.
- 57 An asymptotic formula for Golomb's sum $\sum_{k \geq 1} 1/k \lfloor 1 + \log_n k \rfloor^2$ is derived in (9.54). Find an asymptotic formula for the analogous sum without floor brackets, $\sum_{k \geq 1} 1/k(1 + \log_n k)^2$. *Hint:* Consider the integral $\int_0^\infty u e^{-u} k^{-tu} du = 1/(1 + t \ln k)^2$.

58 Prove that

$$B_m(\{x\}) = -2 \frac{m!}{(2\pi)^m} \sum_{k \geq 1} \frac{\cos(2\pi kx - \frac{1}{2}\pi m)}{k^m}, \quad \text{for } m \geq 2,$$

by using residue calculus, integrating

$$\frac{1}{2\pi i} \oint \frac{2\pi i e^{2\pi i z \theta}}{e^{2\pi i z} - 1} \frac{dz}{z^m}$$

on the square contour $z = x + iy$, where $\max(|x|, |y|) = M + \frac{1}{2}$, and letting the integer M tend to ∞ .

59 Let $\Theta_n(t) = \sum_k e^{-(k+t)^2/n}$, a periodic function of t . Show that the expansion of $\Theta_n(t)$ as a Fourier series is

$$\begin{aligned} \Theta_n(t) = \sqrt{\pi n} & \left(1 + 2e^{-\pi^2 n} (\cos 2\pi t) + 2e^{-4\pi^2 n} (\cos 4\pi t) \right. \\ & \left. + 2e^{-9\pi^2 n} (\cos 6\pi t) + \cdots \right). \end{aligned}$$

(This formula gives a rapidly convergent series for the sum $\Theta_n = \Theta_n(0)$ in equation (9.93).)

60 Explain why the coefficients in the asymptotic expansion

$$\binom{2n}{n} = \frac{4^n}{\sqrt{\pi n}} \left(1 - \frac{1}{8n} + \frac{1}{128n^2} + \frac{5}{1024n^3} - \frac{21}{32768n^4} + O(n^{-5}) \right)$$

all have denominators that are powers of 2.

61 Exercise 45 proves that the discrepancy $D(\alpha, n)$ is $o(n)$ for all irrational numbers α . Exhibit an irrational α such that $D(\alpha, n)$ is *not* $O(n^{1-\epsilon})$ for any $\epsilon > 0$.

62 Given n , let $\{m(n)\} = \max_k \{n_k\}$ be the largest entry in row n of Stirling's subset triangle. Show that for all sufficiently large n , we have $m(n) = \lfloor \overline{m}(n) \rfloor$ or $m(n) = \lceil \overline{m}(n) \rceil$, where

$$\overline{m}(n)(\overline{m}(n) + 2) \ln(\overline{m}(n) + 2) = n(\overline{m}(n) + 1).$$

Hint: This is difficult.

63 Prove that S. W. Golomb's self-describing sequence of exercise 2.36 satisfies $f(n) = \phi^{2-\phi} n^{\phi-1} + O(n^{\phi-1}/\log n)$.

64 Find a proof of the identity

$$\sum_{n \geq 1} \frac{\cos 2n\pi x}{n^2} = \pi^2 \left(x^2 - x + \frac{1}{6} \right) \quad \text{for } 0 \leq x \leq 1,$$

that uses only "Eulerian" (eighteenth-century) mathematics.

- 65 What are the coefficients of the asymptotic series

$$1 + \frac{1}{n-1} + \frac{1}{(n-1)(n-2)} + \cdots + \frac{1}{(n-1)!} = a_0 + \frac{a_1}{n} + \frac{a_2}{n^2} + \cdots ?$$

Research problems

- 66 Find a “combinatorial” proof of Stirling’s approximation. (Note that n^n is the number of mappings of $\{1, 2, \dots, n\}$ into itself, and $n!$ is the number of mappings of $\{1, 2, \dots, n\}$ onto itself.)
- 67 Consider an $n \times n$ array of dots, $n \geq 3$, in which each dot has four neighbors. (At the edges we “wrap around” modulo n .) Let χ_n be the number of ways to assign the colors red, white, and blue to these dots in such a way that no neighboring dots have the same color. (Thus $\chi_3 = 12$.) Prove that

$$\chi_n \sim \left(\frac{4}{3}\right)^{3n^{2/2}} e^{-\pi/6}.$$

- 68 Let Q_n be the least integer m such that $H_m > n$. Find the smallest integer n such that $Q_n \neq \lfloor e^{n-\gamma} + \frac{1}{2} \rfloor$, or prove that no such n exist.

*Th-th-th-that's all,
folks!*