Chapter 1

Stirling's Formula

> The voyage of discovery lies not in seeking new
> horizons, but in seeking with new eyes.
> – Marcel Proust

Surely the most beautiful asymptotic formula in all of mathematics is Stirling's formula:

$$(1.1) \qquad n! \sim n^n e^{-n} \sqrt{2\pi n}.$$

How do the two most important fundamental constants of mathematics, e and π, find their way into an asymptotic formula for the product of integers? We give two very different arguments (one will not show the full formula) that, between them, illustrate a good number of basic asymptotic methods. The formal language of Asymptopia, such as $o(n)$ and $O(n)$, is deferred to Chapter 2. Two further arguments for Stirling's formula are given in §3.2.3.

1.1. Asymptotic Estimation of an Integral

Consider the integral

$$(1.2) \qquad I_n = \int_0^\infty x^n e^{-x} dx.$$

A standard result[1] of freshman calculus, done by integration by parts, is that

$$(1.3) \qquad\qquad I_n = n!$$

Our problem now is to estimate the integral of (1.2).

- Asymptotically, integrals are often dominated by the largest value of the function being integrated.

Let us set

$$(1.4) \qquad y = y_n(x) = x^n e^{-x} \text{ and } z = z_n(x) = \ln y = n \ln x - x.$$

The graph of $y(x)$ when $n = 2$ is unclear, but with $n = 10$ it is looking somewhat like the bell shaped curve. What is going on?
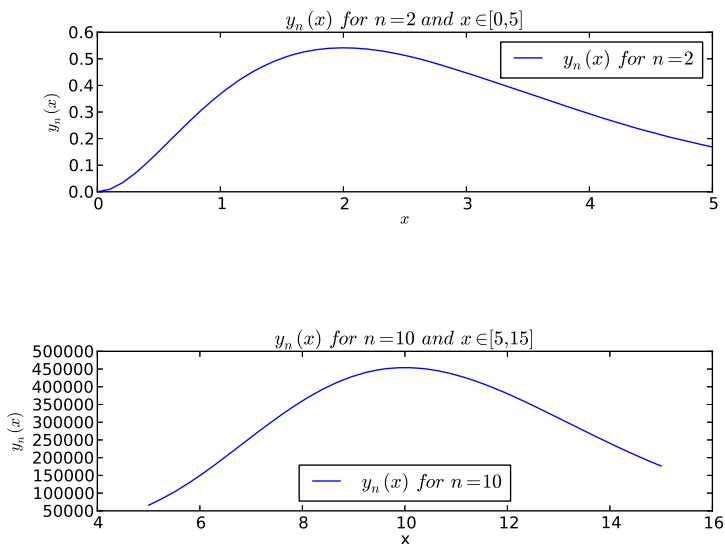

Figure 1. The first plot shows the function $y_n(x) = x^n e^{-x}$ for $n = 2$ and $x \in [0, 5]$, while the second plot shows the same function for $n = 10$ and $x \in [5, 15]$.

---

[1] We shall assume first-year calculus results throughout this work.

Setting $z' = nx^{-1} - 1 = 0$, we find that $z(x)$ (and hence $y(x)$) has a unique maximum at $x = n$ and that $z(x)$ (and hence $y(x)$) is increasing in $[0, n]$ and decreasing in $[n, \infty)$.

Let us compare $y(n) = n^n e^{-n}$ with values of $y(x)$ when x is "near" n. For example, take $x = 1.1n$.

$$(1.5) \qquad y(1.1n) = (1.1n)^n e^{-1.1n} = y(n)(1.1e^{-0.1})^n.$$

But $1.1e^{-0.1} = 0.9953\ldots$. While this number is close to 1, it is a constant less than 1, and so $y(1.1n)$ is exponentially smaller than $y(n)$. Values near $1.1n$ will make a negligible contribution to the integral. Let us move closer and try $x = n + 1$. Now

$$(1.6) \qquad y(n + 1) = (n + 1)^n e^{-n-1} = y(n) \left(1 + \frac{1}{n}\right)^n e^{-1}.$$

As $(1 + \frac{1}{n})^n \sim e$, $y(n + 1) \sim y(n)$, and so values near $x = n + 1$ do contribute substantially to the integral.

Moving from $x = n$ in the positive direction (the negative is similar), the function $y = y(x)$ decreases. If we move by 1 (to $x = n + 1$), we do not yet "see" the decrease, while if we move by $0.1n$ (to $x = 1.1n$), the decrease is so strong that the function has effectively disappeared. (Yes, $y(1.1n)$ is large in an absolute sense, but it is small relative to $y(n)$.) How do we move out from $x = n$ so that we can effectively see the decrease in $y = y(x)$? This is a question of *scaling*.

- Scaling is the art of asymptotic integration.

Let us look more carefully at $z(x)$ near $x = n$. Note that an additive change in $z(x)$ means a multiplicative change in $y(x) = e^{z(x)}$. We have $z'(x) = nx^{-1} - 1 = 0$ at $x = n$. The second derivative $z''(x) = -nx^{-2}$, so that $z''(n) = -n^{-1}$. We can write the first terms of the Taylor series for $z(x)$ about $x = n$:

$$(1.7) \qquad z(n + \epsilon) = z(n) - \frac{1}{2n}\epsilon^2 + \cdots.$$

This gives us a heuristic explanation for our earlier calculations. When $\epsilon = 1$, we have $\frac{1}{2n}\epsilon^2 \sim 0$, so $z(n + \epsilon) = z(n) + o(1)$ and thus $y(n + \epsilon) \sim y(n)$. When $\epsilon = 0.1n$, we have the opposite as $\frac{1}{2n}\epsilon^2$ is large. The middle ground is given when ϵ^2 is on the order of n or

when ϵ is on the order of $\sqrt{n}$. We are thus led to the scaling $\epsilon = \lambda\sqrt{n}$, or

$$(1.8) \qquad x = n + \lambda\sqrt{n}.$$

We formally make this substitution in the integral (1.2). Further, we take the factor $y(n) = n^n e^{-n}$ outside the integral so that now the function has maximum value 1. We have scaled both axes. The scaled function is

$$(1.9) \qquad g_n(\lambda) = \frac{y(n + \lambda\sqrt{n})}{y(n)} = (1 + \lambda n^{-1/2})^n e^{-\lambda\sqrt{n}},$$

and we find (noting that $dx = \sqrt{n}d\lambda$)

$$(1.10) \qquad I_n = \int_0^\infty x^n e^{-x} dx = n^n e^{-n} \sqrt{n} \int_{-\sqrt{n}}^{+\infty} g_n(\lambda) d\lambda.$$

Note that while we have been guided by asymptotic considerations, our calculations up to this point have been exact.
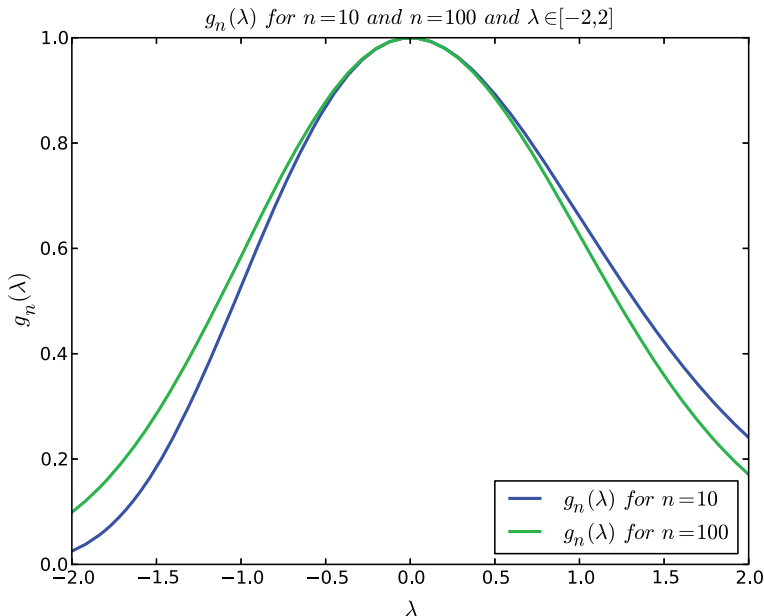

Figure 2. $g_n(\lambda)$ in range $-2 \le \lambda \le +2$ for $n = 10, 100$.

The Taylor series with error term gives

$$(1.11) \qquad \ln(1+\epsilon) = \epsilon - \frac{1}{2}\epsilon^2 + O(\epsilon^3)$$

as $\epsilon \to 0$. Let λ be an arbitrary but fixed real number. Then $\lambda n^{-1/2} \to 0$ so that

(1.12)
$$n\ln(1+\lambda n^{-1/2}) - \lambda n^{1/2} = \lambda n^{1/2} - \frac{1}{2}\lambda^2 + o(1) - \lambda n^{1/2} = -\frac{1}{2}\lambda^2 + o(1)$$

and

$$(1.13) \qquad g_n(\lambda) \to e^{-\lambda^2/2}.$$

That is, when properly scaled, the function $y = x^n e^{-x}$ looks like the bell shaped curve in Figure 2.

Now we would *like* to say

$$(1.14) \qquad \lim_{n\to\infty} \int_{-\sqrt{n}}^{+\infty} g_n(\lambda)d\lambda = \int_{-\infty}^{\infty} e^{-\lambda^2/2}d\lambda = \sqrt{2\pi} \ .$$

Justification of the interchange of limits in the integration of a sequence of functions is one of the most basic and most subtle problems discussed in analysis. Here is a sample theorem: *If $g_n(\lambda)$ are continuous functions on an interval $[a,b]$ and $\lim_{n\to\infty} g_n(\lambda) = g(\lambda)$ for all $\lambda \in [a,b]$, then $\lim_{n\to\infty} \int_a^b g_n(\lambda)d\lambda = \int_a^b g(\lambda)d\lambda$.*

In our example the $g_n(\lambda)$ are indeed continuous and $\lim_{n\to\infty} g_n(\lambda)$ is given by (1.13). But there are three difficulties:

   (1) The left-hand side of the integral in (1.14) is $-\sqrt{n}$.

   (2) The right-hand side of the integral in (1.14) is ∞.

   (3) We will *not* be assuming results from analysis in this book.

A natural approach is to approximate $g_n(\lambda)$ by $e^{-\lambda^2/2}$. The difficulty is that this approximation is not valid throughout the limits of integration. For example, with $\lambda = \sqrt{n}$, $g_n(\lambda) = (2/e)^n$ is not close to $e^{-\lambda^2/2} = e^{-n/2}$. Let us re-examine (1.12) with the error term from

the Taylor series (1.11). Thus *if* $\lambda n^{-1/2} \to 0$, then

(1.15)

$$n \ln(1 + \lambda n^{-1/2}) - \lambda n^{1/2} = \lambda n^{1/2} - \frac{1}{2}\lambda^2 + O(\lambda^3 n^{-1/2}) - \lambda n^{1/2}$$

$$= -\frac{1}{2}\lambda^2 + o(\lambda^2 n^{-1/2}).$$

We now see that the approximation of $g_n(\lambda)$ by $e^{-\lambda^2/2}$ is good as long as $\lambda^2 n^{-1/2} \to 0$, that is, for $\lambda = o(n^{1/4})$. With this in mind let us split the range $[-\sqrt{n}, \infty)$ into a middle range

$$MID = [-L(n), +L(n)],$$

and the two sides

$$LEFT = [-\sqrt{n}, -L(n)]$$

and

$$RIGHT = [L(n), \infty).$$

How should we choose $L(n)$? The middle range should be big enough that most of the integral lies under it but small enough so that the approximation with the bell shaped curve remains valid. The first condition will require that $L(n) \to \infty$ and the second that $L(n) = o(n^{1/4})$. This leaves a lot of room and, indeed, any reasonable $L(n)$ satisfying these criteria would work for our purposes. For definiteness let us set

(1.16) $$L(n) = n^{1/8}.$$

1.1.1. MID. Let us take the most important region, *MID*, first. Guided by the notion that $g_n(\lambda)$ and $e^{-\lambda^2/2}$ will be close, we define an error[2] function

(1.17) $$E_n(\lambda) = g_n(\lambda)/e^{-\lambda^2/2}$$

so that we have the exact expression

(1.18) $$\ln E_n(\lambda) = n \ln(1 + \lambda n^{-1/2}) - \lambda\sqrt{n} + \frac{\lambda^2}{2}.$$

As $\lambda n^{-1/2} \to 0$ in *MID*, we can apply the Taylor series to $\ln(1 + \epsilon)$ with $\epsilon = \lambda n^{-1/2}$. The first two terms cancel the $\lambda\sqrt{n}$ and $\lambda^2/2$ terms,

---

[2]Error does not mean mistake!

which is not so surprising as we *designed* the error to be close to one. We employ the Taylor series to two terms with an error term,

$$(1.19) \qquad n \ln(1 + \lambda n^{-1/2}) = \lambda \sqrt{n} - \frac{\lambda^2}{2} + n \frac{x^3}{3!}.$$

Here x lies somewhere between 0 and $\lambda n^{-1/2}$. As $|\lambda n^{-1/2}| \leq n^{-3/8}$, we can bound

$$(1.20) \qquad |n \frac{x^3}{3}| \leq \frac{1}{3} n^{-1/8}.$$

Thus $|\ln E_n(\lambda)| \leq \frac{1}{3} n^{-1/8}$ throughout λ. Critically, this is a *uniform* bound, which holds for all λ in *MID* simultaneously. As $\ln(E_n(\lambda))$ is small, $E_n(\lambda) - 1$ will also be small. Think of $y = \ln(E_n(\lambda))$, with y small $e^y - 1 \sim y$. But to get a rigorous upper bound, let us use a rougher bound $|e^y - 1| \leq 2y$, valid when $|y|$ is sufficiently small. For n large, $\frac{1}{3} n^{-1/8}$ will be small and so

$$(1.21) \qquad |E_n(\lambda) - 1| \leq \frac{2}{3} n^{-1/8}$$

so that

$$(1.22) \qquad |g_n(\lambda) - e^{-\lambda^2/2}| = e^{-\lambda^2/2} |E_n(\lambda) - 1| \leq \frac{2}{3} n^{-1/8} e^{-\lambda^2/2}$$

and

$$(1.23) \qquad \left| \int_{MID} g_n(\lambda) - e^{-\lambda^2/2} d\lambda \right| \leq \int_{MID} |p_n(\lambda) - e^{-\lambda^2/2}| d\lambda$$
$$\leq \frac{2}{3} n^{-1/8} \int_{MID} e^{-\lambda^2/2} d\lambda.$$

The final integral is less than $\sqrt{2\pi}$, the integral over all λ. The constants are not important, we have bounded the difference in the integrals of $g_n(\lambda)$ and $e^{-\lambda^2/2}$ over *MID* by a constant times $n^{-1/8}$ which in the limit approaches zero.

1.1.2. LEFT. It remains to show that *LEFT* and *RIGHT* give negligible contributions to $\int g_n(\lambda) d\lambda$. Note that we do *not* need asymptotic values of $\int g_n(\lambda) d\lambda$ over *LEFT* or *RIGHT*, only that they approach zero. Thus we can employ a rough (but true) upper bound to $g_n(\lambda)$. The left-hand side is easier. The function $g_n(\lambda)$ is increasing from $-\sqrt{n}$ to $-L(n) = -n^{1/8}$. At $-n^{1/8}$, $g_n(\lambda) \sim e^{-\lambda^2/2} \sim e^{-n^{1/4}/2}$. Since the length of range *LEFT* is less than $\sqrt{n}$, the integral is at

most $\sqrt{n}e^{-n^{1/4}/2}$. The exponential decay dominates the square root growth, and this function goes to zero with n. As this was an upper bound, $\int_{LEFT} g_n(\lambda)d\lambda \to 0$.

1.1.3. RIGHT. The interval $RIGHT$ is more difficult for two reasons: The interval has infinite length so that bounding a single value will not be sufficient. More worrisome, the estimate of $\ln(1 + \epsilon)$ by $\epsilon - \frac{1}{2}\epsilon^2$ is only valid for ϵ small. We require upper bounds that work for the entire range of ϵ. The following specific bounds ((1.24) and (1.25) are included for completeness) are often useful:

$$(1.24) \qquad \ln(1 + \epsilon) \le \epsilon - \frac{1}{2}\epsilon^2 \text{ when } -1 < \epsilon \le 0,$$

$$(1.25) \qquad \ln(1 + \epsilon) \le \epsilon - \frac{1}{4}\epsilon^2 \text{ when } 0 < \epsilon \le 1,$$

$$(1.26) \qquad \ln(1 + \epsilon) \le 0.7\epsilon \text{ when } \epsilon > 1.$$

We break $RIGHT = [n^{1/8}, \infty)$ into two parts. We set
$$NEARRIGHT = [n^{1/8}, n^{1/2}]$$
and
$$FARRIGHT = [n^{1/2}, \infty),$$
reflecting the ranges for the bounds (1.25) and (1.26) with $\epsilon = \lambda n^{-1/2}$. For $NEARRIGHT$ we employ the argument used for $LEFT$. The function $g_n(\lambda)$ is decreasing for λ positive and is $\sim e^{-n^{1/4}/2}$ at $n^{1/8}$. As $NEARRIGHT$ has length less than $\sqrt{n}$, $\int g_n(\lambda)d\lambda$ over $NEARRIGHT$ is at most $\sqrt{n}e^{-n^{1/4}/2}$ which goes to zero.

In $FARRIGHT$, (1.26) gives that

$$(1.27) \quad n\ln(1 + \lambda n^{-1/2}) - \lambda n^{1/2} \le 0.7\lambda\sqrt{n} - \lambda\sqrt{n} \le -0.3\lambda\sqrt{n}.$$

In this interval $g_n(\lambda)$ is thus bounded by the exponentially decaying function $\exp -0.3\lambda\sqrt{n}$. Thus

$$(1.28) \qquad \int_{\sqrt{n}}^{\infty} g_n(\lambda)d\lambda < \int_{\sqrt{n}}^{\infty} e^{-0.3\lambda\sqrt{n}}d\lambda = \frac{1}{0.3\sqrt{n}}e^{-0.3n},$$

and this also goes to zero as $n \to \infty$.

We have shown that the integrals of $g_n(\lambda)$ over *LEFT*, *NEAR-RIGHT*, and *FARRIGHT* all approach zero and that the integral of $g_n(\lambda)$ over *MID* approached $\sqrt{2\pi}$. Putting it all together, the integral of $g_n(\lambda)$ over $[-\sqrt{n}, \infty)$ does indeed approach $\sqrt{2\pi}$.

Whew! Let us take two general principles from this example:

- Crude upper bounds can be used for negligible terms as long as they stay negligible.

- Terms that are extremely small often require quite a bit of work.

1.2. Approximating Sums by Trapezoids

With this method we will not achieve the full Stirling's formula (1.1) but only

$$(1.29) \qquad n! \sim K n^n e^{-n} \sqrt{n}$$

for *some* positive constant K. Our approach follows the classic work [**CR96**] of Richard Courant. We are pleased to reference the eponymous founder of our mathematical home, the Courant Institute.

Our approach is to estimate the logarithm of $n!$ via the formula

$$(1.30) \qquad S_n := \ln(n!) = \sum_{k=1}^{n} \ln(k).$$

The notion is that S_n should be close to the integral of the function $\ln(x)$ between $x = 1$ and $x = n$. We set

$$(1.31) \qquad I_n := \int_1^n \ln(x)dx = [x\ln(x) - x]_1^n = n\ln(n) - n + 1.$$

Let T_n be the value for the approximation of the integral I_n via the trapezoidal rule using step sizes 1. That is, we estimate $\int_i^{i+1} f(x)dx$ by $\frac{1}{2}(f(i) + f(i+1))$. Summing over $1 \le i \le n-1$,

$$(1.32) \qquad T_n = \frac{1}{2}\ln(1) + \sum_{k=2}^{n-1} \ln(k) + \frac{1}{2}\ln(n) = S_n - \frac{1}{2}\ln(n).$$

Set

$$(1.33) \qquad E_n = I_n - T_n$$

to be the error when approximating the integral of $\ln(x)$ by the trape-
zoidal rule. For $1 \le k \le n-1$, let S_k denote the "sliver" of area under
the curve $y = \ln(x)$ for $k \le x \le k+1$ but over the straight line be-
tween $(k, \ln(k))$ and $(k+1, \ln(k+1))$. The curve is over the straight
line as the curve is concave. Then

$$(1.34) \qquad\qquad E_n = \sum_{k=1}^{n-1} \mu(S_k),$$

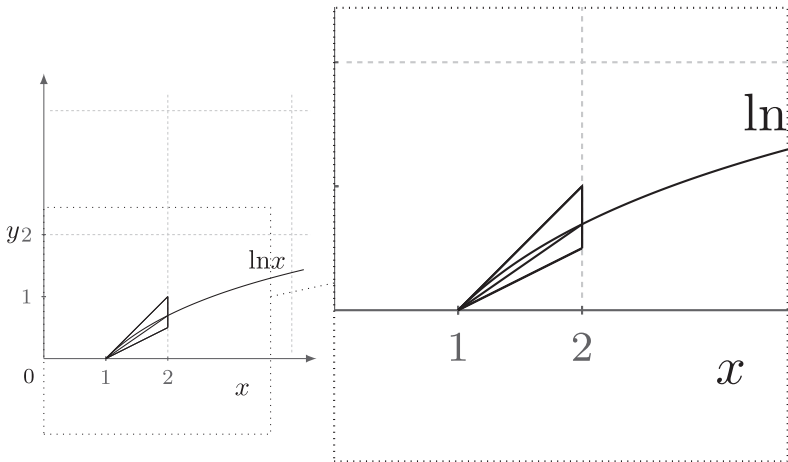where μ denotes the area.

Our goal is to bound the error.

Figure 3. The sliver (shown here with $k = 1$) lies inside
the triangle whose upper and lower lines have slopes $\frac{1}{k}$, $\frac{1}{k+1}$,
respectively.

Theorem 1.1. E_n approaches a finite limit c as $n \to \infty$. Equiva-
lently,

$$(1.35) \qquad\qquad \lim_{n \to \infty} \sum_{k=n}^{\infty} \mu(S_k) = 0.$$

Assuming Theorem 1.1, (1.30)–(1.32) yield

(1.36)
$$\ln(n!) = T_n + \frac{1}{2}\ln n = I_n - E_n + \frac{1}{2}\ln n = n\ln n - n + 1 - c + o(1) + \frac{1}{2}\ln n.$$

Exponentiating both sides

(1.37) $$n! \sim n^n e^{-n} \sqrt{n} e^{1-c}$$

giving the desired (1.29) with $K = e^{1-c}$.

Now, how do we show Theorem 1.1? We consider $\mu(S_k)$ in Asymptopia, as $k \to \infty$. Roughly,[3] $\mu(S_k)$ is the error between the integral from k to $k+1$ of $f(x) = \ln x$ and the straight line approximation of $f(x)$. This error is caused by the *second* derivative of $f(x)$. (Had the second derivative been zero, the straight line would have been the precise function.) Here, the second derivative $f''(x) = -x^{-2}$ is on the order of k^{-2}, and the interval has length 1, so we feel the error should be on the order of k^{-2}. As k^{-2} is decreasing sufficiently quickly, the infinite sum of $\mu(S_k)$ should converge.

Guided by this intuitive approach, we give an explicit upper bound for $\mu(S_k)$. Observe that it need not be a good upper bound. We still would get convergence of $\sum \mu(S_k)$ even if our upper bound were, say, ten times the actual value.

Here is *one* approach that works. Let $P = (k, \ln k)$, and let $Q = (k+1, \ln(k+1))$. Let C denote the curve $f(x) = \ln x$ in the interval $[k, k+1]$. In the interval $[k, k+1]$, our function $f(x) = \ln x$ has derivative between $\frac{1}{k}$ and $\frac{1}{k+1}$. Let U (upper) be the straight line segment starting at P with slope $\frac{1}{k}$, ending at $x = k+1$. Let L (lower) be the straight line segment starting at P with slope $\frac{1}{k+1}$, ending at $x = k+1$. As the derivative of curve C is always between those of U and L, the curve C is under U and over L. At $x = k+1$, L then is below the curve C, so below the point Q. Thus the straight line PQ lies above the line L. We can then bound $\mu(S_k)$, the area between C and the straight line PQ, by the area between U and L. But this latter area is a triangle. Make the base of the triangle the line from U to L at $x = k+1$ to be the distance from U to L at $x = k+1$, which is precisely the difference of the slopes which is $\frac{1}{k} - \frac{1}{k+1}$. The

---

[3]An intuitive feel is very useful, but it must be followed up with a rigorous argument!

height of the triangle is then 1, from $x = k$ to $x = k + 1$. We have thus shown

$$(1.38) \qquad \mu(S_k) \leq \frac{1}{2}\left(\frac{1}{k} - \frac{1}{k+1}\right).$$

This value is $O(k^{-2})$, and so we achieve convergence. Indeed we have the explicit upper bound

$$(1.39) \qquad \sum_{k=1}^{\infty} \mu(S_k) \leq \sum_{k=1}^{\infty} \frac{1}{2}\left(\frac{1}{k} - \frac{1}{k+1}\right) = \frac{1}{2}$$

as the sum telescopes. This yields (1.29), almost Stirling's formula.

1.3. Combining Forces to Estimate the Error

Setting $c = \lim_{n\to\infty} E_n$, define the tail

$$(1.40) \qquad F_n = c - E_n = \sum_{k=n}^{\infty} \mu(S_k).$$

Now (1.36) becomes

$$(1.41) \qquad \ln(n!) = n\ln n - n + 1 - c + \frac{1}{2}\ln n + F_n.$$

From the proof of Stirling's formula in Section 1.1, we know that $e^{1-c} = \sqrt{2\pi}$. Exponentiating both sides, we may express the result as

$$(1.42) \qquad \frac{n!}{n^n e^{-n}\sqrt{2\pi n}} = e^{F_n}.$$

That is, e^{F_n} gives the *error term* in the approximation of Stirling's formula. Since $F_n \to 0$, $e^{F_n} = 1 + F_n(1 + o(1))$ and so

$$(1.43) \qquad \frac{n!}{n^n e^{-n}\sqrt{2\pi n}} = 1 + F_n(1 + o(1)).$$

While (1.42) is exact, we do not have a closed form for F_n. Still, we may find it in Asymptopia.

Consider $\mu(S_k)$ more carefully. Parametrizing $y = k + x$, we have[4]

$$(1.44) \qquad \mu(S_k) = \int_0^1 \ln(k+y) - [(1-y)\ln(k) + y\ln(k+1)]dy$$

---

[4]Moving the region of interest to near zero is often times helpful!

as the bracketed term is the equation of the straight line PQ above. From the Taylor series (the asymptotics here are as $k \to \infty$, uniformly over $y \in [0, 1]$),

$$(1.45) \qquad \ln(k + y) = \ln k + \frac{1}{k}y - \frac{y^2}{2k^2} + O(k^{-3}).$$

As

$$(1.46) \qquad \ln(k + 1) = \ln k + \frac{1}{k} - \frac{1}{2k^2} + O(k^{-3}),$$

we find

$$(1.47) \qquad (1 - y)\ln(k) + y\ln(k + 1) = \ln k + \frac{1}{k}y + \frac{y}{2k^2} + O(k^{-3}).$$

Subtracting[5]

$$(1.48) \qquad \mu(S_k) = \int_0^1 \frac{1}{2k^2}(y - y^2) + O(k^{-3})dy.$$

The main part can be integrated precisely, and

$$(1.49) \qquad \mu(S_k) = \frac{1}{12k^2} + O(k^{-3}).$$

This allows us to estimate F_n:

$$(1.50) \qquad F_n = \sum_{k=n}^{\infty} \mu(S_k) \sim \int_n^{\infty} \frac{1}{12z^2}dx = \frac{1}{12n}.$$

This gives a more precise approximation for $n!$:

$$(1.51) \qquad \frac{n!}{n^n e^{-n}\sqrt{2\pi n}} = \left(1 + \frac{1 + o(1)}{12n}\right).$$

Indeed, with considerably more care one can show that

$$(1.52) \qquad \frac{1}{12n + 1} \le F_n \le \frac{1}{12n},$$

which yields the remarkably close[6] bounds

$$(1.53) \qquad e^{1/(12n+1)} \le \frac{n!}{n^n e^{-n}\sqrt{2\pi n}} \le e^{1/(12n)},$$

which are valid for *all* n.

---

[5] Caution! Subtracting in Asymptopia is tricky! Often times main terms cancel and the secondary terms become paramount. Even worse, occasionally the secondary terms also cancel and it is the tertiary terms that are important.

[6] Try it for $n = 10$.

1.4. Estimating the Integral More Accurately

Note. This section gets quite technical and should be considered optional.

Let us begin again with the precise formula

$$(1.54) \qquad n! = n^n e^{-n} \sqrt{n} \int_{-\sqrt{n}}^{\infty} g_n(\lambda) d\lambda.$$

Our goal is to replicate (1.51) by more accurately estimating $p_n(\lambda)$. Our previous estimate was $e^{-\lambda^2/2}$. Now, however, we will want the estimate to be within an additive $o(n^{-1})$ term. Our previous definition of *MID* will be too broad. Instead we define

$$(1.55) \qquad L(n) = n^{0.01}$$

and

$$\begin{aligned} MID &= [-L(n), +L(n)], \\ LEFT &= [-\sqrt{n}, -L(n)], \\ RIGHT &= [L(n), \sqrt{n}]. \end{aligned}$$

The bounds on $\int p_n(\lambda) d\lambda$ are still (this requires checking!) exponentially small, and thus they are not only $o(1)$ but $o(n^{-1})$. This allows us to concentrate on $\int p_n(\lambda) d\lambda$ over our new *MID*. We have $E_n(\lambda)$ and $\ln(E_n(\lambda))$ as in (1.17) and (1.18). Now, however, we need a more accurate Taylor series estimation for $\ln(1 + \epsilon)$ with $\epsilon = \lambda n^{-1/2}$. A priori, it is unclear just how many terms we will need.

- Experimentation is part of the art of Asymptopia.

After possibly a number of false starts, examine the Taylor series out to four terms with the error term. From Theorem 2.18

$$(1.56) \qquad \ln(1 + \epsilon) = \epsilon - \frac{1}{2}\epsilon^2 + \frac{1}{3}\epsilon^3 - \frac{1}{4}\epsilon^4 + \frac{1}{5}x^5$$

with $|x| \le \epsilon$. Applying this to (1.18), the first two terms cancel as before and

$$(1.57) \qquad \ln(E_n(\lambda)) = \frac{1}{3}\lambda^3 n^{-1/2} - \frac{1}{4}\lambda^4 n^{-1} + \frac{1}{5}n^{-3/2}x^5.$$

With our *MID* now narrower, $|\frac{1}{5}n^{-3/2}x^5| \le n^{-1.45}$ which will be negligible for our purposes here. With $y = \ln E_n(\lambda)$ we want to go

from y to $e^y - 1$. Because we need greater accuracy (and after some experimentation!), we bound

$$(1.58) \qquad e^y - 1 = y + \frac{y^2}{2} + O(y^3).$$

Thus (1.57) becomes

$$(1.59) \quad E_n(\lambda) = 1 + \frac{1}{3}\lambda^3 n^{-1/2} + \frac{1}{18}\lambda^6 n^{-1} - \frac{1}{4}\lambda^4 n^{-1} + O(n^{-1.41}).$$

(The O term in (1.59) contains several terms of which the largest is $(\lambda^3 n^{-1/2})^3$.)

This gives us a good estimate for $\int g_n(\lambda)d\lambda$ over *MID*:

$$
\begin{aligned}
(1.60) \qquad & \int_{-L(n)}^{L(n)} g_n(\lambda)d\lambda \\
& = \int_{-L(n)}^{L(n)} e^{-\lambda^2/2}[1 + \frac{1}{3}\lambda^3 n^{-1/2} + \frac{1}{18}\lambda^6 n^{-1} \\
& \qquad\qquad - \frac{1}{4}\lambda^4 n^{-1} + O(n^{-1.41})].d\lambda
\end{aligned}
$$

The contribution of the $O(n^{-1.41})$ term to the integral is $o(n^{-1})$. This is an acceptable error, so we rewrite

$$(1.61)$$
$$
\begin{aligned}
& \int_{-L(n)}^{L(n)} g_n(\lambda)d\lambda \\
& = o(n^{-1}) + \int_{-L(n)}^{L(n)} e^{-\lambda^2/2}[1 + \frac{1}{3}\lambda^3 n^{-1/2} + \frac{1}{18}\lambda^6 n^{-1} - \frac{1}{4}\lambda^4 n^{-1}]d\lambda.
\end{aligned}
$$

We want to replace the limits of integration to $\pm\infty$, but we must pause for a moment as we require an accuracy of $o(n^{-1})$.

Let us give some very rough upper bounds on $\int_{L(n)}^{\infty} \lambda^4 e^{-\lambda^2/2}d\lambda$, as the other side and the smaller powers are similar. We bound $\lambda^4 \leq e^\lambda$, certainly true for $\lambda \geq L(n)$. Then

$$(1.62) \quad \int_{L(n)}^{\infty} \lambda^4 e^{-\lambda^2/2}d\lambda \leq \int_{L(n)}^{\infty} e^{\lambda - \frac{1}{2}\lambda^2}d\lambda = e^{1/2} \int_{L(n)-1}^{\infty} e^{-y^2/2}dy$$

by substituting $y = \lambda - 1$.[7] Here we substitute $y = L(n) - 1 + z$ and bound $\frac{1}{2}y^2 \geq \frac{1}{2}(L(n) - 1)^2 + z(L(n) - 1)$ so that

$$
(1.63) \qquad \int_{L(n)-1}^{\infty} e^{-y^2/2} dy \leq e^{-(L(n)-1)^2/2} \int_0^{\infty} e^{-z(L(n)-1)} dz
$$

$$
= e^{-(L(n)-1)^2/2)(L(n)-1)^{-1}}.
$$

This is exponentially small in n and, so, certainly $o(n^{-1})$. (Note however that it was important to let $L(n)$ increase fast enough. Had we tried, say, $L(n) = \ln\ln n$, the bound would be $o(1)$ and not the desired $o(n^{-1})$.)

Returning to (1.61) we now have

$$
(1.64)
$$
$$
\int_{-L(n)}^{L(n)} g_n(\lambda) d\lambda
$$
$$
= o(n^{-1}) + \int_{-\infty}^{\infty} e^{-\lambda^2/2}[1 + \frac{1}{3}\lambda^3 n^{-1/2} + \frac{1}{18}\lambda^6 n^{-1} - \frac{1}{4}\lambda^4 n^{-1}] d\lambda.
$$

Fortunately $\int_{-\infty}^{+\infty} \lambda^i e^{-\lambda^2/2} d\lambda$ can be found precisely for each non-negative integer i by elementary[8] calculus. For odd i the integral is zero and for $i = 0, 4, 6$ the integrals are $\sqrt{2\pi}$, $3\sqrt{2\pi}$, and $15\sqrt{2\pi}$, respectively, so that

$$
(1.65) \qquad \int_{-L(n)}^{L(n)} p_n(\lambda) d\lambda = o(n^{-1}) + \sqrt{2\pi}\left(1 + \frac{15}{18n} - \frac{3}{4n}\right),
$$

which is the promised $1 + \frac{1}{12n}$ term.

Remark. We need not stop here. One can take the Taylor series for $\ln(1 + \epsilon)$ out further and redefine *MID* to be narrower. With a considerable amount of effort, one can show

$$
(1.66) \qquad n! = n^n e^{-n} \sqrt{2\pi n}[1 + \frac{1}{12n} + \frac{1}{288n^2} + o(n^{-2})],
$$

and, indeed, one gets an infinite sequence of such approximations.

---

[7]The tail of the normal distribution is more carefully studied in §3.1.
[8]Elementary does not mean easy! Use integration by parts.

1.5. An Application to Random Walks

Here we will apply Stirling's formula to yield a classical result in the study of random walks.

Let G be an arbitrary graph for which each vertex has at least one, but only a finite number of neighbors. Let s (source) be some specified vertex of G. A simple random walk on G begins at s. Each time unit it moves uniformly from its current position v to one of the neighbors of v.

The study of random walks was begun by George Pólya around 1920. There is an essential dichotomy. A random walk is called *recurrent* if with probability 1 it returns to its beginning, here s. Otherwise, the random walk is called *transient*. In this case, while the walk *might* return to s, there is a positive probability that it will *never* return to s. Let $p(t)$ denote the probability (dependent on G and s) that the random walk will be at s at time t. Pólya showed that the dichotomy depended on the decay of $p(t)$. He showed:

Theorem 1.2. *If $\sum_{t=1}^{\infty} p(t)$ is finite, then the random walk is transient, and if $\sum_{t=1}^{\infty} p(t)$ is infinite, then the random walk is recurrent.*

Proof. Suppose there is a probability α that the random walk ever returns to s. Once it returns, it is again beginning a random walk. Hence the probability that it returns at least j times would be α^j. The expected number of times it returns would then be $\sum_{j=1}^{\infty} \alpha^j$. This expected number is also $\sum_{t=1}^{\infty} p(t)$. If $\alpha < 1$, then the sum is finite. If $\alpha = 1$ the sum is infinite. $\qquad\square$

Now let us restrict ourselves to the grid $\mathbb{Z}^d$. The vertices are the vectors $\vec{v} = (a_1, \ldots, a_d) \in \mathbb{Z}^d$, and the neighbors of $\vec{v}$ are those $\vec{w}$ which agree with $\vec{v}$ in all but one coordinate and are one away from $\vec{v}$ in that coordinate. (This is the usual grid for $\mathbb{Z}^d$.) By symmetry, the start matters little so we consider walks beginning at the origin $\vec{0}$. In $\mathbb{Z}^2$, for example, from each (a, b) we move randomly either North $(a, b+1)$, East $(a+1, b)$, South $(a, b-1)$, or West $(a-1, b)$. We

continue forever, giving a sequence $\vec{0} = \vec{w_0}, \vec{w_1}, \ldots$, where $\vec{w_t}$ denote the position at time t.

Is the random walk in $\mathbb{Z}^d$ recurrent or transient? George Pólya gave the surprising solution:

Theorem 1.3. *The random walk in Z^d is recurrent if $d = 1$ or $d = 2$ and is transient if $d \geq 3$.*

From parity considerations one can only return to $\vec{0}$ after an even number of steps. Thus the nature of the random walk depends on whether $\sum_{t=1}^{\infty} p(2t)$ is finite. In Asymptopia we shall find the asymptotics of $p(2t)$ (note that $p(2t)$ depends on the dimension d).

In one dimension we want the probability that out of $2t$ steps precisely t are $+1$ (to the right). This has the formula

$$(1.67) \qquad\qquad p(2t) = 2^{-2t} \binom{2t}{t}.$$

Applying Stirling's formula,

$$(1.68) \qquad p(2t) \sim 2^{-2t} \frac{(2t)^{2t} e^{-2t} \sqrt{2\pi(2t)}}{[t^t e^{-t}\sqrt{2\pi t}]^2} \sim \sqrt{\frac{1}{\pi t}}.$$

As $\sum t^{-1/2}$ diverges, the random walk in $\mathbb{Z}^2$ is recurrent.

For dimension two there is a clever way to compute $p(2t)$. Change the coordinate system[9] to basis $\vec{v_1} = (\frac{1}{2}, \frac{1}{2})$, $\vec{v_2} = (\frac{1}{2}, -\frac{1}{2})$. Now North, South, East, and West have coordinates $(1, -1)$, $(-1, +1)$, $(1, 1)$, and $(-1, -1)$, respectively. One returns to the origin at time $2t$ if and only if each coordinate is zero. The new coordinates are now *independent*, and so we find the closed formula

$$(1.69) \qquad\qquad p(2t) = [2^{-2t}\binom{2t}{t}]^2.$$

From (1.68) we now find

$$(1.70) \qquad\qquad p(2t) \sim \frac{1}{\pi t}.$$

As this series diverges, the random walk in $\mathbb{Z}^2$ is recurrent.

---

[9]Effectively, tilt your head at a 45 degree angle!

Remark. While $\sum \frac{1}{\pi t}$ diverges, it *barely* diverges in the sense that the sum up to t grows only logarithmically. This makes the random walk in $\mathbb{Z}^2$ a strange beast. For example, it is possible to prove that the expected time until the first return to the origin is infinite.

In the remainder we assume that the dimension $d \geq 3$. (The methods apply also to the cases $d = 1, 2$, but there the exact formulae make things easier.) These cases are all quite similar, and the reader may concentrate on $d = 3$. For dimension $d \geq 3$, there does not exist a closed form[10] for $p(2t)$. In Asymptopia, however, that is hardly a stumbling block. In our asymptotics below, $d \geq 3$ is arbitrary but fixed, and $t \to \infty$.

Each step in the random walk is in one of $2d$ directions, each equally likely. However, we split the choice of directions into two parts. First, we decide for each step in which dimension it is moving. Let X_i denote the number of steps in dimension i. Then X_i has binomial distribution $BIN[2t, \frac{1}{d}]$. Note, however, that as $\sum_{i=1}^{d} X_i = 2t$, the X_i are not independent!

Theorem 1.4. *The probability that all X_i, $1 \leq i \leq d$, are even is* $2^{-d+1} + o(1)$.

On an intuitive level, each X_i has probability roughly $1/2$ of being even. However, once $X_1, \ldots, X_{d-1}$ are even, X_d is automatically even. We use a result on binomial distributions which is of independent interest.

Theorem 1.5. *Let $\epsilon > 0$ be arbitrary but fixed. Let $p = p(n)$ with $\epsilon \leq p \leq 1 - \epsilon$ for all n. Let $X = X_n$ have binomial distribution $BIN[n, p(n)]$. Then, the probability that X_n is even approaches $1/2$.*

Proof. The binomial formula gives
(1.71)
$$(px + (1-p)y)^n = \sum_{i=1}^{n} \binom{n}{i} (px)^i ((1-p)y)^{n-i} = \sum_{i=1}^{n} \Pr[X = i] x^i y^{n-i}.$$

---

[10]The often used phrase "closed form" does not have a precise definition. We do not consider an expression involving a summation to be in closed form.

Set $x = -1$, $y = 1$. Then

$$(1.72) \quad (p-(1-p))^n = \sum_{i=1}^{n} \Pr[X = i](-1)^i = \Pr[X \text{ even}] - \Pr[X \text{ odd}].$$

As $1 = \Pr[X \text{ even}] + \Pr[X \text{ odd}]$,

$$(1.73) \qquad \Pr[X \text{ even}] = \frac{1}{2} + \frac{1}{2}(1 - 2p)^n.$$

With p not close to either 0 or 1, $(1 - 2p)^n \to 0$ and $\Pr[X \text{ even}] = \frac{1}{2} + o(1)$. $\qquad\qquad\square$

We could also prove Theorem 1.5 in Asymptopia. Here is the rough[11] idea. The distribution $BIN[n, p]$ takes, with probability $1 - o(1)$, values $i \sim pn$. For such i one computes
$$(1.74)$$
$$\frac{\Pr[BIN[n, p] = i]}{\Pr[BIN[n, p] = i + 1]} = \frac{\binom{n}{i}p^i(1-p)^{n-i}}{\binom{n}{i+1}p^{i+1}(1-p)^{n-i-1}} = \frac{(i+1)(1-p)}{(n-i)p}.$$

For those $i \sim pn$ this ratio is nearly 1. Supposing i even for convenience, the contribution to $\Pr[BIN[n, p] = i]$ to $BIN[n, p]$ being even is asymptotically the same as the contribution $\Pr[BIN[n, p] = i + 1]$ to $BIN[n, p]$ being odd.

Note, however, that the proof of Theorem 1.5 gives a much stronger result. The probability that $BIN[n, p]$ is even minus the probability that $BIN[n, p]$ is odd is exponentially small—something Asymptopia does not yield. Furthermore, though we do not use it here, the proof of Theorem 1.5 shows that the result holds even under the much weaker hypothesis that $p \gg n^{-1}$ and $1 - p \gg n^{-1}$. While we aim for proofs from Asymptopia in this work, we should remain mindful that other techniques can sometimes be even more powerful!

Proof of Theorem 1.4. Formally we show for $1 \le i \le d-1$ that the probability that $X_1, \ldots, X_i$ is all even is $2^{-i} + o(1)$. The case $i = 1$ is precisely Theorem 1.5 with $p = \frac{1}{2}$. By induction suppose the result for i. Set $m = 2t - (X_1 + \cdots + X_i)$. With probability $1 - o(1)$ all $X_1, \ldots, X_i \sim \frac{2t}{d}$ so that $m \sim 2t(1 - \frac{i}{d})$. Thus, with probability $2^{-i} + o(1)$ all $X_1, \ldots, X_i$ are even *and* $m \sim 2t\frac{d-i}{d}$. Conditional on

---

[11] Try to write it in detail yourself!

these values X_{i+1} has distribution $BIN[m, \frac{1}{d-i}]$. From Theorem 1.5, the conditional probability that X_{i+1} is even is $\frac{1}{2} + o(1)$.

In particular, $X_1, \ldots, X_{d-1}$ are all even with probability $2^{1-d} + o(1)$. As $X_1 + \cdots + X_d = 2t$ is even, X_d is then even tautologically. $\square$

Theorem 1.6. *Let $p_d(2t)$ denote the probability that a random walk on $\mathbb{Z}^d$ beginning at the origin is at the origin at time $2t$. Then, for $d \geq 3$ fixed and $t \to \infty$,*

$$(1.75) \qquad p_d(2t) \sim 2^{1-d} \left(\sqrt{\frac{d}{t\pi}} \right)^d.$$

Proof. As above, let X_i denote the number of steps in direction i. With probability 2^{1-d} all X_i are even. From large deviation bounds given later, in particular Theorem 8.3, there is a K so that with probability $1 - o(t^{-d/2})$ all X_i are within $K\sqrt{t \ln t}$ of $\frac{t}{d}$. The $o(t^{-d/2})$ term will not affect (1.75). (We actually only need that the X_i are $\frac{t}{d} + o(t)$.) Condition on the X_i all being even and having values $2s_i$ with $2s_i \sim \frac{2t}{d}$. Now in each dimension the probability that the $+1$ and -1 steps balance is the probability that a random walk in $\mathbb{Z}$ of time $2s_i$ ends at the origin which is from (1.68) $\sim (1/s_i\pi)^{1/2} \sim (d/t\pi)^{1/2}$. Conditioning on the $X_1, \ldots, X_d$ the events that each dimension i balances between $+1$ and -1 are mutually independent, and so the probability that they all balance—precisely what we need to return to the origin—is $\sim ((2d/t\pi)^{1/2})^d$ as claimed. $\square$

We can now easily complete the proof of Polya's Theorem 1.3. When $d \geq 3$, $p_d(2t) = O(t^{-d/2})$. In these cases $d/2 > 1$ and so $\sum_{t=1}^{\infty} p_d(2t)$ is finite, and Theorem 1.2 gives that the random walk is transient.