A MATHEMATICAL OFFERING
PROBLEM 5 SOLUTION

Tim McLarnan

The first part of this problem was to verify the approximation

$$e\left(\frac{n}{e}\right)^n < n! < \left(\frac{e}{2}\right)^2\left(\frac{n+1}{e}\right)^{n+1} \tag{1}$$

for factorials. This approximation and its improvements are called *Stirling's formula*, after an English mathematician roughly of Newton's generation.

In this weak a form, the approximation is easy to get. You start by following the hint and taking the log:

$$\ln(n!) = \ln(1 \cdot 2 \cdot 3 \cdot \cdots \cdot n) = \ln(1) + \ln(2) + \ln(3) + \cdots + \ln(n).$$

Now look at the picture below.

Comparing areas, the left figure shows that

$$\int_1^n \ln x\, dx < \ln(1) + \ln(2) + \ln(3) + \cdots + \ln(n)$$

(remember that $\ln 1 = 0$); and the right figure shows that

$$\ln(1) + \ln(2) + \ln(3) + \cdots + \ln(n) < \int_2^{n+1} \ln x\, dx.$$

Typeset by $\mathcal{AMS}$-TEX

It is a standard example of integration by parts that

$$\int \ln x \, dx = x \ln x - x.$$

Applying this to the inequalities above gives

$$n \ln n - n + 1 < \ln(1) + \ln(2) + \cdots + \ln(n) < (n+1)\ln(n+1) - (n+1) - 2\ln 2 + 2.$$

Raising e to each of these 3 powers and simplifying a bit gives formula (1).

Now, formula (1) roughly situates the factorial function among the exponential functions, but one would still like to have a closer approximation. One reason for this can be seen if we take the ratio of our upper bound to our lower bound for $n!$:

$$\frac{\left(\frac{e}{2}\right)^2 \left(\frac{n+1}{e}\right)^{n+1}}{e\left(\frac{n}{e}\right)^n} = \frac{1}{4}\left(1 + \frac{1}{n}\right)^n (n+1) \to \frac{1}{4}e(n+1).$$

We have used here the fact, which you should be able to derive from l'Hôpital's Rule, that

$$\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n = e.$$

The upshot of this observation is that for large n, our two estimates differ by a factor of roughly $(n+1)e/4$. In partcular, the difference between our upper and lower bounds is much larger than the lower bound itself. Wouldn't it be nice if we could find an estimate whose error was much smaller than the estimate itself? That is, wouldn't it be nice if we could find upper and lower bounds whose ratio approached 1 asymptotically?

One way to find an improved estimate is to think a bit more carefully about the geometry of our approximations so far. Our lower bound was obtained by finding an integral for which $\ln(n!)$ is a left-hand sum. Our upper bound was obtained by finding an integral for which $\ln(n!)$ is a right-hand sum. Might we not get a better estimate for $\ln(n!)$ if we could find an integral for which $\ln(n!)$ is a midpoint sum?[1]

An integral that does this trick is

$$\int_{\frac{1}{2}}^{n+\frac{1}{2}} \ln(x) \, dx,$$

[1]Notice that the process here is an interesting reversal from the usual one. Normally, we have a fixed integral, and are trying to get better and better approximations of it by moving from right- and left-hand sums to midpoint or trapezoid sums to Simpson's Rule. Here, the sitation is reversed. We have the sum $(\ln(1) + \ln(2) + \cdots + \ln(n))$, and we are trying to approximate it with an integral. We are guessing that an integral for which this sum is a midpoint sum will approximate the sum better than an integral for which it is a left- or right-hand sum. Would it be possible to find an integral for which this sum was a Simpson's Rule approximant?

for which $\ln(1) + \ln(2) + \cdots + \ln(n)$ is a midpoint sum with intervals of length 1. To see how closely the sum and the integral agree, look at the slice of length 1 centered at the integer ν. The integral over this slice is

$$\begin{aligned}
\int_{\nu-\frac{1}{2}}^{\nu+\frac{1}{2}} \ln t\, dt &= \int_0^{\frac{1}{2}} (\ln(\nu+t) + \ln(\nu-t))\, dt \\
&= \int_0^{\frac{1}{2}} \ln(\nu^2 - t^2)\, dt \\
&= \int_0^{\frac{1}{2}} \ln\left(\nu^2\left(1 - \frac{t^2}{\nu^2}\right)\right) dt \\
&= \int_0^{\frac{1}{2}} \left(\ln(\nu^2) + \ln\left(1 - \frac{t^2}{\nu^2}\right)\right) dt \\
&= \ln(\nu) + \int_0^{\frac{1}{2}} \ln\left(1 - \frac{t^2}{\nu^2}\right) dt.
\end{aligned}$$

Summing this formula for $1 \le \nu \le n$ gives

$$\int_{\frac{1}{2}}^{n+\frac{1}{2}} \ln t\, dt = \sum_{\nu=1}^{n} \ln \nu + \sum_{\nu=1}^{n} \int_0^{\frac{1}{2}} \ln\left(1 - \frac{t^2}{\nu^2}\right) dt.$$

Working out the integral on the left and replacing the first sum on the right with $\ln(n!)$ gives

$$\begin{aligned}
\ln(n!) &= \left(n+\tfrac{1}{2}\right)\ln\left(n+\tfrac{1}{2}\right) - \left(n+\tfrac{1}{2}\right) - \tfrac{1}{2}\ln\tfrac{1}{2} + \tfrac{1}{2} - \sum_{\nu=1}^{n} \int_0^{\frac{1}{2}} \ln\left(1 - \frac{t^2}{\nu^2}\right) dt \\
&= \left(n+\tfrac{1}{2}\right)\ln\left(n+\tfrac{1}{2}\right) - \tfrac{1}{2}\ln\tfrac{1}{2} - n + C_n.
\end{aligned}$$

If we apply exp to both sides of this equation, we get

$$n! = \left(\frac{n+\frac{1}{2}}{e}\right)^{n+\frac{1}{2}} \sqrt{2e}\, e^{C_n}.$$

In order to obtain the estimate stated in the problem, that the ratio of $n!$ to

$$\left(\frac{n+\frac{1}{2}}{e}\right)^{n+\frac{1}{2}} \sqrt{2\pi}$$

approaches 1 as $n \to \infty$, we would need to show that

$$e^C = \lim_{n\to\infty} e^{C_n} = \sqrt{\pi/e}.$$

Unfortunately, this turns out to be slightly tricky. On the other hand, it is not too hard to show that as $n \to \infty$, at least C_n approaches some constant. This is enough to imply that for large n, the ratio of $n!$ to

$$\left(\frac{n+\frac{1}{2}}{e}\right)^{n+\frac{1}{2}}$$

approaches a constant, but not to specify what constant.

To show that

$$C = \lim_{n\to\infty} C_n = \lim_{n\to\infty} -\sum_{\nu=1}^{n} \int_0^{\frac{1}{2}} \ln\left(1-\frac{t^2}{\nu^2}\right)\, dt$$

exists, we begin by observing that at every point in each integral, $0 \leq t^2/\nu^2 \leq 1/4$. In this interval, a moment's reflection on the derivative of $\ln x$ shows that

$$0 \geq \ln\left(1-\frac{t^2}{\nu^2}\right) \geq -\frac{4t^2}{3\nu^2} \geq -\frac{1}{3\nu^2},$$

which means that

$$0 \geq \int_0^{\frac{1}{2}} \ln\left(1-\frac{t^2}{\nu^2}\right)\, dt \geq -\frac{1}{6\nu^2}.$$

Thus,

$$0 < C_n < \sum_{\nu=1}^{\infty} \frac{1}{6\nu^2} = \frac{\pi^2}{36}.$$

(The numerical value here doesn't matter; though it ought to seem startling and wonderful if you don't know its truth. What does matter is that the series converges, which follows immediately from the integral test.) The consequence is that C_n is a bounded, monotonically increasing function of n, which must therefore have a limit.

Already we have a pretty strong form of Stirling's formula, but let me show how one might take the next step of finding a numerical value for C. The standard way to do this uses identities of the Γ-function, an analytic extension of the factorial function which is defined for all complex numbers. (Did you know one can make unambiguous good sense out of expression like $(-5)!$, or $(1/2)!$, or $i!$?) This is done nicely in Titchmarsh's estimable *The Theory of Functions*, §1.87. I'd like to take a different path, which is the one that occurred to me, and which uses less machinery. Unfortunately, I'll still need both an infinite product identity for $\sin\theta$ which may not be familiar, and some computational craftiness.

We want to show that $e^C = \sqrt{\pi/e}$, i.e., that

$$C = -\sum_{\nu=1}^{n} \int_0^{\frac{1}{2}} \ln\left(1-\frac{r^2}{\nu^2}\right)\, dt = \tfrac{1}{2}(\ln\pi - 1).$$

Formally (though all these steps can be justified rigorously) we have

$$\begin{aligned}
C = -\sum_{\nu=1}^{n} \int_0^{\frac{1}{2}} \ln\left(1 - \frac{t^2}{\nu^2}\right) dt &= -\int_0^{\frac{1}{2}} \sum_{\nu=1}^{n} \ln\left(1 - \frac{t^2}{\nu^2}\right) dt \\
&= -\int_0^{\frac{1}{2}} \ln\left(\prod_{\nu=1}^{n}\left(1 - \frac{t^2}{\nu^2}\right)\right) dt \\
&= -\int_0^{\frac{1}{2}} \ln\left(\frac{\sin(\pi t)}{\pi t}\right) dt \\
&= -\frac{1}{\pi}\int_0^{\frac{\pi}{2}} \ln\left(\frac{\sin\theta}{\theta}\right) d\theta \\
&= -\frac{1}{\pi}\int_0^{\frac{\pi}{2}} \left(\ln(\sin\theta) - \ln\theta\right) d\theta \\
&= -\frac{1}{\pi}\int_0^{\frac{\pi}{2}} \ln(\sin\theta)\, d\theta + \tfrac{1}{\pi}\left[\theta\ln\theta - \theta\right]_0^{\pi/2} \\
&= -\frac{1}{\pi}\int_0^{\frac{\pi}{2}} \ln(\sin\theta)\, d\theta + \tfrac{1}{2}\ln\tfrac{\pi}{2} - \tfrac{1}{2}.
\end{aligned}$$

Here we've used an identity for the sine function,

$$\frac{\sin\theta}{\theta} = \prod_{\nu=1}^{n}\left(1 - \frac{\theta^2}{\pi^2\nu^2}\right),$$

which you may well not have seen[2], but the other steps are routine arithmetic. A moment more arithmetic then shows that in order to have $C = \frac{1}{2}(\ln\pi - 1)$, it suffices to have

$$\int_0^{\frac{\pi}{2}} \ln(\sin\theta)\, d\theta = -\frac{\pi}{2}\log 2. \qquad (2)$$

At this point, we seem to be nearly finished; but there is still one more barrier to overcome. The indefinite integral of $\ln(\sin\theta)$ is a messy expression involving the dilogarithm function, a much more obscure function than the Γ-function. Are we now stymied? No, because we can take advantage of the symmetry of the trig functions over the interval $[0, \pi/2]$ to evaluate the definite integral without ever confronting the indefinite integral!

[2]To prove this identity with modern rigor is not completely trivial. There is a careful argument using Weierstrass' Factor Theorem in volume 2 of Knopp's *Theory of Functions*, and similar proofs both in Titchmarsh and in Whittaker and Watson's *Modern Analysis*. In the old days of Euler and the Bernoullis, though, this identity was both familiar and regarded as easy to prove. Think of both sides of the identity as polynomials of infinite degree. These two polynomials have the same roots (every non-zero multiple of π) and the same constant term (1), so they must be the same polynomial.

Symmetry and the standard properties of the log and trig functions let us write

$$\begin{aligned}
2\int_0^{\frac{\pi}{2}} \ln(\sin\theta)\,d\theta &= \int_0^{\frac{\pi}{2}} \ln(\sin\theta) + \ln(\cos\theta)\,d\theta \\
&= \int_0^{\frac{\pi}{2}} \ln((\sin\theta)(\cos\theta))\,d\theta \\
&= \int_0^{\frac{\pi}{2}} \ln\left(\tfrac{1}{2}\sin(2\theta)\right)\,d\theta \\
&= \int_0^{\frac{\pi}{2}} \ln(\sin(2\theta)) - \ln 2\,d\theta \\
&= -\frac{\pi}{2}\log 2 + \int_0^{\frac{\pi}{2}} \ln(\sin(2\theta)), d\theta \\
&= -\frac{\pi}{2}\log 2 + \frac{1}{2}\int_0^{\pi} \ln(\sin u), du \\
&= -\frac{\pi}{2}\log 2 + \int_0^{\frac{\pi}{2}} \ln(\sin u), du \\
&= -\frac{\pi}{2}\log 2 + \int_0^{\frac{\pi}{2}} \ln(\sin\theta), d\theta.
\end{aligned}$$

Now solve for the integral, and you have (2), and are done. (Does this bit of trickery make you smile, as it does me?)

I liked this problem a lot, since it offered lots of possible stopping points, and lots of possible approaches. It seems to me, as well, that the analytic methods used in the solution offer a lot of interesting lessons. Finally, notice that there are some wonderful ideas here for reading over the break. I still remember fondly the Christmas holiday I spent as an undergraduate reading Knopp's *Theory of Functions*. It's just a suggestion.