A DIRECT PROOF OF STIRLING'S FORMULA

WILLIAM FELLER, Princeton University

1. Introduction. The main excuse for returning once more to the much belabored subject of Stirling's formula is not the simplicity and directness of the following approach, but the circumstance that it provides good illustrations for several topics and techniques of importance. It seems to me that for teaching purposes the discussion is of considerably greater interest than Stirling's formula as such.

The following derivation is self-contained and conceptual. The only formula of calculus used explicitly is

$$I(x) = \int_0^x \log y \, dy = x \log x - x, \tag{1.1}$$

and this can be avoided by expressing the exponent in (2.1) in terms of $I(n)$. Wallis' formula is not used, but the argument leads without artifice to an integral involving the function ϕ defined by

$$\phi(t) = t \prod_{k=1}^{\infty} \left(1 - \frac{t^2}{k^2}\right). \tag{1.2}$$

It is a lucky circumstance that this is the well-known product expansion for the sine:

$$\phi(t) = \frac{\sin \pi t}{\pi} \cdot \tag{1.3}$$

This classical identity is usually proved by relatively deep complex variable methods, but in Section 3 it is shown how the identity can be made plausible, and then proved, by simple elementary methods. The proof merely paraphrases the elegant argument used by E. Artin in his lectures to derive the partial fraction expansion for the cotangents. The whole note is in the spirit of E. Artin to whose cherished memory it is dedicated.

2. Stirling's formula states that

$$n! \sim \sqrt{2\pi}\, n^{n+\frac{1}{2}} e^{-n}. \tag{2.1}$$

The sign $\sim$ indicates that the ratio of the two sides tends to unity. To see how this relation arises naturally we start from the beginning: The problem is to find some estimate for $\log n!$

It is natural to interpret the sum $\log 1 + \cdots + \log n$ as an integral of the step function L which equals $\log k$ in the unit interval centered at k. Our problem is then to appraise the integral of the difference $L(x) - \log x$. In order to deal only with positive integrands we consider separately the several intervals of length $\frac{1}{2}$ in which this difference has a constant sign. Accordingly we put

$$a_k = \tfrac{1}{2}\log k - \int_{k-\frac{1}{2}}^{k} \log x\, dx = \int_{k-\frac{1}{2}}^{k} \log (k/x) dx \tag{2.2}$$

$$b_k = \int_{k}^{k+\frac{1}{2}} \log x\, dx - \tfrac{1}{2}\log k = \int_{k}^{k+\frac{1}{2}} \log (x/k) dx. \tag{2.3}$$

With the abreviation (1.1) we have then

$$a_1 - b_1 + a_2 - b_2 +- \cdots + a_n = \log n! - \tfrac{1}{2}\log n + I(n) - I(\tfrac{1}{2}). \tag{2.4}$$

But

$$a_k = \int_0^{\frac{1}{2}} \log \frac{1}{1-(t/k)}\, dt, \qquad b_k = \int_0^{\frac{1}{2}} \log\left(1 + \frac{t}{k}\right) dt, \tag{2.5}$$

and it is seen trivially that $a_k > b_k > a_{k+1} > 0$. The left side in (2.4) is therefore a partial sum of an alternating series $\Sigma(-1)^k c_k$ with c_k decreasing monotonically to 0. The series therefore converges to a finite sum S, and (2.4) shows that

$$\log n! - (n + \tfrac{1}{2}) \log n + n \to S - I(\tfrac{1}{2}). \tag{2.6}$$

This is Stirling's formula, except that the coefficient $\sqrt{2\pi}$ is replaced by the constant e^C where

$$C = \sum_{k=1}^{\infty} (a_k - b_k) - I(\tfrac{1}{2}). \tag{2.7}$$

[From this C could be calculated to any number of decimals, and in principle our problem is solved. It is also easy to obtain bounds on the error τ_n defined as the difference between the left side and the right side in (2.6). This τ_n is the remainder of our alternating series and hence $0 < \tau_n < b_n$. Since $\log(1+x) \leqq x$ this implies $0 < \tau_n < 1/8n$. This argument can be sharpened, but it is preferable to use the simple and elegant argument of H. E. Robbins which shows that $1/(12n+1) < \tau_n < 1/12n$. (See this MONTHLY, 62 (1955) 26–29)].

It is just a lucky circumstance that C reduces to a familiar constant. Indeed,

$$a_k - b_k = -\int_0^{\frac{1}{2}} \log\left(1 - \frac{t^2}{k^2}\right) dt \tag{2.8}$$

and so (2.7) expresses C as an integral of $-\log \phi(t)$, with ϕ defined in (1.2). In order not to interrupt the argument let us assume (1.3) as known. Then

$$C = -\int_0^{\frac{1}{2}} \log \frac{\sin \pi t}{\pi}\, dt = \frac{1}{2}\log \pi - \int_0^{\frac{1}{2}} \log \sin \pi t\, dt. \tag{2.9}$$

The last integral can be evaluated by considering the familiar symmetries of the sine and cosine. Obviously the value of the integral remains unchanged if $\sin \pi t$ is replaced by $\cos \pi t$, and hence, by the double angle formula

$$\int_0^{\frac{1}{2}} \sin 2\pi t\, dt = \tfrac{1}{2}\log 2 + 2\int_0^{\frac{1}{2}} \sin \pi t\, dt. \tag{2.10}$$

Again, the two integrals in (2.10) are trivially equal, and hence $= -\frac{1}{2}\log 2$. Accordingly $C = \frac{1}{2}\log 2\pi$, as asserted.

3. Proof of the identity (1.3). When the product for ϕ is decomposed into linear factors one notices immediately that $\phi(x+1) = -\phi(x)$. Thus ϕ is an odd periodic function having the same period and same zeros as $\pi^{-1}\sin \pi x$. One may therefore suspect that the two functions are the same. This suspicion becomes virtual certainty when one notices that in the product for $\phi(2x)$ the factor corresponding to $k = 2\nu$ coincides with the νth factor in (1.2), and so

$$\phi(2x) = 2\phi(x)\psi(x) \tag{3.1}$$

with

$$\psi(x) = \prod_{\nu=1}^{\infty}\left(1 - \frac{x^2}{(\nu - \frac{1}{2})^2}\right). \tag{3.2}$$

The decomposition (3.1) resembles the double angle formula for $\sin \pi x$ and leads to the suspicion that $\psi(x) = \cos \pi x$. Now it is easily verified that

$$\psi(x) = \frac{\phi(x + \frac{1}{2})}{\phi(\frac{1}{2})} \tag{3.3}$$

and this is exactly the relation between $\cos \pi x$ and $\pi^{-1}\sin \pi x$. Thus (3.1) reduces to the form of the double angle formula.

After these preparations it is easy to prove the identity (1.3) using Artin's elegant argument. Put

$$f(x) = \log \frac{\pi\phi(x)}{\sin \pi x} \cdot \tag{3.4}$$

The fraction on the right is defined for nonintegral x, but it is easily seen that f is continuous at the origin if we put $f(0) = 0$. Thus f becomes a continuous periodic function. The double angle formulas for ϕ and $\pi^{-1}\sin \pi x$ imply that

$$f(2x) = \frac{f(x) + f(x + \frac{1}{2})}{2} \cdot \tag{3.5}$$

Let m be the maximum of f. It follows from (3.5) that $f(2x) = m$ implies $f(x) = m$, and hence the maximum is assumed at a sequence of points converging to 0. But $f(0) = 0$, and hence f vanishes identically.

This paper is connected with research sponsored by the U. S. Army Research Office (Durham) at Princeton University.