**Stirling formula for $n!$:**

it says that $n!$ is like $\sqrt{2\pi}e^{-n}n^{n+1/2}$.

This is to be interpreted in the folioing sense. Their ratio converges to one. When we say that a sequence $(a_n)$ is like another sequence $(b_n)$ (both are sequences of strictly positive numbers) there are two ways of understanding. Either $a_n - b_n \to 0$ or $a_n/b_n \to 1$ Of course when the first happens, then the second also happens. However the other way is not in general true.

For example $(n)$ is like $(n + 1/n)$ in the first sense. the sequence $(n^2)$ is like $(n^2 + n)$ in the second sense. they are not so in the first sense. In fact the difference between the two sequences is $n$ which becomes larger and larger. But then in what sense are they like each other? well, Both numbers are becoming large, when you replace one by the other, the relative error is going to zero.

If you are measuring length of this room, if you are off by a mile then the error is indeed very huge. On the other hand if you are measuring distance (of earth) to sun, if you are off by a mile or even hundred miles, the error is very very small. So the absolute error is many times unimportant and it is the relative error that matters.

We shall try to approximate the area of the function $f(x) = \log x$ from $x = 1$ to $x = n$ Let us first make a few observations which depend on the fact that

$$f'' = -1/x^2 \le 0.$$

Let $g$ be a twice differentiable function on an interval $(a, b)$ with $g'' \le 0$.

Consider any two points $u < v$ in the interval $(a, b)$. We claim that the chord (or secant) joining the two points $(u, f(u))$ and $(v, f(v))$ lies below the graph of $f$.

To see this, first observe that the equation of the chord is

$$y - f(u) = \frac{f(v) - f(u)}{v - u}(x - u).$$

Consider any point $w \in [u, v]$. We need to show

$$f(u) + \frac{f(v) - f(u)}{v - u}(w - u) \leq f(w).$$

That is,

$$f(u) - f(w) + \frac{f(v) - f(u)}{v - u}(w - u) \leq 0.$$

Or

$$[f(u) - f(w)](v - u) + [f(v) - f(u)](w - u) \leq 0.$$

$$[f(u) - f(w)](v - u) + [f(v) - f(w) + f(w) - f(u)](w - u) \leq 0.$$

$$[f(v) - f(w)](w - u) - [f(w) - f(u)](v - w) \leq 0.$$

By MVT, there are points $\theta \in (u, w)$ and $\eta \in (w, v)$ such that $f(w) - f(u) = f'(\theta)(w - u)$ and $f(v) - f(w) = f'(\eta)(v - w)$. So we need to show

$$f'(\eta)(v - w)(w - u) - f'(\theta)(v - w)(w - u) \leq 0.$$

That is,

$$[f'(\eta) - f'(\theta)](v - w)(w - u) \leq 0.$$

Since $\theta < \eta$ and $f'' \leq 0$ we conclude that the first factor above is negative. Since $u, w < v$, the other two factors are positive and hence the inequality is true.

Consider any point $u$ in the interval $(a, b)$. We claim that the tangent (to the graph of $f$) at $u$ lies above the graph.

The equation of the tangent is

$$y - f(u) = f'(u)(x - u).$$

Let us take any other point $w \in (a, b)$. We need to show

$$f(w) \leq f(u) + f'(u)(w - u).$$

That is,

$$\frac{f(w) - f(u)}{w - u} \leq f'(u).$$

But the left side is $f'(\theta)$ for some $u < \theta < w$ and since $f'$ is decreasing, the inequality is verified.

Thus, for $k \geq 1$, the area under the curve $y = \log x$ is in between the area under the chord joining $(k, \log k)$, $(k+1, \log(k+1))$ and area under the tangent at $x + 1/2$. Thus

$$\frac{1}{2}\log(k+1) + \frac{1}{2}\log k \leq \int_k^{k+1} \log x \; dx \leq \log(k + 1/2).$$

Adding these for $k = 1, 2, \cdots, n-1$ and remembering that $x \log x - x$ is a primitive for $\log x$ we get

$$\log(n!) - \frac{1}{2}\log n \leq n \log n - n + 1 \leq \sum_1^{n-1} \log(k + 1/2).$$

Let

$$a_n = n \log n - n + 1 - [\log(n!) - \frac{1}{2}\log n] = \log\left\{\frac{e^{-n} n^{n+1/2}}{n!}\right\} + 1.$$

Then $a_n$ is the area between the curve $y = \log x$ and the chords explained above, from $x = 1$ to $x = n$ Thus we see

$$a_n \geq 0; \qquad a_n \uparrow . \qquad\qquad (\spadesuit)$$

Also

$$\begin{aligned} a_n &\leq \sum_1^{n-1} \{\log(k + 1/2) - \frac{1}{2}\log(k+1) - \frac{1}{2}\log k\}. \\ &= \frac{1}{2}\sum_1^{n-1}\left\{\log\frac{(k+1/2)}{k} - \log\frac{(k+1)}{k+1/2}\right\} \\ &\leq \frac{1}{2}\sum_1^{n-1}\left\{\log(1 + \frac{1}{2k}) - \log(1 + \frac{1}{2(k+1/2)})\right\} \\ &\leq \frac{1}{2}\sum_1^{n-1}\left\{\log(1 + \frac{1}{2k}) - \log(1 + \frac{1}{2(k+1)})\right\} \\ &= \frac{1}{2}\log\frac{3}{2} - \frac{1}{2}\log(1 + \frac{1}{2n}). \end{aligned}$$

As a consequence $a_n$ is bounded above . So $(\spadesuit)$ implies that $a_n$ converges. Say $a_n \uparrow c$

$$\log\left\{\frac{e^{-n} n^{n+1/2}}{n!}\right\} = a_n - 1 \uparrow c - 1.$$

Or

$$\log\left\{\frac{n!}{e^{-n} n^{n+1/2}}\right\} \to k$$

for some constant $k$. Or

$$\frac{n!}{ke^{-n}n^{n+1/2}} \to 1.$$

We shall now evaluate the constant $k$ by using the above limit in a known case, namely, Walli's product. We know

$$\frac{2^{2n}(n!)^2}{(2n)!\sqrt{n}} \to \sqrt{\pi}.$$

Suppose that we have strictly positive numbers $a_n$ and $b_n$ and $a_n/b_n \to 1$. If $\alpha_n \times a_n \to c$ then $\alpha_n \times b_n \to c$. This is because

$$\alpha_n \times b_n = \alpha_n \times a_n \times \frac{b_n}{a_n} \to c \times 1.$$

Similarly if $\alpha_n/a_n \to c$ then $\alpha_n/b_n \to c$. In other words we can replace $a_n$ by $b_n$. As a consequence the above result of Walli implies

$$\frac{2^{2n}k^2e^{-2n}n^{2n+1}}{ke^{-2n}(2n)^{2n+1/2}\sqrt{n}} \to \sqrt{\pi}.$$

That is,

$$k = \sqrt{2\pi}.$$

Thus

$$\frac{n!}{\sqrt{2\pi}e^{-n}n^{n+1/2}} \to 1.$$

**Improper integrals:**

We have discussed integrals of bounded functions over bounded intervals, both open as well as closed. Integrals where either the function is unbounded or the interval is unbounded are called improper integrals. There is nothing improper about them, just that you can not use Riemann sums blindly. Just as infinite sums are defined as limits of finite (partial) sums, so are these integrals. We shall first discuss unbounded functions over bounded interval.

Of course there are several possibilities. For example, you can consider the function

$$f(x) = \frac{1}{\sqrt{x}} + \frac{1}{\sqrt{x-1}} + \frac{1}{\sqrt{x-2}}; \quad x \in (0,3), x \neq 1; \quad f(1) = 0.$$

Then $f$ is unbounded near three points, namely, zero, one and two. We start with simple situations.

Let us consider function $f$ on an open interval $(a, b)$ which is unbounded only near $a$, that is for any $\epsilon > 0$, it is bounded on $(a + \epsilon, b)$ Then for every $\epsilon > 0$, the integral $\int\limits_{a+\epsilon}^{b} f$ makes sense and is defined by earlier considerations. We say that $\int\limits_{a}^{b} f$ exists if

$$\lim_{\epsilon \downarrow 0} \int_{a+\epsilon}^{b} f$$

exists and then we define this limit as $\int\limits_{a}^{b} f$. You must recall the meaning of limit $\epsilon \downarrow 0$.

Thus $\int\limits_{a}^{b} f$ exists if there is a number $\alpha$ such that for every sequence $a_n \downarrow a$, $a_n > a$ for all $n$ we have

$$\int_{a_n}^{b} f \to \alpha.$$

In that case this number $\alpha$ is defined to be the value of the integral of $f$ over the interval $(a, b)$.

For example consider

$$\int_{a}^{b} \frac{1}{x} dx.$$

This does not exist because for integral over $(\epsilon, 1)$, you get $\log \epsilon$ and does not converge to a finite limit as $\epsilon \downarrow$. Actually the same situation occurs and the integral

$$\int_{0}^{1} \frac{1}{x^{\alpha}} dx$$

does not exist for $\alpha \geq 1$. On the other hand the integral exists for $\alpha < 1$. Of course if $\alpha \leq 0$, then the function is bounded and actually continuous on the closed interval $[0, 1]$ and we need not discuss. So let $0 < \alpha < 1$.

$$\int_{\epsilon}^{1} \frac{1}{x^{\alpha}} dx = \frac{1}{1-\alpha} \left(1 - \epsilon^{-\alpha+1}\right) \to \frac{1}{1-\alpha},$$

as $\epsilon \downarrow 0$.

In a similar way, if the function $f$ defined on an open interval $(a, b)$ is unbounded only near $b$, that is for any $\epsilon > 0$, it is bounded on $(a, b - \epsilon)$

then for every $\epsilon > 0$, the integral $\int\limits_a^{b-\epsilon} f$ makes sense and is defined by earlier considerations. We say that $\int\limits_a^b f$ exists if

$$\lim_{\epsilon \downarrow 0} \int_a^{b-\epsilon} f$$

exists and then we define this limit as $\int\limits_a^b f$. You must recall the meaning of limit $\epsilon \downarrow 0$.

We can, in a fashion analogous to the above, show that the integral

$$\int_a^b \frac{1}{(b-x)^\alpha} dx$$

exists if $\alpha < 0$ and does not exist if $\alpha \geq 1$.

The substitution rule and integration by parts hold good for these integrals. For example suppose $\varphi$ is a continuously differentiable strictly increasing function on a bounded interval $(a, b)$ onto the bounded interval $(c, d)$. Suppose that $f$ is a function integrable on $(a, b)$, possibly unbounded. Then integral of $f(\varphi(x))\varphi'(x)$ exists over $(a, b)$ and we have

$$\int_a^b f(\varphi(x))\varphi'(x) dx = \int_c^d f(y) dy.$$

Suppose that $f$ is unbounded only near $a$, then verify that the equality holds when you consider $a+\epsilon$ to $b$ on left side and from $\varphi(a+\epsilon)$ to $d$ on the right side and then take limits. Similar remark applies if the infinity occurs only near $b$.

Similar argument holds for integration by parts, do it over appropriate interval and take limits.

Let us denote by $\pi$ the area of the region enclosed by circle of radius one. Thus $\pi/4$ is the area of the quarter circle. More precisely,

$$\int_0^1 \sqrt{1-u^2} du = \frac{\pi}{4}.$$

Note that the integrand on the left side is precisely the function describing the quarter circle in the first quadrant. We now show, without using trigonometric functions,

$$\int_0^1 \frac{1}{\sqrt{1-u^2}} du = \frac{\pi}{2}, \qquad (\clubsuit)$$

To prove the claim regarding the improper integral first do integrate by parts to see that for $0 < a < 1$

$$\int_0^a \sqrt{1-u^2}\ \ du = a\sqrt{1-a^2} - \int_0^a u\left[\sqrt{1-u^2}\right]' du$$

$$= a\sqrt{1-a^2} + \int_0^a \frac{u^2}{\sqrt{1-u^2}} du$$

$$= a\sqrt{1-a^2} + \int_0^a \frac{1}{\sqrt{1-u^2}} du - \int_0^a \sqrt{1-u^2} du.$$

Thus

$$\int_0^a \frac{1}{\sqrt{1-u^2}} du = a\sqrt{1-a^2} + 2\int_0^a \sqrt{1-u^2} du.$$

Now taking limit as $a \uparrow 1$ we get ($\clubsuit$).

**Final look at trigonometric functions:**

Let us recall the functions

$$f(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots\cdots;$$

$$g(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots\cdots;$$

We have seen, as a consequence of the fundamental theorem on power series, the following.

($1^0$). $f$, $g$ are defined on all of $R$, differentiable,

$$f' = g; \quad g' = -f; \quad f(0) = 0; \quad g(0) = 1.$$

This implies $(f^2 + g^2)' = ff' + gg' = 0$. Hence $f^2 + g^2$ must be a constant function; since at zero its value is one, we conclude

(2). $f^2 + g^2 \equiv 1$.

As a consequence we have the following.

(3). $-1 \leq f \leq 1; \quad -1 \leq g \leq 1.$ and also that when one of the functions assumes the value zero, then the other must take value $\pm 1$.

We now show the following.

(4). $g(x) = 0$ for some $x > 0$.

Since $g(0) = 1$, and $g$ is continuous, we can fix a $\delta > 0$ so that $g(x) > 0$ for all $x \in [0, \delta]$. Since $f(0) = 0$ and its derivative is strictly positive in $[0, \delta]$ we conclude that $f$ is strictly increasing in this interval and $f(\delta) = c > 0$. We claim that $g$ must assume value zero in the interval $[\delta, \delta + \frac{2}{c}$ Otherwise, $g(\delta) > 0$ implies that $g$ is positive through this interval. But then $f$ must be strictly increasing and hence $f(x) > c$ throughout this interval. But then by MVT, there is a $\theta$ in this interval such that

$$\left|g(\delta + \frac{2}{c}) - g(\delta)\right| = \frac{2}{c}|f(\theta)| > 2.$$

But from (3) we have

$$\left|g(\delta + \frac{2}{c}) - g(\delta)\right| \leq 2.$$

This contradiction proves the claim.

(5). Let $\alpha = \min\{x > 0 : g(x) = 0\}$. Then $\alpha > 0$.

In other words, $\alpha$ is the smallest positive zero of $g$. Indeed from (4), we see that the set on right side is non-empty. By continuity of $g$ it follows that $g(\alpha) = 0$, recall that there is a sequence of points in the set decreasing to its infimum. Hence $\alpha$ is again in the set and hence it is minimum. Since $g(0) = 1$ we see that $\alpha > 0$.

(6). $f(\alpha) = +1$

In fact as noticed in (3), $f(\alpha) = \pm 1$. But $f$ is strictly increasing through out in $[0, \alpha]$ with $f(0) = 0$. Hence $f(\alpha) == 1$.

(7). $\alpha = \pi/2$ where $\pi$ is the area enclosed by the unit circle.

As noted above $f$ is strictly increasing differentiable function on $(0, \alpha)$ onto $(0, 1)$. If $\varphi(y) = 1/\sqrt{1 - y^2}$ for $0 < y < 1$, then as discussed in the improper integrals,

$$\int_0^1 \varphi(y) dy = \pi/2.$$

But by substitution rule this integral equals

$$\int_0^\alpha \varphi(f(x)) f'(x) dx = \int_0^\alpha g(x)/\sqrt{1 - f^2(x)} = \int_0^\alpha 1 dx = \alpha.$$

Thus $\alpha = \pi/2$.

(8) The functions $\varphi = f$ and $\psi = g$ are the only solutions of

$$\varphi' = \psi; \quad \psi' = -\varphi; \quad \varphi(0) = 0; \quad \psi(0) = 1.$$

from (1), we know that $f$ and $g$ are indeed solutions. If there is another pair; $\varphi = \tilde{f}$ and $\psi = \tilde{g}$, then the functions

$$\eta = f - \tilde{f}; \qquad \zeta = g - \tilde{g}$$

satisfy

$$\eta' = \zeta; \quad \zeta' = -\eta; \quad \eta(0) = 0 = \zeta(0).$$

But then the argument leading to (2) applies to show that $\eta^2 + \zeta^2 \equiv 0$ showing that $f = \tilde{f}$ and $g = \tilde{g}$.

(9). $f(x + \alpha) = g(x); \qquad g(x + \alpha) = -f(x).$

Define the functions $f_1(x) = -g(x + \alpha)$ and $g_1(x) = f(x + \alpha)$. Then by definition of $\alpha$ we see $f_1(0) = 0$ and by (6) we see $g(0) = 1$. Moreover, by chain rule fro derivatives we see

$$f_1'(x) = -g'(x + \alpha) = -[-f(x + \alpha)] = f(x + \alpha) = g_1(x).$$

similarly, $g_1'(x) = -f_1(x)$.. But then (8) shows that these functions $f_1$ and $g_1$ are same as the functions $f$ and $g$ proving the claim.

(10). $f(x + 2\alpha) = -f(x); \qquad g(x + 2\alpha) = -g(x).$

Repeated application of (9) proves this. Again a repeated application of this proves

(11). $f(x + 4\alpha) = f(x); \qquad g(x + 4\alpha) = g(x).$

Thus the functions $f$ and $g$ are periodic of period $2\pi$. We have also seen as an application of the Cauchy product of power series the following.

(12). $f(x + y) = f(x)g(y) + g(x)f(y); \quad g(x + y) = g(x)g(y) - f(x)f(y).$

This concludes our discussion of the trigonometric functions. This identifies the functions we introduced as power series with the sine and cosine functions that you learnt in high school. Of course, the uniqueness result (8) is enough for such an identification. But this discussion also clarifies the role of $\pi$.

**return to improper integrals:**

We have discussed the definition of $\int\limits_a^b f(x)dx$ if $f$ is unbounded at one of the points $a$ or $b$. What if $f$ is unbounded at both ends. For example

$$\int_0^1 \frac{1}{\sqrt{x}\sqrt{1-x}}dx$$

Suppose that $f$ is unbounded at both end points, but is bounded otherwise, that is if $a < u < v < b$ then $f$ is bounded on the interval $[u, v]$. We say that the integral $\int\limits_a^b f$ exists if there is a number $\alpha$ such that

$$\int_{a_n}^{b_n} f \to \alpha$$

whenever we have $a < a_n < b_n < b$ and $a_n \to a$ and $b_n \to b$. In that case we write

$$\int_a^b f(x)dx = \alpha.$$

Note that this is not same as saying that the limit

$$\lim_{\epsilon \to 0, \epsilon > 0} \int_{a+\epsilon}^{b-\epsilon} f$$

exists.

Let us consider the case of unbounded interval. Suppose that $f$ is a function defined on the interval $(0, \infty)$ which is bounded on every interval $(0, a)$. We say that the integral

$$\int_0^\infty f$$

exists if the limit

$$\lim_{a_n \to \infty} \int_0^{a_n} f$$

exists and then we define this limit as the value of the integral.

for example

$$\int_0^\infty e^{-x}dx$$

exists and equals one.

$$\int_0^\infty x^n e^{-x} dx$$

exists and equals $n!$ if $n$ is a positive integer. you prove it buy induction on $n$ and integration by parts over $(0, a)$. You need to use $a^k e^{-a} \to 0$ as $a \to \infty$.