Math 151C (Advanced Calculus)
Solution to Rudin 8.20

Problem 1. The following simple computation yields a good approximation to Stirling's formula.

For $m = 1, 2, 3, \ldots$, define

$$f(x) = (m + 1 - x)\log(m) + (x - m)\log(m + 1)$$

if $m \leq x \leq m + 1$, and define

$$g(x) = \frac{x}{m} - 1 + \log(m)$$

if $m - \frac{1}{2} \leq x < m + \frac{1}{2}$. Draw the graphs of f and g. Note that $f(x) \leq \log(x) \leq g(x)$ if $x \geq 1$ and that

$$\int_1^n f(x)\,\mathrm{d}x = \log(n!) - \frac{1}{2}\log(n) > -\frac{1}{8} + \int_1^n g(x)\,\mathrm{d}x.$$

Integrate $\log(x)$ over $[1, n]$. Conclude that

$$\frac{7}{8} < \log(n!) - \left(n + \frac{1}{2}\right)\log(n) + n < 1$$

for $n = 2, 3, 4, \ldots$ (*Note*: $\log(2\pi) \approx 0.918\ldots$). Thus

$$\mathrm{e}^{7/8} < \frac{n!}{(n/\mathrm{e})^n\sqrt{n}} < \mathrm{e}.$$

While grading the most recent homework, I noticed that many of you had difficulty with problem 20 in chapter 8. The argument presented in that problem is quite opaque—it is not at all clear what the intuition behind the estimate should be, nor how it came about. So, instead of relying on intuition to guide computation, many of you just turned the crank on the integrals and obtained a result. Instead of simply following, let's see if we can build some intuition.

The fundamental idea in this exercise is that the logarithm is concave. This means that any line tangent to the graph of $\log(x)$ will remain entirely above the graph, and any chord will remain entirely below the graph. See Figure 1. As such, the logarithm may be estimated from above and below, from which one may obtain a Stirling-like formula which gives a reasonable approximation to the factorial function in terms of exponentials.

The first step in this process is to verify the claims that tangent lines lie above the graph of $\log(x)$, and that chords lie below. This is the content of the following two lemmata:

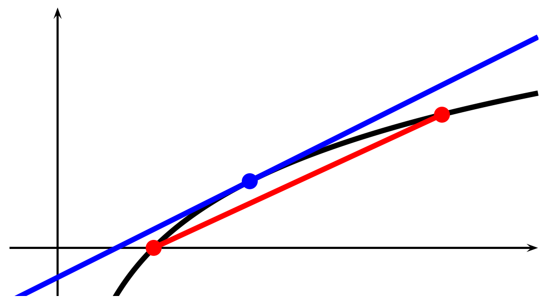

Figure 1: The graph of $\log(x)$ is shown in black. The blue line is tangent to the graph of $\log(x)$ at the blue point, while the red segment is a chord of the graph.

Lemma 1. *Fix some $a \in (0, \infty)$ and define*

$$\ell(x) = \frac{1}{a}(x - a) + \log(a). \tag{1}$$

Then

$$\log(x) \leq \ell(x)$$

for all $x \in (0, \infty)$.

Remark 2. Note that the graph of ℓ is a line which is tangent to the graph of $\log(x)$ at the point $(a, \log(a))$. Hence this lemma could be stated simply as "tangent lines lie above the graph of log."

Proof. The goal is to show that $0 \leq \ell(x) - \log(x)$ for all $x \in (0, \infty)$. Observe that

$$\ell(x) - \log(x) = \frac{1}{a}(x - a) + \log(a) - \log(x) = \left(\frac{x}{a} - 1\right) - \log\left(\frac{x}{a}\right).$$

Making the change of variables $\xi = x/a$, it is sufficient to show that

$$0 \leq (\xi - 1) - \log(\xi) \tag{2}$$

for all $\xi \in (0, \infty)$. Recall that, by defintion,

$$\log(\xi) = \int_1^\xi \frac{1}{t}\, \mathrm{d}t.$$

This next bit is non-obvious. The intuition is that the logarithm represents an area, and we may only add measurements which have like units. Hence $\xi - 1$ must also represent an area. The simplest area which could be represented by this expression is a rectangle which is 1 unit tall and $\xi - 1$ units wide. But this is precisely the area under the graph of $f(x) = 1$ on the interval $[1, \xi - 1]$. Thus

$$\xi - 1 = \int_1^\xi \mathrm{d}t.$$

Therefore

$$\xi - 1 - \log(\xi) = \int_1^\xi \mathrm{d}t + \int_1^\xi \frac{1}{t}\,\mathrm{d}\xi = \int_1^\xi 1 - \frac{1}{t}\,\mathrm{d}t.$$

If $\xi > 1$, then the integrand is positive, and so the integral is positive. If $\xi < 1$, then

$$\int_1^\xi 1 - \frac{1}{t}\,\mathrm{d}t = \int_\xi^1 \frac{1}{t} - 1\,\mathrm{d}t.$$

Again, the integrand on the right-hand side is positive, so the integral is positive. Equality occurs when $\xi = 1$. It therefore follows that

$$0 < (\xi - 1) - \log(\xi)$$

for any $\xi \in (0, \infty)$, from which the desired result follows. □

Remark 3. The inequality at (2) is a well-known estimate, which is (essentially) Bernoulli's inequality in disguise. This is something worth Googling.

Lemma 4. *The logarithm function is concave, in the sense that*

$$\log(\lambda a + (1-\lambda)b) \geq \lambda \log(a) + (1-\lambda)\log(b) \tag{3}$$

for any $a, b \in (0, \infty)$ *and* $\lambda \in [0, 1]$.

Remark 5. The intuition here is that if $c \in (a, b)$, then there exists some $\lambda \in (0, 1)$ such that $c = \lambda a + (1-\lambda)b$. That is, we may think of c as a weighted average of a and b. Hence the left-hand side of (3) is the value of the logarithm at c.

To understand the right-hand side, recall a line through two points (x_1, y_1) and (x_2, y_2) may be parameterized by

$$\lambda\langle x_2 - x_1, y_2 - y_1\rangle + (x_1, y_1) = (\lambda x_2 + (1-\lambda)x_1, \lambda y_2 + (1-\lambda)y_1),$$

where λ can range over the reals. Taking

$$(x_1, y_1) = (a, \log(a)) \qquad \text{and} \qquad (x_2, y_2) = (b, \log(b)),$$

this gives the line through two points on the graph of log. In particular, the points correspond to the values $\lambda = 0$ and $\lambda = 1$. The chord is then the collection of points parameterized by $\lambda \in [0, 1]$. Thus the right-hand side of 3 corresponds to a parameterization of chord—specifically, it gives the y-coordinate of the point on the chord with x-coordinate $c = \lambda a + (1-\lambda)b$.

Combining these two ideas, the inequality states that if $c \in [a, b]$, then then point on the graph of log with x-coordinate c lies above the point on the chord with the same x-coordinate. In other words, the lemma may be simply stated as "a chord lies below the graph of log."

Proof. As log is continuous, it is sufficient to show that the inequality holds for $\lambda = 1/2$. If this can be done, then an induction argument shows that the result holds for any dyadic rational $\lambda = k/2^n$ natural numbers $0 \leq k \leq n$. Any real in $[0, 1]$ can then be approximated by a sequence of dyadic rationals. If you care to work out a more complete explanation of this reduction, examine Problem 4.24 in Rudin.

With $\lambda = \frac{1}{2}$, the inequality at (3) becomes

$$\log\left(\frac{a+b}{2}\right) \geq \frac{1}{2}\log(ab).$$

By the monotonicity of the logarithm, this holds for positive a and b if and only if

$$\frac{a+b}{2} \geq \sqrt{ab}.$$

But

$$\frac{a+b}{2} = \sqrt{\left(\frac{a+b}{2}\right)^2} = \sqrt{\frac{a^2+b^2}{4} + ab} \geq \sqrt{ab},$$

where the inequality follows from the monotonicity of the square root function, which is the required inequality. □

Thus far, we have established that tangents lie above the graph of the logarithm, and chords lie below. This puts us in a position to understand the construction of the functions f and g which Rudin provides. Both functions are piecewise linear, and defined so that they are continuous on intervals with integer endpoints. I think that this can be made a bit more clear by working as follows:

Fix some natural number m (for the sake of this document, the natural numbers do not include zero). On the interval $[m - \frac{1}{2}, m + \frac{1}{2})$, take g to be the function whose graph is a line tangent to the graph of $\log(x)$ at the midpoint of the interval, i.e. at the point m. Taking $a = m$ in (1) gives

$$g(x) = \frac{1}{m}(x - m) + \log(m) = \frac{x}{m} - 1 + \log(m).$$

This is precisely the definition of g given by Rudin. It now follows immediately from Lemma 1 that

$$\log(x) \leq g(x)$$

on the interval $[m - \frac{1}{2}, m + \frac{1}{2})$, which further implies that the inequality holds for all $x \geq 1$.

Continuing to work on the interval $[m, m+1)$, define f so that the graph of f is a line which intersects the graph of log at the endpoints of the interval. That is,

$$f(m) = \log(m), \qquad f(m+1) = \log(m+1),$$

and f is a linear function (in the sense that $f(x) = \alpha x + \beta$ for appropriately chosen real constants α and β). Taking $a = m$, $b = m + 1$, and $\lambda = m + 1 - x$ in (3), Remark 5 implies that

$$f(x) = (m+1-x)\log(m)+(1-(m+1-x))\log(m+1) = (m+1-x)\log(m)+(x-m)\log(m+1).$$

It follows immediately from Lemma 4 that

$$f(x) \leq \log(x)$$

on the interval $[m, m + 1)$, from which it further implies that the inequality holds for all $x \geq 1$.

Having established the inequalities suggested by Rudin, the next step is to integrate. This process is made much easier by noting that for each natural number m, the integrals

$$\int_{m}^{m+1} f(x)\,\mathrm{d}x \qquad \text{and} \qquad \int_{m-\frac{1}{2}}^{m+\frac{1}{2}} g(x)\,\mathrm{d}x$$

correspond to the areas of trapezoids. The area of a trapezoid is given by

$$\text{Area} = \frac{1}{2}(b_1 + b_2)h,$$

where b_1 and b_2 are the lengths of the parallel sides, and h is the distance between these two sides (the height of the trapezoid). For those of you that are uncomfortable with this geometric interpretation of the integrals, we can prove it as a lemma:

Lemma 6. *Suppose that*

$$\ell(x) = \alpha x + \beta$$

Fix $a \in \mathbb{R}$ and $h > 0$, and define

$$b_1 = \ell(a) \qquad \textit{and} \qquad b_2 = \ell(a + h).$$

Then

$$\int_{a}^{a+h} \ell(x)\,\mathrm{d}x = \frac{1}{2}(b_1 + b_2)h.$$

That is, the integral of a linear function over an interval is the are of a trapezoid.

Proof. This is a routine computation: first, observe that

$$\frac{1}{2}(b_1 + b_2)h = \frac{1}{2}((\alpha a + \beta) + (\alpha(a + h) + \beta))h = \frac{\alpha}{2}\left[(2a + h) + \beta\right] h.$$

Then, integrating,

$$\int_{a}^{a+h} \ell(x)\,\mathrm{d}x = \int_{a}^{a+h} \alpha x + \beta\,\mathrm{d}x$$

$$
\begin{aligned}
&= \left[\frac{\alpha}{2}x^2 + \beta x\right|_{x=a}^{a+h} \\
&= \frac{\alpha}{2}\left((a+h)^2 + \beta(a+h) - a^2 - \beta a\right) \\
&= \frac{\alpha}{2}\left[(2a+h) + \beta\right] h.
\end{aligned}
$$

The two displayed quantities agree, which completes the proof. □

In any case, using this "trapezoid integral rule",

$$\int_{m-\frac{1}{2}}^{m+\frac{1}{2}} g(x)\,\mathrm{d}x = \frac{1}{2}\left(g(m-\tfrac{1}{2}) + g(m+\tfrac{1}{2})\right)$$

for any $m \in \mathbb{N}$. Expanding this using the defintion of g gives

$$
\begin{aligned}
\frac{1}{2}\left(g(m-\tfrac{1}{2}) + g(m+\tfrac{1}{2})\right) &= \frac{1}{2}\left(\left(\frac{m-\frac{1}{2}}{m} - 1 + \log(m)\right) + \left(\frac{m+\frac{1}{2}}{m} - 1 + \log(m)\right)\right) \\
&= \log(m).
\end{aligned}
$$

As the ultimate goal is to integrate over $[1,n]$, the cases where $m = 1$ and $m = n$ needs to be handled slightly differently:

$$\int_{1}^{1+\frac{1}{2}} g(x)\,\mathrm{d}x = \frac{1}{4}\left((g(1) + g(\tfrac{3}{2})\right) = \frac{1}{4}\left(\left(\frac{1}{1} - 1 + \log(1)\right) + \left(\frac{\frac{3}{2}}{1} - 1 + \log(1)\right)\right) = \frac{1}{8},$$

and

$$\int_{n-\frac{1}{2}}^{n} g(x)\,\mathrm{d}(x) = \frac{1}{4}\left(g(n-\tfrac{1}{2}) + g(n)\right) = -\frac{1}{8n} + \frac{1}{2}\log(n)$$

Therefore

$$
\begin{aligned}
\int_{1}^{n} g(x)\,\mathrm{d}x &= \int_{1}^{1+\frac{1}{2}} g(x)\,\mathrm{d}x + \sum_{m=2}^{n-1}\int_{m-\frac{1}{2}}^{m+\frac{1}{2}} g(x)\,\mathrm{d}x + \int_{n-\frac{1}{2}}^{n} g(x)\,\mathrm{d}(x) \\
&= \frac{1}{8} + \sum_{m=2}^{n-1}\log(m) - \frac{1}{8n} + \frac{1}{2}\log(n) \\
&= \frac{1}{8} + \log(n!) - \frac{1}{8n} - \frac{1}{2}\log(n).
\end{aligned}
$$

By a similar style of reasoning,

$$
\begin{aligned}
\int_{1}^{n} f(x)\,\mathrm{d}x &= \sum_{n=1}^{n-1}\int_{m}^{m+1} f(x)\,\mathrm{d}x \\
&= \frac{1}{2}\sum_{m=1}^{n-1}(f(m) + f(m+1)) \qquad \text{(trapezoid rule)}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{2}\sum_{m=1}^{n-1}(\log(m) + \log(m+1)) \\
&= \frac{1}{2}\sum_{m=1}^{n-1}\log(m) + \frac{1}{2}\sum_{m=2}^{n}\log(n) \\
&= \left[\sum_{m=1}^{n-1}\log(m)\right] + \frac{1}{2}\log(n) && \text{(since } \log(1) = 0\text{)} \\
&= \log(n!) - \frac{1}{2}\log(n). && \text{(add and subtract } \tfrac{1}{2}\log(n)\text{, combine logs)}
\end{aligned}
$$

Summarizing, we now have

$$\int_1^n f(x)\,\mathrm{d}x = \log(n!) - \frac{1}{2}\log(n) \geq \log(n!) - \frac{1}{2}\log(n) - \frac{1}{8n} = -\frac{1}{8} + \int_1^n g(x)\,\mathrm{d}x,$$

as Rudin claims. By the monotonicity of integration (i.e. if f and g are two nonnegative functions with $f \leq g$ everywhere, then $\int f \leq \int g$), we have

$$
\begin{aligned}
&\int_1^n f(x)\,\mathrm{d}x \leq \int_1^n \log(x)\,\mathrm{d}x \leq \int_1^n g(x)\,\mathrm{d}x \leq \frac{1}{8} + \int_1^n f(x)\,\mathrm{d}x \\
&\qquad \implies \log(n!) - \frac{1}{2}\log(n) \leq n\log(n) - n + 1 \leq \frac{1}{8} + \log(n!) - \frac{1}{2}\log(n) \\
&\qquad \implies 0 \leq \left(n + \frac{1}{2}\right)\log(n) - \log(n!) - n + 1 \leq \frac{1}{8} \\
&\qquad \implies 1 \geq -\left(n + \frac{1}{2}\right)\log(n) + \log(n!) + n \geq \frac{7}{8}.
\end{aligned}
$$

Exponentiating the middle term gives

$$\exp\left(-n\log(n) - \frac{1}{2}\log(n) + \log(n) + \log(n!) + n\right) = \frac{1}{n^n}\cdot\frac{1}{\sqrt{n}}\cdot n!\cdot \mathrm{e}^n = \frac{n!}{(n/\mathrm{e})^n\sqrt{n}}.$$

Therefore

$$\mathrm{e}^{7/8} \leq \frac{n!}{(n/\mathrm{e})^n\sqrt{n}} \leq \mathrm{e},$$

as claimed.