Step #1

> Do not laugh at notations; invent them, they are powerful. In fact, mathematics is, to a large extent, invention of better notations.
>
> Richard Feynman

Hello to Big-Oh

If f and g are complex-valued functions, we say "f is big-Oh of g", and write $f = O(g)$, to mean that there is a constant $C \geq 0$ such that $|f| \leq C|g|$ for all indicated (or implied) values of the variables. We refer to C as the "implied constant". For instance,

$$x = O(x^2) \quad \text{on} \quad [1, \infty), \qquad \text{with } C = 1 \text{ an acceptable implied constant,}$$

while

$$x \neq O(x^2) \quad \text{on } [0, 1].$$

As a more complicated example,

$$\log(1 + x) = x - \frac{1}{2}x^2 + O(x^3) \quad \text{on } [-9/10, 9/10],$$

meaning: there is a function $E(x)$ with $\log(1 + x) = x - \frac{1}{2}x^2 + E(x)$ on $[-9/10, 9/10]$ with $E(x) = O(x^3)$ on $[-9/10, 9/10]$. You can prove this using the Maclaurin series for $\log(1 + x)$. (Really; try it!)

1.1. Basic properties

(a) For any constant c, we have $c \cdot O(g) = O(g)$.
Note. Interpret this to mean: "If $f = O(g)$, then $c \cdot f = O(g)$." Parts (b)–(e) should be interpreted similarly.
(b) $O(g) \cdot O(h) = O(gh)$,
(c) $O(f) + O(g) = O(|f| + |g|)$,

(d) If $f = O(g)$ then $O(f) + O(g) = O(g)$,
(e) If $f = O(g)$ and $g = O(h)$, then $f = O(h)$.

1.2. Prove: $\log(1+x) = x + O(x^2)$ for all $x \geq 0$. Is the same estimate true on $(-0.99, \infty)$? on $(-1, \infty)$?

1.3. We say that $f(x) = O(g(x))$ "as $x \to \infty$" or "for all large x" if $\exists\, x_0$ such that $f(x) = O(g(x))$ on (x_0, ∞). Prove: If $\lim_{x\to\infty} g(x) = 0$, then as $x \to \infty$,

$$\frac{1}{1+O(g(x))} = 1 + O(g(x)), \quad \mathrm{e}^{O(g(x))} = 1 + O(g(x)),$$
$$\text{and} \quad \log(1 + O(g(x))) = O(g(x)).$$

Note. Interpret the first claimed equation to mean that if $f(x) = O(g(x))$ as $x \to \infty$, then $1/(1 + f(x)) = 1 + O(g(x))$, as $x \to \infty$. Similarly for the others.

1.4. As $x \to \infty$,

$$\left(1 + \frac{1}{x}\right)^x = \mathrm{e} - \frac{\mathrm{e}}{2x} + O\left(\frac{1}{x^2}\right).$$

1.5. If f and g are positive-valued, then $(f+g)^2 \leq 2(f^2 + g^2)$. More generally, for any real $\kappa > 0$, we have $(f+g)^\kappa = O_\kappa(f^\kappa + g^\kappa)$. Here and elsewhere, a subscripted parameter indicates that you are allowed to choose your implied constant to depend on this parameter.

Asymptotic Analysis

1.6. For $n \in \mathbb{Z}^+$, define

$$a_n = \frac{1}{n} - \int_n^{n+1} \frac{\mathrm{d}t}{t}.$$

Interpret a_n as an area and explain, from this geometric perspective, how to see that $\sum_{n=1}^\infty a_n$ converges.

1.7. There is a real number γ (the "Euler–Mascheroni constant") such that for all positive integers N,

$$0 \geq \sum_{n \leq N} \frac{1}{n} - \big(\log(N+1) + \gamma\big) \geq -\frac{1}{N+1}.$$

1.8. For all real $x \geq 1$: $\displaystyle\sum_{n \leq x} \frac{1}{n} = \log x + \gamma + O(1/x).$

Ingenuity

1.9. (NEWMAN) Let $a_1 = 1$, and let $a_{n+1} = a_n + \frac{1}{a_n}$, for all $n \in \mathbb{Z}^+$. Then $a_n = \sqrt{2n} + O(n^{-1/2} \log n)$, as $n \to \infty$.

Step #2

> Mathematicians have tried in vain to this day to discover some order in the sequence of prime numbers, and we have reason to believe that it is a mystery into which the human mind will never penetrate.
>
> Leonhard Euler

Asymptotic Analysis

If f is strictly decreasing on $[n, n+1]$, then $f(n) > \int_n^{n+1} f(t)\,\mathrm{d}t > f(n+1)$ (draw a picture!). If f is strictly increasing, then the inequalities reverse. Use these observations to establish the following estimates.

2.10. For $s > 1$: $\quad \frac{1}{s-1} < \sum_{n=1}^{\infty} n^{-s} < \frac{s}{s-1}$.

2.11. For $s > 1$ and $x \geq 1$: $\quad \sum_{n>x} n^{-s} < x^{-s} + \frac{1}{s-1} x^{1-s} \leq \frac{s}{s-1} x^{1-s}$.

2.12. For $x \geq 1$: $\quad \log \lfloor x \rfloor ! = x \log x - x + O(\log(\mathrm{e}x))$. Why do we write $\mathrm{e}x$ and not x?

Infinitely Many Primes

Prove each statement and deduce the infinitude of primes.

2.13. (Stieltjes) If $p_1, \ldots, p_k$ is any finite list of distinct primes, with product P, and ab is any factorization of P into positive integers, then $a + b$ has a prime factor not among $p_1, \ldots, p_k$.

2.14. (Goldbach) The "Fermat numbers" $2^{2^n}+1$, for $n = 0, 1, 2, 3, \ldots$, are pairwise relatively prime.

2.15. (Perott) For some constant $c > 0$, and each $N \in \mathbb{Z}^+$, the count of squarefree integers in $[1, N]$ is

$$> N - \sum_{m \geq 2} N/m^2 \geq cN.$$

Thus, there are infinitely many squarefree integers.

2.16. (Ramanujan, Pillai, Ennola, Rubinstein) Let $\mathcal{P} = \{p_1, \ldots, p_k\}$ be a set of k primes, where $k < \infty$. For each $x \geq 1$, the number of integers in $[1, x]$ divisible by no primes outside of $\mathcal{P}$ coincides with the number of nonnegative integer solutions $e_1, \ldots, e_k$ to the inequality

$$e_1 \log p_1 + \cdots + e_k \log p_k \leq \log x. \tag{*}$$

The number of such solutions is

$$\frac{(\log x)^k}{k! \prod_{i=1}^k \log p_i} + O_{\mathcal{P}}((\log(\mathrm{e}x))^{k-1}).$$

Hint. Here is a way to start on the upper bound. To each nonnegative integer solution $e_1, \ldots, e_k$ of (*), associate the $1 \times 1 \times \cdots \times 1$ (hyper)cube in $\mathbb{R}^k$ having $(e_1, \ldots, e_k)$ as its "leftmost" corner. Show that all of these cubes sit inside the k-dimensional (hyper)tetrahedron defined by '$x_1 \log p_1 + \cdots + x_k \log p_k \leq \log(xp_1 \cdots p_k)$, all $x_i \geq 0$'. What is the volume of that tetrahedron? How does this volume compare to the number of cubes? It might help to first assume that $k = 2$ and draw some pictures.

Combinatorial Methods

2.17. For all $n \in \mathbb{Z}^+$, and all $0 \leq r \leq n$:

$$\binom{n}{0} - \binom{n}{1} + \cdots + (-1)^r \binom{n}{r} = (-1)^r \binom{n-1}{r}.$$

2.18. For a finite set A, and subsets $A_1, \ldots, A_k$ of A, state and prove the "inclusion-exclusion formula" for $|A \setminus (A_1 \cup A_2 \cup \cdots \cup A_k)|$. Why is it called "inclusion–exclusion"?

2.19. (Legendre)

$$\begin{aligned}\pi(x) - \pi(\sqrt{x}) + 1 \\ = \lfloor x \rfloor - \sum_{p_1 \leq \sqrt{x}} \left\lfloor \frac{x}{p_1} \right\rfloor + \sum_{p_1 < p_2 \leq \sqrt{x}} \left\lfloor \frac{x}{p_1 p_2} \right\rfloor - \sum_{p_1 < p_2 < p_3 \leq \sqrt{x}} \left\lfloor \frac{x}{p_1 p_2 p_3} \right\rfloor + \ldots\end{aligned}$$

Ingenuity

2.20. (Goldbach) If $f(T) \in \mathbb{Z}[T]$ and $f(n)$ is prime for all $n \in \mathbb{Z}^+$, then $f(T)$ is constant.

2.21. (Reiner) If k is an integer larger than 1, then the sequence $\{2^{2^n} + k\}_{n=0}^{\infty}$ contains infinitely many composite terms.

Note. It is an open problem to prove this also when $k = 1$.

Step #3

The worst thing you can do to a problem is solve it completely.

Daniel Kleitman

Asymptotic Analysis

The "Euler–Riemann zeta function" $\zeta(s)$ is defined, for $s > 1$, by $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$.

3.22. Justify the "Euler product representation" of the Euler–Riemann zeta function: For $s > 1$,

$$\zeta(s) = \prod_{p \text{ prime}} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \dots\right) = \prod_p \frac{1}{1 - \frac{1}{p^s}}.$$

3.23. For $s > 1$: $\quad \log \zeta(s) = \sum_p \sum_{k \geq 1} \frac{1}{kp^{ks}} = \sum_p \frac{1}{p^s} + O(1).$

3.24. For $1 < s < 2$: $\quad \sum_p \frac{1}{p^s} = \log \frac{1}{s-1} + O(1)$. It follows (why?) that $\sum_p \frac{1}{p}$ diverges. (Euler)

3.25. Find a sequence $\{c(n)\}_{n=1}^{\infty}$ with the property that

$$\zeta(s) \sum_{n=1}^{\infty} \frac{c(n)}{n^s} = 1$$

(for all $s > 1$), and describe $c(n)$ in terms of the prime factorization of n. (We will see later that there is a unique such sequence $\{c(n)\}$.)

Combinatorial Methods

3.26. (Jordan, Bonferroni) If one halts the inclusion-exclusion formula after an inclusion, one always overshoots (in the sense of obtaining an estimate at least as large as correct). If one stops after an exclusion, one always undershoots.

3.27. Let $\mathcal{A}$ be a set of positive integers. If $\sum_{a\in\mathcal{A}} \frac{1}{a}$ converges, then $\mathcal{A}$ contains 0% of the positive integers, in the sense that

$$\lim_{x\to\infty}\left(\sum\nolimits_{n\le x,\, n\in\mathcal{A}} 1 \Big/ \sum\nolimits_{n\le x} 1\right) = 0.$$

3.28. Let $\mathcal{A}$ be a set of positive integers for which $\sum_{a\in\mathcal{A}} \frac{1}{a}$ diverges. List the elements of $\mathcal{A}$: $a_1 < a_2 < a_3 < \dots$. Then there are infinitely many m for which $a_m < m(\log m)^{1.01}$. It follows that there are arbitrarily large values of x for which

$$\sum_{n\le x,\, n\in\mathcal{A}} 1 > x/(\log x)^{1.01}.$$

Can you think of other functions that can replace $x/(\log x)^{1.01}$ here?

Arithmetic Functions and the Anatomy of Integers

3.29. Suppose that f, g, h are arithmetic functions related by an identity $f(n) = \sum_{d\mid n} g(d)h(n/d)$, valid for all $n \in \mathbb{Z}^+$. Explain why

$$\sum_{n\le x} f(n) = \sum_{a\le x} g(a) \sum_{b\le x/a} h(b) = \sum_{b\le x} h(b) \sum_{a\le x/b} g(a).$$

3.30. For $x \ge 1$: $\displaystyle\sum_{n\le x} \tau(n) = x\log x + O(x)$. (Thus, a number $n \le x$ has $\approx \log x$ divisors "on average".)

3.31. Large values of the divisor function

(a) The numbers $n = 2^k$ all satisfy $\tau(n) > \log n$.
(b) For every real A, there are infinitely many $n \in \mathbb{Z}^+$ with $\tau(n) > (\log n)^A$.

3.32. For all $n \in \mathbb{Z}^+$: $\tau(n) \le 2n^{1/2}$.

Ingenuity

3.33. For every $N \in \mathbb{Z}^+$, there is a $d \in \mathbb{Z}^+$ for which the following holds: There are at least N primes p for which $p + d$ is also prime.

Step #4

What did the analytic number theorist say when they were drowning? Log-log-log-log-log.

Anonymous

Variations on a Theme of Euler

4.34.

(a) For all $x > 0$, and every $\epsilon \in (0,1)$:

$$\sum_{p \le x} \frac{1}{p} \le \sum_{p \le x} \frac{1}{p} \left(\frac{x}{p}\right)^{\epsilon} = x^{\epsilon} \sum_{p \le x} \frac{1}{p^{1+\epsilon}} \le x^{\epsilon} \log \frac{1}{\epsilon} + O(x^{\epsilon}).$$

(b) For all sufficiently large values of x:

$$\sum_{p \le x} \frac{1}{p} \le \log\log x + 2\log\log\log x.$$

Hint. Use (a) with $\epsilon = \frac{1}{\log x \cdot \log\log x}$. (But how did we come up with this choice of ϵ?)

4.35.

(a) For all $x > 0$, and every $\epsilon \in (0,1)$:

$$\sum_{p \le x} \frac{1}{p} \ge \sum_{p} \frac{1}{p^{1+\epsilon}} - \sum_{p > x} \frac{1}{p^{1+\epsilon}} \ge \log \frac{1}{\epsilon} - \sum_{n > x} \frac{1}{n^{1+\epsilon}} + O(1).$$

(b) For all sufficiently large values of x:

$$\sum_{p \le x} \frac{1}{p} \ge \log\log x - 2\log\log\log x.$$

From Problems 4.34 and 4.35, we conclude that as $x \to \infty$: $\sum_{p \le x} \frac{1}{p} = \log\log x + O(\log\log\log x)$. Later we will prove sharper estimates for this sum.

Arithmetic Functions and the Anatomy of Integers

4.36. Recall that Euler's ϕ-function satisfies

$$\phi(n) = n \prod_{p \mid n} \left(1 - \frac{1}{p}\right) = n \sum_{d \mid n} \frac{\mu(d)}{d}.$$

Here $\mu(n)$ is the Möbius function, which appeared as the solution sequence $c(n)$ in Problem 3.25. Deduce from Problem 3.29 that for $x \ge 1$:

$$\sum_{n \le x} \phi(n) = \frac{1}{2} x^2 \sum_{a \le x} \frac{\mu(a)}{a^2} + O(x \log(\mathrm{e}x)).$$

4.37. (Dirichlet, Mertens) For $x \ge 1$:

$$\sum_{n \le x} \phi(n) = \frac{1}{2\zeta(2)} x^2 + O(x \log(\mathrm{e}x)).$$

4.38. (Dirichlet) A lattice point is chosen uniformly at random from the square $(0, N] \times (0, N]$, where $N \in \mathbb{Z}^+$. As $N \to \infty$, the probability its coordinates are relatively prime tends to $\frac{1}{\zeta(2)}$.

Computing with Roots of Unity

4.39. Let $m \in \mathbb{Z}^+$. For $a \in \mathbb{Z}$:

$$\frac{1}{m} \sum_{k \bmod m} \mathrm{e}^{2\pi \mathrm{i} k a/m} = \begin{cases} 1 & \text{if } a \equiv 0 \pmod{m}, \\ 0 & \text{otherwise.} \end{cases}$$

Here the sum on k is taken over any set of integer representatives of $\mathbb{Z}_m$.

4.40. (Counting square roots mod m) Let $m \in \mathbb{Z}^+$. For $n \in \mathbb{Z}$:

$$\#\{a \bmod m : a^2 \equiv n \pmod{m}\} = \frac{1}{m} \sum_{k \bmod m} \mathrm{e}^{2\pi \mathrm{i} k n/m} \sum_{a \bmod m} \mathrm{e}^{-2\pi \mathrm{i} k a^2/m}.$$

Dirichlet Series

By now we have seen multiple expressions of the form $\sum_{n=1}^{\infty} \frac{a(n)}{n^s}$, where the $a(n)$ are complex numbers. These are known as "Dirichlet series".

4.41. Suppose that $\sum_{n=1}^{\infty} \frac{a(n)}{n^s}$ is a Dirichlet series that converges for some real number $s = s_0$. Then for some real number C, we have $|a(n)| \leq Cn^{s_0}$ for all n. Hence, $\sum_{n=1}^{\infty} \frac{a(n)}{n^s}$ converges absolutely for every $s > s_0 + 1$. Furthermore, for every $m \in \mathbb{Z}^+$:

$$\lim_{s \to \infty} m^s \sum_{n=m}^{\infty} \frac{a(n)}{n^s} = a(m).$$

Mathematical Masterpieces: The Identity as Art Form

4.42. (Goldbach) Find the sum of the infinite series

$$\frac{1}{3} + \frac{1}{7} + \frac{1}{8} + \frac{1}{15} + \frac{1}{24} + \frac{1}{26} + \frac{1}{31} + \frac{1}{35} + \frac{1}{48} + \frac{1}{63} + \dots$$

whose denominators, increased by 1, are the distinct numbers of the form n^m with $n, m \geq 2$ (the perfect powers).

Step #5

> I have had my results for a long time: but I do not yet know how I am to arrive at them.
>
> Carl Friedrich Gauss

Distribution of Squares mod p

Let p be an odd prime.

5.43. (Gauss) The "Gauss sum" associated to p is

$$G = \sum_{a \bmod p} \mathrm{e}^{2\pi \mathrm{i} a^2/p}.$$

Show that for $k \in \mathbb{Z}$, $p \nmid k$: $\displaystyle\sum_{a \bmod p} \mathrm{e}^{2\pi \mathrm{i} k a^2/p} = \left(\frac{k}{p}\right) G.$

Here $\left(\frac{k}{p}\right)$ is the Legendre symbol: 0 when $p \mid k$, and otherwise 1 or -1, according to whether or not k is a square mod p.

5.44. For $n \in \mathbb{Z}$:

$$\#\{a \bmod p : a^2 \equiv n \pmod{p}\} = 1 + \frac{G}{p} \sum_{k \bmod p} \mathrm{e}^{2\pi \mathrm{i} k n/p} \left(\frac{-k}{p}\right).$$

Deduce: $\displaystyle\left(\frac{n}{p}\right) = \frac{G}{p} \sum_{k \bmod p} \mathrm{e}^{2\pi \mathrm{i} k n/p} \left(\frac{-k}{p}\right).$

5.45. Prove: $G \cdot \overline{G} = p$. (Here the bar denotes complex conjugation.) Deduce that G is a square root of $\left(\frac{-1}{p}\right)p$.
[Thus, $G = \pm\sqrt{p}$ when $p \equiv 1 \pmod 4$ and $G = \pm \mathrm{i}\sqrt{p}$ when $p \equiv 3 \pmod 4$. Gauss worked for years to determine which sign to take, eventually proving that the $+$ sign is always correct.]

Hint. Start from the expression for $\left(\frac{n}{p}\right)$ proved in Problem 5.44. Take the modulus squared of both sides and sum on n mod p.

Variations on a Theme of Euler

Below, we write $\omega(n)$ for the number of distinct prime factors of n and we use $\Omega(n)$ for the number of prime factors of n, counted with multiplicity. For example, $\omega(45) = 2$, while $\Omega(45) = 3$. Equivalently,

$$\omega(n) = \sum_{p \mid n} 1, \qquad \Omega(n) = \sum_{p^k \mid n} 1.$$

5.46. For every nonnegative integer k, and real $x \geq 1$:

$$\sum_{\substack{n \leq x \\ n \text{ squarefree} \\ \omega(n)=k}} \frac{1}{n} \leq \frac{1}{k!}\left(\sum_{p \leq x} \frac{1}{p}\right)^k .$$

5.47. For $x > 1$:

$$\exp\left(\sum_{p \leq x} \frac{1}{p}\right) \geq \sum_{\substack{n \leq x \\ n \text{ squarefree}}} \frac{1}{n}.$$

Also:

$$\zeta(2) \sum_{\substack{n \leq x \\ n \text{ squarefree}}} \frac{1}{n} \geq \sum_{n \leq x} \frac{1}{n} > \log x.$$

Deduce:

$$\sum_{p \leq x} \frac{1}{p} > \log\log x - 1.$$

This improves the lower bound of Problem 4.35.

Arithmetic Functions and the Anatomy of Integers

5.48. (Dirichlet) For $x \geq 1$: $\displaystyle\sum_{n \leq x} \sigma(n) = \frac{1}{2}\zeta(2)x^2 + O(x \log(\mathrm{e}x))$.

5.49. For all $n \in \mathbb{Z}^+$: $2^{\omega(n)} \leq \tau(n) \leq 2^{\Omega(n)}$.

Dirichlet Series

5.50. (Kalmár) A "multiplicative composition" of n is a representation of n as a product of integers > 1, where order matters. We let $g(n)$ denote the number of multiplicative compositions of n. For instance, $g(1) = 1$ (the empty composition has all parts > 1), while $g(6) = 3$ (for $2 \cdot 3$, $3 \cdot 2$, 6).

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$g(n)$	1	1	1	2	1	3	1	4	2	3	1	8	1	3	3	8	1	8	1	8

Let $\rho = 1.72864\ldots$ be the solution in $(1, \infty)$ to $\zeta(\rho) = 2$.

Prove: For all $s > \rho$,

$$\sum_{n=1}^{\infty} \frac{g(n)}{n^s} = \frac{1}{2 - \zeta(s)}.$$

5.51. If $\sum_{n=1}^{\infty} \frac{a(n)}{n^s}$ and $\sum_{n=1}^{\infty} \frac{b(n)}{n^s}$ converge and are equal for all large real numbers s, then each $a(n) = b(n)$. (This implies the uniqueness of the sequence $\{c(n)\}$ in Problem 3.25.)

Mathematical Masterpieces: The Identity as Art Form

5.52. For every nonnegative integer n,

$$\int_0^{\pi/2} \sin^{2n} x \,\mathrm{d}x = \frac{\pi}{2} \cdot \frac{1 \cdot 3 \cdots (2n-1)}{2 \cdot 4 \cdots (2n)}, \quad \text{while}$$

$$\int_0^{\pi/2} \sin^{2n+1} x \,\mathrm{d}x = \frac{2 \cdot 4 \cdots (2n)}{3 \cdot 5 \cdots (2n+1)}.$$

Here, as usual, empty products are to be understood to equal 1.

5.53. (Wallis) Show that as $n \to \infty$,

$$\frac{\int_0^{\pi/2} \sin^{2n} x \,\mathrm{d}x}{\int_0^{\pi/2} \sin^{2n+1} x \,\mathrm{d}x} \to 1.$$

Conclude that $\displaystyle\frac{\pi}{2} = \prod_{k=1}^{\infty} \left(\frac{2k}{2k-1} \cdot \frac{2k}{2k+1} \right)$.