A GENTLE INTRODUCTION TO

NUMBER THEORY

LUKE COLLINS
luke.collins@um.edu.mt ◦ lc.mt

LECTURE I*
29TH AUGUST, 2022

Mathematics is the queen of the sciences, and number theory is the queen of mathematics.
— Carl Friedrich Gauß

INTRODUCTION

Let $\mathbb{N}$ denote the set $\{1, 2, \dots, \}$ of natural numbers, and let $\mathbb{Z}$ denote the ring of integers. We say that an integer a *divides* another integer b, written $a \mid b$, if there exists an integer d such that $b = ad$. An integer n is said to be *square-free* if $a^2 \nmid n$ for all $a \geqslant 2$.

We say that an integer $p \geqslant 2$ is *prime* if its only divisors are ± 1 and $\pm p$. Throughout these notes, we reserve the letter p for primes. (For instance, $\sum_p \frac{1}{p}$ denotes the sum of reciprocals of all the primes.)

The *highest common factor* of two integers a and b (not both zero), denoted by $\mathrm{hcf}(a, b)$ or just (a, b), is the largest integer d such that $d \mid a$ and $d \mid b$. Two integers a and b are said to be *relatively prime* or *coprime* if $(a, b) = 1$.

Theorem 1.1. *Every integer $n \geqslant 2$ is prime or a product of primes.*

Proof. By induction on n. For the base case with $n = 2$, it is easily seen that only $\pm 1, \pm 2$ divide 2, so 2 is prime. Now suppose the theorem holds for all integers smaller than n; and suppose n is not prime (otherwise we are done).

*A series of lectures for the Malta Mathematical Society. If you find any mathematical, grammatical or typographical errors whilst reading these notes, please let the author know via email: luke.collins@um.edu.mt.

Then there exists a divisor a of n with $1 \neq a \neq n$, such that $n = ad$ for some d. But $a, d < n$, so they are both either prime or products of primes by the hypothesis; thus we have expressed n as a product of primes. □

Theorem 1.2. *There are infinitely many primes.*

Euclid's proof. For contradiction, suppose there are finitely many primes, and let $p_1, \ldots, p_n$ denote all of them. Define $N := 1 + \prod_{i=1}^{n} p_i$. Clearly $N \geqslant 2$, and it is not prime since it is larger than each p_i, but it is neither a product of primes since if $p_i \mid N$, then $p_i \mid (N - \prod_{i=1}^{n} p_i) = 1$ which is impossible. This contradicts theorem 1.1. □

We also have the following pair of familiar results which we will not prove here, but can be found in any standard textbook on elementary number theory or basic algebra.

Theorem 1.3 (Fundamental Theorem of Arithmetic)**.** *Every non-zero integer n has a factorisation as a product of primes (possibly the empty product), multiplied by ± 1. The factorisation is unique up to the order in which the primes are written.*

In algebraic terms, the fundamental theorem of arithmetic simply says that $\mathbb{Z}$ is a unique factorisation domain.

Theorem 1.4 (Bézout's Lemma)**.** *Let $a, b \in \mathbb{Z}$, not both zero. Then there exist $x, y \in \mathbb{Z}$ such that $(a, b) = ax + by$.*

Bézout's lemma is key to proving many facts about the integers, such as the following, which usually plays a role in the proof of the fundamental theorem of arithmetic. The Bézout coefficients x and y can be computed using Euclidean division.

Proposition 1.5 (Euclid's Lemma)**.** *Let p be a prime, and let $a, b \in \mathbb{Z}$. Then*

$$p \mid ab \implies p \mid a \quad \text{or} \quad p \mid b.$$

Proof. Suppose $p \nmid a$, so $(p, a) = 1$. Then by Bézout's lemma, we may express $1 = ax + py$, which implies that $b = abx + pby$. Since p divides ab, it divides the right-hand side, so it must divide b. □

In terms of algebra, this last result shows that our definition of "prime" agrees with the algebraic one. Bézout's lemma immediately gives us a way to solve Diophantine equations of the form $ax + by = c$.

Corollary 1.6. *Let $a, b, c \in \mathbb{Z}$. Then the equation*

$$ax + by = c$$

has solutions in integers if and only if c is a multiple of (a, b).

Proof. Write $(a, b) = ax' + by'$ by Bézout's lemma. If c is a multiple of (a, b), say, $c = k(a, b)$, then taking $x = kx'$ and $y = ky'$ clearly gives a solution.

Conversely, if $ax + by = c$ has a solution, then c must be a multiple of (a, b), since $(a, b) \mid (ax + by)$. □

SOME FUNCTIONS OF INTEREST

An *arithmetic function* is a sequence $f \colon \mathbb{N} \to \mathbb{C}$. Such functions are said to be *multiplicative* if $f(ab) = f(a)f(b)$ for all a, b with $(a, b) = 1$, and *completely multiplicative* if this holds for all $a, b \in \mathbb{N}$. Similarly we say that f is *additive* if $f(ab) = f(a) + f(b)$ for all a, b with $(a, b) = 1$, and *completely additive* if this holds for all $a, b \in \mathbb{N}$.

Some arithmetic functions we study in number theory are the following.

(i) $\omega(n)$ denotes the number of distinct prime divisors of n.

(ii) $\Omega(n)$ denotes the number of prime divisors of n (with multiplicity).

(iii) $\mu(n)$ denotes the *Möbius function*, defined by

$$\mu(n) \coloneqq \begin{cases} (-1)^{\omega(n)} & \text{if } n \text{ is square-free,} \\ 0 & \text{otherwise.} \end{cases}$$

(iv) $\varphi(n)$ denotes the (*Euler*) *totient function*, which counts the number of integers between 1 and n that are coprime to n, i.e.,

$$\varphi(n) \coloneqq \sum_{\substack{k \leqslant n \\ (k,n)=1}} 1.$$

(v) $\Lambda(n)$ denotes the (*von*) *Mangoldt function*, defined by

$$\Lambda(n) \coloneqq \begin{cases} \log p & \text{if } n = p^k \text{ for some integer } k \geqslant 1, \\ 0 & \text{otherwise.} \end{cases}$$

The sum $\sum_{n\leqslant x}\Lambda(n)$ is known as *Chebychev's psi counting function*, and is usually denoted by $\Psi(x)$.

Observe that the function ω is additive, Ω is completely additive, μ and φ are both multiplicative, and Λ is neither multiplicative nor additive. (The multiplicativity of φ is not obvious, but follows easily from proposition 3.3.)

Another important function (which is not arithmetic) is the *prime counting function* $\pi(x)$, which counts the number of primes $p\leqslant x$ for any $x\in\mathbb{R}$, i.e.,

$$\pi(x) \coloneqq \sum_{p\leqslant x} 1.$$

One of the most famous results in analytic number theory is the celebrated prime number theorem (PNT). This states that $\pi(x)$ is asymptotic to $x/\log x$, i.e.,

$$\lim_{x\to\infty}\frac{\pi(x)}{x/\log x} = 1.$$

Later on we will give a proof of this theorem. We will actually prove that $\Psi(x)$ is asymptotic to x; this is equivalent to PNT (which we will also show), and turns out to make the working simpler.

A LOOK AT $\mu(n)$ AND $\varphi(n)$

The first few values of μ and φ are given in table 1. Notice that the only

n	1	2	3	4	5	6	7	8	9	10	11	12
$\mu(n)$	1	-1	-1	0	-1	1	-1	0	0	1	-1	0
$\varphi(n)$	1	1	2	2	4	2	6	4	6	4	10	4

TABLE 1: The first few values of $\mu(n)$ and $\varphi(n)$.

non-zero values the Möbius function takes on are ±1, but the pattern is quite erratic. It might be informative to study the cancellation which we obtain when we sum over its values. Instinctively, we might try to study the sum

$$M(x) = \sum_{n\leqslant x}\mu(n)$$

of its consecutive values, but this turns out to be a difficult sum to understand. In fact, a long standing conjecture, known as Merten's conjecture, was that $|M(x)| < \sqrt{x}$ for all x, which seems reasonable if we look at a plot

of $M(x)$ and $\pm\sqrt{x}$ on the same axes (figure 1). This conjecture was of particular interest because it has been shown to imply the Riemann hypothesis; but we now know that the conjecture is false, i.e., there exists a value of x such that $|M(x)| > \sqrt{x}$. No explicit value of x for which this happens is known, but we know that there is an x between 10^{16} and $10^{10^{40}}$ which does the job.

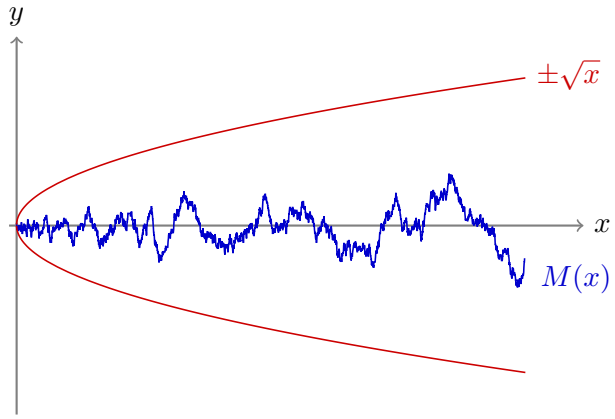

FIGURE 1: The Mertens function $M(x)$ and $\pm\sqrt{x}$ plotted for $1 \leqslant x \leqslant 10\,000$

It turns out that, since μ is multiplicative, the sum we should study is actually that over the *divisors* of a given n. Let $\varepsilon(n)$ denote the arithmetic function which equals 1 when $n = 1$, and 0 for all other values of n. Then we have the following.

Proposition 3.1. *Let $n \geqslant 1$. Then*

$$\sum_{d|n} \mu(d) = \varepsilon(n).$$

For example, when $n = 6$, we have

$$\mu(1) + \mu(2) + \mu(3) + \mu(6) = 1 - 1 - 1 + 1 = 0 = \varepsilon(6).$$

Proof. When $n = 1$, we have that $\sum_{d|1} \mu(d) = \mu(1) = 1$. For $n \geqslant 2$, factorise $n = \prod_{i=1}^{k} p_i^{n_i}$ by the fundamental theorem of arithmetic, where $k = \omega(n)$. Observe that the sum has non-zero contributions from square-free divisors only, i.e.,

$$\sum_{d|n} \mu(d) = \mu(1) + \mu(p_1) + \cdots + \mu(p_k) + \mu(p_1 p_2) + \cdots + \mu(p_1 \cdots p_k)$$

$$= 1 + (-1)^1 + \cdots + (-1)^1 + (-1)^2 + \cdots + (-1)^k$$
$$= 1 + \binom{k}{1}(-1)^1 + \binom{k}{2}(-1)^2 + \cdots + \binom{k}{k}(-1)^k = (1-1)^k = 0,$$

as required. □

We have a similarly nice result for the divisor sum of the totient function.

Proposition 3.2. *Let $n \geqslant 1$. Then*

$$\sum_{d|n} \varphi(d) = n.$$

For instance, with $n = 6$, we have

$$\varphi(1) + \varphi(2) + \varphi(3) + \varphi(6) = 1 + 1 + 2 + 2 = 6.$$

Proof. Consider the n fractions $1/n, \ldots, n/n$, and simplify them so that they are in reduced form. Each fraction, when reduced, will have a denominator which is a divisor of n. But for each divisor d of n, there are precisely $\varphi(d)$ fractions with denominator d. Thus the total number of fractions is $\sum_{d|n} \varphi(d)$. □

For a more formal proof, one should partition $\mathbb{Z}/n\mathbb{Z}$ according to the value of (n, r) for each residue r, obtaining $n = \#(\mathbb{Z}/n\mathbb{Z}) = \sum_{d|n} \varphi(d)$. We also have a nice formula for $\varphi(n)$ which makes it easier to evaluate.

Proposition 3.3 (Euler product formula)**.** *Let $n \geqslant 1$. Then*

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right).$$

Proof. Suppose we randomly pick a number in $\{1, \ldots, n\}$. The probability that it is coprime to n is $\varphi(n)/n$. We can also work out this probability in a different way. The probability that a number in $\{1, \ldots, n\}$ is divisible by some prime divisor p of n is $1/p$; thus the probability that it is *not* divisible by p is $1 - 1/p$. Finally, since we want our random number to be coprime to n, this happens precisely when no prime divisor of n is a divisor of the randomly chosen number; i.e., this occurs with probability $\prod_{p|n}(1-\frac{1}{p})$. Thus we have

$$\frac{\varphi(n)}{n} = \prod_{p|n} \left(1 - \frac{1}{p}\right).$$ □

Example 3.4. This allows us to evaluate $\varphi(n)$ more easily, e.g.,

$$\varphi(40) = 40(1 - \tfrac{1}{2})(1 - \tfrac{1}{5}) = 16.$$

Corollary 3.5. *Let $a, b \in \mathbb{Z}$, $n \geqslant 1$, and p be a prime. Then:*

(i) $\varphi(p^n) = p^{n-1}(p-1)$,

(ii) $\varphi(ab) = \varphi(a)\,\varphi(b)$ *for* $(a,b) = 1$,

(iii) $\varphi(ab) = \varphi(a)\,\varphi(b) \cdot (a,b)/\varphi((a,b))$.

Proof. (i) follows immediately from proposition 3.3, and (ii) follows from (iii). To show (iii), recall by Euclid's lemma that a prime divides ab implies it divides a or b (or both). In particular, those which divide both a and b are precisely those which divide (a,b). Thus

$$\frac{\varphi(ab)}{ab} = \prod_{p \mid ab} \left(1 - \frac{1}{p}\right) = \frac{\prod_{p \mid a} (1 - \frac{1}{p}) \prod_{p \mid b} (1 - \frac{1}{p})}{\prod_{p \mid (a,b)} (1 - \frac{1}{p})} = \frac{\frac{\varphi(a)}{a} \cdot \frac{\varphi(b)}{b}}{\frac{\varphi((a,b))}{(a,b)}},$$

which implies (iii). □

The next proposition gives us an interesting relation between μ and φ.

Proposition 3.6. *Let $n \geqslant 1$. Then*

$$\varphi(n) = \sum_{d \mid n} \mu(d)\, \frac{n}{d}.$$

Proof. Let $\mathbb{1}_{\mathscr{P}} = 1$ if $\mathscr{P}$ is true, and 0 otherwise. By definition of φ and proposition 3.1,

$$\begin{aligned}
\varphi(n) = \sum_{\substack{k \leqslant n \\ (k,n)=1}} 1 = \sum_{k=1}^{n} \mathbb{1}_{(k,n)=1} &= \sum_{k=1}^{n} \sum_{d \mid (k,n)} \mu(d) = \sum_{k=1}^{n} \sum_{\substack{d \mid n \\ d \mid k}} \mu(d) \\
&= \sum_{k=1}^{n} \sum_{d \mid n} \mu(d)\, \mathbb{1}_{d \mid k} \\
&= \sum_{d \mid n} \mu(d) \sum_{k=1}^{n} \mathbb{1}_{d \mid k} = \sum_{d \mid n} \mu(d)\, \frac{n}{d},
\end{aligned}$$

which completes the proof. □

MÖBIUS INVERSION AND DIRICHLET CONVOLUTION

We saw in the last section (propositions 3.2 and 3.6) that

$$n = \sum_{d|n} \varphi(d) \qquad \text{and} \qquad \varphi(n) = \sum_{d|n} \mu(d)\, \frac{n}{d}.$$

We can regard these two together as an instance of the following.

Theorem 4.1 (Möbius inversion)**.** *Let $f, g\colon \mathbb{N} \to \mathbb{C}$ be a pair of arithmetic functions. Then*

$$f(n) = \sum_{d|n} g(d) \qquad \textit{if and only if} \qquad g(n) = \sum_{d|n} \mu(d)\, f\Big(\frac{n}{d}\Big).$$

In the case of propositions 3.2 and 3.6, we have $f(n) = n$ and $g(n) = \varphi(n)$.

Proof of theorem 4.1. Suppose $f(n) = \sum_{d|n} g(d)$. Then

$$\begin{aligned}
\sum_{d|n} \mu(d)\, f\Big(\frac{n}{d}\Big) &= \sum_{d|n} \mu(d) \sum_{e|\frac{n}{d}} g(e) \\
&= \sum_{d|n} \mu(d) \sum_{e \leqslant n} g(e)\, \mathbb{1}_{e|\frac{n}{d}} \\
&= \sum_{d|n} \mu(d) \sum_{e \leqslant n} g(e)\, \mathbb{1}_{d|\frac{n}{e}} \\
&= \sum_{e \leqslant n} g(e) \sum_{d|n} \mu(d)\, \mathbb{1}_{d|\frac{n}{e}} = \sum_{e \leqslant n} g(e) \sum_{d|\frac{n}{e}} \mu(d) = g(n)
\end{aligned}$$

by proposition 3.1. Conversely, if $g(n) = \sum_{d|n} \mu(d)\, f(\frac{n}{d})$, then

$$\begin{aligned}
\sum_{d|n} g(d) &= \sum_{d|n} \sum_{e|d} \mu(e)\, f\Big(\frac{d}{e}\Big) \\
&= \sum_{abe=n} \mu(e)\, f(a) = \sum_{a|n} f(a) \sum_{e|\frac{n}{a}} \mu(e) = f(n)
\end{aligned}$$

again by similar reasoning with proposition 3.1, as required. □

In general, sums of the form

$$\sum_{d|n} f(d)\, g\Big(\frac{n}{d}\Big) \qquad \text{or, written differently,} \qquad \sum_{ab=n} f(a)\, g(b),$$

are quite common in number theory. This leads to the following definition.

Definition 4.2 (Dirichlet convolution)**.** Let $f, g\colon \mathbb{N} \to \mathbb{C}$ be two arithmetic functions. Their (*Dirichlet*) *convolution*, denoted by $f * g$, is the function defined by

$$(f * g)(n) \coloneqq \sum_{d|n} f(d)\, g\Big(\frac{n}{d}\Big).$$

In terms of this notation, we may express proposition 3.6 as

$$\varphi = \mu * \mathrm{id},$$

where id denotes the identity function $\mathrm{id}(n) \coloneqq n$ for all n. Similarly, if $\mathbb{1}$ denotes the function $\mathbb{1}(n) = 1$ for all n, then Möbius inversion may be expressed as

$$g = f * \mathbb{1} \iff f = \mu * g.$$

Proposition 4.3 (Abelian group structure)**.** *Let $f, g, h\colon \mathbb{N} \to \mathbb{C}$ be three arithmetic functions, and let $\varepsilon(n)$ denote the function which is 1 at 1 and 0 everywhere else, i.e., $\varepsilon(n) \coloneqq \mathbb{1}_{n=1}$. Then we have the following:*

(i)	$f * g$ *is an arithmetical function,*	(CLOSURE)
(ii)	$(f * g) * h = f * (g * h)$,	(ASSOCIATIVITY)
(iii)	$f * \varepsilon = \varepsilon * f$,	(IDENTITY)
(iv)	*if* $f(1) \neq 0$, *then there exists* f^{-1} *such that* $f * f^{-1} = f^{-1} * f = \varepsilon$,	(INVERSE)
(v)	$f * g = g * f$.	(ABELIAN)

Proof. (i)–(iii) and (v) are straightforward, the interesting one is (iv). It turns out the inverse f^{-1} is given by the recursive formula

$$f^{-1}(1) \coloneqq \frac{1}{f(1)} \qquad \text{and} \qquad f^{-1}(n) \coloneqq -\frac{1}{f(1)} \sum_{\substack{d|n \\ d\neq n}} f\Big(\frac{n}{d}\Big) f^{-1}(d),$$

and we can prove this by induction. Clearly for the base case, we want f^{-1} to satisfy $(f * f^{-1})(1) = \varepsilon(1)$, which expands to $f(1)f^{-1}(1) = 1$, and this forces us to take $f^{-1}(1)$ as above.

Now for the inductive step, let $n \geqslant 2$, and suppose f^{-1} is well-defined for all values less than n. Then we want to ensure $(f * f^{-1})(n) = \varepsilon(n) = 0$, i.e.,

$$\sum_{d|n} f\Big(\frac{n}{d}\Big) f^{-1}(d) = 0,$$

or, taking out the term where $d = n$,

$$f(1)f^{-1}(n) + \sum_{\substack{d|n \\ d\neq n}} f\left(\frac{n}{d}\right) f^{-1}(d) = 0.$$

Since we assumed $f^{-1}(d)$ is defined for all $d < n$, we can solve this to get a unique value for $f^{-1}(n)$ as defined above, this establishes the existence and uniqueness of f^{-1}. □

Remark 4.4. By proposition 3.1, we have $\mu * \mathbb{1} = \varepsilon$, so μ and $\mathbb{1}$ are Dirichlet inverses of each other; i.e., $\mu^{-1} = \mathbb{1}$ and $\mathbb{1}^{-1} = \mu$.

This fact, together with the group structure, gives us a much more concise proof of Möbius inversion: simply multiply $g = f * \mathbb{1}$ by μ both sides to get $g * \mu = f * \mathbb{1} * \mu = f$, and for the converse, just multiply $f = g * \mu$ by $\mathbb{1}$.

Let us now give a result on the von Mangoldt function, and use Möbius inversion to get a corollary for free.

Proposition 4.5. *Let $n \geqslant 1$. Then*

$$\sum_{d|n} \Lambda(d) = \log n,$$

or in terms of convolutions, $\mathbb{1} * \Lambda = \log$.

Proof. Simply factorise $n = \prod_{i=1}^{k} p_i^{n_i}$ to get

$$\sum_{d|n} \Lambda(d) = \sum_{i=1}^{k} \sum_{\substack{\ell \geqslant 1 \\ p_i^\ell | n}} \log p_i = \sum_{i=1}^{k} n_i \log p_i = \log n,$$

as required. □

By Möbius inversion, we also get:

Corollary 4.6. *Let $n \geqslant 1$. Then*

$$\Lambda(n) = -\sum_{d|n} \mu(d) \log d.$$

Proof. By Möbius inversion, $\mathbb{1} * \Lambda = \log \implies \Lambda = \log * \mu$, i.e.,

$$\Lambda(n) = \sum_{d|n} \mu(d) \log\left(\frac{n}{d}\right) = \log n \sum_{d|n} \mu(d) - \sum_{d|n} \mu(d)\ \log d,$$

and by proposition 3.1 the first sum is $\varepsilon(n) \log n$ which equals 0 for all n. □

PARTIAL SUMMATION

One of the most useful techniques on obtaining asymptotic formulæ for sums is so called *partial summation*, which involves writing a function as the integral of its derivative, and swapping the sum and the integral. As an example, take the series

$$\sum_{n \leqslant x} \log n.$$

We can write $\log n$ as $\int_1^n \frac{dt}{t}$ to get

$$\begin{aligned}
\sum_{n \leqslant x} \int_1^n \frac{dt}{t} &= \sum_{n \leqslant x} \int_1^x \frac{\mathbb{1}_{t \leqslant n}}{t}\, dt \\
&= \int_1^x \left(\sum_{n \leqslant x} \mathbb{1}_{t \leqslant n} \right) \frac{dt}{t} \\
&= \int_1^x \frac{\lfloor x \rfloor - \lceil t \rceil + 1}{t}\, dt \\
&= \int_1^x \frac{x - t + O(1)}{t}\, dt \\
&= x \log x - (x - 1) + O(\log x) \\
&= x \log x - x + O(\log x).
\end{aligned}$$

Another example, notice that $\frac{1}{n} = 1 - \int_1^n \frac{dt}{t^2}$, so

$$\begin{aligned}
\sum_{n \leqslant x} \frac{1}{n} &= \sum_{n \leqslant x} \left(1 - \int_1^n \frac{dt}{t^2} \right) \\
&= \lfloor x \rfloor - \sum_{n \leqslant x} \int_1^x \frac{\mathbb{1}_{t \leqslant n}}{t^2}\, dt \\
&= \lfloor x \rfloor - \int_1^x \left(\sum_{n \leqslant x} \mathbb{1}_{t \leqslant n} \right) \frac{dt}{t^2}
\end{aligned}$$

$$\begin{aligned}
&= \lfloor x \rfloor - \int_1^x \frac{\lfloor x \rfloor - t + O(1)}{t^2}\, dt \\
&= \lfloor x \rfloor - \lfloor x \rfloor \Big(1 - \frac{1}{x}\Big) + \log x - \int_1^x \frac{O(1)}{t^2}\, dt \\
&= \frac{\lfloor x \rfloor}{x} + \log x - O(1) \\
&= \log x + O(1),
\end{aligned}$$

and if we are a bit more careful, for $t \geqslant 0$, we can write $\lfloor t \rfloor = t - \{t\}$, so

$$\begin{aligned}
\sum_{n \leqslant x} \frac{1}{n} &= \lfloor x \rfloor - \int_1^x \frac{\lfloor x \rfloor - t + \{t\}}{t^2}\, dt \\
&= \lfloor x \rfloor - \lfloor x \rfloor \Big(1 - \frac{1}{x}\Big) + \log x - \int_1^x \frac{\{t\}}{t^2}\, dt \\
&= \log x + \frac{\lfloor x \rfloor}{x} - \int_1^x \frac{\{t\}}{t^2}\, dt \\
&= \log x + \frac{x + O(1)}{x} - \int_1^\infty \frac{\{t\}}{t^2}\, dt + \underbrace{\int_x^\infty \frac{\{t\}}{t^2}\, dt}_{\ll \int \frac{dt}{t^2} = \frac{1}{x}} \\
&= \log x + 1 - \int_1^\infty \frac{\{t\}}{t^2}\, dt + O\Big(\frac{1}{x}\Big) \\
&= \log x + \gamma + O\Big(\frac{1}{x}\Big).
\end{aligned}$$

Example 5.1. Here we use summation by parts to answer a probabilistic question. We show that the probability that two randomly chosen numbers from $\{1, \dots, n\}$ add up to a perfect square is

$$\frac{4\sqrt{2} - 4}{3\sqrt{n}} + O\Big(\frac{1}{n}\Big).$$

The probability is s_n/n^2, where

$$s_n = \#\{(a, b) \in \{1, \dots, n\}^2 : a + b = k^2 \text{ for some } k\}.$$

We can construct all valid pairs (a, b) which contribute to s_k as follows. First, we pick a in the range $1 \leqslant a \leqslant n$, and then choose k such that $b = a - k^2$

lies in the range $1 \leqslant k^2 - a \leqslant n$, i.e., such that

$$1 + a \leqslant k^2 \leqslant n + a,$$

i.e.,

$$\lceil \sqrt{1+a} \rceil \leqslant k \leqslant \lfloor \sqrt{n+a} \rfloor.$$

For each $1 \leqslant a \leqslant n$, there are precisely $\lfloor \sqrt{n+a} \rfloor - \lceil \sqrt{1+a} \rceil + 1$ choices of k, and therefore precisely this number of corresponding choices for b. Thus we have

$$\begin{aligned} s_n &= \sum_{a \leqslant n} \big(\lfloor \sqrt{n+a} \rfloor - \lceil \sqrt{1+a} \rceil + 1 \big) \\ &= \sum_{a \leqslant n} \big(\sqrt{n+a} - \sqrt{1+a} + O(1) \big) \\ &= \sum_{a \leqslant n} \sqrt{n+a} - \sum_{a \leqslant n} \sqrt{1+a} + O(n), \end{aligned} \tag{‡}$$

and we are led to estimating a pair of sums of the form $\sum_{a \leqslant n} \sqrt{X+a}$. Summing by parts, we see that

$$\begin{aligned} \sum_{a \leqslant n} \sqrt{X+a} &= n\sqrt{X+n} - \frac{1}{2} \int_1^n \frac{\lfloor t \rfloor}{\sqrt{X+t}} \, dt \\ &= n\sqrt{X+n} - \frac{1}{2} \int_1^n \frac{t}{\sqrt{X+t}} \, dt + O\Big(\int_1^n \frac{dt}{\sqrt{X+t}} \Big) \\ &= n\sqrt{X+n} + \frac{2X(\sqrt{X+n} - \sqrt{X+1}) - n\sqrt{X+n}}{3} + O(\sqrt{X+n}), \end{aligned}$$

and so by (‡),

$$\begin{aligned} s_n &= \left(n\sqrt{2n} + \frac{2n(\sqrt{2n} - \sqrt{n+1}) - n\sqrt{2n}}{3} \right) \\ &\qquad - \left(n\sqrt{1+n} + \frac{2(\sqrt{1+n} - \sqrt{2}) - n\sqrt{1+n}}{3} \right) + O(n) \\ &= \frac{4(n\sqrt{2n} - n\sqrt{1+n})}{3} + O(n) = \frac{4(\sqrt{2}-1)}{3} \, n\sqrt{n} + O(n), \end{aligned}$$

thus the probability is

$$\frac{4\sqrt{2}-4}{3\sqrt{n}} + O\Big(\frac{1}{n}\Big),$$

as required. □