5. Fifth Lecture : 10/11

A more careful analysis of Euclid's argument gives an explicit lower bound for $\pi(x)$.

Proposition 5.1. *If $x \geq 2$ then*

$$\pi(x) \geq \lfloor \log_2 \log_2 x \rfloor + 1. \tag{5.1}$$

Proof. Let p_n denote the n:th prime. Then it follows from Euclid's argument (see eq. (4.19)) that

$$p_{n+1} \leq \left(\prod_{i=1}^{n} p_i \right) + 1. \tag{5.2}$$

Let $p_n = 2^{a_n}$. Then substituting into (5.2) gives

$$2^{a_{n+1}} \leq 2^{\sum_{i=1}^{n} a_i} + 1, \tag{5.3}$$

which implies at the very least that

$$a_{n+1} \leq \left(\sum_{i=1}^{n} a_i \right) + 1. \tag{5.4}$$

Note also that $p_1 = 2 = 2^1$, so $a_1 = 1$. If we had equality in (5.4) then it's easy to check that the solution to the recurrence would be $a_n = 2^{n-1}$. Thus we may deduce that, in fact, $a_n \leq 2^{n-1}$.

Thus $p_n \leq 2^{2^{n-1}}$. In other words, if $x = 2^{2^{n-1}}$ then $\pi(x) \geq \log_2 \log_2 x + 1$. Then (5.1) follows for general x, since $\pi(x)$ is an increasing function of x. □

The estimate (5.1) is very, very far from the truth. For example, $2^{19} < 10^6 < 2^{20}$, so (5.1) says there are at least 20 primes up to a million. In fact, there are 78,498 primes up to a million. The computation of the function $\pi(x)$ was one of the main preoccupations of eminent mathematicians after the rebirth of number theory in the 17th and 18th centuries. A combination of numerical and heuristic evidence led people[1] to believe that $\pi(x)$ behaved something like $x/\log x$. One of the great achievements of 19th century mathematics was the rigorous proof of this fact.

To state the result concisely, it is convenient to use the following notation :

Notation. Let $f(x), g(x)$ be two real-valued functions. We write $f(x) \sim g(x)$ to denote that $\lim_{x\to\infty} f(x)/g(x) = 1$.

Now we have

Theorem 5.2. (Hadamard, de la Vallee Poussin 1896)[2]

$$\pi(x) \sim \frac{x}{\log x}. \tag{5.5}$$

[1] Notably Legendre and Gauss, amongst others, made explicit conjectures as to the behaviour of the function $\pi(x)$.

[2] These two guys seem to have worked independently of one another.

The proof of this theorem uses much of the basic machinery of real and complex analysis which was also developed throughout the 19th century, applying it to the so-called *Riemann zeta function* (which we define below). This function was actually first studied by Euler (mid 18th century), who used it to make the first real progress beyond Euclid's work on the distribution of the primes. We will present Euler's results below. The reason the zeta function is named after Riemann was because of a seminal paper he wrote on this function in 1859[3], which vastly expanded overnight the state of knowledge regarding its complex-analytical properties. It was this paper which laid the groundwork for the final assault on the prime number theorem in the latter part of the 19th century. For a comprehensive presentation of the proof, see for example the book

H. Davenport, *Multiplicative Number Theory*, Springer GTM Series.

There are two interesting historical footnotes :

1. In 1852, Chebyshev got 'close' to the prime number theorem, using only 'elementary' methods, which in this context means not using the zeta function. We will present Chebyshev's result in the next lecture. It is curious that his relatively simple techniques got so close to the PNT, but then seem to have hit a brick wall.
2. In 1949, Erdős and Selberg, in part working together, shocked most of the mathematical community by publishing 'elementary' proofs of PNT. The issue of who contributed what to their efforts has in fact been the source of a lot of controversy (they published their work separately, for example). For an account of their methods, the historical background and the controversy, see for example the article

http://www.math.columbia.edu/∼goldfeld/ErdosSelbergDispute.pdf

Note that the methods employed by Erdős and Selberg do not seem to have had a significant influence on the further development of number theory, for example on efforts to prove the *Riemann hypothesis*, which we will come back to later. Thus it is still very much an open question as to how much of 'analytic' number theory really does require the machinery of real and complex analysis, and how much knowledge about the distribution of the primes can be gleaned by other, theoretically simpler methods.

The remainder of this lecture is devoted to presenting Euler's contribution to this topic.

Definition. Let $s \in \mathbb{C}$ with Re$(s) > 1$. Define

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{5.6}$$

[3] B. Riemann, *Über die Anzahl der Primzahlen unter einer gegebenen Grösse*, Monatsberichte der Berliner Akademie (1859), 671-680. Note that the paper is only 10 pages long ! It is also the only paper Riemann ever wrote on number theory !! And, in Fermat-like fashion, it doesn't contain rigorous proofs of most of the main results !!! However, what it does contain is all the most important facts about the zeta function which were needed to prove the PNT. It also contains what has become known as the *Riemann hypothesis*.

The function $\zeta(s)$ is called the *(Riemann) zeta function*.

We all know from one-variable calculus that the sum in (5.6) is absolutely convergent when Re$(s) > 1$, so the zeta function is well-defined in this range. In fact, for any $\delta > 0$, the sum converges uniformly in the half-plane Re$(s) > 1 + \delta$, hence, by a theorem of Weierstraβ, the zeta function is analytic for Re$(s) > 1$[4]. Since the machinery of complex analysis was not properly developed during Euler's lifetime, he was probably only dimly aware of these facts, and in his work he did not make any use of the properties of $\zeta(s)$ as an ANALYTIC function. That would come later, first with Dirichlet (see Lecture 7 below), but primarily with Riemann. Euler did possess the tools to prove the following two results, however.

Theorem 5.3. *Let Re$(s) > 1$. Then*

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}. \tag{5.7}$$

In other words, the infinite product over all the primes converges absolutely in the half-plane Re$(s) > 1$ and coincides with $\zeta(s)$.

Proof. In a rigorous presentation of the proof, one must pay careful attention to issues of convergence, but here I just wish to give the central ideas. One way of proving (5.7) is to show that, for every integer $N > 0$,

$$\left[\prod_{p \leq N} \left(1 - \frac{1}{p^s}\right)\right] \cdot \zeta(s) = \sum_{p \dagger n \text{ for any } p \leq N} \frac{1}{n^s}. \tag{5.8}$$

This can be done 'one prime at a time' so to speak, by exploiting the fact, for any individual prime p,

$$\left(1 - \frac{1}{p^s}\right) \cdot \zeta(s) = \sum_{p \dagger n} \frac{1}{n^s}. \tag{5.9}$$

As $N \to \infty$ the RHS of (5.8) will converge to 1, implying (5.7).

An alternative method is to start with the infinite product and note that, by the binomial theorem, each factor can be expanded as an infinite series

$$\left(1 - \frac{1}{p^s}\right)^{-1} = \sum_{m=0}^{\infty} \frac{1}{p^{ms}}. \tag{5.10}$$

Thus the infinite product becomes

$$\prod_p \left(\sum_{m=0}^{\infty} \frac{1}{p^{ms}}\right), \tag{5.11}$$

and when one multiplies this out, one sees that, for every $n > 0$, the term $1/n^s$ appears exactly once, by FTA. □

[4] If you haven't taken a course in complex analysis, and hence don't understand the meaning of this sentence, it doesn't matter for the sake of this course.

Remark 5.4. An expression of the form

$$\prod_p f_p(p^s), \tag{5.12}$$

where each $f_p(x)$ is a rational function, is called an *Euler product*.

Corollary 5.5. (Euler) *There are infinitely many primes and the sum of their reciprocals diverges.*

Proof. Taking log of both sides of (5.7) we get, for $\mathrm{Re}(s) > 1$,

$$\log \zeta(s) = -\sum_p \log\left(1 - \frac{1}{p^s}\right). \tag{5.13}$$

Next recall that the Taylor series for the log function, valid when $|z| < 1$, is given by

$$-\log(1-z) = \sum_{m=1}^{\infty} \frac{1}{m} z^m. \tag{5.14}$$

Substituting (5.14) into (5.13) we get, also for $\mathrm{Re}(s) > 1$,

$$\log \zeta(s) = \sum_p \sum_{m=1}^{\infty} \frac{1}{mp^{ms}}. \tag{5.15}$$

Now group the terms on the RHS into two groups, those with $m = 1$ and those with $m > 1$. Note that we are changing the order of summation here, but that is okay because the series is absolutely convergent when $\mathrm{Re}(s) > 1$[5]. We obtain

$$\log \zeta(s) = \sum_p p^{-s} + \sum_p \sum_{m=2}^{\infty} \frac{1}{mp^{ms}}. \tag{5.16}$$

The idea now is to look at what happens when $s \to 1^+$. The sum over $m \geq 2$ does not blow up - in fact, it's value at $s = 1$ can easily be shown to be less than, for example, $2\zeta(2) = \pi^2/3$. On the other hand, we all know from envariabelanalys (use, say, the integral test) that $\zeta(s) \to +\infty$ as $s \to 1^+$. This means that the LHS of (5.16) goes to infinity as $s \to 1^+$. It follows that the same is true of the RHS, and hence that

$$\lim_{s \to 1^+} \sum_p p^{-s} = \infty. \tag{5.17}$$

In particular, the sum must contain infinitely many terms (i.e.: there are infinitely many primes) and $\sum p^{-1}$ diverges. This proves the corollary. □

[5] Recall from envariabelanalys that the sum of an absolutely convergent series is independent of the ordering of the terms. This is not true for conditionally convergent series (Riemann's theorem).

6. SIXTH LECTURE : 10/11

The purpose of this lecture is twofold : (1) using Euler's results, to present some heuristic arguments for PNT (2) to prove Chebyshev's 1852 result which, until Erdős and Selberg came along, was as close as anyone got to proving PNT by 'elementary' means.

FIRST HEURISTIC FOR PNT. Let p_n denote the n:th prime. Corollary 5.5 says that $\sum 1/p_n$ diverges. This certainly implies that p_n cannot grow too quickly with n. How quickly can it grow ? Well, trivially $p_n \geq n$ and $\sum 1/n$ diverges. Recall that the latter is proven by comparing with the integral

$$\int \frac{1}{x}\, dx = \log x, \tag{6.1}$$

which diverges as $x \rightarrow \infty$. Similarly, since for any $\epsilon > 0$, the integral

$$\int_1^\infty \frac{1}{x^{1+\epsilon}}\, dx = \frac{1}{\epsilon} \tag{6.2}$$

converges, this suggests that p_n can't grow like $n^{1+\epsilon}$, for arbitrarily small $\epsilon > 0$. This still leaves some room for manoeuvre, though. Let's use the following notation :

$$log^{(0)}x := x, \quad log^{(k)}x := \log(\log^{(k-1)} x) \; \forall \; k > 0. \tag{6.3}$$

Then one easily shows that, for any $k \geq 0$ (the case $k = 0$ being (6.1)),

$$\int \frac{1}{\prod_{i=0}^{k} \log^{(i)} x}\, dx = \log^{(k+1)} x. \tag{6.4}$$

Hence all these integrals diverge, which leaves the possibility, for any $k \geq 0$, that

$$p_n \sim \prod_{i=0}^{k} \log^{(k)} n. \tag{6.5}$$

Now $k = 0$ is simply unreasonable : it's not the case that 'most' numbers are prime. For purely aesthetic reasons (the Ockham's Razor principle), we might now settle on the simplest remaining alternative, namely $k = 1$, i.e.:

$$p_n \sim n \log n. \tag{6.6}$$

It's easy to check that (6.6) implies (5.5).

There are several obvious problems in turning this heuristic argument into a proof. We mention a few glaring ones :

1. There's nothing in the argument to rule out a constant multiplicative factor, say $p_n \sim 176\, n \log n$.
2. We are choosing one of infinitely many alternatives on purely aesthetic grounds.
3. More subtly, an expression like (6.6) implies some 'regularity' in the distribution of the primes. There is nothing in our heuristic which might explain such regularity in the first place and which would rule out dense clustering of the primes in certain intervals interspaced with long prime-free gaps, in such a way that (6.6) only held 'on average'.

All of these concerns can be addressed by either arguing on aesthetic grounds or by collecting large volumes of numerical evidence. Note that a couple of hundred years ago, the latter option wasn't so readily available. Hence, the reader must decide for him/herself how convincing the heuristic is, but it should be clear there are serious problems in turning it into a rigorous proof.

SECOND HEURISTIC FOR PNT. Let p be a prime. Choose a positive integer at random[6] and let A_p denote the event that the chosen integer is a multiple of p. Clearly,

$$\mathbb{P}(A_p) = \frac{1}{p} \text{ and } \mathbb{P}(\overline{A_p}) = 1 - \frac{1}{p}. \tag{6.7}$$

Now let p, q be two distinct primes. We claim that

$$\mathbb{P}(A_p \cap A_q) = \frac{1}{pq} \text{ and } \mathbb{P}(\overline{A_p} \cap \overline{A_q}) = \left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{q}\right). \tag{6.8}$$

Note that the second statement in (6.8) follows from the first by an inclusion-exclusion argument. The first statement is essentially a reformulation of FTA, which implies that an integer is divisible by both p and q if and only if it is divisible by pq[7]. For future reference, we note an elegant probabilistic interpretation of (6.8), namely

For distinct primes p and q, the events A_p and A_q are independent. (6.9)

Now back to the heuristic. Let N be a large positive integer, and imagine we choose an integer at random from among $\{1, ..., N\}$. Let E be the event that this integer is prime. The event E occurs if and only if our chosen number is not a proper multiple of any prime up to N. If we ignore the word 'proper' (thus already introducing a delicate error in our estimates) then, by (6.8), the probability of the event E is given by

$$\mathbb{P}(E) = \prod_{p \le N} \left(1 - \frac{1}{p}\right). \tag{6.10}$$

Now compare this with (5.6) and (5.7). In fact, if we truncate the series for $\zeta(s)$ at $n = N$, set $s = 1$ and wave our hands a lot we can postulate that

$$\prod_{p \le N} \left(1 - \frac{1}{p}\right) \approx \frac{1}{\sum_{n=1}^{N} 1/n.} \tag{6.11}$$

The sum is about $\log N$ (see (6.1)). Thus we have an argument which suggests that the probability of an integer from among $\{1, ..., N\}$ being prime is about $1/\log N$. Clearly, this implies something like (5.5).

[6] If you've taken a course in probability theory, you may appreciate that this statement is, strictly speaking, meaningless. However, I am going to gloss over such technicalities here.

[7] More generally, FTA implies that if GCD$(a, b) = 1$ then an integer is divisible by both a and b if and only if it is divisible by ab. Note that this fact really does rely on unique prime factorisation. For consider eq. (4.17). It implies that, in the ring $\mathbb{Z}[\sqrt{-5}]$, the number 6 is divisible by each of the primes 2 and $1 + \sqrt{-5}$. However, it is not divisible by $2(1 + \sqrt{-5})$.

This heuristic also suffers many flaws[8], though of a somewhat different nature to the first one. It is important to note that the two arguments are significantly different in spirit. The latter one can be considered as an example of 'probabilistic' reasoning, a strategy which is often very helpful (though sometimes misleading) for seeing beyond the technical difficulties of a particular problem to get a feeling for 'what's actually going on'. We will encounter more examples of this type of heuristic reasoning later in the course.

We now turn to the proof of Chebyshev's theorem. First some further notation. We write $f(x) \lesssim g(x)$ to denote that $\limsup_{x\to\infty} f(x)/g(x) \leq 1$. We write $f(x) = O(g(x))$ to denote that $\limsup_{x\to\infty} f(x)/g(x) < \infty$.

Theorem 6.1. (Chebyshev 1852) *There exist constants* $0 < c_1 < c_2$ *such that*

$$c_1 \frac{x}{\log x} \lesssim \pi(x) \lesssim c_2 \frac{x}{\log x}. \tag{6.14}$$

In the proof presented below, we will determine explicit constants $c_1 = \log 2 \approx 0.693...$, $c_2 = 2c_1 \approx 1.386...$. By refining this basic approach (at the cost of a lot of extra technical details), Chebyshev was able to give better constants, his best being $c_1 \approx 0.92...$, $c_2 \approx 1.105...$ However, his method does not seem to lead to the 'truth', i.e.: $c_1 = c_2 = 1$. We will make use of the following function :

Definition. The function $\Lambda : \mathbb{N} \to \mathbb{R}$ defined by

$$\Lambda(n) := \begin{cases} \log p, & \text{if } n > 1 \text{ and } n \text{ is a power of the prime } p, \\ 0, & \text{if } n = 1 \text{ or } n \text{ is not a prime power,} \end{cases} \tag{6.15}$$

is called the *von Mangoldt function*[9].

[8] One such flaw is that the approximation (6.11) is wrong by a constant multiplcative factor. In 1874, Mertens proved that

$$\prod_{p \leq N} \left(1 - \frac{1}{p}\right) \sim \frac{1}{e^\gamma \log N}, \tag{6.12}$$

where γ is the so-called *Euler-Mascheroni constant*

$$\gamma := \lim_{n\to\infty} \left(\sum_{i=1}^{n} \frac{1}{i} - \log n \right). \tag{6.13}$$

I might return to Mertens work later on (haven't decided yet !)

[9] This function was introduced by the German mathematician Hans von Mangoldt some time around 1878. It is an important technical tool in the proof of the PNT. More precisely, von Mangoldt studied the function

$$\psi(x) := \sum_{n \leq x} \Lambda(n), \tag{6.16}$$

which had already appeared in Chebyshev's work. It can be considered as a weighted version of the prime counting function $\pi(x)$. Actually, it also assigns non-zero weights to all prime powers, but one can show that the major contribution comes from the primes themselves. For technical reasons, the function $\psi(x)$ is easier to study by complex-analytical methods than $\pi(x)$ and the standard approach to proving PNT first of all leads to an asymptotic estimate for $\psi(x)$. From this it is not too hard to deduce an estimate for $\pi(x)$. For some further information, consult Homework 2.

Proof. of Theorem 6.1. The starting point of the proof is the identity

$$\sum_{m|n} \Lambda(m) = \log n. \tag{6.17}$$

To see this, consider a prime factorisation of n,

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}. \tag{6.18}$$

Then the definition of Λ is easily seen to imply that the LHS of (6.17) is just

$$\sum_{i=1}^{k} \alpha_i \log p_i = \log \left(\prod_{i=1}^{k} p_i^{\alpha_i} \right) = \log n,$$

which proves (6.17). Next, consider the function $T : \mathbb{R}^+ \to \mathbb{R}^+$ given by

$$T(x) = \sum_{n \leq x} \Lambda(n) \lfloor \frac{x}{n} \rfloor. \tag{6.19}$$

I claim that

$$T(x) = x \log x - x + O(\log x). \tag{6.20}$$

To show this, we first assert that

$$\sum_{n \leq x} \Lambda(n) \lfloor \frac{x}{n} \rfloor = \sum_{n \leq x} \left(\sum_{m|n} \Lambda(m) \right). \tag{6.21}$$

This identity follows from the observation that, for each $m \leq x$, the quantity $\Lambda(m)$ appears once on each side of (6.21) for every $n \leq x$ such that $m|n$. From (6.17) and (6.21) we deduce that

$$T(x) = \sum_{n \leq x} \log n = \log(\lfloor x \rfloor !) \tag{6.22}$$

Then (6.20) follows from Stirling's formula[10]. In order to complete the proof of Chebyshev's theorem, the function which will actually need to study is

$$S(x) := T(x) - 2T(x/2) = \sum_{n \leq x} \Lambda(n) \left(\lfloor \frac{x}{n} \rfloor - 2 \lfloor \frac{x}{2n} \rfloor \right). \tag{6.25}$$

On the one hand, (6.20) is easily checked to imply that

$$S(x) = (\log 2)x + O(\log x). \tag{6.26}$$

On the other hand, we shall obtain both upper and lower bounds for $S(x)$ in terms of $\pi(x)$. These will yield, respectively, the lower and upper bounds for $\pi(x)$ in (6.14). First

[10] Stirling's formula is usually presented as the statement that

$$n! \sim n^n e^{-n} \sqrt{2\pi n}. \tag{6.23}$$

Taking logarithms, this implies that

$$\log n! = n \log n - n + O(\log n). \tag{6.24}$$

This logarithmic version of Stirling's formula is actually much easier to prove than the formula itself. One can do so by a careful comparison of $\log n!$ with the integral $\int_1^n \log x \, dx$. The details are left as an exercise to the reader.

to the upper bound. One readily verifies that the bracketed difference of $\lfloor\ \rfloor$-functions in (6.25) is less than or equal to 1, for any x and n. Hence

$$S(x) \le \sum_{n\le x} \Lambda(n) = \sum_{p\le x} (\log p) \cdot \lfloor \frac{\log x}{\log p} \rfloor \le (\log x) \cdot \sum_{p\le x} 1 = (\log x)\pi(x). \tag{6.27}$$

This and (6.26) immediately imply the left-hand inequality in (6.14). The lower bound for $S(x)$ will require a little more work. First observe that if $x/2 < n \le x$ then the bracketed term in (6.25) is equal to 1. Hence

$$S(x) \ge \sum_{x/2<n\le x} \Lambda(n) \ge \sum_{x/2<p\le x} \log p \ge (\log x/2) \left[\pi(x) - \pi(x/2)\right]. \tag{6.28}$$

From this and (6.26) we deduce that

$$\pi(x) \le \pi(x/2) + (\log 2)\frac{x}{\log x - \log 2} + O(1). \tag{6.29}$$

Iterating, we have for any $k < \lfloor \log_2 x \rfloor - 1$ that

$$\pi\left(\frac{x}{2^k}\right) \le \pi\left(\frac{x}{2^{k+1}}\right) + (\log 2)\frac{x}{2^k(\log x - k\log 2)} + O(1). \tag{6.30}$$

Adding over all such k then yields

$$\pi(x) \le (c_x \log 2)\frac{x}{\log x} + O(\log x), \tag{6.31}$$

where

$$c_x = \sum_{k=0}^{\lfloor \log_2 x\rfloor - 1} \frac{1}{2^k}\frac{\log x}{\log x - k\log 2}. \tag{6.32}$$

I leave it as a (really ugly !) exercise to the reader to verify that $c_x \to 2$ as $x \to \infty$. This implies the right-hand inequality of (6.14) and completes the proof of the theorem. □

NOTATION: For two functions $f,\ g : \mathbb{R}_+ \to \mathbb{R}_+$ we write $f(x) = \Theta(g(x))$ if $f(x) = O(g(x))$ and $g(x) = O(f(x))$ both hold. We write $f(x) = o(g(x))$ if $\lim_{x\to\infty} f(x)/g(x) = 0$.

Chebyshev's theorem says that $\pi(x) = \Theta\left(\frac{x}{\log x}\right)$. In Corollary 5.5 we showed that $\sum 1/p$ diverges. We will now use Chebyshev's theorem to deduce a more precise result (Theorem 6.4 below).

DEFINITION: The function $\Psi : \mathbb{R}_+ \to \mathbb{R}_+$ is defined as

$$\Psi(x) = \sum_{n\le x} \Lambda(n). \tag{6.33}$$

The function $\Psi(x)$ is a kind of weighted version of the prime-counting function $\pi(x)$, which turns out to be useful to work with.

Proposition 6.2.

$$\Psi(x) = \Theta(x). \tag{6.34}$$

Proof. By the definition of the von Mangoldt function, we have

$$\Psi(x) = \sum_{p \leq x} \log p + \sum_{p^k \leq x,\, k \geq 2} \log p. \tag{6.35}$$

It is easy to see that the second sum is $O(\sqrt{x}\,\log x) = o(x)$, so it suffices to show that the first sum is $\Theta(x)$. On the one hand,

$$\sum_{p \leq x} \log p \leq \sum_{p \leq x} \log x = (\log x) \sum_{p \leq x} 1 = (\log x)\pi(x) = \Theta(x), \tag{6.36}$$

by Theorem 6.1. On the other hand,

$$\sum_{p \leq x} \log p \geq \sum_{\sqrt{x} < p \leq x} \log p > \left(\frac{1}{2} \log x\right)(\pi(x) - \pi(\sqrt{x})) = \Theta(x), \tag{6.37}$$

also by Theorem 6.1. This proves the proposition. □

Remark 6.3. The full Prime Number Theorem implies by the same argument that $\Psi(x) \sim x$. In fact the usual analytic proof of PNT proceeds the other way round, by first establishing the estimate for $\Psi(x)$ and then deducing that for $\pi(x)$.

The classical "good" estimate for the partial sum of the reciprocals of the primes is the following:

Theorem 6.4.

$$\sum_{p \leq x} \frac{1}{p} = \log \log x + b + O\left(\frac{1}{\log x}\right), \tag{6.38}$$

where b is some constant.

Proof. The proof involves studying another function $L : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ defined as

$$L(x) = \sum_{p \leq x} \frac{\log p}{p}. \tag{6.39}$$

Note that the relationship between $L(x)$ and the sum in (6.38) is somehat akin to that between $\Psi(x)$ and $\pi(x)$. The precise relationship between the two will be given in Claim 2 below. First we verify

CLAIM 1:

$$L(x) = \log x + O(1). \tag{6.40}$$

Proof of Claim 1: We return to the function $T(x)$ defined in (6.19). Let $\{t\}$ denote the fractional part of a real number t. Obviously, $0 \leq \{t\} < 1$ for every t. Thus,

$$T(x) = \sum_{n \leq x} \Lambda(n) \left(\frac{x}{n} - \left\{\frac{x}{n}\right\}\right) = x \left(\sum_{n \leq x} \frac{\Lambda(n)}{n}\right) + O\left(\sum_{n \leq x} \Lambda(n)\right). \tag{6.41}$$

The second bracketed sum is what we defined as $\Psi(x)$ and hence is $\Theta(x)$ by Proposition 6.2. Dividing (6.41) across by x and using (6.20), we find that

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \log x + O(1). \tag{6.42}$$

But, by definition of the functions Λ and L,

$$\sum_{n\le x}\frac{\Lambda(n)}{n}=L(x)+\sum_{p^k\le x,\,k\ge 2}\frac{\log p}{p^k}. \tag{6.43}$$

The second sum is certainly bounded by the same eexpression, but summing instead over all prime powers p^k, not just those up to x. But the latter is clearly convergent since

$$\sum_{p^k,\,k\ge 2}\frac{\log p}{p^k}=\sum_p \log p\sum_{k=2}^{\infty}\frac{1}{p^k}=\sum_p\frac{\log p}{p(p-1)}\le\sum_{n=1}^{\infty}\frac{\log n}{n(n-1)}=O(1). \tag{6.44}$$

Claim 1 follows from (6.42)-(6.44).

CLAIM 2:

$$\sum_{p\le x}\frac{1}{p}=\frac{L(x)}{\log x}+\int_2^x\frac{L(u)}{u(\log u)^2}\,du. \tag{6.45}$$

To prove this, we start with the integral and use the definition of L to write

$$\int_2^x\frac{L(u)}{u(\log u)^2}\,du=\int_2^x\left(\sum_{p\le x}\frac{\log p}{p}\right)\frac{1}{u(\log u)^2}\,du. \tag{6.46}$$

We can interchange the sum and the integral and use the elementary fact that $\int\frac{du}{u(\log u)^2}=-\frac{1}{\log u}+C$ to obtain

$$\begin{aligned}\int_2^x\left(\sum_{p\le x}\frac{\log p}{p}\right)\frac{1}{u(\log u)^2}\,du&=\sum_{p\le x}\frac{\log p}{p}\int_p^x\frac{du}{u(\log u)^2}= \\ &=\sum_{p\le x}\frac{\log p}{p}\left(\frac{1}{\log p}-\frac{1}{\log x}\right)=\sum_{p\le x}\frac{1}{p}-\frac{L(x)}{\log x},\end{aligned} \tag{6.47}$$

which proves Claim 2.

Let $E(x):=L(x)-\log x$. Claim 1 asserts that $E(x)=O(1)$. Substituting this into (6.45) we have

$$\sum_{p\le x}\frac{1}{p}=1+\frac{E(x)}{\log x}+\int_2^x\frac{1}{u(\log u)}\,du+\int_2^x\frac{E(u)}{u(\log u)^2}\,du. \tag{6.48}$$

The first integral is $\log\log x-\log\log 2$. Hence if we let

$$b:=1-\log\log 2+\int_2^{\infty}\frac{E(u)}{u(\log u)^2}\,du, \tag{6.49}$$

then

$$\sum_{p\le x}\frac{1}{p}=\log\log x+b+\mathcal{E}(x), \tag{6.50}$$

where

$$\mathcal{E}(x)=\frac{E(x)}{\log x}+\int_x^{\infty}\frac{E(u)}{u(\log u)^2}\,du. \tag{6.51}$$

Since $E(x) = O(1)$ it is clear that $\mathcal{E}(x) = O\left(\frac{1}{\log x}\right)$, so the proof of Theorem 6.4 is complete. □

Remark 6.5. See the lecture notes for the course in Arithmetic Combinatorics for a nice application of Theorem 6.4.

From Theorem 6.4 and following the same strategy as in the proof of Corollary 5.5, it is easy to deduce that

$$\prod_{p \leq x}\left(1 - \frac{1}{p}\right) = \frac{c}{\log x}\left(1 + O\left(\frac{1}{\log x}\right)\right), \tag{6.52}$$

for some constant $c > 0$. It requires more work though to show that $c = e^{-\gamma}$, see (6.12). We will not do this here, but see for example Chapter 8 of Niven-Zuckerman-Montgomery.

7. Seventh Lecture : 12/11

The next few lectures will involve somewhat more algebra again. First, let us remind ourselves of some standard facts and notations with which we are hopefully all familiar.

Let $n \in \mathbb{N}$. We define the relation 'congruence modulo n' on $\mathbb{Z}$ by

$$a \equiv b \ (\mathrm{mod}\, n) \quad \Leftrightarrow \quad n | a - b. \tag{7.1}$$

This is clearly an equivalence relation. There are n equivalence classes, represented most naturally by the numbers $0, 1, ..., n-1$. The set of equivalence classes is denoted $\mathbb{Z}/n\mathbb{Z}$ or $\mathbb{Z}_n$. I will primarily use the latter notation. It is called the set of *congruence/residue classes* modulo n. The most important basic fact about the sets $\mathbb{Z}_n$ is that they inherit the structure of an algebraic ring from $\mathbb{Z}$:

Proposition 7.1. *Addition and multiplication of congruence classes modulo n is well-defined, so that $\mathbb{Z}_n$ is a commutative ring with unity. In other words, if $a, b, c, d \in \mathbb{Z}$ satisfy $a \equiv c \ (mod\, n)$ and $b \equiv d \ (mod\, n)$, then also*

$$a + b \equiv c + d \ (mod\, n) \ \ \textit{and} \ \ ab \equiv cd \ (mod\, n). \tag{7.2}$$

Now we start a new track for our investigations by noting that a slight modification of Euclid's proof of Theorem 1.2 yields a stronger result, namely :

Theorem 7.2. *Let $n > 2$. Then there are infinitely many primes not congruent to $1 \ (mod\, n)$.*

Proof. Suppose the contrary and let $p_1, ..., p_k$ be a full list of the primes not congruent to 1 (mod n). Consider the number

$$T := n \left(\prod_{i=1}^{k} p_i \right) - 1. \tag{7.3}$$

Then T is clearly not divisible by any p_i. Also $T \equiv -1 \ (\mathrm{mod}\, n)$, thus $T \not\equiv 1 \ (\mathrm{mod}\, n)$, since $n > 2$. It follows from Proposition 7.1 that at least one prime factor of T cannot be congruent to 1 (mod n) either. This contradicts the completeness of our list. □

Corollary 7.3. *(i) There are infinitely many primes congruent to 2 (mod 3).*
(ii) There are infinitely many primes congruent to 3 (mod 4).
(iii) There are infinitely many primes congruent to 5 (mod 6).

For $n = 5$ or $n > 6$, Theorem 7.2 does not tell us whether there are infinitely many primes in any particular congruence class modulo n. It also doesn't tell us anything at all about primes congruent to 1 (mod n). Now some congruence classes obviously cannot be full of primes, for

Proposition 7.4. *If $n \in \mathbb{N}$ and $a \in \mathbb{Z}$ satisfy GCD$(a, n) > 1$, then there is at most one prime congruent to $a \ (mod\, n)$, namely GCD(a, n) itself, if this happens to be a prime.*

It seems reasonable to expect that nothing else can go wrong, and this is indeed the case :

Theorem 7.5. (Dirichlet 1829) *Let $n \in \mathbb{N}$ and $a \in \mathbb{Z}$ satisfy GCD$(a, n) = 1$. Then there are infinitely many primes congruent to a (mod n). In fact, the sum of the reciprocals of these primes diverges, for any a and n.*

Dirichlet's theorem is regarded as the first major success of methods which nowadays would be referred to as 'analytic number theory'. His proof (which is highly non-trivial : see Davenport's book for a full treatment) is loosely based on the techniques of Euler, but he needed to introduce a generalisation of the zeta function, nowadays known as *Dirichlet L-functions*. I will define these later on, to give you a jumping-off point in case you want to study Dirichlet's proof.

Remark 7.6. As a companion to his elementary proof of the PNT, Selberg gave an elementary proof of Dirichlet's theorem. Both proofs are in Volume 50 of the Annals of Mathematics (1949).

Dirichlet's theorem still leaves a deeper question unanswered, namely : is it the case that, for a fixed modulus n, the primes are *equidistributed* amongst the congruence classes $a \pmod n$ for which GCD$(a, n) = 1$. The answer is yes, and was proven by applying to L-functions the same methods used to prove the PNT. Before stating the result, we remind ourselves of another piece of standard notation :

Notation. The *Euler ϕ-function* is the function $\phi : \mathbb{N} \rightarrow \mathbb{N}$ defined by

$$\phi(n) := \#\{a \in \{0, 1, ..., n-1\} : \text{GCD}(a, n) = 1\}. \tag{7.4}$$

Then we have

Theorem 7.7. (Extended PNT) *Let $n \in \mathbb{N}$ and $a \in \mathbb{Z}$ satisfy GCD$(a, n) = 1$. For $x > 0$ let $\pi_{a,n}(x)$ denote the number of primes up to x which are congruent to a (mod n). Then*

$$\pi_{a,n}(x) \sim \frac{1}{\phi(n)} \frac{x}{\log x}. \tag{7.5}$$

The proofs of Theorems 7.5 and 7.7 are beyond the scope of this course. I want to prove at least one special case of the former, however[11], namely that there are infinitely many primes congruent to 1 (mod 4). We need some more algebraic background for this and much else.

Basically, I am interested in the structure of $\mathbb{Z}_n$ as a ring and of that of $\mathbb{Z}_n^*$, the subset of multiplicative units, as a (multiplicative) group. Note that, by Euclid's Lemma, $|\mathbb{Z}_n^*| = \phi(n)$. The ring structure is fairly straightforward :

Proposition 7.8. *Let $n \in \mathbb{N}$ have prime factorisation $n = \prod_{i=1}^{k} p_i^{\alpha_i}$. Then there is an isomorphism of rings*

$$\mathbb{Z}_n \cong \prod_{i=1}^{k} \mathbb{Z}_{p_i^{\alpha_i}}, \tag{7.6}$$

where the RHS denotes a direct product of the rings $\mathbb{Z}_{p_i^{\alpha_i}}$, $i = 1, ..., k$.

[11] There is a book in the library, *Elementary Theory of Numbers* by Wacław Sierpiński, which contains 'elementary' proofs of a number of other special cases.

Proof. There is a natural map f from the LHS to the RHS of (7.6), namely

$$f[x \text{ (mod } n)] = [x \text{ (mod } p_1^{\alpha_1}), \cdots, x \text{ (mod } p_k^{\alpha_k})]. \tag{7.7}$$

It is trivial that f respects the operations of addition and multiplication in the respective rings and is thus a ring homomorphism. To prove it is an isomorphism, it just remains to show it is a bijection of sets. Since it is a map between finite sets, it even suffices to show that f is injective. And since we already know it is a ring homomorphism, it suffices to show that ker$(f) = \{0\}$. So suppose $x \text{ (mod } n)$ is in the kernel of f. This means, by definition, that $x \equiv 0 \text{ (mod } p_i^{\alpha_i})$, for $i = 1, ..., k$. Thus $p_i^{\alpha_i}$ divides x for each $i = 1, ..., k$ and, by FTA (see Footnote 7 in Lecture 6), this implies that $\prod_{i=1}^{k} p_i^{\alpha_i}$, namely n, also divides x. Thus $x \equiv 0 \text{ (mod } n)$, as required. □

Remark 7.9. The fact that the map f above is a bijection of sets can be formulated more concretely as follows : Let $p_1, ..., p_k$ be distinct primes, $\alpha_1, ..., \alpha_k$ non-negative integers and $a_1, ..., a_k$ any integers. Let $n = \prod_{i=1}^{k} p_i^{\alpha_i}$. Then there is a unique solution $x \in \{0, 1, ..., n-1\}$ to the system of congruences

$$x \equiv a_i \text{ (mod } p_i^{\alpha_i}), \quad i = 1, ..., k. \tag{7.8}$$

This way of formulating Proposition 7.8 is what is usually referred to as the *Chinese Remainder Theorem*.

The structure of $\mathbb{Z}_n^*$ as a multiplicative abelian group is more interesting. Note, to begin with, that a immediate consequence of (7.6) is that

$$\mathbb{Z}_n^* \cong \prod_{i=1}^{k} \mathbb{Z}_{p_i^{\alpha_i}}^*, \tag{7.9}$$

where the RHS now denotes a direct product of abelian groups. This reduces the study of the structure of $\mathbb{Z}_n^*$ to the case where n is a prime power. The most important case then is when n is actually prime, and the following is a fundamental result in abstract algebra :

Theorem 7.10. *Let p be a prime. Then $\mathbb{Z}_p^*$ is a cyclic group.*

Proof. Assuming you have taken some course in abstract algebra, this is a result whose proof you should have seen already, so I only want to remind you of the outline of it. The theorem is a special case of the fact that the multiplicative group of non-zero elements in any finite field is cyclic. This is a consequence of the following two facts :

1. In a finite field, for any $k \in \mathbb{N}$ the equation $x^k = 1$ has at most k solutions. This fact is easily established, since we can consider the equation as a polynomial equation and, since we're working in a field, the number of roots of a polynomial cannot exceed its degree.
2. Let G be a finite multiplicative group. If, for each $k \in \mathbb{N}$, there are at most k solutions in G to the equation $x^k = 1$, then G is cyclic. Note that this statement holds for all finite groups, though it's easier to prove for abelian groups, which is the only case we need here. The point is, if G is a non-cyclic, but abelian finite group, then the Fundamental Theorem for Finite Abelian Groups is easily seen to imply that there must

be some prime p for which G contains a subgroup isomorphic to $C_p \times C_p$. Then already in this subgroup, we have $p^2 - 1 > p$ solutions to $x^p = 1$. □

Teminology. Let p be a prime. A generator of the cyclic group $\mathbb{Z}_p^*$ is called a *primitive root modulo* p. For example, 3 is a primitive root modulo 7, since (mod 7),

$$3^1 \equiv 3,\ 3^2 \equiv 2,\ 3^3 \equiv 6,\ 3^4 \equiv 4,\ 3^5 \equiv 5,\ 3^6 \equiv 1. \tag{7.10}$$

Let C_r denote an abstract cyclic group of order r. For general prime powers, we have the following result :

Theorem 7.11. *Let p be a prime and $\alpha \in \mathbb{N}$. If p is odd then $\mathbb{Z}_{p^\alpha}^*$ is cyclic. If $p = 2$ and $\alpha \geq 2$ then $\mathbb{Z}_{2^\alpha}^* \cong C_2 \times C_{2^{\alpha-2}}$, the direct product of cyclic groups of orders 2 and $2^{\alpha-2}$. Moreover, the two factors are always generated by -1 and 5.*

Proof. The proof is technical and uninspiring and I did not present it at the lecture. However, in case you are interested, here it is for completeness. I will leave out some of the more gory details of the calculations for you to check yourself.

First suppose p is odd. Let g be a primitive root modulo p (which exists by Theorem 7.10). We shall show that for an appropriate choice of an integer x, the integer $g + px$ is a primitive root modulo p^α for every $\alpha > 1$. It is required to choose x such that

$$(g + px)^d \equiv 1 \pmod{p^\alpha} \quad \Rightarrow \quad p^{\alpha-1}(p-1) \mid d. \tag{7.11}$$

Note that the order of $g + px$ modulo p^α divides $p^{\alpha-1}(p-1)$ a priori, since the order of any element in a group divides the group order (Lagrange's Theorem).

First, for any choice of x, the fact that g is a primitive root modulo p already implies that $p - 1$ must divide d, since $g + px \equiv g \pmod{p}$ and so

$$(g + px)^d \equiv 1 \pmod{p^\alpha} \Rightarrow (g+px)^d \equiv 1 \pmod{p} \Leftrightarrow g^d \equiv 1 \pmod{p} \Leftrightarrow p - 1 \mid d. \tag{7.12}$$

Since $g^{p-1} \equiv 1 \pmod{p}$, we have $g^{p-1} = 1 + py$ for some integer y. The binomial theorem states that

$$(g + px)^{p-1} = \sum_{i=0}^{p-1} \binom{p-1}{i} g^{p-1-i}(px)^i. \tag{7.13}$$

Modulo p^2 only the terms $i = 0, 1$ contribute, and we have that $(g + px)^{p-1} = 1 + pz$ where

$$z \equiv y + (p-1)g^{p-2}x \pmod{p}. \tag{7.14}$$

Since the coefficient of x in (7.14) is not divisible by p, we can choose x such that z is not divisible by p. We now claim that this is sufficient for $g + px$ to be the required primitive root. It needs to be shown that, if GCD$(z, p) = 1$, then

$$(1 + pz)^{p^m} \equiv 1 \pmod{p^\alpha} \quad \Rightarrow \quad \alpha - 1 \leq m. \tag{7.15}$$

Once again, this follows immediately from the binomial theorem, which in this case states that

$$(1 + pz)^{p^m} = \sum_{i=0}^{p^m} \binom{p^m}{i} (pz)^i. \tag{7.16}$$

Since $(z, p) = 1$, one sees immediately that p^{m+1} is the highest power of p dividing the $i = 1$ term. With a little more care one checks that, since p is odd, p^{m+2} divides each term for $i > 1$. Hence

$$(1 + pz)^{p^m} \equiv 1 + p^{m+1} \ (\text{mod } p^{m+2}), \quad \text{for any } m > 0. \tag{7.17}$$

And (7.15) follows immediately from (7.17). This completes the proof of the theorem for odd p.

Now suppose $p = 2$. The argument is similar to the above, in particular the binomial theorem is used. We omit details, but just note that, in order to prove the theorem for $\alpha > 3$ (it may be proven for $\alpha \leq 3$ by inspection), one writes $5 = 1 + 2^2$ and uses the binomial theorem to prove that

$$(1 + 2^2)^{2^{\alpha-2}} \equiv 1 \ (\text{mod } 2^\alpha), \tag{7.18}$$

$$(1 + 2^2)^{2^{\alpha-3}} \equiv 1 + 2^{\alpha-1} \ (\text{mod } 2^\alpha). \tag{7.19}$$

Eq. (7.18) implies that the cyclic subgroup of $\mathbb{Z}^*_{2^\alpha}$ generated by 5 has order $2^{\alpha-2}$, and (7.19) implies that -1 is not an element of this subgroup. Then elementary group theory implies that $\mathbb{Z}^*_{2^\alpha}$ is the internal direct product of the subgroups generated by -1 and 5. □

Remark 7.12. Note, in particular, that $\mathbb{Z}^*_{2^\alpha}$ is non-cyclic for all $\alpha \geq 3$. For example, $\mathbb{Z}^*_8$ is isomorphic to the Klein-4 group $C_2 \times C_2$. Observe that $\mathbb{Z}^*_8 = \{1, 3, 5, 7\}$ and check directly that $1^2 \equiv 3^2 \equiv 5^2 \equiv 7^2 \equiv 1$ (mod 8).

We shall start applying the above results next day.

8. Eighth Lecture : 14/11

Proposition 8.1. *Let $n \in \mathbb{N}$. Then*

$$\phi(n) = n \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right), \tag{8.1}$$

where the product is taken over the distinct prime divisors of n (i.e.: each distinct prime divisor is counted only once).

Proof. Let $n = \prod_{i=1}^{k} p_i^{\alpha_i}$. It follows from (7.9) that

$$\phi(n) = \prod_{i=1}^{k} \phi(p_i^{\alpha_i}). \tag{8.2}$$

Thus, in order to prove (8.1), it suffices to show that, if $n = p^\alpha$ is a prime power then

$$\phi(p^\alpha) = p^\alpha \left(1 - \frac{1}{p}\right) = p^\alpha - p^{\alpha-1}. \tag{8.3}$$

But this is clear, since an integer is relatively prime to p^α if and only if it is not divisible by p. Since every p:th integer is a multiple of p, it follows that $1 - \frac{1}{p}$ of the p^α integers among $\{0, 1, ..., p^\alpha - 1\}$ are relatively prime to p. □

Remark 8.2. A function $f : \mathbb{N} \to \mathbb{C}$ is said to be *multiplicative* if

$$f(ab) = f(a)f(b), \quad \text{whwnever GCD}(a, b) = 1. \tag{8.4}$$

Thus Proposition 8.1 implies, in particular, that the Euler ϕ-function is multiplicative.

Eq. (8.1) implies that, in order to compute $\phi(n)$, it suffices to factorise n. Hence, determination of the ϕ-function is certainly no more computationally challenging than integer factorisation. As far as I am aware, it is still an open problem to prove the converse. One needs to be more precise as to what one actually means here, but one way of posing the problem is as follows :

Question. Can computation of the ϕ-function be reduced to integer factorisation in polynomial time ? In other words, assuming one has an infinitely fast algorithm for computing ϕ, is there a polynomial time algorithm which takes an integer n as input and outputs the prime factorisation of n ?

Note that it is quite easy to show that the answer is 'yes', if it is known that the input n is a product of exactly two distinct primes (see Homework 2). This is the situation which arises in RSA cryptography, for example, where the security of the cryptosystem is, strictly speaking, dependent on the difficulty of computing $\phi(n)$ for such n. Hence, the security is indeed dependent on the difficulty of integer factorisation.

Remark 8.3. It also follows from (8.1) that the quotient $\phi(n)/n$ can be become arbitrarily small (see Homework 2). Clearly, it can also become arbitrarily close to 1, since $\phi(p) = p - 1$ when p is a prime. It's an interesting question to ask what the 'average' behaviour of the quotient $\phi(n)/n$ is. One very nice result in this direction is that

$$\sum_{n \leq x} \phi(n) = \frac{3}{\pi^2} x^2 + O(x \log x). \tag{8.5}$$

For a proof of this and similar results, see either the Supplementary Notes to this lecture or, for example, Chapter 2 of the book *A Concise Introduction to the Theory of Numbers*, by Alan Baker. Eq. (8.5) has the nice interpretation that the probability of a pair of 'randomly chosen integers'[12] being relatively prime is $6/\pi^2$.

Lemma 8.4. *If G is a finite group and $x \in G$, then $x^{|G|} = 1$.*

Proof. Lagrange's Theorem states that if H is a subgroup of G, then $|H|$ divides $|G|$. Applying this to the cyclic subgroup generated by x, we conclude that $x^n = 1$ for some n dividing $|G|$. Hence $x^{|G|} = 1$ also. □

Proposition 8.5. *Let $n \in \mathbb{N}$ and $a \in \mathbb{Z}$ such that GCD$(a, n) = 1$. Then*

$$a^{\phi(n)} \equiv 1 \ (\textit{mod}\ n). \tag{8.6}$$

In particular, if p is a prime and a is not a multiple of p, then

$$a^{p-1} \equiv 1 \ (\textit{mod}\ p). \tag{8.7}$$

Proof. Apply Lemma 8.4 to the group $\mathbb{Z}_n^*$. □

Remark 8.6. Eq. (8.6) is usually designated *Euler's Theorem* and the special case (8.7) referred to as *Fermat's (Little) Theorem*.

For the remainder of this lecture, we shall discuss the subject of *primality testing*. The problem is to find an efficient algorithm for deciding whether an input $n \in \mathbb{N}$ is prime or not. Clearly, primality testing is no more difficult than factorisation, but it is certainly conceivable that it might be easier. If so, that would be a significant finding, since factorisation seems to be a hard problem. Efficient primality testing was an issue of great concern already to people like Euler, Lagrange, Gauss etc., who were interested for example in compiling long lists of primes and thereby investigating numerically the behaviour of the function $\pi(x)$. Since they didn't have computers, efficiency was at a premium.

One of the oldest ideas for testing primality which by-passes the need to factorise a number is to use (8.7). Given an input n, the strategy can be summarised as follows :

1. Pick a random $x \in \{1, ..., n-1\}$. Compute GCD(x, n) using Euclid's algorithm. If GCD$(x, n) > 1$, then n is not prime. Otherwise go to step 2.
2. Compute x^{n-1} (mod n). If the answer is not 1 (mod n), then n is not prime. Otherwise, pick another random number $y \in \{1, ..., n-1\}$ and go back to step 1.
3. If, after a 'large' number of trials, we still have not been able to conclude that n is prime, then abort the algorithm and output that n is prime.

There is an obvious problem with this strategy : it can give the wrong answer ! The problem is we don't know 'how many' trials have to fail before we can be sure that n is indeed prime. This matter requires a much deeper analysis, but it turns out that nowadays there exist very fast, *non-deterministic* primality tests, which are essentially based on Fermat's Little Theorem (though not so simple as the one described above). What

[12] Again, one has to be more precise about what one actually means here, but I will leave that as an exercise for you to figure out yourselves.

'non-deterministic' means is that the tests involve some randomness and give the wrong answer with a very small, but non-zero probability. For all practical purposes, state-of-the-art primality tests are foolproof, the probability of an error being so small that one would not expect one to be made during the lifetime of the universe, for example !

But important theoretical considerations remain. First of all, it is important to note that the simple test described above can fail spectacularly :

Definition. A composite number $n \in \mathbb{N}$ is called a *Carmichael number* or *Euler pseudoprime* if $x^{n-1} \equiv 1 \pmod{n}$ whenever GCD$(x, n) = 1$.

Example. 561 is a Carmichael number. We have $561 = 3 \cdot 11 \cdot 17$. Thus, by eq. (7.9) and Theorem 7.10,

$$\begin{aligned} \mathbb{Z}_{561}^* \cong \mathbb{Z}_3^* \times \mathbb{Z}_{11}^* \times \mathbb{Z}_{17}^* &\cong C_2 \times C_{10} \times C_{16} \qquad (8.8) \\ \cong C_2 \times (C_2 \times C_5) \times C_{16} \cong (C_2 \times C_2) \times (C_5 \times C_{16}) &\cong C_2 \times C_2 \times C_{80}, \end{aligned}$$

where we have used the fact[13] that $C_r \times C_s \cong C_{rs}$ whenever GCD$(r, s) = 1$.

By (8.8), the group $\mathbb{Z}_{561}^*$ has *exponent* 80, i.e.: $x^{80} \equiv 1 \pmod{561}$ for all $x \in \mathbb{Z}_{561}^*$ and 80 is the smallest positive integer for which this is the case. Since 80 divides $561 - 1 = 560$, it follows that 561 is a Carmichael number.

It turns out that there are infinitely many Carmichael numbers, though this wasn't proven until 1994 ! See the Wikipedia article on them for more information[14].

Hence, simple-minded primality tests based on Fermat's Little Theorem, like the one described above, cannot be fully deterministic, though they are certainly fast. Indeed we have already shown in an earlier lecture that Euclid's algorithm runs in polynomial time, and so does a computation of the form : given $a, b, c \in \mathbb{N}$, compute $a^b \pmod{c}$. A fast algorithm for performing this latter computation is the so-called *square and multiply algorithm*[15].

So the important remaining theoretical question is : Does there exist a polynomial time primality test which is fully deterministic ? The answer is 'yes'. More precisely,

1. In 1976, G.L. Miller and M.O. Rabin presented such a test. However, in order to prove that it ran in polynomial time, they needed to assume the so-called *Generalised Riemann Hypothesis*. This is a generalisation of the classical Riemann Hypothesis to Dirichlet L-functions. For a presentation of the Miller-Rabin algorithm, and many other interesting computational problems, see the book

N. Koblitz, *A Course in Number Theory and Cryptography,* Springer GTM Series.

[13] Go back and check your notes from some abstract algebra course if this fact confuses you.

[14] Including an estimate for the number $C(x)$ of Carmichael numbers up to x which involves the function $\log \log \log x$!!

[15] I did an example at the lecture, and I think it would be too tedious to rehash it again here.

2. In 2002, two Indian CS students presented as part of their M.Sc. thesis, along with their advisor, the first deterministic primality test which could be proved unconditionally to run in polynomial time. In practice, their algorithm is too slow (despite improvements in the intervening years) to compete with state-of.the-art probabilistic algorithms, but it is a historical theoretical breakthrough since it implies that primality testing is in the class $\mathscr{P}$. A description of their method, known as the *AKS algorithm*, can be found in

M. Agrawal, N. Kayal and N. Saxena, *Primes in* $\mathscr{P}$, Ann. Math. (2) **160** (2004), 781-793.