TOPICS IN NUMBER THEORY (PART B)

BEN GREEN

Contents

1. Distribution of primes and Mertens' estimates 3
2. Arithmetic functions 7
3. The Riemann ζ-function 15
4. Dirichlet characters and L-functions. 18
5. Binary quadratic forms 25
6. Sums of four squares 33
7. Irrationals and transcendental numbers, approximation 35
8. Continued fractions 38
9. Continued fractions and approximation by rationals 44
10. Periodic continued fractions and quadratic irrationals 46
Appendix A. Results from real analysis 47
Appendix B. Results from complex analysis 49
Appendix C. Real quadratic characters 49
References 50

Overview. Notes for Part B Topics in Number Theory, a new course being taught for the first time in 2026/7. The synopsis below is the same as the one advertised, up to ordering, except for items explicitly added in red. These were added during the process of writing the course.

Synopsis.

Elementary distribution of primes. Upper and lower bounds on $\pi(X)$ (Chebyshev bounds). Mertens' estimates [2 lectures]

Arithmetic functions and their distribution. Möbius inversion and Dirichlet convolution. The functions τ (number of divisors), ω (number of prime factors) and ϕ (Euler's totient function). Mean value of τ and the hyperbola method. Hardy-Ramanujan law and normal order of ω and τ. Divisor bound, mean value of ϕ and the probability that random integers are coprime. Introduction to Dirichlet series. [3 lectures]

The Riemann ζ-function and Euler's proof of infinitely many primes. Analytic continuation to $\mathrm{Re}\, s > 0$ and statement of the Riemann Hypothesis. Discussion of the prime number theorem and consequences of the Riemann

Hypothesis. Dirichlet characters and L-functions. Dirichlet's proof that there are infinitely many primes $a \mod q$ when a, q are coprime, assuming $L(1,\chi) \neq 0$. Proof that $L(1,\chi) \neq 0$ in special cases. [3 lectures]

Binary quadratic forms and sums of squares. Basic reduction theory of positive definite binary forms, class number. Review of quadratic reciprocity. Representation of primes by positive definite binary quadratic forms. Brief discussion on the Kronecker symbol, quadratic characters and the class number formula. Lagrange's theorem on sums of four squares. [3 lectures]

Irrationality, approximation and continued fractions. Continued fractions of rationals and Euclid's algorithm. Dirichlet's theorem. Best approximations. Continued fractions of quadratic irrationals. Pell's equation. Irrationality of e. Construction of a transcendental number. [5 lectures]

Reading list. Full notes will be provided for the course but there are several excellent books which provide extra discussion and context. Perhaps the most relevant are

Davenport, H. *The higher arithmetic*, 8th edition, Cambridge University Press 2008.

Hardy, G. H. and Wright, E. M. *An introduction to the theory of numbers*, 6th edition, Oxford University Press 2008.

Further reading. The following further texts or parts of text are at an appropriate level for the course and provide further context.

A nice introduction to probabilistic number theory is the recent book

Kowalski, E. *An introduction to probabilistic number theory*, Cambridge Studies in Advanced Mathematics **192**, Cambridge University Press 2021.

Another recommended book is

Erdős, P. and Surányi, J., *Topics in the theory of numbers*, Undergraduate Texts in Mathematics, Springer-Verlag 2003.

Another highly-recommended text is Serre's classic

Serre, J.-P. *A course in arithmetic*, Graduate Texts in Mathematics **7**, Springer-Verlag 1973

Montgomery and Vaughan's book *Multiplicative Number Theory I. Classical Theory.* Cambridge Advanced Studies in Math **97** is a useful text for some of the details.

Sheet X. Sheet X is an exercise sheet which goes over the details of various items we will use without proof in the course, for instance basic analytic estimates for sums. These are all the sort of things at the level of prelims or Part A courses. To follow the present course, all you need is the statements of these results, but it should be helpful to refresh your memory of the proofs.

Notation. If m, n are two positive integers, then I like to write (m, n) for the highest common factor of m and n. This is pretty standard notation, and is almost ubiquitous in research papers. Some authors write $\mathrm{hcf}(m, n)$ or $\gcd(m, n)$; you may find this in past exam papers.

Although we will not use it in this course, I write $[m, n]$ for the lowest common multiple of m and n. Some authors use $\mathrm{lcm}(m, n)$ or, worse still, $l.c.m.(m, n)$.

The notation $\lfloor x \rfloor$ 'the floor of x' means the greates integer less than or equal to x. For example, $\lfloor \pi \rfloor = 3$.

Throughout the course we will use number 1 to mean the indicator of some event, for example $1_{n \leqslant N}$ equals 1 if $n \leqslant N$ and 0 otherwise.

In analytic number theory, $\log x$ always means natural log and we never write $\ln x$.

An important part of this course will be introducing so-called *asymptotic notation*: the big-oh notation $O(\cdot)$ and similar things. This has its own section, Section 1 below.

I use the abbreviations LHS and RHS for left-hand side and right-hand side respectively.

Use of AI tools. I have written these notes myself using published resources and books, as I always have done. Completed versions of each section were put through an LLM to pick up typos and minor inaccuracies, which were then corrected. Suggestions by the LLM to change content or exposition were ignored.

1. Distribution of primes and Mertens' estimates

Our first topic of discussion is the distribution of prime numbers. The main highlights of this section are bounds for $\pi(X)$, the number of primes less than or equal to X. We will also prove two estimates of Mertens, the second of which gives an asymptotic formula for $\sum_{p \leqslant X} 1/p$. (Whenever we write $\sum_p$ it will mean a sum over primes.)

Chebyshev bounds. Our first main result is that the number of primes $\leqslant X$ is on the order of $X/\log X$.

Theorem 1.1. *There are absolute constants $0 < c_1 < c_2$ such that*

$$c_1 \frac{X}{\log X} \leqslant \pi(X) \leqslant c_2 \frac{X}{\log X}$$

for all $X \geqslant 2$.

Proof. We begin with the lower bound. We look at the prime factorisation of $\binom{2n}{n}$, which we write as

$$\binom{2n}{n} = \prod_{p \leqslant 2n} p^{w_p(n)}.$$

It follows from Legendre's formula for the number of powers of p dividing $m!$ (see Sheet X) and the definition of binomial coefficients that

$$w_p(n) = \sum_{i=1}^{\infty} \Big\lfloor \frac{2n}{p^i} \Big\rfloor - 2\Big\lfloor \frac{n}{p^i} \Big\rfloor. \tag{1.1}$$

Now each term in (1.1) is at most 1, since $\lfloor 2x \rfloor - 2\lfloor x \rfloor \in \{0, 1\}$ for all real x. Moreover, the terms can only be non-zero for $i \leqslant \log(2n)/\log p$. It follows that

$$p^{w_p(n)} \leqslant p^{\log(2n)/\log p} = 2n.$$

Taking products over p (noting that all primes dividing $\binom{2n}{n}$ are $\leqslant 2n$) we obtain

$$\binom{2n}{n} \leqslant (2n)^{\pi(2n)}. \tag{1.2}$$

Now $\binom{2n}{n}$ is almost as big as 4^n; note that $\sum_{r=0}^{2n} \binom{2n}{r} = 4^n$ and, in this sum, the middle term $\binom{2n}{n}$ is the largest one. It follows from this observation and (1.2) that, for n sufficiently large,

$$2^n < \frac{4^n}{2n+1} \leqslant \binom{2n}{n} \leqslant (2n)^{\pi(2n)}.$$

Taking logs gives

$$\pi(2n) \geqslant \frac{n \log 2}{\log(2n)}.$$

This is the lower bound in Theorem 1.1 in the case $X = 2n$ an even integer (with $c_1 = \frac{1}{2}\log 2$). The general case follows easily by considering the greatest even integer $\leqslant X$, and reducing c_1 slightly. We leave the details as an exercise.

We turn now to the upper bound. At the cost of increasing c_2 by a factor of 2, it is enough to handle the case $X = 2^k$, k a positive integer (since $t/\log t$ is an increasing function of t). We again consider the binomial coefficient $\binom{2n}{n}$, noting now that if n is an integer and if $n < p \leqslant 2n$ then p divides $\binom{2n}{n}$ precisely once. Therefore

$$n^{\pi(2n)-\pi(n)} \leqslant \prod_{n<p\leqslant 2n} p \leqslant \binom{2n}{n} \leqslant 4^n,$$

and so, taking logs,

$$\pi(2n) - \pi(n) \leqslant \frac{n \log 4}{\log n}.$$

Applying this with $n = 2, 2^2, 2^3 \ldots, 2^{k-1}$ and summing gives

$$\pi(2^k) \leqslant 2^k\Big(\frac{1}{k-1} + \frac{1}{2(k-2)} + \frac{1}{2^2(k-3)} + \cdots + \frac{1}{2^{k-2}}\Big).$$

We claim that the bracketed expression is bounded by C/k for some absolute constant C. There are many ways to see this. For instance, the ratio between

the terms $1/2^{i-1}(k-i)$ and $1/2^i(k-i-1)$ is $2(k-i-1)/(k-i)$, which is $\geqslant 4/3$ for $i \leqslant k-3$, so

$$\sum_{i=2}^{k-1} \frac{1}{2^{i-1}(k-i)} \leqslant \frac{1}{k-1}\Big(1+\frac{3}{4}+(\frac{3}{4})^2+\dots\Big)+\frac{1}{2^{k-2}},$$

which implies the stated bound that $\pi(2^k) \leqslant C2^k/k$ for some C. □

The famous *prime number theorem* is the assertion that one can take both c_1 and c_2 arbitrarily close to 1 for large X, namely

$$\lim_{X\to\infty} \frac{\pi(X)}{X/\log X} = 1 \tag{1.3}$$

This is quite a difficult theorem and we will not prove it in this course, though we will make some further comments later on.

Asymptotic notation. Before continuing, we introduce some notation which is commonplace throughout the more analytic part of number theory, and which we will use quite extensively in the course. This is the use of 'asymptotic' notations $\ll, \gg$ and O, which are used to describe the rough size of error terms. The statements $X = O(Y)$, $X \ll Y$ and $Y \gg X$ all mean the same thing, namely that for some absolute constant C we have $|X| \leqslant CY$. The main point of the notation is that we do not specify the constant C. Usually, this is because we do not need to for applications. It might be tedious to produce an explicit C that works, or occasionally impossible. Different instances of the notation may imply different values of C.

One often sees the $O()$ notation used by itself for error terms. An example is that for $X \geqslant 1$ we have

$$\sum_{n\leqslant X} \log n = X\log X + O(X), \tag{1.4}$$

which is (a weak) form of Stirling's formula, which we will use below. What this means is that

$$\Big|\sum_{n\leqslant X} \log n - X\log X\Big| \leqslant CX$$

for some absolute constant C, and for all $X \geqslant 1$. For the proof of (1.4), see Sheet X.

Mertens estimates. We now turn to two estimates of Mertens, which give interesting information about the distribution of primes. Here, and throughout the rest of the course, $\sum_p$ means a sum over primes, whilst $\prod_p$ denotes a product over primes. (We will discuss infinite products a little more when we get to them.)

Proposition 1.2. *We have Mertens' first estimate*

$$\sum_{p\leqslant X}\frac{\log p}{p} = \log X + O(1) \tag{1.5}$$

and Mertens' second estimate[1]

$$\sum_{p\leqslant X}\frac{1}{p} = \log\log X + O(1). \tag{1.6}$$

Remark. It is at first sight surprising that it is not too hard to prove these two statements, whereas it is much more difficult (the prime number theorem) to get an asymptotic for $\sum_{p\leqslant X}1$. What makes this so is that the 'weights' $\log p/p$ and $1/p$ in (1.5) and (1.6) are much 'smoother' than the weight $1_{p\leqslant X}$.

Proof. We first prove (1.5). It suffices to handle the case X an integer. The idea is to look at the prime factorization of $X!$. By Legendre's formula (see Sheet X), the power of p dividing $X!$ is $\sum_{i\geqslant 1}\lfloor\frac{X}{p^i}\rfloor$. Thus, taking logs,

$$\log X! = \sum_{p\leqslant X}\log p\big(\lfloor\frac{X}{p}\rfloor + \lfloor\frac{X}{p^2}\rfloor + \dots\big) \tag{1.7}$$

By Stirling's formula (in the weak form (1.4)) we have

$$\log X! = \sum_{n\leqslant X}\log n = X\log X + O(X). \tag{1.8}$$

Turning to the right-hand side of (1.7), we first look at the contribution from the terms $\lfloor\frac{X}{p^i}\rfloor$ with $i\geqslant 2$. This contribution is bounded by

$$X\sum_{p\leqslant X}(\log p)\big(\frac{1}{p^2}+\frac{1}{p^3}+\dots\big) \leqslant 2X\sum_{p\leqslant X}\frac{\log p}{p^2} \ll X;$$

in fact the sum over *integers* $\sum_{n\leqslant X}\frac{\log n}{n^2}$ is convergent (see Sheet X). Now we look at the contribution to the RHS of (1.7) from the terms $\lfloor X/p\rfloor$. We have $\lfloor X/p\rfloor = \frac{X}{p}+O(1)$, and the contribution from the first of these terms is precisely $X\sum_{p\leqslant X}\frac{\log p}{p}$, that is to say X times the sum we are interested in. The contribution from the $O(1)$ terms is

$$\ll \sum_{p\leqslant X}\log p \leqslant \pi(X)\log X \ll X,$$

where for the second bound we used the upper Chebyshev estimate (that is, the upper bound in Theorem 1.1). Putting all this together gives that the RHS of (1.7) is

$$X\sum_{p\leqslant X}\frac{\log p}{p} + O(X).$$

[1]This is actually a weak form of the result usually known as Mertens' second estimate, which is that $\sum_{p\leqslant X}\frac{1}{p} = \log\log X + B + O(\frac{1}{\log X})$ for some constant B. See, for instance, Montgomery and Vaughan Theorem 2.7 (d).

Since the LHS of (1.7) is given as in (1.8), comparing the two gives the desired bound (1.5).

We will deduce the second Mertens estimate (1.6) from the first one using a technique called partial summation. What makes this work is that $\frac{1}{p}$ is a 'smoother' weight than $\frac{\log p}{p}$, which allows us to express $\sum_{p\leqslant X}\frac{1}{p}$ in terms of an average

$$\int_2^X f(x)\Big(\sum_{p\leqslant x}\frac{\log p}{p}\Big)\mathrm{d}x, \tag{1.9}$$

for a suitably chosen f. To motivate the choice of f, we swap the order of summation and integration in (1.9) to get

$$\sum_{p\leqslant X}\frac{\log p}{p}\int_p^X f(x)\,\mathrm{d}x.$$

This suggests the choice $f(x)=\frac{1}{x(\log x)^2}$ (that is, the derivative of $-\frac{1}{\log x}$), since then

$$\int_p^X f(x)\,\mathrm{d}x=\frac{1}{\log p}-\frac{1}{\log X}.$$

With this choice of f, we therefore have

$$\sum_{p\leqslant X}\frac{1}{p}=\frac{1}{\log X}\sum_{p\leqslant X}\frac{\log p}{p}+\int_2^X\frac{\mathrm{d}x}{x(\log x)^2}\Big(\sum_{p\leqslant x}\frac{\log p}{p}\Big). \tag{1.10}$$

Using the first Mertens estimate (1.5), this is

$$O(1)+\int_2^X\frac{\mathrm{d}x}{x(\log x)^2}\big(\log x+O(1)\big)=O(1)+\int_2^X\frac{\mathrm{d}x}{x\log x},$$

using here that the integral $\int_2^X\frac{\mathrm{d}x}{x(\log x)^2}=\frac{1}{\log 2}-\frac{1}{\log X}=O(1)$. To complete the proof, we note that $\int_2^X\frac{\mathrm{d}x}{x\log x}=\log\log X-\log\log 2$. □

2. Arithmetic functions

In this section and the next we study *arithmetic functions*, which are simply functions $f:\mathbf{N}\to\mathbf{C}$.

The main functions. Divisor bound. We will particularly focus on functions of particular interest in number theory, and these are

- The divisor function τ, where $\tau(n)$ is defined to be the number of positive integer divisors of n (including 1 and n);
- The number of distinct prime factors of n, denoted by $\omega(n)$;
- The number of prime factors of n counted with multiplicity, denoted by $\Omega(n)$;
- Euler's ϕ function, where $\phi(n)$ is defined to be the number of integers x with $1\leqslant x\leqslant n$ which are coprime to n, or equivalently the order of the multiplicative group $(\mathbf{Z}/n\mathbf{Z})^\times$;

- The Möbius function $\mu(n)$, defined to be $(-1)^k$ if n is a product of k distinct primes, and 0 otherwise.

Multiplicativity. We say that an arithmetic function f is multiplicative if $f(mn) = f(m)f(n)$ whenever m, n are coprime. (Note carefully that we do not require $f(mn) = f(m)f(n)$ for all m, n; that property, when it holds, is called *complete multiplicativity*.)

Lemma 2.1. *The functions τ, ϕ, μ are multiplicative.*

Proof. For μ this follows immediately from the definition. That $\phi(mn) = \phi(m)\phi(n)$ when m, n are coprime follows from the Chinese remainder theorem; in fact $(\mathbf{Z}/mn\mathbf{Z})^\times \cong (\mathbf{Z}/m\mathbf{Z})^\times \times (\mathbf{Z}/n\mathbf{Z})^\times$. This was shown in ASO Number Theory. Finally, for τ we note that if m and n are coprime then the divisors of mn are precisely the integers dd' with $d \mid m$ and $d' \mid n$, and that all such integers are distinct. □

A surprisingly powerful fact, which comes up everywhere in number theory, is that the number of divisors of n is rather small. This is called the *divisor bound*.

Proposition 2.2 (Divisor bound)**.** *Let $\varepsilon > 0$. Then there is a constant $C_\varepsilon > 0$ such that*

$$\tau(n) \leqslant C_\varepsilon n^\varepsilon$$

for all n.

Remark. Quite often one will see this bound written as $\tau(n) \ll_\varepsilon n^\varepsilon$; it is understood here that $\varepsilon > 0$ can be arbitrary. Another notation one will sometimes see is $\tau(n) = n^{o(1)}$.

Proof. We first note a formula for $\tau(n)$ in terms of the prime factorisation $n = p_1^{a_1} \cdots p_k^{a_k}$, namely

$$\tau(n) = (a_1 + 1) \cdots (a_k + 1). \tag{2.1}$$

This follows immediately from the multiplicativity of τ and the fact that $\tau(p^a) = a + 1$, since the divisors of p^a are precisely the elements of the set $\{1, p, p^2, \dots, p^a\}$.

We have $a + 1 \leqslant e^a \leqslant p^{a\varepsilon}$ provided $p > e^{1/\varepsilon}$.

Since $\lim_{a\to\infty} \frac{a+1}{2^{a\varepsilon}} = 0$, there is some constant B (depending on ε) such that $a + 1 \leqslant B2^{a\varepsilon}$ for all a. We of course have $a + 1 \leqslant Bp^{a\varepsilon}$ for all $p \geqslant 2$.

By (2.1), the result now follows with $C_\varepsilon := B^m$, where m is the number of primes $\leqslant e^{1/\varepsilon}$. □

Dirichlet convolution and Möbius inversion. If $f, g : \mathbf{N} \to \mathbf{C}$ are two arithmetic functions then we define

$$f \star g(n) := \sum_{d|n} f(d)g(n/d) = \sum_{ab=n} f(a)g(b).$$

This operation is known as *Dirichlet convolution*. It is easy to check that it is commutative (that is, $f \star g = g \star f$) and associative (that is, $f \star (g \star h) = (f \star g) \star h$). It also has an identity, which is the function δ such that $\delta(1) = 1$ and $\delta(n) = 0$ for all other n. For this function, we have $f \star \delta = f$ for any function f.

The following result is known as *Möbius inversion*. Here, 1 denotes the arithmetic function taking the value 1 at all inputs.

Proposition 2.3. *Let $f, g : \mathbf{N} \to \mathbf{C}$ be two arithmetic functions. Then $g = f \star 1$ if and only if $f = g \star \mu$.*

Proof. Note that $\mu \star 1 = \delta$. This is a routine check: if $n = p_1^{a_1} \dots p_k^{a_k}$ with $k \geqslant 1$ (and the p_i primes) then

$$\begin{aligned}\sum_{d|n} \mu(d) &= \sum_{\varepsilon_i \in \{0,1\}} \mu(p_1^{\varepsilon_1} \dots p_k^{\varepsilon_k}) \\ &= \sum_{\varepsilon_i \in \{0,1\}} \mu(p_1^{\varepsilon_1}) \cdots \mu(p_k^{\varepsilon_k}) = (1-1) \cdots (1-1) = 0.\end{aligned}$$

Here we used the multiplicativity of μ. Now if $g = f \star 1$ then

$$g \star \mu = (f \star 1) \star \mu = f \star (1 \star \mu) = f \star \delta = f.$$

Conversely if $f = g \star \mu$ then

$$f \star 1 = (g \star \mu) \star 1 = g \star (\mu \star 1) = g \star \delta = g.$$

This concludes the proof. □

Dirichlet series. Suppose that $f : \mathbf{N} \to \mathbf{C}$ is an arithmetic function. Then we define the *Dirichlet series*

$$D_f(s) := \sum_{n=1}^{\infty} f(n) n^{-s}. \tag{2.2}$$

The definition is not really meaningful unless the series is known to converge (though there is a notion of a 'formal Dirichlet series'). A condition on f which comes up in practice is that, for some $a \in \mathbf{R}$, we have

$$|f(n)| \leqslant C_\varepsilon n^{a+\varepsilon} \tag{2.3}$$

for all $\varepsilon > 0$ (often this would be written as $|f(n)| \leqslant n^{a+o(1)}$). For example, the divisor bound implies that the divisor function τ satisfies this condition with $a = 0$, and clearly the constant function 1 satisfies this condition, also with $a = 0$. We have the following analytic result.

Lemma 2.4. *Suppose that f is an arithmetic function satisfying (2.3). Then $D_f(s)$ converges to a holomorphic function in the domain $\{s \in \mathbf{C} : \operatorname{Re} s > a + 1\}$.*

Proof. Suppose that $\operatorname{Re} s \geqslant a+1+\delta$. Then for $N > N_0(\delta)$ sufficiently large we have

$$\Big|\sum_{n=N+1}^{\infty} f(n)n^{-s}\Big| \leqslant \sum_{n=N+1}^{\infty} n^{a+\delta/2}n^{-a-1-\delta} \ll N^{-\delta/2}.$$

(See Sheet X for the last estimate.) Thus if we set

$$D_f^{(N)}(s) := \sum_{n=1}^{N} f(n)n^{-s} \tag{2.4}$$

then $D_f^{(N)}(s) \to D_f(s)$ uniformly in $\operatorname{Re} s \geqslant a+1+\delta$, and so certainly on $\operatorname{Re} s > a+1+\delta$. Since a uniform limit of holomorphic functions on a domain is holomorphic on that domain (Proposition B.1), it follows that $D_f(s)$ is holomorphic on $\operatorname{Re} s > a+1+\delta$. Since $\delta > 0$ was arbitrary, the result follows. □

We now establish a result linking Dirichlet convolution and Dirichlet series. Formally, the result is that $D_{f_1\star f_2} = D_{f_1}D_{f_2}$, and the proof is morally

$$\begin{aligned} D_{f_1\star f_2}(s) &= \sum_n \Big(\sum_{n_1n_2=n} f_1(n_1)f_2(n_2)\Big)n^{-s} \\ &= \Big(\sum_{n_1} f_1(n_1)n_1^{-s}\Big)\Big(\sum_{n_2} f_2(n_2)n_2^{-s}\Big) = D_{f_1}(s)D_{f_2}(s). \end{aligned}$$

However, one must worry about convergence. The following lemma handles this more carefully.

Lemma 2.5. *Suppose that f_1, f_2 satisfy (2.3) with exponents $a_1, a_2 \geqslant 0$ respectively. Then $f := f_1 \star f_2$ satisfies (2.3) with exponent a_1+a_2, and so D_{f_1}, D_{f_2}, $D_{f_1\star f_2}$ are holomorphic on $\{s \in \mathbf{C} : \operatorname{Re} s > a_1+a_2+1\}$. On this domain, we have $D_{f_1\star f_2}(s) = D_{f_1}(s)D_{f_2}(s)$.*

Proof. For the proof (which is not examinable) see Sheet 1. □

In the remainder of this section we look in greater detail at the three specific functions ω, τ, ϕ.

The number of prime factors. In this section we look at $\omega(n)$, the number of distinct prime factors of n. We will study the typical behaviour of $\omega(n)$ for $n \leqslant N$.

The second Mertens estimate (1.6) quite easily gives the following result on the average behaviour of $\omega(n)$.

Proposition 2.6. *We have*

$$\frac{1}{N}\sum_{n\leqslant N} \omega(n) = \log\log N + O(1).$$

Proof. Given n, write $X_p(n) = 1$ if $p \mid n$, and 0 otherwise. Then

$$\omega(n) = \sum_{p \leqslant N} X_p(n). \tag{2.5}$$

Since the number of $n \leqslant N$ which are divisible by p is $\frac{N}{p} + O(1)$, we have using (1.6) that

$$\sum_{n \leqslant N} \omega(n) = \sum_{p \leqslant N} \sum_{n \leqslant N} X_p(n) = \sum_{p \leqslant N} \big(\frac{N}{p} + O(1)\big) = N(\log\log N + O(1)).$$

This completes the proof. □

We will go further than this and show that $\omega(n)$ is in fact strongly concentrated about $\log\log N$ for $n \leqslant N$, a result known as the Hardy-Ramanujan Theorem. The key to this is the following second moment estimate.

Proposition 2.7. *We have*

$$\frac{1}{N} \sum_{n \leqslant N} \omega(n)^2 \leqslant (\log\log N)^2 + O(\log\log N).$$

Proof. Using (2.5) again, we have

$$\sum_{n \leqslant N} \omega(n)^2 = \sum_{p,p' \leqslant N} \#\{n \leqslant N : p, p' \mid n\}.$$

We split the sum into $p = p'$ and $p \neq p'$ and use that

$$\#\{n \leqslant N : p \mid n\} \leqslant \frac{N}{p}$$

and that, when $p \neq p'$,

$$\#\{n \leqslant N : p, p' \mid n\} \leqslant \frac{N}{pp'}.$$

It follows that

$$\frac{1}{N} \sum_{n \leqslant N} \omega(n)^2 \leqslant \sum_{p \leqslant N} \frac{1}{p} + \sum_{\substack{p,p' \leqslant N \\ p \neq p'}} \frac{1}{pp'} = \big(\sum_{p \leqslant N} \frac{1}{p}\big)^2 - \sum_{p \leqslant N} \frac{1}{p^2} + \sum_{p \leqslant N} \frac{1}{p}.$$

The result follows from the second Mertens estimate, (1.6). □

Remarks. This argument is an example of the *second moment method*. It is possible to think of this probabilistically. The prime factors counting function $\omega(n)$ is a sum $\sum_p X_p(n)$. If we imagine that $\mathbf{n}$ is sampled uniformly at random from $\{1, \dots, N\}$, each $X_p(\mathbf{n})$ is a random variable, and $\omega(\mathbf{n})$ is a sum of these. The intuition is then that the $X_p(\mathbf{n})$ are approximately pairwise independent. Proposition 2.7 then amounts to verifying that these variables are indeed close to pairwise independent by computing the variance.

One could run a similar analysis with $\omega(n)$ replaced by $\Omega(n)$, which counts the number of prime factors of n *with multiplicity*. However, we also have the following result which states that these two functions are close on average.

Proposition 2.8. *We have*

$$\frac{1}{N}\sum_{n\leqslant N}(\Omega(n)-\omega(n))\leqslant 4.$$

Proof. The left-hand side is equal to

$$\frac{1}{N}\sum_{n\leqslant N}\sum_{j\geqslant 2}\sum_{p}1_{p^j|n}.$$

We have $\sum_{n\leqslant N}1_{d|n}=\lfloor N/d\rfloor\leqslant N/d$ for any d, so this is bounded above by

$$\sum_{p}\sum_{j\geqslant 2}p^{-j}=\sum_{p}\frac{1}{p^2}\frac{1}{1-1/p}\leqslant 2\sum_{p}\frac{1}{p^2}<4.$$

□

The following corollary is a statement that ω and Ω are concentrated about their means (much more precise statements are possible).

Corollary 2.9. *Let $\varepsilon>0$. Then, for N large enough in terms of ε, at least $(1-\varepsilon)N$ integers $n\leqslant N$ have $(1-\varepsilon)\log\log N\leqslant\omega(n)\leqslant(1+\varepsilon)\log\log N$. A similar result holds for $\Omega(n)$.*

Proof. Using Proposition 2.6 and Proposition 2.7 we compute

$$\frac{1}{N}\sum_{n\leqslant N}(\omega(n)-\log\log N)^2\ll\log\log N.$$

It follows that, writing

$$S(\varepsilon,N):=\{n\leqslant N:\omega(n)\notin[(1-\varepsilon)\log\log N,(1+\varepsilon)\log\log N]\},$$

we have

$$|S(\varepsilon,N)|(\varepsilon\log\log N)^2\ll N\log\log N,$$

which means that $|S(\varepsilon,N)|\ll(\varepsilon^2\log\log N)^{-1}N$. In particular, if N is large enough then $|S(\varepsilon,N)|\leqslant\varepsilon N$. This establishes the result for ω.

To derive the result for Ω, observe that from Proposition 2.8 it follows that if we write

$$T(\varepsilon,N):=\{n\leqslant N:\Omega(n)-\omega(n)\geqslant\varepsilon\log\log N\}$$

then

$$|T(\varepsilon,N)|\leqslant\frac{4}{\varepsilon\log\log N}N\leqslant\varepsilon N,$$

if N is big enough in terms of ε. Finally, note that if $n\notin S(\varepsilon/2,N)\cup T(\varepsilon/2,N)$ (a set of size $\leqslant\varepsilon N$) then $(1-\varepsilon)\log\log N\leqslant\Omega(n)\leqslant(1+\varepsilon)\log\log N$. □

Here is a fun application of the above bound for $\Omega(n)$. It concerns the $N \times N$ multiplication table (where the 10×10 or 12×12 table is the one you learned at school).

Theorem 2.10. *Most integers $\leqslant N^2$ do not lie in the $N \times N$ multiplication table.*

Proof. (Sketch) A typical integer $n \leqslant N^2$ has $\Omega(n) \approx \log\log N^2 \approx \log\log N$. A typical element $n' = n_1 n_2$ in the multiplication table has $\Omega(n') = \Omega(n_1) + \Omega(n_2) \approx 2\log\log N$. □

We also have the following result on the typical size (or 'normal order') of the divisor function.

Theorem 2.11. *Let $\varepsilon > 0$. Then, if N is sufficiently large in terms of ε, we have*

$$(\log N)^{\log 2-\varepsilon} \leqslant \tau(n) \leqslant (\log N)^{\log 2+\varepsilon}$$

for at least $(1-\varepsilon)N$ values of $n \leqslant N$.

Proof. This follows quickly from Corollary 2.9 and the inequalities

$$2^{\omega(n)} \leqslant \tau(n) \leqslant 2^{\Omega(n)},$$

which are immediate from (2.1) and the fact that $2 \leqslant a+1 \leqslant 2^a$ for all integers $a \geqslant 1$. □

Remark. Sometimes one sees this result written in the following shorthand: $\tau(n) = (\log n)^{\log 2+o(1)}$ almost surely. This means precisely the statement of Theorem 2.11.

More on the divisor function. We have already established the *divisor bound*, which is the statement that $\tau(n) \ll_\varepsilon n^\varepsilon$. In Theorem 2.11 we established the typical size of $\tau(n)$. In this section we study the *average* size of $\tau(n)$ which, surprisingly at first sight, is much larger than the typical size of $\tau(n)$. This is not a contradiction: the bulk of the contribution to the average value $\frac{1}{N}\sum_{n\leqslant N}\tau(n)$ turns out to be from (exceptional) integers n with more than the typical number of prime factors.

We will need Euler's constant γ, defined by

$$\gamma := \lim_{Y\to\infty} \sum_{n\leqslant Y} \frac{1}{n} - \log Y.$$

This constant exists and has the value $\approx 0.577\ldots$. Moreover, we have

$$\sum_{n\leqslant Y} \frac{1}{n} = \log Y + \gamma + O(\frac{1}{Y}) \tag{2.6}$$

For these two facts, see Sheet X.

Theorem 2.12. *We have*

$$\frac{1}{N}\sum_{n\leqslant N}\tau(n) = \log N + 2\gamma - 1 + O(\frac{1}{\sqrt{N}}).$$

In particular, the average size of $\tau(n)$ on $[1,N]$ is essentially $\log N$.

Proof. The argument is known as Dirichlet's hyperbola method. Since $\tau(n)$ is the number of pairs (a,b) of positive integers with $ab=n$, we have[2]

$$\sum_{n\leqslant N}\tau(n) = \sum_{a,b} 1_{ab\leqslant N}.$$

We divide the sum up into (1) $a,b\leqslant\sqrt{N}$, (2) $a>\sqrt{N}$ and (3) $b>\sqrt{N}$, noting that these possibilities cover everything and are mutually exclusive.

On the range (1), we always have $ab\leqslant N$ and so the sum is simply the number of pairs (a,b) with $a\leqslant\sqrt{N}, b\leqslant\sqrt{N}$, which is $(\sqrt{N}+O(1))^2 = N+O(\sqrt{N})$.

The contributions from (2) and (3) are the same, and each is

$$\sum_{a>\sqrt{N}}\sum_{b\leqslant N/a} 1 = \sum_{b\leqslant\sqrt{N}}\sum_{\sqrt{N}<a\leqslant N/b} 1 = \sum_{b\leqslant\sqrt{N}}\sum_{a\leqslant N/b} 1 - N + O(\sqrt{N}).$$

Since

$$\sum_{a\leqslant N/b} 1 = \frac{N}{b} + O(1),$$

this is

$$N\sum_{b\leqslant\sqrt{N}}\frac{1}{b} - N + O(\sqrt{N}).$$

We therefore see that the contribution from (2) (and also from (3)) is

$$\frac{1}{2}N\log N + (\gamma-1)N + O(\sqrt{N}).$$

Adding the contributions from (1), (2) and (3) gives the result. □

Euler's ϕ function. The main theorem of this section is the following.

Theorem 2.13. *We have*

$$\sum_{n\leqslant N}\phi(n) = \frac{3}{\pi^2}N^2 + O(N\log N).$$

Consequently, the probability that two random integers are coprime is $6/\pi^2$.

[2]Here $1_{ab\leqslant N}$ denotes the indicator of the event $ab\leqslant N$, that is to say it equals 1 if $ab\leqslant N$ and 0 otherwise.

Remark. Let us explain the second statement (both what it means, and why it follows from the first). One mathematically precise version of the claim is that if one picks two integers $a, b \in \{1, \dots, N\}$ uniformly at random, then the probability they are coprime tends to $\frac{6}{\pi^2}$ as $N \to \infty$. But $\sum_{n \leqslant N} \phi(n)$ is exactly the number of pairs (a, b) of integers with $1 \leqslant a \leqslant b \leqslant N$ and a and b coprime.

Proof. Write F for the arithmetic function $F(n) = n$. We first note that $1 \star \phi = F$, that is to say $\sum_{d|n} \phi(n/d) = n$; this is by dividing up the integers $x \leqslant n$ according to the highest common factor $d = (n, x)$.

By Möbius inversion, it follows that $\phi = \mu \star F$, that is to say

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

Summing over $n \leqslant N$ and swapping the order of summation yields

$$\sum_{n \leqslant N} \phi(n) = \sum_{d \leqslant N} \frac{\mu(d)}{d} \sum_{\substack{n \leqslant N \\ d|n}} n = \sum_{d \leqslant N} \mu(d) \sum_{k \leqslant N/d} k.$$

Since

$$\sum_{k \leqslant N/d} k = \frac{1}{2} \lfloor \frac{N}{d} \rfloor \big(\lfloor \frac{N}{d} \rfloor + 1 \big) = \frac{N^2}{2d^2} + O(\frac{N}{d})$$

(for $d \leqslant N$) we have

$$\sum_{n \leqslant N} \phi(n) = \frac{1}{2} N^2 \sum_{d \leqslant N} \frac{\mu(d)}{d^2} + O\big(N \sum_{d \leqslant N} \frac{1}{d}\big).$$

The error term here is $\ll N \log N$ which is acceptable. The error incurred by completing the sum $\sum_{d \leqslant N} \frac{\mu(d)}{d^2}$ to infinity is $\ll N^2 \sum_{d > N} \frac{1}{d^2} \ll N$ (see Sheet X). Thus it suffices to evaluate $\sum_d \frac{\mu(d)}{d^2} = D_\mu(2)$ (the Dirichlet series of μ at 2, see (2.2)).

However we have $\mu \star 1 = \delta$ and so, by Lemma 2.5, $D_\mu(2) D_1(2) = D_\delta(2) = 1$. It follows that

$$D_\mu(2) = D_1(2)^{-1} = \big(\sum_{n=1}^{\infty} n^{-2} \big)^{-1} = 6/\pi^2.$$

The result follows. □

3. The Riemann ζ-function

In this section we introduce the famous Riemann ζ-function (zeta function) and explore its basic properties as well as the connection to primes.

The ζ-function: basic definitions and facts. Let $s \in \mathbf{C}$ have $\operatorname{Re} s > 1$. Then we define

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}. \tag{3.1}$$

This is the same thing as the Dirichlet series for the constant function 1, and so by Lemma 2.4 it converges in $\operatorname{Re} s > 1$ and defines a holomorphic function there.

Euler product. Unique factorization into primes convinces us of the identity

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots = \left(1 + \frac{1}{2^s} + \frac{1}{2^{2s}} + \dots\right)\left(1 + \frac{1}{3^s} + \frac{1}{3^{2s}} + \dots\right)\cdots,$$

where there is one product for each prime $p = 2, 3, 5, \dots$. Summing each of the geometric series suggests then that

$$\zeta(s) = \prod_{p} (1 - p^{-s})^{-1}, \tag{3.2}$$

where the infinite product means $\lim_{N\to\infty} \prod_{p\leqslant N}(1-p^{-s})^{-1}$. Let us pause to justify this formally. Let s, $\operatorname{Re} s > 1$, be fixed. Suppose $\operatorname{Re} s = 1 + \delta$. Fix an N, and suppose that $m > \log_2 N$. Then the product

$$P(m, N) := \prod_{p\leqslant N} \Big(\sum_{j=0}^{m} p^{-js}\Big)$$

may be expanded, and it equals $1 + 2^{-s} + \cdots + N^{-s}$ plus a number of further, distinct, terms n^{-s} with $n > N$. Thus

$$|P(m, N) - \zeta(s)| \leqslant \sum_{n=N+1}^{\infty} |n^{-s}| \ll N^{-\delta}$$

uniformly in m. Letting $m \to \infty$, we thus obtain

$$\Big|\prod_{p\leqslant N} (1 - p^{-s})^{-1} - \zeta(s)\Big| \ll N^{-\delta}.$$

Finally, letting $N \to \infty$, we confirm the equality (3.2).

Equation (3.2) is called the Euler product for ζ and it is valid for $\operatorname{Re} s > 1$.

An amusing consequence of the above is yet another proof (our third) that there are infinitely many primes. Indeed, suppose not. Then, letting $s \to 1^+$ in (3.2) we see that $\lim_{s\to 1^+} \zeta(s) < \infty$. However this is not the case, since

$$\lim_{s\to 1^+} \sum_{n=1}^{N} n^{-s} = \sum_{n=1}^{N} n^{-1},$$

and the harmonic series on the right diverges.

By Möbius inversion, we have $1 \star \mu = \delta$. By Lemma 2.5, it follows that

$$1 = D_\delta(s) = D_{1\star\mu}(s) = D_1(s)D_\mu(s) = \zeta(s)\Big(\sum_{n=1}^{\infty} \mu(n)n^{-s}\Big)$$

for $\operatorname{Re} s > 1$ (the required analytic condition (2.3) is trivial to verify in this case, with $a = 0$). An immediate consequence of this is the following nonvanishing result.

Corollary 3.1. *We have $\zeta(s) \neq 0$ when $\operatorname{Re} s > 1$.*

Analytic continuation. The real power of the ζ-function comes from considering $\zeta(s)$ for $\operatorname{Re} s \leqslant 1$. The definition (3.1) only makes sense for $\operatorname{Re} s > 1$, but we have the following remarkable fact.

Theorem 3.2. *There is a meromorphic[3] continuation of $\zeta(s)$ to the whole complex plane $\mathbf{C}$, holomorphic except for a simple pole at $s = 1$.*

By the identity principle in complex analysis, this meromorphic continuation is unique. We will not prove this theorem in full in this course, but we will show how to meromorphically continue $\zeta(s)$ to the domain $\operatorname{Re} s > 0$.

Proposition 3.3. *The Riemann zeta function $\zeta(s)$ has a (unique) meromorphic continuation to the right half-plane $\operatorname{Re}(s) > 0$, holomorphic except for a simple pole at $s = 1$.*

Proof. For $\operatorname{Re} s > 1$ we have

$$\begin{aligned}
\zeta(s) = \sum_{n=1}^{\infty} n^{-s} &= \sum_{n=1}^{\infty} n(n^{-s} - (n+1)^{-s}) \\
&= s\sum_{n=1}^{\infty} n \int_n^{n+1} x^{-s-1}\,\mathrm{d}x \\
&= s\int_1^{\infty} \lfloor x \rfloor x^{-s-1}\,\mathrm{d}x \\
&= \frac{s}{s-1} - s\int_1^{\infty} \{x\}x^{-s-1}\,\mathrm{d}x.
\end{aligned}$$

(Here, $\{x\} := x - \lfloor x \rfloor$.) The integral here defines a holomorphic function on $\operatorname{Re} s > 0$ by differentiating under the integral. The rigorous justification of this last step is not examinable in this course but may be found in Proposition A.3. □

With this in hand, we can state the famous Riemann hypothesis (RH).

Conjecture 3.4. *The only zeros of $\zeta(s)$ with $0 < \operatorname{Re} s < 1$ are those with $\operatorname{Re} s = \frac{1}{2}$.*

[3] You may wish to revisit A2.2 Complex Analysis when studying this section, though we will not need any detailed proofs from that course. Meromorphic means holomorphic except for a discrete set of isolated poles.

The Riemann Hypothesis remains an unsolved problem. Even the fact that $\zeta(s) \neq 0$ for $\operatorname{Re} s = 1$ is not trivial; this fact is essentially equivalent to the prime number theorem (1.3). The proof of the equivalence involves a lot of mathematics and in fact is most of the content of the course C3.8: Analytic Number Theory.

In a remarkable October 2026 breakthrough, an internal model of OpenAI gave a proof that $\zeta(s) \neq 0$ for $\operatorname{Re} s > \frac{7}{8}$. This result, known as the 'Quasi Riemann hypothesis', massively surpasses previous bounds in this direction.

4. Dirichlet characters and L-functions.

In this section we introduce Dirichlet characters and their Dirichlet series, which are called Dirichlet L-functions. We show how these may be used to prove the following celebrated theorem of Dirichlet.

Theorem 4.1. *Let $q \in \mathbf{N}$ and suppose that a is coprime to q. Then there are infinitely many primes p with $p \equiv a (\mathrm{mod}\, q)$.*

Characters on finite abelian groups. An important object of study in this section will be *Dirichlet characters*. A Dirichlet character to modulus q is the lift to $\mathbf{Z}$ of a character on $(\mathbf{Z}/q\mathbf{Z})^*$, the multiplicative group of $\mathbf{Z}/q\mathbf{Z}$. Before saying exactly what this means, we give the basic theory of characters on finite abelian groups G.

Definition 4.2. Let G be a finite abelian group. Then a *character* on G is a homomorphism $\psi : G \to \mathbf{C}^* = \mathbf{C} \setminus \{0\}$.

Write e_G for the identity in G, and use addition for the group operation on G. If n is the order of G and $g \in G$ then $\psi(g)^n = \psi(ng) = \psi(e_G) = 1$, so ψ in fact takes values in the nth roots of unity. Here are the basic results on characters.

Proposition 4.3. *Let G be a finite abelian group. Then the characters on G form a group $\widehat{G}$ under pointwise multiplication of the same order as G (and isomorphic to it) with identity element ψ_0, the trivial character which takes the value 1 everywhere. Moreover we have*

$$\sum_{g \in G} \psi(g) = 0 \tag{4.1}$$

unless $\psi = \psi_0$ and

$$\sum_{\psi \in \widehat{G}} \psi(g) = 0 \tag{4.2}$$

unless $g = e_G$.

Proof. (Sketch) For the proof we write G with the group operation being addition, which is standard when dealing with abstract abelian groups. Later, when we specialise to $G = (\mathbf{Z}/q\mathbf{Z})^*$, we will use multiplication.

That the characters form a group is an easy exercise in the definition. To show that $|\widehat{G}| = |G|$ (and to establish that $\widehat{G} \cong G$, though we will

not actually need this fact) it suffices to handle the case $G = \mathbf{Z}/N\mathbf{Z}$ (with the group operation being addition!), since every abelian group is a direct product of such groups (in fact with $N = p^j$). On this group, each character is completely determined by the value of $\psi(1)$, which must be an Nth root of unity $e^{2\pi ir/N}$ (the whole character is then $\psi(x) = e^{2\pi irx/N}$) and we can see by inspection that $G \cong \widehat{G}$.

To prove (4.1), pick some g_0 such that $\psi(g_0) \neq 1$. Then, since $g + g_0$ ranges over G as g does,

$$\sum_{g\in G} \psi(g) = \sum_{g\in G} \psi(g + g_0) = \psi(g_0) \sum_{g\in G} \psi(g),$$

from which the result follows.

The proof of (4.2) is very similar. Pick some ψ_1 such that $\psi_1(g) \neq 1$. It is not obvious that such as ψ_1 exists. However, it is true when $G = \mathbf{Z}/N\mathbf{Z}$ (by inspection using the above description of characters on $\mathbf{Z}/N\mathbf{Z}$) and then it follows for arbitrary G by writing G as a direct product of such cyclic groups.

Since $\psi\psi_1$ ranges over $\widehat{G}$ as ψ does,

$$\sum_{\psi\in\widehat{G}} \psi(g) = \sum_{\psi\in\widehat{G}} (\psi\psi_1)(g) = \psi_1(g) \sum_{\psi\in\widehat{G}} \psi(g),$$

and the result follows. □

Remark. For students taking B2.1 Representation Theory, this proposition is basically the abelian case of a large part of that course.

Dirichlet characters. We now specialise to $G = (\mathbf{Z}/q\mathbf{Z})^*$ which, as was shown in ASO Number Theory, is an abelian group of order $\phi(q)$. We write the group operation as multiplication (juxtaposition). Let $\psi : (\mathbf{Z}/q\mathbf{Z})^* \to \mathbf{C}^*$ be a character, thus $\psi(ab) = \psi(a)\psi(b)$ for all $a, b \in (\mathbf{Z}/q\mathbf{Z})^*$. Then we can 'lift' ψ to a function $\chi : \mathbf{Z} \to \mathbf{C}$ by defining

$$\chi(n) = \begin{cases} \psi(n (\mathrm{mod}\, q)) & \text{if } (n, q) = 1; \\ 0 & \text{if } (n, q) \neq 1. \end{cases}$$

Such a function χ is called a *Dirichlet character* to modulus q.

Equivalently, a Dirichlet character to modulus q is a map $\chi : \mathbf{Z} \to \mathbf{C}$ such that

- $\chi(n + q) = \chi(n)$ for all n;
- $\chi(n) = 0$ unless $(n, q) = 1$;
- $|\chi(n)| = 1$ if $(n, q) = 1$, and
- $\chi(ab) = \chi(a)\chi(b)$ whenever $(a, q) = (b, q) = 1$.

Note that $\chi(ab) = \chi(a)\chi(b)$ actually holds *for all* a, b.

Since the Dirichlet characters are in 1-1 correspondence with characters $\psi : (\mathbf{Z}/q\mathbf{Z})^* \to \mathbf{C}^*$, the number of Dirichlet characters to modulus q is $\phi(q)$.

Example. The constant function 1 is a Dirichlet character to modulus 1 (in fact the only one).

The function χ defined by $\chi(n) = 0$ if n is even and $\chi(n) = 1$ if n is odd is a Dirichlet character to modulus 2, the lift of the trivial character from $(\mathbf{Z}/2\mathbf{Z})^*$ (which is the group of order 1) to $\mathbf{C}^*$.

A slightly less trivial example is the Dirichlet character χ_4 defined by $\chi_4(n) = 0$ if n is even and $\chi_4(1) = 1$, $\chi_4(3) = -1$, and χ_4 periodic with period 4, which is the lift of the nontrivial character on $(\mathbf{Z}/4\mathbf{Z})^*$, a group of order 2.

Given a modulus q, there is one Dirichlet character χ_0 which has a special status, namely the *principal* character, which is the lift of the trivial character ψ_0 on $(\mathbf{Z}/q\mathbf{Z})^*$. Thus it is defined by

$$\chi_0(n) = \begin{cases} 1 & (n,q) = 1; \\ 0 & (n,q) \neq 1. \end{cases}$$

We record the following facts, which are immediate consequences of Proposition 4.3 translated to this setting.

Lemma 4.4. *Let χ be a nonprincipal Dirichlet character to modulus q. Then*

$$\sum_{i=1}^{q} \chi(n+i) = 0 \tag{4.3}$$

for any integer n. We also have

$$\sum_{\chi} \chi(n) = \begin{cases} \phi(q) & \text{if } n \equiv 1 (\mathrm{mod}\, q) \\ 0 & \text{otherwise,} \end{cases} \tag{4.4}$$

where the sum is over all the $\phi(q)$ Dirichlet characters to modulus q.

Proof. (4.3) follows from (4.1) and (4.4) follows from (4.2). □

We can now define the Dirichlet L-function $L(s,\chi)$, which is a generalisation of the Riemann ζ-function.

Definition 4.5. Let χ be a Dirichlet character. Then for $\mathrm{Re}\, s > 1$ we define

$$L(s,\chi) := \sum_{n=1}^{\infty} \chi(n) n^{-s}.$$

Like the ζ-function, Dirichlet L-functions have an Euler product. This (and the proof) is a direct generalisation.

$$L(s,\chi) = \prod_{p} (1 - \chi(p)p^{-s})^{-1}. \tag{4.5}$$

As for the ζ-function, this is well-defined and holomorphic on $\mathrm{Re}\, s > 1$. For principal characters, $L(s,\chi_0)$ is in fact very closely related to the ζ-function.

Proposition 4.6. *Let χ_0 be the principal Dirichlet character to modulus q. Let $\operatorname{Re} s > 1$. Then*

$$L(s, \chi_0) = \zeta(s) \prod_{p|q} (1 - p^{-s}).$$

Proof. This is immediate from the Euler product. □

Since $\prod_{p|q}(1 - p^{-s})$ is an entire function which does not vanish at 1, it follows immediately from the preceding section that $L(s, \chi_0)$ extends to a meromorphic function on $\operatorname{Re} s > 0$ with a simple pole at $s = 1$.

Proposition 4.7. *Suppose that χ is a nonprincipal character to some modulus q. Then the series $\sum_{n=1}^{\infty} \chi(n) n^{-s}$ is (conditionally) convergent for $\operatorname{Re} s > 0$, and it defines a holomorphic function in this domain. We denote this $L(s, \chi)$.*

Proof. The key fact we will use is that

$$\Big| \sum_{n \leqslant x} \chi(n) \Big| \leqslant q \tag{4.6}$$

for all x, which follows immediately from the fact that χ sums to zero over any interval of length q (that is, (4.3)).

Now, using

$$N^{-s} - n^{-s} = s \int_N^n x^{-s-1} \, \mathrm{d}x$$

we have we have

$$\sum_{N < n \leqslant N'} \chi(n) n^{-s} = N^{-s} \sum_{N < n \leqslant N'} \chi(n) - s \int_N^{N'} x^{-s-1} \Big(\sum_{x < n \leqslant N'} \chi(n) \Big) \, \mathrm{d}x.$$

Therefore, using (4.6), if $\operatorname{Re} s = \sigma$ we have

$$\begin{aligned} \Big| \sum_{N < n \leqslant N'} \chi(n) n^{-s} \Big| &\leqslant q N^{-\sigma} + |s| q \int_N^{N'} x^{-\sigma-1} \, \mathrm{d}x \\ &= q N^{-\sigma} + \frac{|s| q}{\sigma} \big(N^{-\sigma} - (N')^{-\sigma} \big) \ll q \big(1 + \frac{|s|}{\sigma} \big) N^{-\sigma}. \end{aligned}$$

It follows that $(\sum_{n \leqslant N} \chi(n) n^{-s})_N$ is uniformly Cauchy in any domain $\{ s \in \mathbf{C} : \operatorname{Re} s > s_0, |s| \leqslant X \}$ with $s_0 > 0$ and finite X. This implies the result using standard facts from complex analysis. □

Remarks. We caution that we have not shown the Euler product relation (4.5) holds other than for $\operatorname{Re} s > 1$. The above proof is an instance of 'partial summation' or 'Abel summation' but my personal preference is to proceed from first principles rather than quote a general result.

Suppose that χ is a nonprincipal Dirichlet character. The value $L(1,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n}$ (which is finite) is arithmetically significant. One famous evaluation is

$$L(1,\chi_4) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4}. \tag{4.7}$$

We have the following result, which we will not prove fully in lectures, though a complete proof will be given across lectures and the example sheets.

Proposition 4.8. *Let χ be a nonprincipal Dirichlet character. Then $L(1,\chi) \neq 0$.*

This result can be proved by hand in particular cases, as you are encouraged to do for characters of moduli 3 and 5 on the second example sheet.

Primes in arithmetic progressions. Dirichlet L-functions, and in particular their non-vanishing at 1, can be used to show the infinitude of primes in fixed arithmetic progressions.

Theorem 4.9. *Assume that $L(1,\chi) \neq 0$ for all nonprincipal χ to modulus q. Then, for any a coprime to q, we have*

$$\sum_{p \equiv a (\mathrm{mod}\, q)} \frac{1}{p} = \infty.$$

In particular, there are infinitely many primes congruent to $a (\mathrm{mod}\, q)$.

The first step in the proof is to express the condition $1_{n \equiv a (\mathrm{mod}\, q)}$ in terms of characters.

Lemma 4.10. *Suppose that $(a,q) = 1$. Then we have*

$$\frac{1}{\phi(q)} \sum_{\chi (\mathrm{mod}\, q)} \overline{\chi(a)} \chi(n) = 1_{n \equiv a (\mathrm{mod}\, q)}.$$

Proof. If n is not coprime to q, both sides are zero. By abuse of notation, consider χ as a character on $(\mathbf{Z}/q\mathbf{Z})^*$. Then for $x \in (\mathbf{Z}/q\mathbf{Z})^*$ we have $\overline{\chi(a)}\chi(x) = \chi(xa^{-1})$. Now we use (4.4). □

Let $s > 1$ be real. It follows using Lemma 4.10 that

$$\sum_{p \equiv a (\mathrm{mod}\, q)} p^{-s} = \frac{1}{\phi(q)} \sum_{\chi} \overline{\chi(a)} \sum_{p : p \nmid q} \chi(p) p^{-s}.$$

If $\chi = \chi_0$, the principal character mod q, and if $p \nmid q$ then $\chi_0(p) = 1$ and the inner sum is $\sum_{p : p \nmid q} p^{-s}$. We have $\lim_{s \to 1^+} \sum_p p^{-s} = \infty$ by the second Mertens' estimate.

It therefore suffices to prove the following.

Lemma 4.11. *Suppose that χ is a nonprincipal character mod q and that $L(1,\chi) \neq 0$. Then $\lim_{s \to 1^+} \sum_p \chi(p) p^{-s}$ exists and is finite.*

We first sketch an incorrect 'proof' which nonetheless serves to convey the main idea. We 'have'

$$L(1,\chi) = \prod_p (1-\chi(p)p^{-1})^{-1}$$

and so

$$\log L(1,\chi) = -\sum_p \log(1-\chi(p)p^{-1}) = \sum_p \frac{\chi(p)}{p} + \sum_p \frac{\chi(p)^2}{2p^2} + \dots$$

Since $L(1,\chi) \neq 0$, $\log L(1,\chi)$ is finite. Since all sums $\sum_p \frac{\chi(p)^2}{p^2}, \dots$ are finite, $\sum_p \frac{\chi(p)}{p}$ is finite.

There are two important issues with this attempted proof, namely

- We have not shown the validity of the Euler product expansion at $s = 1$, and
- Complex numbers do not have a unique logarithm; it is well-defined only up to integer multiples of $2\pi i$.

Another issue (more easily circumvented) is that there are infinitely many sums $\sum_p \frac{\chi(p)^j}{p^j}$, $j = 2, 3, \dots$, that we need to consider. We now give an actual proof, motivated by the above flawed sketch.

Proof. First note that, since $L(s,\chi)$ is holomorphic on $\operatorname{Re} s > 0$ and $L(1,\chi) \neq 0$, there is a holomorphic branch of logarithm $\log L(s,\chi)$ on the ball $B_\varepsilon(1)$. (See Proposition B.3 for a reminder of this fact from complex analysis.)

Now for $s > 1$ we have the Euler product

$$L(s,\chi) = \prod_p (1-\chi(p)p^{-s})^{-1}.$$

Consider the function

$$F(s,\chi) := \sum_{j=1}^{\infty} \frac{1}{j} \sum_p \chi(p)^j p^{-js}.$$

The purpose of this definition is that $F(s,\chi)$ is a logarithm of $L(s,\chi)$ on $\operatorname{Re} s > 1$, due to the series expansion $\log(1-z) = -\sum_{j=1}^{\infty} \frac{1}{j} z^j$ (see Proposition B.4.) Since complex logarithms are unique up to integer multiples of $2\pi i$, it follows that

$$F(s,\chi) = \log L(s,\chi) + 2\pi i n(s)$$

on the intersection $\mathcal{D} := B_\varepsilon(1) \cap \{s \in \mathbf{C} : \operatorname{Re} s > 1\}$, where $n : \mathcal{D} \to \mathbf{Z}$ is an integer-valued function. Since both $F(s,\chi)$ and $\log L(s,\chi)$ are continuous on the connected domain $\mathcal{D}$, $n(s)$ must be constant on $\mathcal{D}$. Thus

$$F(s,\chi) = \log L(s,\chi) + C_\chi$$

for $s \in \mathcal{D}$, and so

$$\lim_{s \to 1^+} F(s, \chi) = \log L(1, \chi) + C_\chi$$

is finite. Now observe that

$$F(s, \chi) = \sum_p \chi(p) p^{-s} + F_*(s, \chi),$$

where

$$F_*(s, \chi) = \sum_{j=2}^{\infty} \frac{1}{j} \sum_p \chi(p)^j p^{-sj}.$$

To conclude the proof, it suffices to check that $\lim_{s \to 1^+} F_*(s, \chi)$ exists and is finite. This follows from dominated convergence and the bound

$$\sum_{j=2}^{\infty} \frac{1}{j} \sum_p p^{-j} < \sum_p \sum_{j=2}^{\infty} p^{-j} \leqslant 2 \sum_p p^{-2} < 4.$$

(This final bound is exactly the same as the one in Proposition 2.8.) □

Nonvanishing of $L(1, \chi)$. In the last subsection, we proved Dirichlet's theorem on primes in arithmetic progressions on the assumption that $L(1, \chi) \neq 0$ for nonprincipal characters χ, which we stated as Proposition 4.8.

Here, we prove this statement in the case that χ is *complex*, which means it is not real-valued (equivalently, $\chi(p) \notin \mathbf{R}$ for at least one prime p). The real-valued case is harder, and is outlined as Question 12 on Sheet 2.

In the proof, we will use the fact that if χ is a Dirichlet character to modulus q then so is $\overline{\chi}$, defined by $\overline{\chi}(n) := \overline{\chi(n)}$. If χ is complex then χ and $\overline{\chi}$ are distinct characters. For $\operatorname{Re} s > 1$, we have

$$L(s, \overline{\chi}) = \sum_{n=1}^{\infty} \overline{\chi(n)} n^{-s} = \overline{L(\overline{s}, \chi)}.$$

Taking limits as $s \to 1$, it follows that

$$L(1, \overline{\chi}) = \overline{L(1, \chi)}. \tag{4.8}$$

Now we prove that $L(1, \chi) \neq 0$ for complex characters.

Proposition 4.12. *We have $L(1, \chi) \neq 0$ for complex χ.*

Proof. Consider the function

$$F(s, \chi) := \sum_{j=1}^{\infty} \frac{1}{j} \sum_p \chi(p)^j p^{-js} = \sum_{j=1}^{\infty} \frac{1}{j} \sum_p \chi(p^j) p^{-js}.$$

As noted above, this is a logarithm of $L(s, \chi)$ for $\operatorname{Re} s > 1$. Now sum this over all characters χ to modulus q; we obtain

$$\sum_\chi F(s, \chi) = \sum_{j=1}^{\infty} \frac{1}{j} \sum_p p^{-js} \sum_\chi \chi(p^j) = \phi(q) \sum_{j=1}^{\infty} \frac{1}{j} \sum_{p^j \equiv 1 (\mathrm{mod}\, q)} p^{-js},$$

where we used (4.4) for the final step. In particular, if s is a real number greater than 1 then $\sum_\chi L(s,\chi)$ is real and non-negative, and so by exponentiating we see that

$$\prod_\chi L(s,\chi) \tag{4.9}$$

is a real number $\geqslant 1$.

Note that the $L(s,\chi)$ are all holomorphic on $\operatorname{Re} s > 0$ except for the L-function $L(s,\chi_0)$ attached to the principal character χ_0 mod q, which has a simple pole at $s=1$ like the ζ-function, as we saw in Proposition 4.6. Considering limits of (4.9) as $s \to 1^+$ through reals, it follows that $\prod_\chi L(s,\chi)$ is holomorphic in a neighbourhood of $s=1$ except possibly for a simple pole at $s=1$, and is not zero at $s=1$.

Since only $L(s,\chi_0)$ has a pole at $s=1$, we see that at most one of the $L(1,\chi)$, $\chi \neq \chi_0$, can be zero. By (4.8), this means that we cannot have $L(1,\chi)$ when χ is complex, because otherwise we would additionally have $L(1,\overline{\chi}) = 0$. □

5. Binary quadratic forms

We now turn to a different topic, the question of representing integers by binary quadratic forms.

Review of quadratic reciprocity. First we recall the basic properties of the Legendre symbol as taught in (for example) ASO Number Theory. This symbol $\left(\frac{a}{p}\right)$ is defined for odd prime p and for all integers a, and it takes the value 1 if a is a quadratic residue mod p and -1 otherwise, except when $a \equiv 0 (\operatorname{mod} p)$ in which case we define $\left(\frac{a}{p}\right) = 0$.

This has the following properties:

(i) Multiplicativity: $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$;

(ii) Periodicity: $\left(\frac{a+p}{p}\right) = \left(\frac{a}{p}\right)$;

(iii) Reciprocity $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$, if q is an odd prime different from p.

It should also be noted that

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}} = \begin{cases} 1 & p \equiv 1 \pmod 4; \\ -1 & p \equiv 3 \pmod 4 \end{cases} \tag{5.1}$$

and that

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = \begin{cases} 1 & p \equiv 1, 7 \pmod 8; \\ -1 & p \equiv 3, 5 \pmod 8. \end{cases} \tag{5.2}$$

Definite binary forms. Given $a, b, c \in \mathbf{Z}$, we would like to know which integers can be represented as $\psi(\mathbf{x})$ for $\mathbf{x} = (x_1, x_2) \in \mathbf{Z}^2$, where

$$\psi(\mathbf{x}) = ax_1^2 + bx_1x_2 + cx_2^2. \tag{5.3}$$

We will particularly focus on the representation of primes, though see Example Sheet 3 for more on the representation of general integers. For instance, we will establish the following famous theorem of Fermat[4]:

Theorem 5.1 (Fermat)**.** *All primes $p \equiv 1 (\mathrm{mod}\, 4)$ can be written as $x_1^2 + x_2^2$.*

Before proving this theorem (and related results for other quadratic forms) we develop some theory.

Definition 5.2. Let ψ be a quadratic form as in (5.3). Then the discriminant $\mathrm{Disc}(\psi)$ is defined to be $b^2 - 4ac$. If $a > 0$ and $\mathrm{Disc}(\psi) < 0$ then we say that ψ is *positive definite*.

We will restrict attention to positive definite quadratic forms in this course. It is useful to observe that if we write $\psi(\mathbf{x}) = \mathbf{x}^T M \mathbf{x}$ with $M = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$ then $\mathrm{Disc}(\psi) = -4 \det M$.

Lemma 5.3. *If ψ is positive definite then $\psi(x_1, x_2) > 0$ for all $(x_1, x_2) \neq (0, 0)$.*

Proof. Set $\Delta = \mathrm{Disc}(\psi)$. We note that

$$4a\psi(x_1, x_2) = (2ax_1 + bx_2)^2 - \Delta x_2^2,$$

from which the result is clear. □

We also introduce the notion of being properly represented by a form.

Definition 5.4. We say that n is *properly* represented by ψ if there exist *coprime* integers x_1, x_2 with $n = \psi(x_1, x_2)$.

The modular group Γ. Equivalent forms. It is clear that the set of integers properly represented by $\psi(\mathbf{x}) = x_1^2 + x_2^2$ is the same as the set of integers properly represented by $\psi'(\mathbf{x}) = x_1^2 + 2x_1x_2 + 2x_2^2 = (x_1 + x_2)^2 + x_2^2$, since (x_1, x_2) ranges over coprime pairs in $\mathbf{Z}^2$ as $(x_1 + x_2, x_2)$ does. The notion of *equivalent* forms captures this observation in proper generality.

In this definition and for the rest of the section, a prominent role will be played by the *modular group* $\Gamma = \mathrm{SL}_2(\mathbf{Z})$, that is to say the group of 2-by-2 integer matrices with determinant 1.

We are going to see two different actions of the modular group:

- The usual action of Γ on $\mathbf{C}^2$ from linear algebra, that is to say if $g = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \in \Gamma$ and $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ then $g\mathbf{x} = \begin{pmatrix} px_1 + qx_2 \\ rx_1 + sx_2 \end{pmatrix}$;

[4]For more accurate history see Davenport.

- An action of Γ on binary quadratic forms, introduced in Lemma 5.6 below.

There should be no danger of confusion in terms of which action we are talking about, because the actions are on different spaces.

Definition 5.5. Let ψ be a binary quadratic form. Then any form ψ' defined by $\psi'(\mathbf{x}) := \psi(g^{-1}\mathbf{x})$ for some $g \in \Gamma$, is said to be equivalent to ψ.

Remark. Since Γ is a group, the definition would have been equivalent had we written g in place of g^{-1}, but the g^{-1} is convenient in Lemma 5.6 below.

To reconcile this with the above example, one may see that if $g = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ then $\psi'(\mathbf{x}) = \psi(g^{-1}\mathbf{x})$. Equivalence as just defined is indeed an equivalence relation. Equivalent forms have the same discriminant and properly represent the same integers. The 'right' way to express this properly is the following.

Lemma 5.6. *There is an action of* $\mathrm{SL}_2(\mathbf{Z})$ *on binary quadratic forms of discriminant* Δ *given by* $(g\psi)(\mathbf{x}) := \psi(g^{-1}\mathbf{x})$. *If* $\Delta < 0$, *this action preserves the space of positive definite forms.*

Proof. The statement that this is an action is essentially the proof as well:

$$(gh)\psi(\mathbf{x}) = \psi((gh)^{-1}\mathbf{x}) = \psi(h^{-1}(g^{-1}\mathbf{x})) = (h\psi)(g^{-1}\mathbf{x}) = g(h\psi)(\mathbf{x}).$$

To see that the discriminant is preserved, we use the fact that if we write $\psi(\mathbf{x}) = \mathbf{x}^T M \mathbf{x}$ with $M = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$ then $\Delta = \mathrm{Disc}(\psi) = -4 \det M$. We have

$$(g\psi)(\mathbf{x}) = \psi(g^{-1}\mathbf{x}) = \mathbf{x}^T g^{-T} M g^{-1} \mathbf{x},$$

and so

$$\mathrm{Disc}(g\psi) = -4 \det(g^{-T} M g^{-1}) = -4 \det M = \mathrm{Disc}(\psi),$$

using here that $\det g = 1$. Finally we check that the action preserves the space of positive definite forms. For this, suppose $g = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$ and that $\psi(\mathbf{x}) = ax_1^2 + bx_1x_2 + cx_2^2$. Then one may compute that the x_1^2-coefficient of ψ' is $as^2 - brs + cr^2 = \psi(s, -r)$, which is positive by Lemma 5.3. □

Lemma 5.7. *Suppose that* ψ, ψ' *are equivalent forms. Then the set of integers* n *which are properly represented by* ψ *and by* ψ' *are the same.*

Proof. This follows from the above and the fact that $g^{-1}\mathbf{x}$ ranges over all coprime elements of $\mathbf{Z}^2$ as $\mathbf{x}$ does. □

Representation by some form of a given discriminant. The question of whether a positive integer n is properly represented by some particular form ψ is quite subtle. However, there is a relatively simple criterion for deciding whether n is properly represented by *some* form of discriminant Δ.

Proposition 5.8. *Let $\Delta < 0$ and $n > 0$. There exists some form ψ of discriminant Δ properly representing n if and only if the congruence $u^2 \equiv \Delta \pmod{4n}$ has a solution u.*

Proof. Suppose first that n is properly represented by some form $\psi(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2$ of discriminant Δ, thus $ap^2 + bpq + cq^2 = n$ with p, q coprime. By Euclid's algorithm, there exist r, s such that $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ lies in $\mathrm{SL}_2(\mathbf{Z})$. Consider the (equivalent) form

$$\psi'(x_1, x_2) = \psi(px_1 + rx_2, qx_1 + sx_2) = nx_1^2 + ux_1x_2 + vx_2^2,$$

using here that the coefficient of x_1^2 is (by direct computation) $ap^2 + bpq + cq^2 = n$. We have

$$\Delta = \mathrm{Disc}(\psi) = \mathrm{Disc}(\psi') = u^2 - 4vn,$$

so the stated congruence has a solution.

Conversely, suppose the stated congruence has a solution; then there exist integers u, v such that $u^2 - 4vn = \Delta$, which means that the form $\psi'(x, y) = nx^2 + uxy + vy^2$ has discriminant Δ. Manifestly, it properly represents n by taking $(x, y) = (1, 0)$. □

Reduced forms and the class number. It turns out that there are only finitely many forms of a given discriminant Δ, up to equivalence.

Definition 5.9. Let $\psi(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2$ be a positive definite quadratic form. We say that ψ is *primitive* if $\gcd(a, b, c) = 1$ and *reduced* if either

- $c > a$ and $-a < b \leqslant a$ or
- $c = a$ and $0 \leqslant b \leqslant a$.

It is very easy to check that equivalence of quadratic forms preserves the property of being primitive.

Proposition 5.10. *Every primitive, positive definite binary quadratic form is equivalent to a primitive reduced form.*

Proof. For notational brevity, we write $[a, b, c]$ for the quadratic form $\psi(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2$, and $[a, b, c] \sim [a', b', c']$ if two forms are equivalent. We use the following two particular types of equivalence:

(i) $[a, b, c] \sim [a, b + 2a, c + a + b]$;
(ii) $[a, b, c] \sim [c, -b, a]$.

To explain these equivalences, note that

$$\begin{aligned}\psi(x_1+x_2,x_2) &= a(x_1+x_2)^2+b(x_1+x_2)x_2+cx_2^2\\ &= ax_1^2+(2a+b)x_1x_2+(c+a+b)x_2^2\end{aligned}$$

and that

$$\psi(x_2,-x_1) = cx_1^2-bx_1x_2+ax_2^2.$$

Note that (i) also implies

$$[a,b,c]\sim[a,b-2a,c+a-b] \tag{5.4}$$

Now we explain why every (primitive) $[a,b,c]$ is equivalent to a (primitive) reduced form. First, by repeated use of (i) (or (5.4)), we may find a form $[a,b',c']\sim[a,b,c]$ with $-a<b\leqslant a$. If $c'>a$, we are done: the form $[a,b',c']$ is reduced. Suppose $c'=a$. If $b'\geqslant 0$, we are done. If not, we may apply (ii) once to get a reduced form. If $c'<a$, we apply (ii), obtaining a form $[a_1,b_1,c_1]=[c',-b',a]$ which is equivalent to $[a,b,c]$, and with $0<a_1<a$. Iterating this process leads to a reduced form after finitely many steps, since the value of a is always a positive integer. □

In fact, every primitive form is equivalent to precisely one reduced form, or in other words the reduced forms are inequivalent. We will not prove this in lectures but some exercises about it are on Example Sheet 3.

The class number. The number of primitive reduced forms of discriminant Δ is called the class number and we denote it $h(\Delta)$.

Proposition 5.11. *The class number $h(\Delta)$ is finite; in fact $h(\Delta)\ll|\Delta|$.*

Proof. Suppose that (a,b,c) is reduced and has discriminant $\Delta=b^2-4ac<0$. We have

$$|\Delta| = 4ac-b^2\geqslant 4a^2-a^2=3a^2,$$

so $a\leqslant\sqrt{|\Delta|/3}$. Now, given a we must have $|b|\leqslant a$ for a reduced form, giving $\leqslant 1+2\sqrt{|\Delta|/3}$ possibilities, and c is uniquely determined. □

Sums of two squares. Consider $\psi(\mathbf{x})=x_1^2+x_2^2$. This is a quadratic form of discriminant -4.

Lemma 5.12. *The class number $h(-4)$ is 1.*

Proof. We need to enumerate all triples (a,b,c) satisfying the conditions for being reduced (Definition 5.9) and with $b^2-4ac=-4$. By (the proof of) Proposition 5.11, we must have $a\leqslant\sqrt{4/3}$ and so $a\leqslant 1$. There are no triples with $a=0$. With $a=1$ we must have either $b=0$ or $b=1$. The case $b=0$ gives the triple $(1,0,1)$, corresponding to the form $\psi(\mathbf{x})=x_1^2+x_2^2$. By working mod 4 one immediately sees that we cannot have $b=1$ and $b^2-4ac=-4$. □

The class number $h(-4)$ being 1 means that all forms of discriminant -4 are equivalent. Thus if n is properly represented by some form of discriminant -4, it is properly represented by all of them, and in particular by $x_1^2 + x_2^2$. Therefore we deduce from Proposition 5.8 that n is properly represented by $x_1^2 + x_2^2$ if and only if the congruence $u^2 \equiv -4 \pmod{4n}$ has a solution.

Corollary 5.13. *Let $p \equiv 1 \pmod 4$ be a prime. Then p is a sum of two squares.*

Proof. We know from ASO Number Theory (see (5.1)) that -1 is a square mod p, say $t^2 \equiv -1 (\mathrm{mod}\, p)$. Then $u := 2t$ satisfies $u^2 \equiv -4 (\mathrm{mod}\, p)$ and $u^2 \equiv -4 (\mathrm{mod}\, 4)$, so satisfies the required congruence. □

The Kronecker symbol; real quadratic characters. In the next two sections we are going to link the discussion of Dirichlet characters in the last section to the class number. We will be looking at the Dirichlet characters χ which take real values - these had a special place in the last section, since it was for these characters that proving $L(1, \chi) \neq 0$ was particularly difficult. Such characters are traditionally called *real quadratic characters* because χ^2 is principal.

The Kronecker symbol $\left(\frac{a}{n}\right)$ is generalisation of the Legendre symbol. Different sources take different points of view on the set of values at which $\left(\frac{a}{n}\right)$ ought to be defined. For example Wikipedia defines it for *all* integer a, n. Iwaniec and Kowalski [4] define it for any a which is what they call a *discriminant*, that is to say any $a \neq 0$ with $a \equiv 0, 1 \pmod 4$. We will follow the still more restrictive approach[5] of Montgomery and Vaughan [5], where the Kronecker symbol $\left(\frac{a}{n}\right)$ is defined only when a is a *fundamental discriminant*.

Definition 5.14. A fundamental discriminant is an integer D of the following type: either (i) $D \equiv 1 \pmod 4$ and squarefree, or (ii) $D = 4m$ with $m \equiv 2, 3 \pmod 4$ and m squarefree.

Now let us give the definition of the Kronecker symbol. Suppose that D is a fundamental discriminant. Then we define

$$\chi_D(-1) = \left(\frac{D}{-1}\right) = \mathrm{sgn}(D) = \begin{cases} 1 & D > 0 \\ -1 & D < 0. \end{cases} \tag{5.5}$$

We also define

$$\chi_D(2) = \left(\frac{D}{2}\right) = \begin{cases} 0 & D \text{ even} \\ 1 & D \equiv 1 \pmod 8 \\ -1 & D \equiv 5 \pmod 8. \end{cases} \tag{5.6}$$

[5]To add further potential confusion, note that Montgomery and Vaughan use the term *quadratic discriminant* rather than fundamental discriminant.

We define $\chi_D(p) = \left(\frac{D}{p}\right)$ to agree with the Legendre symbol when p is an odd prime. We define $\chi_1(0) = \left(\frac{1}{0}\right) = 1$ and $\chi_D(0) = \left(\frac{D}{0}\right) = 0$ for $D \neq 1$.

Finally, we define χ_D to be the unique completely multiplicative function which is compatible with these values, that is to say

$$\chi_D(mn) = \chi_D(m)\chi_D(n). \tag{5.7}$$

To spell it out, if $n = up_1^{e_1} \cdots p_k^{e_k}$ with $u \in \{\pm 1\}$, we define

$$\chi_D(n) = \left(\frac{D}{n}\right) = \left(\frac{D}{u}\right) \prod_i \left(\frac{D}{p_i}\right)^{e_i}.$$

Proposition 5.15. *Suppose that D is a fundamental discriminant. Then χ_D is a Dirichlet character to modulus $|D|$.*

Proof. (Partial) Observe that $\chi_D(n) = 0$ unless $(n, |D|) = 1$, and that $\chi_D(n)$ takes the values ± 1 otherwise. By definition, χ_D is completely multiplicative. Therefore, by the discussion after the definition of Dirichlet characters in the last section, all that needs to be shown is that χ_D has $|D|$ as a period, that is to say that

$$\chi_D(n + |D|) = \chi_D(n) \tag{5.8}$$

for all integers n.

We will only prove the case D odd in detail; the case D even is similar, but fiddly around even numbers and I have chosen not to include it. Assume, then, that $D \equiv 1 \pmod 4$ and that D is squarefree. We have

$$D = (-1)^t p_1 \cdots p_k \tag{5.9}$$

where the p_i are distinct odd primes, and t is equal to the number of p_i which are $3 \pmod 4$. We claim that

$$\chi_D(n) = \prod_{i=1}^{k} \left(\frac{n}{p_i}\right) \tag{5.10}$$

for all $n \in \mathbf{Z}$. Once this is proven, the desired periodicity statement (5.8) follows immediately, since $\left(\frac{n}{p_i}\right)$ has period p_i.

It remains to establish (5.10). We first suppose that $n = p$, an odd prime not dividing D. By quadratic reciprocity we have

$$\chi_D(p) = \left(\frac{(-1)^t p_1 \cdots p_k}{p}\right) = \left(\frac{-1}{p}\right)^t \prod_{i=1}^{k} \left(\frac{p_i}{p}\right).$$

If $p \equiv 1 \pmod 4$ then by quadratic reciprocity this is $\prod_{i=1}^{k} \left(\frac{p}{p_i}\right)$. The same is true for $p \equiv 3 \pmod 4$ as well, since the number of sign changes in switching the $\left(\frac{p_i}{p}\right)$ to $\left(\frac{p}{p_i}\right)$ is exactly t. Thus we have (5.10) in these cases.

Now suppose $n = -1$. The RHS of (5.10) is exactly $(-1)^t$, where t is the number of p_i which are $3 \pmod 4$. By comparing with (5.9), this is exactly $\operatorname{sgn}(D)$ and so agrees with the definition (5.5).

Now suppose $n = 2$. By the expression for $\left(\frac{2}{p}\right)$ (see (5.2)), the RHS of (5.10) is $(-1)^u$, where u is the number of p_i which are $\pm 3 \pmod 8$. Note that D is $\pm 1 \pmod 8$ if and only if u is even. Since D is already known to be $1 \pmod 4$, $D \equiv 1 \pmod 8$ if and only if u is even, and so (5.10) follows from the definition of $\chi_D(2)$.

Finally, we note that both sides of (5.10) are zero if n is a prime dividing D, and both sides are completely multiplicative. This is enough to establish (5.10) in all cases. □

The converse of this result is almost true, but with an important technical caveat. Suppose that $q < \tilde{q}$ are two natural numbers with $q \mid \tilde{q}$. Then a Dirichlet character χ to modulus q *induces* a Dirichlet character $\tilde{\chi}$ to modulus $\tilde{q}$ by setting

$$\tilde{\chi}(n) = \begin{cases} \chi(n) & \text{if } (n, \tilde{q}) = 1 \\ 0 & \text{otherwise} \end{cases} \tag{5.11}$$

If a Dirichlet character χ is not induced from a character of (strictly) smaller modulus, we say that χ is *primitive*. Every Dirichlet character is induced from some primitive character.

Remark. If $\tilde{\chi}$ is induced from χ then $L(s, \chi)$ and $L(s, \tilde{\chi})$ are very closely related; see Example Sheet 3. For most of the theory of Dirichlet L-functions one should restrict to the primitive case (for example, the functional equation, which we have not discussed in this course, only works out nicely in this case – see [2]).

Here, then, is a more complete version of Proposition 5.16 with converse.

Proposition 5.16. *Let D be a fundamental discriminant. Then χ_D is a primitive Dirichlet character to modulus $|D|$. Moreover, all real, primitive, Dirichlet characters arise (uniquely) in this way.*

Proof. The proof will not be lectured (and is not examinable). □

The class number formula. In this section we describe some remarkable formulæ for $L(1, \chi_D)$, which may be used to show that $L(1, \chi_D) \neq 0$. In view of Proposition 5.16, this implies that $L(1, \chi) \neq 0$ for all real, primitive nonprincipal χ. The same result holds without the primitivity assumption (see Example Sheet 2, Q10 for relevant comments on this). This is the key fact we needed in Section 4. The case $D < 0$ is closely related to the work of this section.

Theorem 5.17. *Suppose that D is a fundamental discriminant and that $D < 0$. Then we have*

$$L(1, \chi_D) = \frac{2\pi h(D)}{w|D|^{1/2}}$$

where $w = 6$ if $D = -3$, $w = 4$ if $D = -4$ and $w = 2$ for all other D.

(The number w arises from forms with nontrivial automorphisms, that is to say where there are $g \neq I \in \Gamma$ with $\psi = g\psi$.)

**Positive D.* The theorem above covers negative fundamental discriminants D. There is an analogous formula for positive discriminants. It can be described either in terms of binary quadratic forms (but now indefinite ones) or in terms of class numbers of quadratic fields. For the former interpretation, see [5, Section 4.4]. In terms of class numbers of real quadratic fields, the formula is

$$L(1, \chi_D) = \frac{2h(\mathbf{Q}(\sqrt{D}))\log \varepsilon}{\sqrt{D}},$$

where ε is the *fundamental unit* in $\mathbf{Q}(\sqrt{D})$, that is to say the unique $\varepsilon > 1$ which generates the group of units up modulo ± 1.

Remarks. *These remarks are mostly aimed at those attending B3.4 Algebraic Number Theory, a truly remarkable fact is that, when $\Delta < 0$ is a fundamental discriminant, the class number $h(\Delta)$ is the same thing as the class number $h(K)$ of the imaginary quadratic field $K = \mathbf{Q}(\sqrt{\Delta})$. For example, we have seen that $h(-4) = 1$, and this corresponds to the fact that the class number of $\mathbf{Q}(i)$ is 1, or in other words that the ring of integers $\mathbf{Z}[i]$ has unique factorization into irreducibles.

Other examples of class number 1 were given on Example Sheet 3. It turns out that having $h(\Delta) = 1$ is a rare phenomenon, and in fact there are just nine fundamental discriminants Δ for which this is so: $-\Delta \in \{3, 4, 7, 8, 11, 19, 43, 67, 163\}$.

Not all discriminants arising from binary quadratic forms are fundamental discriminants. For example, the form $\psi(x_1, x_2) = x_1^2 + 3x_2^2$, has discriminant -12, which is not a fundamental discriminant. We saw on Example Sheet 3 that $h(-12) = 1$. There are three more examples, namely $-\Delta = 16, 27, 28$. (There is an algebraic number theory connection here too; for the form $\psi(x_1, x_2)$, the connection is that $\mathcal{O} = \mathbf{Z}[\sqrt{-3}]$ has class number 1. However, $\mathcal{O}$ is not the whole ring of integers in $\mathbf{Q}(\sqrt{-3})$, it is instead a *non-maximal order*. The fact that $\mathbf{Z}[\sqrt{-3}]$ has class number 1 does not mean it is a PID. We leave further exploration to the reader.)

Enumerating the reduced forms of discriminant Δ is a sensible way to compute $h(K)$ by hand for such fields. For more information, see my notes from that course [1].

6. Sums of four squares

In the last section we studied sums of two squares as part of a more general theory. A self-contained elementary proof that every prime $p \equiv 1 (\mathrm{mod}\, 4)$ is the sum of two squares was given on Example Sheet 3. In this section we prove the famous theorem of Lagrange using an argument quite similar to the latter.

Theorem 6.1. *Every natural number is the sum of four squares.*

Proof. We note the identity

$$(x_1^2+x_2^2+x_3^2+x_4^2)(y_1^2+y_2^2+y_3^2+y_4^2) = (x_1y_1+x_2y_2+x_3y_3+x_4y_4)^2+ \\ +(x_1y_2-x_2y_1+x_3y_4-x_4y_3)^2+(x_1y_3-x_3y_1+x_4y_2-x_2y_4)^2 \\ +(x_1y_4-x_4y_1+x_2y_3-x_3y_2)^2. \tag{6.1}$$

This means that the set of numbers which are the sum of four squares is closed under multiplication. Since $2 = 1^2+1^2+0^2+0^2$, it suffices to show that any odd prime p is in this set.

We now claim that there is some $m > 0$ such that

$$mp = x_1^2+x_2^2+x_3^2+x_4^2.$$

To see this, first observe that every element x of $\mathbf{Z}/p\mathbf{Z}$ is a sum of two squares, since the set S of squares $(\mathrm{mod}\, p)$ has size $\frac{1}{2}(p+1)$, and hence S and $x-S$ must intersect. Writing 1 and $-1(\mathrm{mod}\, p)$ as sums of two squares and then adding gives a sum of four squares, not all zero, which is a multiple of p.

Assume that m is minimal with this property, and suppose as a hypothesis for contradiction that $m > 1$.

Reducing each $x_i(\mathrm{mod}\, p)$ to lie in the interval $(-p/2, p/2)$, we may assume that $|x_i| < p/2$ (note that p is odd, so the inequality is indeed strict). It follows that

$$mp < 4(\frac{p}{2})^2 = p^2,$$

and so $0 < m < p$.

If m is even, then the x_i may be grouped into two pairs in which the parities are equal, say $x_1 \equiv x_2(\mathrm{mod}\, 2)$, $x_3 \equiv x_4(\mathrm{mod}\, 2)$. But then

$$\frac{1}{2}mp = (\frac{x_1+x_2}{2})^2+(\frac{x_1-x_2}{2})^2+(\frac{x_3+x_4}{2})^2+(\frac{x_3-x_4}{2})^2,$$

contrary to the minimality of m.

Suppose, then, that m is odd. Not all of the x_i are divisible by m, as this would imply $m^2 | x_1^2+x_2^2+x_3^2+x_4^2 = mp$ and so $m|p$. Since we are assuming $m > 1$, this forces $m = p$, but we have already proved that $m < p$. Pick y_i with $|y_i| < m/2$ and $x_i \equiv y_i(\mathrm{mod}\, m)$, $i = 1, \dots, 4$. This is possible with *strict* inequality, as claimed, since m is odd. Then $y_1^2+y_2^2+y_3^2+y_4^2$ is positive, and also a multiple of m since it is congruent to $x_1^2+x_2^2+x_3^2+x_4^2$. Suppose that $y_1^2+y_2^2+y_3^2+y_4^2 = rm$. Then

$$rm < 4(\frac{m}{2})^2 = m^2$$

and so $r < m$. Now from (6.1), we have

$$rm^2p = (x_1y_1+x_2y_2+x_3y_3+x_4y_4)^2+\dots,$$

where the $\dots$ comprises the three other terms in (6.1). One may easily check, using $x_i \equiv y_i(\mathrm{mod}\, m)$, that all four of the bracketed terms are multiples of

m. Therefore

$$rp = (\frac{x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4}{m})^2 + \dots,$$

with all the bracketed terms integers. Since $r < m$, this contradicts the supposed minimality of m. □

We end with some further (nonexaminable) remarks about sums of s squares for $s \notin \{2, 4\}$.

Theorems about sums of three squares lie a little deeper, at least partly because there is no analogue of the multiplicativity identity (6.1). However, any student of number theory should certainly be aware of the main result on the topic, which is often called the Gauss-Legendre three squares theorem.

Theorem 6.2. *All numbers not of the form $4^m(8k+7)$ are the sum of three squares.*

Sums of squares have a rich theory. Sums of three squares are connected to class numbers. Gauss showed that if $d > 3$ is squarefree then the number of representations of d as a sum of 3 squares is[6] $ch(\mathbf{Q}(\sqrt{-d}))$, where $c = 12$ if $d \equiv 1, 2 (\mathrm{mod}\, 4)$, $c = 24$ if $d \equiv 3 (\mathrm{mod}\, 8)$, and $c = 0$ if $d \equiv 7 (\mathrm{mod}\, 8)$.

Representations by sums of squares are intimately connected to the theory of modular forms of half integral weight. This leads to beautiful results: for example, the number of ways to write n as a sum of four squares is 8 times the sum of the divisors d of n with $4 \nmid d$.

A good source for more information is [3, Section 11.3], where one may find explicit formulae for the number of representations of n as a sum of s squares, $s = 4, 6, 8, 10, 12$. The formula for $s = 8$ is particularly clean, the number of representations in this case being $16 \sum_{d|n} (-1)^{n-d} d^3$, a formula of Jacobi.

7. Irrationals and transcendental numbers, approximation

To start the course we briefly discuss irrational numbers and their approximation by rationals. We will say more on this topic in Section 9.

Recall that

- $\mathbf{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ denotes the integers;
- $\mathbf{N} = \{1, 2, \dots\}$ denotes the natural numbers;
- $\mathbf{Q} = \{\frac{a}{b} : a, b \in \mathbf{Z}, b \neq 0\}$ denotes the rationals;
- $\mathbf{R}$ denotes the reals.

The set of irrational (real) numbers is $\mathbf{R} \setminus \mathbf{Q}$.

The rationals are dense in the reals, and in fact the following simple result of Dirichlet shows that any real number can be well approximated by rationals.

[6] This remark is intended for those who have attended B3.4.

Theorem 7.1. *Let α be irrational. Then there are infinitely many rationals $\frac{a}{q}$ with $q > 0$ such that $|\alpha - \frac{a}{q}| \leqslant \frac{1}{q^2}$.*

Proof. Write $\mathbf{R}/\mathbf{Z}$ for the reals taken modulo 1. We use the notation $\|x\|_{\mathbf{R}/\mathbf{Z}}$ to mean the distance from x to the nearest integer. Let $Q \in \mathbf{N}$. Consider the numbers $0, \alpha, 2\alpha, \dots, Q\alpha$ modulo 1. Divide $\mathbf{R}/\mathbf{Z}$ into Q intervals of length $1/Q$. By the pigeonhole principle there must be distinct i, j such that $i\alpha, j\alpha$ lie in the same interval, which then means that $\|\alpha(i-j)\|_{\mathbf{R}/\mathbf{Z}} \leqslant Q^{-1}$. Without loss of generality, $i > j$. Set $q := i - j$. Then $0 < q \leqslant Q$ and $\|\alpha q\|_{\mathbf{R}/\mathbf{Z}} \leqslant Q^{-1}$. Equivalently, there is a such that $|\alpha - \frac{a}{q}| \leqslant 1/qQ$. This implies the result. □

This result is close to best possible, as the following shows.

Proposition 7.2. *There exists an α such that $|\alpha - \frac{a}{q}| \geqslant \frac{1}{4q^2}$ for all q.*

Proof. We take $\alpha = \sqrt{2}$. Let $\frac{a}{q}$ be a rational. If $|\alpha - \frac{a}{q}| \geqslant 1$ then the result is trivial, so suppose this is not the case. Then $|\alpha + \frac{a}{q}| \leqslant 2\sqrt{2} + 1 < 4$, and so

$$\big|\alpha - \frac{a}{q}\big| = \frac{|2 - \frac{a^2}{q^2}|}{|\alpha + \frac{a}{q}|} > \frac{1}{4}\big|2 - \frac{a^2}{q^2}\big| = \frac{1}{4q^2}|2q^2 - a^2| \geqslant \frac{1}{4q^2},$$

where the last step follows from the fact that $2q^2 - a^2$ is a non-zero integer (since $\sqrt{2}$ is irrational).

□

When we look at this topic in more detail below, we will use continued fractions to give an algorithm for finding such rational approximations to α.

Many famous constants in mathematics are known to be irrational, for example the base of natural logarithms e.

Proposition 7.3. *The number e, the base of natural logarithms, is irrational.*

Proof. We of course have

$$e = 1 + 1 + \frac{1}{2!} + \dots + \frac{1}{n!} + \dots$$

Suppose that $e = \frac{p}{q}$ with $p, q \in \mathbf{N}$. By multiplying numerator and denominator by 2 if necessary, we may[7] assume $q \geqslant 2$. Then $q!e \in \mathbf{N}$. However,

$$q!e = q! + q! + \frac{q!}{2!} + \dots + \frac{q!}{q!} + \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \dots$$

and therefore

$$\frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \dots \in \mathbf{N}.$$

[7]Alternatively, e is 'obviously' not an integer, though you may care to consider why it is 'obvious' that $e \neq 3$.

However,

$$\frac{1}{q+1}+\frac{1}{(q+1)(q+2)}+\cdots<\frac{1}{q+1}+\frac{1}{(q+1)^2}+\cdots=\frac{1}{q}<1,$$

contradiction. □

The irrationality of π is a little harder - you can follow a guided proof of this on Example Sheet 4.

Algebraic and transcendental numbers.

Definition 7.4. A complex number α is *algebraic* if it is the solution to some polynomial equation $f(\alpha)=0$, where $f\in\mathbf{Z}[X]$ is a non-zero polynomial with integer coefficients. If f has degree at most d, then we say that α has degree at most d. If α is not algebraic, we say that it is *transcendental*.

Remark. One actually defines the degree of α to be the degree of the monic irreducible f which has α as a root.[8]

It is not obvious that transcendental numbers exist. There is a classic argument of Cantor which shows that they do, because the set of algebraic numbers is countable, whereas the real (or complex) numbers are uncountable[9].

One may also construct an explicit transcendental number using the following result, which states that an algebraic number cannot be *too* well approximated by rationals. The following theorem is due to Liouville.

Proposition 7.5. *Suppose that α is an algebraic number of degree at most d which is not rational. Then there is some constant c (depending on α) such that $|\alpha-\frac{a}{q}|\geqslant cq^{-d}$ for every rational $\frac{a}{q}$ with $q>0$.*

Proof. The proof is a generalisation of the proof of Proposition 7.2. Suppose that $f(\alpha)=0$, where $f(X)=b_dX^d+\cdots+b_0$ with $b_0,\dots,b_d\in\mathbf{Z}$. We may further assume (reducing d if necessary) that f is irreducible over $\mathbf{Q}$, and so $f(a/q)\neq 0$ for all rationals a/q.

Now let $\frac{a}{q}$ be a rational. If $|\alpha-\frac{a}{q}|\geqslant 1$ then the result is trivial (provided we choose $c\leqslant 1$), so assume from now on that $|\alpha-\frac{a}{q}|<1$. Since $q^df(a/q)\in\mathbf{Z}$ (and $f(a/q)\neq 0$), this means that

$$\big|f\big(\frac{a}{q}\big)\big|\geqslant q^{-d}. \tag{7.1}$$

By the factor theorem we have

$$f(X)-f\big(\frac{a}{q}\big)=\big(X-\frac{a}{q}\big)\Big(b_d\big(X^{d-1}+\cdots+(\frac{a}{q})^{d-1}\big)+\cdots+b_1\Big).$$

[8]See B3.4 Algebraic number theory for much more detail on algebraic numbers.

[9]Those students taking B1.2 Set Theory may find details of this argument on Sheet 2, Section C, Q11.

Setting $X = \alpha$ and applying (7.1), it follows that

$$q^{-d} \leqslant |\alpha - \frac{a}{q}| \cdot \Big| b_d\big(\alpha^{d-1} + \cdots + (\frac{a}{q})^{d-1}\big) + \cdots + b_1 \Big|.$$

We have $|\frac{a}{q}| \leqslant 1 + |\alpha|$, and so the second expression is bounded by some $C = C(\alpha)$ (there is no point in computing this exactly). Therefore $|\alpha - \frac{a}{q}| \geqslant \frac{1}{C} q^{-d}$, and the result follows upon taking $c = \min(1, 1/C)$. □

To construct a transcendental number, we only need write down an α which is not rational, but which is extremely well-approximated by rationals.

Theorem 7.6. *The number $\alpha = \sum_{n=1}^{\infty} 2^{-n!}$ is transcendental.*

Proof. Consider the rational $\frac{a}{q} = \sum_{n=1}^{N} 2^{-n!}$, where $q := 2^{N!}$. We have

$$|\alpha - \frac{a}{q}| = \sum_{n=N+1}^{\infty} 2^{-n!} < 2 \cdot 2^{-(N+1)!} = 2 \cdot q^{-(N+1)}. \tag{7.2}$$

We first observe that this implies α irrational. Suppose, by contrast, that $\alpha = \frac{a_0}{q_0}$. Then $\frac{a}{q} < \frac{a_0}{q_0}$ (since $\frac{a}{q}$ is a truncation of the series for $\alpha = \frac{a_0}{q_0}$, all of whose terms are positive). But then

$$|\alpha - \frac{a}{q}| = |\frac{a_0}{q_0} - \frac{a}{q}| \geqslant \frac{1}{qq_0},$$

which contradicts (7.2) for N large enough.

By (7.2), taking the sequence of such qs would contradict Proposition 7.5 if α were algebraic of degree at most some d. □

Clearly, the above construction can be varied in several ways. Numbers of this type are called *Liouville numbers*.

8. Continued fractions

Finite continued fractions. Let us consider an example of Euclid's algorithm. With initial numbers 10 and 17, the algorithm runs as follows:

$$17 = 1 \cdot 10 + 7$$
$$10 = 1 \cdot 7 + 3$$
$$7 = 2 \cdot 3 + 1$$

From this, we conclude that the greatest common divisor of 17 and 10 is 1 (and, by tracing back through the algorithm, we have a way of writing 1 as a linear combination of 10 and 17:

$1 = 7 - 2\cdot 3 = 7 - 2\cdot(10 - 1\cdot 7) = 3\cdot 7 - 2\cdot 10 = 3\cdot(17 - 1\cdot 10) - 2\cdot 10 = 3\cdot 17 - 5\cdot 10.$)

Dividing by appropriate denominators, the relations above are equivalent to

$$\frac{17}{10} = 1 + \frac{7}{10}$$

$$\frac{10}{7} = 1 + \frac{3}{7}$$
$$\frac{7}{3} = 2 + \frac{1}{3}.$$

Substituting backwards from the bottom leads to

$$\frac{10}{7} = 1 + \frac{1}{2 + \frac{1}{3}}$$

and then

$$\frac{17}{10} = 1 + \frac{1}{1 + \frac{1}{2+\frac{1}{3}}}.$$

Write

$$[a_0; a_1, \dots, a_n] := a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots + \frac{1}{a_n}}}}. \tag{8.1}$$

With this notation, we have shown that

$$\frac{17}{10} = [1; 1, 2, 3].$$

Expressions like this are called (finite) *continued fractions*. The numbers a_i, which must always be positive integers, are called the *partial quotients*.

Write $\frac{p_0}{q_0} = 1$, $\frac{p_1}{q_1} = [1;1] = 1 + \frac{1}{1} = 2$, $\frac{p_2}{q_2} = [1;1,2] = 1 + \frac{1}{1+\frac{1}{2}} = \frac{5}{3}$. These (and $\frac{p_3}{q_3} = \frac{17}{10}$ itself) are called the *convergents* to $\frac{17}{10}$, and they are particularly good approximations to it. We will see much more on this topic later.

The above works in the same way for any starting numbers (u, v) with $u \geqslant v$ in place of $(17, 10)$. The fact that the Euclidean algorithm terminates then implies the following result. For a careful proof, see Sheet 4 Q1.

Proposition 8.1. *Suppose that $u \geqslant v \geqslant 1$ are integers. Then we have $\frac{u}{v} = [a_0; a_1, \dots, a_n]$ for some $n \geqslant 0$.*

That is, any element of $\mathbf{Q} \cap [1, \infty)$ has a finite continued fraction expansion. Conversely, any finite continued fraction gives an element of $\mathbf{Q} \cap [1, \infty)$.

The representation of a rational $\frac{u}{v}$ as a continued fraction is not quite unique. For example, $[1;1,1] = 1 + \frac{1}{1+\frac{1}{1}} = [1;2] = \frac{3}{2}$. (This can be shown to be the only form of non-uniqueness)

Infinite continued fractions – introduction and motivation. The rest of this section will be about *infinite* continued fractions. For example, can we make sense of

$$[1; 2, 2, 2, 2, 2, \dots] = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2+\dots}}}? \tag{8.2}$$

Let us assume that this object exists and defines a real number α. Then we have

$$\alpha = 1 + \frac{1}{1+\alpha}.$$

Rearranging gives $\alpha = \pm\sqrt{2}$ and so, since $\alpha > 0$, $\alpha = \sqrt{2}$.

Our first main aim in this section will be to give a rigorous underpinning to the notation $[a_0; a_1, a_2, \dots]$. Then, we will explore periodic continued fractions in more detail. A particular highlight will be showing that $\sqrt{N}$ has a particularly nice type of periodic continued fraction expansion, for any $N \geqslant 2$. Specifically, we will prove the following theorem.

Theorem 8.2. *Let N be a positive integer which is not a square. Then the continued fraction expansion of $\sqrt{N}$ is periodic after the initial term. In fact,*

$$\sqrt{N} = [a_0; \overline{a_1, a_2, \dots, a_n, 2a_0}]$$

for some n, where the bar denotes that the continued fraction repeats; thus $\sqrt{N} = [a_0; a_1, \dots, a_n, 2a_0, a_1, a_2, \dots]$.

Remark. For example,

$$\sqrt{31} = [5; \overline{1, 1, 3, 5, 3, 1, 1, 10}].$$

In fact, the forwards/backwards symmetry you see here is a universal feature, but we will not prove this fact in the notes (though it is not especially hard).

Returning to (8.2), consider the truncations

$$[1;2] = \frac{3}{2}, \quad [1;2,2] = \frac{7}{5}, \quad [1;2,2,2] = \frac{17}{12}, \dots$$

One may notice some patterns. First of all, these are all rather good rational approximations to $\sqrt{2}$. For example, $\sqrt{2} = 1.4142\dots$ and $\frac{17}{12} = 1.4167\dots$. Later on we will show that, in a certain sense, the elements of this sequence are the best rational approximations to $\sqrt{2}$.

Second, observe that

$$3^2 - 2 \cdot 2^2 = 1,$$

$$7^2 - 2 \cdot 5^2 = -1,$$

$$17^2 - 2 \cdot 12^2 = 1.$$

The equation $x^2 - 2y^2 = 1$ is an instance of *Pell's equation*, and we will investigate how continued fractions may be used to solve these equations.

Analytic properties of continued fractions. Fix some sequence $a = (a_0, a_1, a_2, \dots)$ of natural numbers, all $\geqslant 1$. The aim of this section is to attach rigorous meaning to the infinite continued fraction $[a_0; a_1, a_2, \dots]$. For this part of the discussion it is convenient to allow the a_i to be real numbers, all $\geqslant 1$. Our main theorem will then be the following.

Theorem 8.3. *Suppose that $a_0, a_1, \ldots \geqslant 1$ are real numbers. Then*

$$[a_0; a_1, a_2, \ldots] := \lim_{n\to\infty} [a_0; a_1, \ldots, a_n]$$

exists.

Before giving the proof, we assemble various preliminaries which will be used in their own right later on. Write

$$\mathrm{GL}_2(\mathbf{Z}) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbf{Z}, ad - bc = \pm 1\}.$$

This acts on the *projective real line* $\mathbb{P}^1(\mathbf{R}) = \mathbf{R} \cup \{\infty\}$ by Möbius maps, that is to say

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot x := \frac{ax + b}{cx + d}.$$

(The right way to think of this is by identifying $\mathbb{P}^1(\mathbf{R})$ with non-zero pairs $\binom{t}{u}$, which are identified by projective equivalents $\binom{t}{u} \sim \binom{\lambda t}{\lambda u}$. Then $x \in \mathbf{R}$ corresponds to the equivalence class of $\binom{x}{1}$, and ∞ to the equivalence class of $\binom{1}{0}$. The Möbius action then corresponds to the usual action of matrices on vectors in $\mathbf{R}^2$. Alternatively, you will lose nothing in this course by assuming that ∞ behaves the way you think it should, for instance $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \infty = \frac{a}{c}$ if $c \neq 0$, and ∞ if $c = 0$.)

Consider the specific matrices $M(a) := \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix}$, which all lie in $\mathrm{GL}_2(\mathbf{Z})$ when a is an integer. The Möbius action is

$$M(a) \cdot x = a + \frac{1}{x}.$$

This suggests a link to continued fractions, and indeed we have the following lemma.

Lemma 8.4. *Let $a_0, a_1, \ldots$ reals. We have*

$$[a_0; a_1, \ldots, a_n] = M(a_0)M(a_1)\cdots M(a_n) \cdot \infty.$$

Proof. This is a very simple induction on n. When $n = 0$ we have $M(a_0) \cdot \infty = a_0 = [a_0]$. By inductive hypothesis we have $[a_1; \ldots, a_n] = M(a_1)\cdots M(a_n) \cdot \infty$, so

$$\begin{aligned}[a_0; a_1, \ldots, a_n] &= a_0 + \frac{1}{[a_1; \ldots, a_n]} \\ &= M(a_0) \cdot [a_1; \ldots, a_n] = M(a_0)M(a_1)\cdots M(a_n) \cdot \infty,\end{aligned}$$

as required. □

Note carefully that here we are using the group action property, namely that $M \cdot (M' \cdot x) = (MM') \cdot x$. A consequence of this is the following.

Lemma 8.5. *We have*

$$[a_0; a_1, \dots, a_n] = \frac{p_n(a)}{q_n(a)},$$

where for $n \geqslant 1$, $p_n(a), q_n(a)$ satisfy the relations

$$\begin{pmatrix} p_n(a) & p_{n-1}(a) \\ q_n(a) & q_{n-1}(a) \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \dots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}. \tag{8.3}$$

Proof. A restatement of Lemma 8.4 (using the definition of $\mathbb{P}^1(\mathbf{R})$ as equivalence classes of non-zero vectors in $\mathbf{R}^2$) is that

$$\begin{pmatrix} p_n(a) \\ q_n(a) \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \dots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Moreover, noting that

$$\begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

we have

$$\begin{pmatrix} p_{n-1}(a) \\ q_{n-1}(a) \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \dots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The result follows. □

Corollary 8.6. *We have*

$$p_n(a)q_{n-1}(a) - p_{n-1}(a)q_n(a) = (-1)^{n-1}.$$

Proof. Immediate from (8.3) by taking determinants of both sides. (This most naturally gives $(-1)^{n+1}$ but of course this is the same.) □

Another important consequence of Lemma 8.5 are the following recurrence relations.

Lemma 8.7. *We have*

$$p_n(a) = a_n p_{n-1}(a) + p_{n-2}(a)$$

and

$$q_n(a) = a_n q_{n-1}(a) + q_{n-2}(a).$$

Suppose that $a_i \geqslant 1$ for all $i \geqslant 0$. Then we have $q_n(a) \geqslant F_{n+1}$ and $p_n(a) \geqslant F_{n+2}$ where F_i is the ith Fibonacci number. Moreover we have

$$1 = q_0 \leqslant q_1 \leqslant q_2 \leqslant \dots \tag{8.4}$$

and

$$q_{i+2} \geqslant 2q_i \tag{8.5}$$

for $i \geqslant 0$.

Proof. Write $p_n = p_n(a)$ and $q_n = q_n(a)$ for brevity. It is immediate from (8.3) that

$$\begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = \begin{pmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{pmatrix} \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix},$$

which is equivalent to the stated recurrences. All the inequalities on the p_i, q_i are immediate by induction. □

We are now ready to give the proof of Theorem 8.3.

Proof. [Proof of Theorem 8.3] As in the statement of the theorem, suppose that $a_0, a_1, \ldots \geqslant 1$. It follows immediately by induction from Lemma 8.7 that $q_n(a) \geqslant F_{n+1}$ and $p_n(a) \geqslant F_{n+2}$ where F is the Fibonacci sequence. Now by Lemma 8.6 we have

$$\begin{aligned} [a_0; a_1, \ldots, a_n] - [a_0; a_1, \ldots, a_{n-1}] &= \frac{p_n(a)}{q_n(a)} - \frac{p_{n-1}(a)}{q_{n-1}(a)} \\ &= (-1)^{n-1} \frac{1}{q_{n-1}(a) q_n(a)}. \end{aligned}$$

and so

$$\left|[a_0; a_1, \ldots, a_n] - [a_0; a_1, \ldots, a_{n-1}]\right| \leqslant \frac{1}{F_n F_{n+1}}. \tag{8.6}$$

The result follows (any crude bound such as $F_n \geqslant n$ for large n will do here). □

By the same mode of argument we have the following.

Lemma 8.8. *Suppose that a_i, $i = 0, \ldots, n$ and a_n' are all real and $\geqslant 1$. Then*

$$\left|[a_0; a_1, \ldots, a_n'] - [a_0; a_1, \ldots, a_n]\right| \leqslant \frac{2}{F_n F_{n+1}}$$

Proof. This is immediate from (8.6). □

Continued fraction expansions of irrationals. From now on we will be interested in the case that all a_i are integers. As a supplement to Theorem 8.3, we have the following.

Theorem 8.9. *A real number $\alpha > 1$ may be expressed as $[a_0; a_1, a_2, \ldots]$ with $a_i \in \mathbf{N}$ if and only if α is irrational. The representation is unique.*

In the proof, and later on in the section, we will use the *continued fraction map* $T : (1, \infty) \setminus \mathbf{Z} \to (1, \infty)$ defined by

$$T(x) := \frac{1}{x - \lfloor x \rfloor}. \tag{8.7}$$

The relevance of this to continued fractions is that if we define $a_i := \lfloor T^i(\alpha) \rfloor$ then we have the identity

$$\alpha = [a_0; a_1, \ldots, a_{n-1}, T^n(\alpha)], \tag{8.8}$$

which follows by an easy induction and the relation

$$T^n(\alpha) = a_n + \frac{1}{T^{n+1}(\alpha)}.$$

Note that the assumption that α is irrational guarantees that $T^n(\alpha)$ is never an integer (since otherwise, by (8.8), α would be given by a finite continued fraction) and so all the iterates $T^i(\alpha)$ are well-defined.

Proof. [Proof of Theorem 8.9] Define $a_i := \lfloor T^i(\alpha) \rfloor$ as above. From (8.8) and Lemma 8.8, we have

$$\left|\alpha - [a_0; a_1, \dots, a_n]\right| \leqslant \frac{2}{F_n F_{n+1}} \to 0.$$

The fact that $\alpha = [a_0; a_1, a_2, \dots]$ follows.

For uniqueness, observe that $a_0 < [a_0; a_1, a_2, \dots] < a_0 + \frac{1}{a_1}$, and so $a_0 = \lfloor \alpha \rfloor$. We may then proceed by induction. □

As a consequence of Theorem 8.9 we may talk about *the* continued fraction expansion of an irrational $\alpha > 1$. The integers $a_0, a_1, a_2, \dots$ are known as the partial quotients of α.

9. Continued fractions and approximation by rationals

Earlier in the course we briefly discussed approximation by rationals, proving Dirichlet's theorem that for any α there exist infinitely many a/q with $|\alpha - a/q| \leqslant 1/q^2$. In this section we will use continued fractions to go into more detail on this topic, for example giving an algorithm to find such approximants to a given α.

Convergents. In this section, suppose that α is irrational with continued fraction expansion

$$\alpha = [a_0; a_1, \dots]$$

where the a_i, $i \geqslant 1$ are all natural numbers and $a_i \in \mathbf{Z}$ (we allow this so we are not restricted to $\alpha > 1$). Set $p_n := p_n(a)$, $q_n = q_n(a)$. The fractions $\frac{p_n}{q_n}$ are called the *convergents* to α. Here are some basic properties of the sequence of convergents.

Lemma 9.1. *p_n, q_n are coprime. The* convergents *p_n/q_n satisfy*

$$a_0 = \frac{p_0}{q_0} < \frac{p_2}{q_2} < \frac{p_4}{q_4} < \dots < \alpha < \dots < \frac{p_5}{q_5} < \frac{p_3}{q_3} < \frac{p_1}{q_1} = a_0 + \frac{1}{a_1},$$

and

$$\alpha = \lim_{n\to\infty} \frac{p_n}{q_n}.$$

We have

$$\left|\alpha - \frac{p_k}{q_k}\right| < \frac{1}{q_k q_{k+1}}. \tag{9.1}$$

Finally, the convergents get progressively closer to α:

$$\big|\alpha - \frac{p_k}{q_k}\big| < \big|\alpha - \frac{p_{k-1}}{q_{k-1}}\big|. \tag{9.2}$$

Proof. Lemma 8.6 immediately implies that p_n, q_n are coprime. Moreover, that lemma gives

$$\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} = \frac{(-1)^{n-1}}{q_n q_{n-1}}, \tag{9.3}$$

which is > 0 if n is odd and < 0 if n is even. Applying (9.3) with $n-1$ and adding we obtain

$$\frac{p_n}{q_n} - \frac{p_{n-2}}{q_{n-2}} = \frac{(-1)^{n-2}}{q_{n-1}}\Big(\frac{1}{q_{n-2}} - \frac{1}{q_n}\Big),$$

using here the fact that using the fact that $q_n > q_{n-2}$. This is > 0 if n is even and < 0 if n is odd. These two observations give the claimed ordering of the convergents The inequality (9.1) follows from (9.3) with $n = k+1$, and the fact that α lies (strictly) between $\frac{p_k}{q_k}$ and $\frac{p_{k+1}}{q_{k+1}}$.

Finally we turn to (9.2). Since $\frac{p_{k-1}}{q_{k-1}}$ and $\frac{p_k}{q_k}$ lie on opposite sides of α, it is enough to show that $\big|\frac{p_{k-1}}{q_{k-1}} - \frac{p_k}{q_k}\big| > 2\big|\alpha - \frac{p_k}{q_k}\big|$. However we have $\big|\frac{p_{k-1}}{q_{k-1}} - \frac{p_k}{q_k}\big| = \frac{1}{q_{k-1}q_k}$ and by (9.1) we know that $|\alpha - \frac{p_k}{q_k}| < \frac{1}{q_k q_{k+1}}$. The required bound then follows from the fact that $q_{k+1} \geqslant 2q_{k-1}$, which we proved in Lemma 8.7 by a simple induction using the recurrence $q_{k+1} = a_{k+1}q_k + q_{k-1}$. □

A particular consequence of (9.1) (and the fact that $q_k \leqslant q_{k+1}$) is that $|\alpha - \frac{p}{q}| < \frac{1}{q^2}$ for any convergent p/q, so the convergents verify the bound of Dirichlet's theorem and give a constructive way of finding such approximants.

Our next result is that the convergents p_k/q_k are 'the best' rational approximations to α in the following precise sense.

Theorem 9.2. *Suppose that α is an irrational number and that $\frac{p_k}{q_k}$ is a convergent to α. Let $k \geqslant 1$. Then the only rational $\frac{p}{q}$ with $q \leqslant q_k$ for which $|\alpha - \frac{p}{q}| \leqslant |\alpha - \frac{p_k}{q_k}|$ is $\frac{p_k}{q_k}$ itself.*

Proof. (May not be lectured depending on time.) We suppose for notational simplicity that k is odd; the proof in the case k even is essentially identical. By (9.2), if $\frac{p}{q}$ is a better rational approximation to α then it must lie in the interval $(\frac{p_{k-1}}{q_{k-1}}, \frac{p_k}{q_k})$. We claim that any such rational must have $q > q_k$, which concludes the proof.

We divide into two cases, according to whether $\frac{p}{q} > \alpha$ or not.

Case 1: $\alpha < \frac{p}{q} < \frac{p_k}{q_k}$. Then by (9.1) we have

$$\frac{1}{q_k q_{k+1}} > \big|\alpha - \frac{p_k}{q_k}\big| > \big|\frac{p}{q} - \frac{p_k}{q_k}\big| = \frac{|pq_k - qp_k|}{qq_k} \geqslant \frac{1}{qq_k},$$

so in this case $q > q_{k+1}$.

Case 2: $\frac{p_{k-1}}{q_{k-1}} < \frac{p}{q} < \alpha$. Then (very similarly)

$$\frac{1}{q_{k-1}q_k} > \left|\alpha - \frac{p_{k-1}}{q_{k-1}}\right| > \left|\frac{p}{q} - \frac{p_{k-1}}{q_{k-1}}\right| = \frac{|p_{k-1}q - q_{k-1}p|}{qq_{k-1}} \geqslant \frac{1}{qq_{k-1}},$$

which implies that $q > q_k$. □

10. Periodic continued fractions and quadratic irrationals

The main aim of this section is to establish Theorem 8.2, about the continued fraction expansion of $\sqrt{N}$, and related results. In this version of the notes we state the main results, enough for you to attempt all the questions on Example Sheet 4. Some or perhaps all of the proofs will be lectured, depending on time.

First, we recall Theorem 8.2.

Theorem 8.2. *Let N be a positive integer which is not a square. Then the continued fraction expansion of $\sqrt{N}$ is periodic after the initial term. In fact,*

$$\sqrt{N} = [a_0; \overline{a_1, a_2, \dots, a_n, 2a_0}]$$

for some n, where the bar denotes that the continued fraction repeats; thus $\sqrt{N} = [a_0; a_1, \dots, a_n, 2a_0, a_1, a_2, \dots]$.

Consider an (infinite) continued fraction $[a_0; a_1, a_2, \dots]$. Throughout this section, the a_i are positive integers. We say that this is

- Periodic if there is some $m \geqslant 1$ for which $a_n = a_{n+m}$ for n sufficiently large;
- Purely periodic if there is some $m \geqslant 1$ for which $a_n = a_{n+m}$ for all $n \geqslant 0$.

In the first case we write $[a_0; a_1, \dots, a_n, \overline{a_{n+1}, \dots, a_{n+m}}]$ (for some n suitably large), where the bar denotes that this portion of the continued fraction repeats. Thus, for example, $[1; \overline{2}] = [1; 2, 2, 2, \dots] = \sqrt{2}$, as we showed earlier on.

A *quadratic irrational* is any irrational (real) number satisfying a quadratic equation with coefficients in $\mathbf{Z}$. Our first result shows that all periodic continued fractions are quadratic irrationals.

Proposition 10.1. *Suppose that $\alpha = [a_0; a_1, \dots, a_n, \overline{a_{n+1}, \dots, a_{n+m}}]$ is a periodic continued fraction. Then α is a quadratic irrational.*

Proof. Set $P := M(a_0) \cdots M(a_n)$ and $Q := M(a_{n+1}) \cdots M(a_{n+m})$; thus $\alpha = P \cdot \beta$, where $\beta = \lim_{r\to\infty} Q^r \cdot \infty$, and this limit exists by Theorem 8.3. We have $Q \cdot \beta = \beta$, and therefore $M \cdot \alpha = \alpha$, where $M := PQP^{-1}$. Thus, if $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then $\alpha = \frac{a\alpha+b}{c\alpha+d}$, which gives a quadratic equation satisfied by α. □

The main highlight of this section is the following famous theorem of Lagrange, which states that the converse is also true.

Theorem 10.2. *Suppose that $\alpha > 1$ is a quadratic irrational. Then the continued fraction expansion of α is periodic.*

This is quite a substantial theorem. We will discuss at least some of the proof in lectures. A consequence of the theory we will develop in the proof is that continued fractions may be used to find solutions to *Pell's equation* $x^2 - Ny^2 = 1$ for any given N. Here is the main theorem.

Theorem 10.3. *Let N be a positive integer, not a square. Let the continued fraction expansion of $\sqrt{N}$ be $[a_0; \overline{a_1, \dots, a_n, 2a_0}]$, and consider the convergent p_n/q_n. Then we have*

$$p_n^2 - Nq_n^2 = (-1)^{n-1}.$$

In fact, $(x, y) = (p_n, q_n)$ is the smallest positive integer solution to $x^2 - Ny^2 = \pm 1$ (provided the period $\overline{a_1, a_2, \dots, a_n, 2a_0}$ is the shortest one), but we will not prove this fact.

If n is odd, $(x, y) = (p_n, q_n)$ is already a solution to Pell's equation. If n is even, one may observe that

$$(x^2 - Ny^2)^2 = (x^2 + Ny^2)^2 - N(2xy)^2$$

and so $x' = x^2 + Ny^2$, $y' = 2xy$ gives a solution.

Appendix A. Results from real analysis

This section is non-examinable.

In this section we collect some results from analysis which are necessary for the rigorous justification of some of the results of the course.

We begin with some results on differentiation and limits. The next result is Proposition 5.1 in my notes from the Oxford first-year course Analysis III: Integration.

Proposition A.1. *Suppose that $f_n : [a, b] \to \mathbf{C}$, $n = 1, 2, \dots$ is a sequence of functions with the property that f_n is continuously differentiable on (a, b), that f_n converges pointwise to some function f on $[a, b]$, and that f_n' converges uniformly to some bounded function g on (a, b). Then f is differentiable and $f' = g$.*

Remark. In my notes, the result was stated for functions with values in $\mathbf{R}$, but it applies to complex-valued functions by taking real and imaginary parts.

Corollary A.2. *Suppose that $u_n : [a, b] \to \mathbf{C}$ is a sequence of functions with the property that u_n is continuously differentiable on (a, b), and such that $|u_n(x)|, |u_n'(x)| \leqslant M_n$ with $\sum_{n=1}^{\infty} M_n < \infty$. Then $\sum_{n=1}^{\infty} u_n$ is differentiable, with derivative $\sum_{n=1}^{\infty} u_n'$.*

Proof. Follows by applying the above with $f_n = u_1 + \cdots + u_n$ and the Weierstrass M-test. □

We note that the result also holds if n ranges over $\mathbf{Z}$ (with $\sum_{n\in\mathbf{Z}} u_n$ defined to be $\sum_{n=-\infty}^{\infty} u_n = \lim_{N\to\infty}\sum_{n=-N}^{N} u_n$). It also holds for functions from $\mathbf{R}/\mathbf{Z}$ to $\mathbf{C}$, as can be seen by lifting such a function to $[0,2]$ (say).

Results of the following kind were used several times in the course to argue that a function defined by an integral is holomorphic by 'differentiating under the integral'.

Proposition A.3. *Let $I \subseteq [0,\infty)$ be a (possibly infinite) interval. Let $f : I \to \mathbf{R}$ be a function such that*

$$\int_I |f(x)|x^t dx < \infty \tag{A.1}$$

for all t in some (possibly infinite) open interval $T \subset \mathbf{R}$. For z with $\operatorname{Re} z \in T$, define

$$F(z) := \int_I f(x)x^z dx.$$

Then $F(z)$ is holomorphic in the domain $\mathcal{D} := \{z : \operatorname{Re} z \in T\}$ and

$$F'(z) = \int_I f(x)x^z \log x\, dx.$$

Proof. We verify the definition of derivative directly. Let $z \in \mathcal{D}$, and let δ be small enough that the ball $B_\delta(z)$ is contained within $\mathcal{D}$. Thus, the task is to show that

$$\lim_{h\to 0}\int_I f(x)x^z\Big(\frac{x^h-1}{h} - \log x\Big)dx = 0.$$

Since the derivative of $x^h = e^{h\log x}$ at $h = 0$ is $\log x$, we do have

$$\lim_{h\to 0}\frac{x^h-1}{h} = \log x$$

pointwise. Therefore, by the dominated convergence theorem, it is enough to show that

$$\int_I |f(x)x^z| \sup_{|h|\leqslant\delta}\Big|\frac{x^h-1}{h} - \log x\Big|dx < \infty. \tag{A.2}$$

Now $(x^z-1)/h$ is a holomorphic function of h (it has a removable singularity at $h = 0$) and so by the maximum modulus principle we have

$$\sup_{|h|\leqslant\delta}\Big|\frac{x^h-1}{h}\Big| = \sup_{|h|=\delta}\Big|\frac{x^h-1}{h}\Big| \ll_\delta 1 + x^\delta,$$

here assuming $x \geqslant 0$. It follows that the LHS of (A.2) is bounded by

$$\ll_\delta \int_I |f(x)x^z|(1 + x^\delta + |\log x|)dx \ll \int_I |f(x)x^z|(1 + x^\delta + x^{-\delta})dx.$$

(For this last step, note that $|\log x|$ is either $\ll x^{\delta}$ (for $x \geqslant 1$) or $\ll x^{-\delta}$ (for $0 < x \leqslant 1$). This is finite by the assumption (A.1) and the fact that $z \pm \delta \in \mathcal{D}$. □

Appendix B. Results from complex analysis

This section is in initial draft. Only the statements of the results in this section are examinable.

Proposition B.1. *The uniform limit of holomorphic functions on a domain is holomorphic.*

Proposition B.2. *Suppose that $(f_n(z))_{n\in\mathbf{N}}$ is uniformly Cauchy on some domain $\mathcal{D} \subset \mathbf{C}$. Then there is some holomorphic function $f : \mathcal{D} \to \mathbf{C}$ such that $f_n \to f$ uniformly.*

Proposition B.3. *Suppose that $f(z)$ is holomorphic and nonzero on some ball $B_r(z_0)$. Then there is a holomorphic branch of $\log f$, that is to say there is a holomorphic function $g(z)$ on $B_r(z_0)$ such that $e^{g(z)} = f(z)$.*

Proposition B.4. *The series $-\sum_j \frac{z^j}{j}$ defines a holomorphic function on $|z| < 1$ which is a logarithm of $1 - z$ on this domain.*

Appendix C. Real quadratic characters

This is cut and paste from my Kronecker notes, which can be referred back to if necessary.Montgomery–Vaughan Theorem 9.13 is one resource

For the rest of this section we will refer to the two different types of fundamental discriminant as type (i) and (ii) respectively. This is not standard nomenclature and we will only use it in this section.

Apart from $D = 1$, these are precisely the discriminants of quadratic fields over $\mathbf{Q}$, and indeed the discriminant of $\mathbf{Q}(\sqrt{D})$ is D. Equivalently, if m is square-free, the quadratic field $\mathbf{Q}(\sqrt{m})$ has discriminant $4m$ if $m \equiv 2, 3 \pmod 4$ and m if $m \equiv 1 \pmod 4$.

Definition C.1. A *prime discriminant* is a fundamental discriminant in the following list: $-4, 8, -8$, p for $p \equiv 1 \pmod 4$ and $-p$ for $p \equiv 3 \pmod 4$.

Lemma C.2. *The product of two coprime fundamental discriminants is a fundamental discriminant. Every fundamental discriminant is a product of coprime prime discriminants.*

Proof. For the first statement, there are three cases to consider. If D, D' are both of type (i), that is to say both are squarefree and $1 \pmod 4$, then clearly their product is also.

If D, D' are both of type (ii), they cannot be coprime.

The remaining case is that D is of type (i) and D' is of type (ii), thus $D \equiv 1 \pmod 4$ and $D' = 4m$ with $m \equiv 2, 3 \pmod 4$ and m squarefree. Then

$DD' = 4m'$ where $m' = mD$ is squarefree and congruent to m, and hence to 2 or 3, modulo 4.

Turning now to the second statement, we again split into the two types of D. If D is of type (i), thus D is squarefree and $1 (\mathrm{mod}\, 4)$, write $D = (-1)^\eta p_1 \cdots p_k$ with $\eta \in \{0,1\}$ and the p_i distinct odd primes, with $p_1, \dots, p_\ell$ being those congruent to $3 (\mathrm{mod}\, 4)$. Since $D \equiv 1 (\mathrm{mod}\, 4)$, $\ell \equiv \eta (\mathrm{mod}\, 2)$ and so $D = (-p_1)\cdots(-p_\ell)p_{\ell+1}\cdots p_k$ is a product of prime fundamental discriminants.

If D is of type (ii), write $D = 4m$ with $m = (-1)^\eta 2^\nu p_1 \cdots p_k$, with $\eta, \nu \in \{0,1\}$ and the p_i distinct odd primes, with $p_1, \dots, p_\ell$ being those congruent to $3 (\mathrm{mod}\, 4)$. Then, since $m \equiv 2, 3 (\mathrm{mod}\, 4)$, we have $(-1)^{\eta+\ell} 2^\nu \equiv 2, 3 (\mathrm{mod}\, 4)$, that is to say either (1) $\nu = 0$ and $\eta + \ell \equiv 1 (\mathrm{mod}\, 2)$ or (2) $\nu = 1$. In case (1), we have $D = (-4)(-p_1)\cdots(-p_\ell)p_{\ell+1}\cdots p_k$. In case (2), we have $D = (\pm 8)(-p_1)\cdots(-p_\ell)p_{\ell+1}\cdots p_k$ (for an appropriate choice of $\pm$). □

Lemma C.3. *Any nonzero integer $d \equiv 0, 1 (\mathrm{mod}\, 4)$ can be written as $l^2 D$ for some positive integer l and some fundamental discriminant D.*

Proof. Let e^2 be the largest square dividing d, and set $d' := d/e^2$. Thus d' is squarefree. If e is odd, then $d \equiv d' (\mathrm{mod}\, 4)$, and so $d' \equiv 1 (\mathrm{mod}\, 4)$ (it cannot be $0 (\mathrm{mod}\, 4)$, being squarefree). Thus we are done in this case, taking $l = e$ and $D = d'$.

Suppose that e is even. We must have $d' = 1, 2, 3 (\mathrm{mod}\, 4)$. If $d' \equiv 1 (\mathrm{mod}\, 4)$, we are done. If $d' = 2$ or $3 (\mathrm{mod}\, 4)$ then we have $d = l^2 D$ with $l = e/2$ and $D = 4d'$. □

Remark. In fact l and D are unique.

Corollary C.4. *Any nonzero integer d can be written as $vl^2 D$, where $v \in \{\pm 1, \pm 2\}$ and D is a fundamental discriminant.*

Proof. If $d \equiv 0, 1 (\mathrm{mod}\, 4)$ then this follows from Lemma C.3. If $d \equiv 3 (\mathrm{mod}\, 4)$ then we can take $v = -1$, and then apply Lemma C.3. If $d \equiv 2 (\mathrm{mod}\, 4)$ then we can take $v = \pm 2$ and apply Lemma C.3. □

References

[1] A. A. Author, *B3.4 algebraic number theory: notes.*

[2] Harold Davenport, *Multiplicative number theory*, third ed., Graduate Texts in Mathematics, vol. 74, Springer-Verlag, New York, 2000, Revised and with a preface by Hugh L. Montgomery.

[3] Henryk Iwaniec, *Topics in classical automorphic forms*, Graduate Studies in Mathematics, vol. 17, American Mathematical Society, Providence, RI, 1997.

[4] Henryk Iwaniec and Emmanuel Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004.

[5] Hugh L. Montgomery and Robert C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007.

Mathematical Institute, Oxford
Email address: ben.green@maths.ox.ac.uk