13. THIRTEENTH LECTURE : 24/11

We now finish off the proof of the reciprocity law.

Proof. of Theorem 11.7. By Gauss' lemma, $\left(\frac{p}{q}\right) = (-1)^l$ where l is the number of integers $x \in [1, \frac{q-1}{2}]$ such that $[px]_q < 0$. The latter inequality holds if and only if there is an integer y such that

$$-\frac{q}{2} < px - qy < 0. \tag{13.1}$$

Hence, l equals the number of integer solutions to the pair of inequalities (13.1) which in turn satisfy

$$0 < x < \frac{q}{2}. \tag{13.2}$$

Note that the right-hand inequality in (13.1) implies that $y > (\frac{p}{q})x > 0$, whereas the left-hand inequality, together with (13.2), imply that

$$qy < px + \frac{q}{2} < p\left(\frac{q}{2}\right) + \frac{q}{2} = \left(\frac{p+1}{2}\right)q \Rightarrow y < \frac{p+1}{2} \Rightarrow y < \frac{p}{2}, \tag{13.3}$$

since y is an integer. In other words, every integer solution to (13.1) and (13.2) also satisfies

$$0 < y < \frac{p}{2}. \tag{13.4}$$

A similar analysis gives that $(\frac{q}{p}) = (-1)^m$, where m is the number of intgeer solutions to (13.2), (13.4) and the double-inequality (got by simultaneously interchanging $p \leftrightarrow q$, $x \leftrightarrow y$ in (13.1),(13.2) and (13.4))

$$0 < px - qy < \frac{p}{2}. \tag{13.5}$$

Hence the LHS of (11.13) equals $(-1)^{l+m}$, where $l + m$ is the total number of integer solutions to (13.2), (13.4) and (got by combining (13.1) and (13.5))

$$-\frac{q}{2} < px - qy < \frac{p}{2}. \tag{13.6}$$

Eqs. (13.2) and (13.4) define a rectangle $\mathcal{R}$ containing $\frac{1}{4}(p-1)(q-1)$ integer points. Hence, to complete the proof, it suffices to show that there are an even number of integer points in this rectangle which do not satisfy (13.6). These points are contained in two disjoint subsets A and B of $\mathcal{R}$, where

$$A := \{(x,y) \in \mathcal{R} : px - qy < -q/2\}, B := \{(x,y) \in \mathcal{R} : px - qy > p/2\}. \tag{13.7}$$

To prove that the number of integer points in $A \cup B$ is even, it suffices to establish a 1-1 correspondence between the integer points in A and those in B. One may now tediously verify that such a correspondence is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \leftrightarrow \begin{pmatrix} \frac{1}{2}(q+1) - x \\ \frac{1}{2}(p+1) - y \end{pmatrix}. \tag{13.8}$$

□

Before moving onto a new topic, we wish to describe Jacobi's extension of Theorem 11.7.

Definition. Let n be any odd integer, and let

$$n = \prod_{i=1}^{k} p_i^{\alpha_i} \tag{13.9}$$

be its' prime factorisation. For any integer a, we define the *Jacobi symbol* $\left(\frac{a}{n}\right)$ by

$$\left(\frac{a}{n}\right) := \prod_{i=1}^{k} \left(\frac{a}{p_i}\right)^{\alpha_i}, \tag{13.10}$$

where each of the terms on the RHS is an ordinary Legendre symbol.

Remark 13.1. Note that $(\frac{a}{n}) = 0$ if and only if GCD$(a, n) > 1$. Otherwise, $(\frac{a}{n}) = \pm 1$. The Jacobi symbol is multiplicative, i.e.:

$$\left(\frac{a}{n}\right)\left(\frac{b}{n}\right) = \left(\frac{ab}{n}\right), \tag{13.11}$$

and thus defines a real character modulo n. This follows directly from the corresponding property of Legendre symbols. However, if n is not a prime, it is not necessarily the case that $(\frac{a}{n}) = 1 \Leftrightarrow a$ is a quadratic residue modulo n. Indeed, by (7.9) and Theorems 7.10 and 7.11, a is a quadratic residue mod n if and only if $(\frac{a}{p}) = 1$ for every prime p dividing n. But if, for example, n is a product of two distinct primes p and q, each congruent to 3 (mod 4), then $(\frac{-1}{p}) = (\frac{-1}{q}) = -1$, but $\left(\frac{-1}{n}\right) = \left(\frac{-1}{pq}\right) = \left(\frac{-1}{p}\right)\left(\frac{-1}{q}\right) = +1$.

Jacobi formulated the following extensions of Proposition 12.2, Theorem 12.4 and Theorem 11.7 respectively. The third part is called the *Jacobi reciprocity law.*

Theorem 13.2. *Let m, n be any two odd integers. Then*
(i)

$$\left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}} = \begin{cases} 1, & \text{if } n \equiv 1 \pmod 4, \\ -1, & \text{if } n \equiv 3 \pmod 4. \end{cases} \tag{13.12}$$

(ii)

$$\left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}} = \begin{cases} 1, & \text{if } n \equiv \pm 1 \pmod 8, \\ -1, & \text{if } n \equiv \pm 3 \pmod 8. \end{cases} \tag{13.13}$$

(iii)

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = (-1)^{\frac{1}{4}(m-1)(n-1)}. \tag{13.14}$$

Proof. The proofs of the various parts of Theorem 13.2 employ the corresponding results for Legendre symbols, the definition of the Jacobi symbol and repeated use of the fact that, if n_1, n_2 are any two odd integers, then

$$\frac{1}{2}(n_1 - 1) + \frac{1}{2}(n_2 - 1) \equiv \frac{1}{2}(n_1 n_2 - 1) \pmod 2. \tag{13.15}$$

We omit the mind-numbingly boring details. □

Example. Jacobi reciprocity provides an efficient method for computing Legendre symbols $\left(\frac{a}{p}\right)$ which involves neither factorising a (so as to reduce the problem to Gauss reciprocity) nor Euler's criterion. As an example, note that 997 is a prime and let's compute

$$\left(\frac{366}{997}\right). \tag{13.16}$$

Since $366 = 2 \cdot 183$ we have, by (13.11),

$$\left(\frac{366}{997}\right) = \left(\frac{2}{997}\right)\left(\frac{183}{997}\right). \tag{13.17}$$

Since $997 \equiv 5 \pmod 8$, eq. (13.13) implies that

$$\left(\frac{2}{997}\right) = -1. \tag{13.18}$$

Since $997 \equiv 1 \pmod 4$, eq. (13.14) implies that

$$\left(\frac{183}{997}\right) = \left(\frac{997}{183}\right). \tag{13.19}$$

Next, note that the value of a Jacobi symbol $\left(\frac{a}{n}\right)$ only depends on $a \pmod n$. Since $997 = 5 \cdot 183 + 82$, we thus have

$$\left(\frac{997}{183}\right) = \left(\frac{82}{183}\right). \tag{13.20}$$

Time-out here : at this point we have

$$\left(\frac{366}{997}\right) = -\left(\frac{82}{183}\right). \tag{13.21}$$

Now we keep applying the same type of simplifications. The reader is encouraged to check the validity of each of the following steps :

$$\left(\frac{82}{183}\right) = \left(\frac{2}{183}\right)\left(\frac{41}{183}\right) = \left(\frac{41}{183}\right) = \left(\frac{183}{41}\right) = \left(\frac{19}{41}\right) = \left(\frac{41}{19}\right) = \left(\frac{3}{19}\right) = -\left(\frac{19}{3}\right) = -\left(\frac{1}{3}\right) = -1.$$

From this and (13.21), we finally conclude that $\left(\frac{366}{997}\right) = +1$.

Remark 13.3. The above algorithm for computing a Legendre symbol $\left(\frac{a}{p}\right)$ has roughly the same complexity as Euclid's algorithm for computing $\mathrm{GCD}(a, p)$, since both involve repeated divisions.

Remark 13.4. Gauss put a lot of effort into finding generalisations of his quadratic reciprocity law to higher degree congruences. He had a lot of success in particular with cubic reciprocity, but even here the theory becomes a lot more complicated. A vast and abstract generalisation of Gauss' law was formulated by Emil Artin in the 1920s. It is called the *Artin reciprocity law* and is one of the jewels in the crown of algebraic number theory, more especially that part of the field called *Class Field Theory*. For an

introduction to modern algebraic number theory, including class field theory[1], see for example one of the following texts :

S. Lang, *Algebraic Number Theory*, Springer GTM Series.
J.W.S. Cassels and A. Fröhlich, *Algebraic Number Theory*, Academic Press (1967).

[1]You should learn both *Galois theory* and *commutative algebra* before you attempt to learn this stuff. Parts of Cassels-Fröhlich require even more algebraic prerequisites, mainly *homological algebra*.

14. FOURTEENTH LECTURE : 26/11

In this and the next lecture, I want to show how Dirichlet used characters to prove his theorem on primes in arithmetic progressions. We will not give a complete proof - this would take too long and is quite difficult - but will reduce the problem to a statement about the zeroes of certain meromorphic functions. This idea of recasting statements about the distribution of the primes as assertions about the distribution of the zeroes of certain functions of a complex variable has been the central guiding philosophy in analytic number theory since the proof of the prime number theorem in the late 19th century. Though that represents the most spectacular single success of the philosophy, Dirichlet's theorem already hinted at its power. Nowadays, its most famous expression is the still wide-open Riemann Hypothesis.

Dirichlet introduced the following generalisation of the zeta-function :

Definition. Let $d \in \mathbb{N}$ and χ be an extended Dirichlet character modulo d. The *Dirichlet L-function* for χ is the function of a complex variable s defined by

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}, \tag{14.1}$$

whenever the series converges.

Note that if $d = 1$ and $\chi = \chi_0$ then $L(s,\chi) = \zeta(s)$. The basic issue is now for which $s \in \mathbb{C}$ the series converges. The answer is

Proposition 14.1. *If χ is a trivial character, then the series converges for Re$(s) > 1$. If χ is non-trivial, then it converges for Re$(s) > 0$.*

The first statement is proven by simply noting that $|L(s,\chi)| \leq |\zeta(s)|$, whenever Re$(s) > 1$. What's really new here is the second statement. To prove this, we need some lemmas, namely : a general fact about characters, plus the so-called Dirichlet convergence test. Regarding the former, the result we need is

Lemma 14.2. *Let G be a finite abelian group.*
(i) Let $g \in G$. Then

$$\sum_{\chi \in \hat{G}} \chi(g) = \begin{cases} |G|, & \text{if } g = 1_G, \\ 0, & \text{if } g \neq 1_G. \end{cases} \tag{14.2}$$

(ii) Let $\chi \in \hat{G}$. Then

$$\sum_{g \in G} \chi(g) = \begin{cases} |G|, & \text{if } \chi = \chi_0, \\ 0, & \text{if } \chi \neq \chi_0. \end{cases} \tag{14.3}$$

Proof. (i) If $g = 1_G$ then $\chi(g) = 1$ for every $\chi \in \hat{G}$. Hence $\sum_\chi \chi(g) = |\hat{G}| = |G|$, by Theorem 11.4. If $g \neq 1_G$ then there exists at least one character - let's pick one and denote it χ_g - with the property that $\chi_g(g) \neq 1$. This follows from the proof of Theorem 11.4. Then, since $\hat{G}$ is a group, we have that

$$\sum_{\chi} \chi(g) = \sum_{\chi} (\chi \cdot \chi_g)(g) = \chi_g(g) \cdot \sum_{\chi} \chi(g). \tag{14.4}$$

Since $\chi_g(g) \neq 1$, it follows that the sum must be zero, v.s.v.

(ii) The proof is similar. If $\chi = \chi_0$ then $\chi_0(g) = 1$ for every $g \in G$, thus $\sum_g \chi_0(g) = |G|$. If $\chi \neq \chi_0$ then there is some $g_* \in G$ such that $\chi(g_*) \neq 1$. Now, since G is a group,

$$\sum_g \chi(g) = \sum_g \chi(g_* g) = \chi(g_*) \cdot \sum_g \chi(g). \tag{14.5}$$

Since $\chi(g_*) \neq 1$, the sum must equal zero, v.s.v. □

The convergence criterion we need is a classical one which many of you may have already seen in some other course.

Lemma 14.3. (Dirichlet's convergence test) *Let $(a_n)_1^\infty$ and $(b_n)_1^\infty$ be sequences of complex and positive real numbers respectively. Suppose the following two conditions are satisfied :*

(i) the sequence $(A_N)_1^\infty$ is bounded, where $A_N := \sum_{n=1}^N a_n$,
(ii) the b_n form a non-increasing sequence and $\lim_{n\to\infty} b_n = 0$.

Then the sequence $\sum_{n=1}^\infty a_n b_n$ converges.

Proof. I will not give full details but just indicate the idea. One uses the so-called *Abel partial summation formula*. This states that, for any sequences (a_n), (b_n) of complex numbers, with the sequence (A_N) defined as above we have, for any $N > 0$, that

$$\sum_{n=1}^{\infty} a_n b_n = \sum_{n=1}^{N} A_n(b_n - b_{n+1}) + A_{N+1} b_{N+1} + \sum_{n=N+2}^{\infty} a_n b_n. \tag{14.6}$$

Specifically, one uses (14.6) to show that, if the conditions of Dirichlet's test are met, then the tails

$$\tau_N = \sum_{n=N}^{\infty} a_n b_n \tag{14.7}$$

of the product series form a Cauchy sequence, which suffices to prove that the series converges. □

Proof. of Proposition 14.1. We can now complete the proof of the assertion that the series defining $L(s,\chi)$ converges when $\mathrm{Re}(s) > 0$, for any non-trivial character χ. First, to keep things simple, let me assume that $s \in \mathbb{R}$. Set $a_n = \chi(n)$, $b_n = n^{-s}$. The second condition in Dirichlet's test is clearly satisfied when $s > 0$. So is the first condition by Lemma 14.2(ii), applied to the group $G = \mathbb{Z}_d^*$, where d is the modulud of χ. By definition of a Dirichlet character, the sequence a_n is periodic with period d and, when χ is non-trivial, the lemma implies that the sum of the a_n over any single period is zero.

The proof for non-real s requires a little more careful analysis, and I don't want to go into it. It is left as an (optional !) exercise. □

Remark 14.4. Using Abel's summation formula it can be shown that for non-trivial χ (resp. trivial χ) and any $\delta > 0$, the series $L(s,\chi)$ converges uniformly in $\mathrm{Re}(s) > \delta$ (resp. $\mathrm{Re}(s) > 1+\delta$) and hence that $L(s,\chi)$ is analytic in $\mathrm{Re}(s) > 0$ (resp. $\mathrm{Re}(s) > 1$).

We continue next day ...

15. FIFTEENTH LECTURE : 27/11

Once we've introduced the right generalisation of the ζ-function, the idea is to imitate Euler's approach to proving Corollary 5.5. The first step is a generalisation of Theorem 5.3 :

Lemma 15.1. *For any character χ we have, when Re$(s) > 1$, that*

$$L(s,\chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}. \tag{15.1}$$

Proof. Follow exactly the same approach as in the proof of Theorem 5.3. Further details omitted. □

We are now ready to state our main result :

Theorem 15.2. *Let $a, d > 0$ and GCD$(a,d) = 1$. If*

$$L(1,\chi) \neq 0, \tag{15.2}$$

for every non-trivial character χ modulo d, then there are infinitely many primes $p \equiv a \pmod d$. Moreover, the sum of their reciprocals diverges.

Proof. Let χ be a character modulo d and consider the L-function $L(s,\chi)$. Suppose Re$(s) > 1$. Take log of both sides of (15.1) and expand the RHS in a Taylor series to obtain, for Re$(s) > 1$, that

$$\log L(s,\chi) = \sum_p \sum_{m=1}^{\infty} \frac{\chi(p^m)}{mp^{ms}}. \tag{15.3}$$

The difference between this and the situation with Corollary 5.5 is that now we are only interested in those primes $p \equiv a \pmod d$, for some a with GCD$(a,d) = 1$. So how do we isolate these primes in the sum (15.3) ? The trick is to use Lemma 14.2(i) this time, again applied to the group $G = \mathbb{Z}_d^*$. It yields that

$$\sum_\chi \bar\chi(a)\chi(n) = \begin{cases} \phi(d), & \text{if } n \equiv a \pmod d, \\ 0, & \text{otherwise,} \end{cases} \tag{15.4}$$

where the sum is taken over all extended Dirichlet characters modulo d.

So what we do now is to take a weighted sum of both sides of (15.3) over all such characters. By (15.4), we get that

$$\frac{1}{\phi(d)} \sum_\chi \bar\chi(a) \log L(s,\chi) = \sum_{m=1}^{\infty} \sum_{p^m \equiv a \pmod q} \frac{1}{mp^{ms}}. \tag{15.5}$$

Here we're still assuming that Re$(s) > 1$, and the sum is taken over all characters modulo d. Next, as in the proof of Corollary 5.5, we split the terms of the right-hand sum into two groups, those with $m = 1$ and those with $m > 1$. We observe as before that the latter sum is bounded as $s \to 1$ and conclude that

$$\lim_{s\to 1^+} \frac{1}{\phi(d)} \sum_\chi \bar\chi(a) \log L(s,\chi) = \lim_{s\to 1^+} \sum_{p \equiv a \pmod q} p^{-s} + O(1). \tag{15.6}$$

Dirichlet's theorem is precisely the statement that the limit of the right-hand sum is $+\infty$. Hence we have reduced the proof of the theorem to showing that

$$\lim_{s\to 1^+} \frac{1}{\phi(d)} \sum_{\chi} \bar{\chi}(a) \log L(s,\chi) = +\infty. \tag{15.7}$$

If $\chi = \chi_0$, the trivial character, then it is easy to see that $L(s,\chi_0) \to +\infty$ as $s \to 1^+$. Indeed, the same method as for the proof of (15.1) can be used to show that

$$L(s,\chi_0) = \left[\prod_{p|d} \left(1 - \frac{1}{p^s}\right)\right] \cdot \zeta(s), \quad \text{when } \mathrm{Re}(s) > 1. \tag{15.8}$$

Hence, (15.7) would be proven if we could show that $\log L(s,\chi)$ were bounded, as $s \to 1^+$, for every $\chi \neq \chi_0$.

But we know from Proposition 14.1 and Remark 14.4 that if $\chi \neq \chi_0$, then $L(s,\chi)$ is analytic in the range $\mathrm{Re}(s) > 0$. In particular, $L(s,\chi)$ is bounded as $s \to 1$. Hence, by choosing a suitable branch of the logarithm, the same is true of $\log L(s,\chi)$ unless $L(1,\chi) = 0$. This completes the proof of Theorem 15.2. □

This is as far as we shall go with 'classical analytic number theory' in this course. Before moving on in a new direction, I want to prove one final result which has been mentioned earlier, namely the famous Four Squares Theorem of Lagrange (Theorem 10.2). A full proof will be presented next day. Today, I will just prove a lemma, which is reminiscent of Lemma 9.5 :

Lemma 15.3. *The set of non-negative integers which can be expressed as a sum of four squares is closed under multiplication.*

Proof. This follows immediately from an algebraic identity analagous to (9.5), namely

$$\begin{aligned}(a^2+b^2+c^2+d^2)(e^2+f^2+g^2+h^2) \\ = (ae+bf+cg+dh)^2 + (af-be+dg-ch)^2 \\ +(ag-ce+bh-df)^2 + (ah-de+cf-bg)^2.\end{aligned} \tag{15.9}$$

Like previously, the 'right' way to think about this identity is as an expression of the fact that

$$|z_1 z_2| = |z_1||z_2|, \tag{15.10}$$

for certain kinds of 'numbers' z_1, z_2. The question is, what kinds of numbers ? The answer to this was really only worked out many years after Lagrange by the Irish mathematician and physicist Hamilton, and the numbers in question are called *quaternions*. The set of quaternions is denoted $\mathbb{H}$. First of all, $\mathbb{H}$ is a 4-dimensional vector space over $\mathbb{R}$, which we can write formally as

$$\mathbb{H} = \mathbb{R} \oplus \mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}k. \tag{15.11}$$

In particular, quaternions can be added componentwise. We can make $\mathbb{H}$ into a ring by means of the following multiplication rules :

$$i^2 = j^2 = k^2 = -1, \tag{15.12}$$

$$ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j. \tag{15.13}$$

Note that the fact that $i^2 = -1$ means that we can identify the subspace $\mathbb{R} \oplus \mathbb{R}i$ with $\mathbb{C}$, so that $\mathbb{H}$ is also a 2-dimensional vector space over $\mathbb{C}$. As such, it cannot be a field, since $\mathbb{C}$ is *algebraically closed* by the Fundamental Theorem of Algebra, and hence has no finite-dimensional field extensions. And, sure enough, $\mathbb{H}$ isn't a field, since (15.13) tells us that multiplication is not commutative. However, it turns out that all the other field axioms are satisfied, so $\mathbb{H}$ is a so-called *division ring*[2]. In particular, this means that every non-zero quaternion has a multiplicative inverse. This is, in fact, demonstrated in the same way as for complex numbers. Let

$$z := a + bi + cj + dk \tag{15.14}$$

be a quaternion. We can define its *conjugate* by

$$\overline{z} = a - bi - cj - dk \tag{15.15}$$

and its *absolute value*, a non-negative real number, by

$$|z| = \sqrt{a^2 + b^2 + c^2 + d^2}. \tag{15.16}$$

Then one can check that, if $z \neq 0$, then

$$\frac{1}{z} = \frac{\overline{z}}{|z|^2}. \tag{15.17}$$

Most importantly, the identity (15.9) is just the explicit form of (15.10) for quaternionic absolute values. □

[2]A famous theorem of Frobenius states that the only division rings which are finite-dimensional as vector spaces over $\mathbb{R}$ are $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$ (where the dimensions are 1,2 and 4 respectively). Hence, though you can go beyond $\mathbb{C}$ to $\mathbb{H}$ as long as you're willing to do without commutativity of multiplication, in some very strong sense, you really, really cannot go beyond $\mathbb{H}$.

16. SIXTEENTH LECTURE : 28/11

Proof. of Theorem 10.2. By Lemma 15.3, it suffices to prove that every prime is a sum of four squares. And since $2 = 1^2+1^2+0^2+0^2$, we may confine ourselves to odd primes.

So let p be an odd prime. Let $l > 0$ be such that lp is the smallest non-zero multiple of p expressible as the sum of four squares. Our aim is to show that $l = 1$. We achieve this in several steps :

Step 1 : $l < p$.

As x runs over all residue classes modulo p, so x^2 runs over $\frac{p+1}{2}$ distinct classes (this follows from (11.12)). Similarly, as y runs over all classes mod p, so $-1-y^2$ runs over $\frac{p+1}{2}$ different classes. By the Pigeonhole principle, and the fact that $a^2 \equiv (-a)^2 \pmod p$, there exist $x, y \in [0, p/2)$ such that $x^2 \equiv -1 - y^2 \pmod p$, hence $x^2 + y^2 + 1 = rp$, for some integer r. Thus rp is a sum of four squares. But $x, y \in [0, p/2) \Rightarrow r < p$ (in fact, $r < p/2$). Hence $l < p$, as required.

Step 2 : l is odd.

Suppose

$$x^2 + y^2 + z^2 + w^2 = rp, \tag{16.1}$$

where r is even. Then rp is even, hence an even number of x, y, z and w have to be even. Hence, WLOG, $x \equiv y \pmod 2$ and $z \equiv w \pmod 2$. But then

$$\left(\frac{r}{2}\right) p = \left(\frac{x+y}{2}\right)^2 + \left(\frac{x-y}{2}\right)^2 + \left(\frac{z+w}{2}\right)^2 + \left(\frac{z-w}{2}\right)^2, \tag{16.2}$$

and the RHS is a sum of four integer squares. This proves that l must be odd.

Step 3 : We now suppose that $l > 1$ and obtain a contradiction. Let

$$x^2 + y^2 + z^2 + w^2 = lp \tag{16.3}$$

be any representation of lp as a sum of four integer squares. Let a, b, c, d be the numerically least residues of x, y, z and w respectively, modulo l, as defined in the statement of Gauss' lemma. Then $a^2 + b^2 + c^2 + d^2 \equiv x^2 + y^2 + z^2 + w^2 \equiv 0 \pmod l$, say

$$a^2 + b^2 + c^2 + d^2 = kl. \tag{16.4}$$

Since l is odd, each of a, b, c and d lies in the OPEN interval $(-l/2, l/2)$ and hence $k < l$. By (15.9), the number $(kl)(lp) = l^2(kp)$ can be written as the sum of four integer squares, which we denote E, F, G and H. By inspection of (15.9) and the fact that $x \equiv a, y \equiv b, z \equiv c$ and $w \equiv d \pmod l$, we see that each of E, F, G and H is divisible by l. Hence, dividing across by l^2, we find that

$$kp = \left(\frac{E}{l}\right)^2 + \left(\frac{F}{l}\right)^2 + \left(\frac{G}{l}\right)^2 + \left(\frac{H}{l}\right)^2, \tag{16.5}$$

is a sum of four integer squares. Since $k < l$, this contradicts the definition of l unless $k = 0$. But if $k = 0$ then, by (16.4), each of a, b, c and d equals zero, hence each

of x, y, z and w is divisible by l. But then the LHS of (16.3) is divisible by l^2, which implies that $l \mid p$. But p is a prime so either $l = p$, which is impossible by ***Step 1***, or $l = 1$, v.s.v. $\square$

For the remainder of the course, we are going to ask questions of a different character than previously. Basically, one can say that a lot of classical number theory is concerned with the properties of certain specific sets of numbers, like primes, squares, numbers satisfying specific Diophantine equations etc. In the twentieth century it has gradually become more popular to ask questions about general, or 'random' sets of numbers. This is, of course, a vague statement, though it does reflect a fundamentally different viewpoint. Tackling such questions has in turn brought with it new methods to number theory, many combinatorial or probabilistic, as well as finding new applications of more classical techniques like complex analysis and Fourier analysis. I want to introduce two subjects, which I will designate *additive number theory* and *Ramsey theory*[3].

ADDITIVE NUMBER THEORY

The fundamental concept in this area is the following :

Definition. Let $A \subseteq \mathbb{Z}$. The *sumset* $A + A$ is the set consisting of all integers which can be expressed as a sum of two elements of A, i.e.:

$$A + A = \{n \in \mathbb{Z} : n = a_1 + a_2 \text{ for some } a_1, a_2 \in A\}. \tag{16.6}$$

Note that it is allowed to have $a_1 = a_2$. An alternative notation for $A + A$ is $2A$. Don't confuse this with $2 * A$, which denotes the *dilation*

$$2 * A = \{2a : a \in A\}. \tag{16.7}$$

In fact, observe that $2 * A \subseteq 2A$.

Example. $A = \{0, 1, 3, 4, 7\}$. Then one just computes by hand that

$$A + A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 14\}. \tag{16.8}$$

Example. $A = \{n^2 : n \in \mathbb{N}\}$. Then, by Theorem 9.6, $A + A$ consists of all $n \in \mathbb{N}$ such that, if p is a prime dividing n and $p \equiv 3 \pmod 4$, then $p^{2k} || n$, for some $k > 0$.

Definition. Let $h \in \mathbb{N}$ and $A \subseteq \mathbb{Z}$. The *h-fold sumset* hA is defined recursively by

$$1A = A, \quad 2A = A + A, \quad hA = (h-1)A + A, \text{ for } h \geq 3. \tag{16.9}$$

In other words,

$$hA = \{n \in \mathbb{Z} : n = a_1 + \cdots + a_h, \text{ for some } a_1, ..., a_h \in A.\} \tag{16.10}$$

One of the classical notions of additive number theory is given by the next definition. In what follows, $\mathbb{N}_0$ denotes the set of non-negative integers.

[3]Another popular heading is *additive combinatorics*, which captures large parts of both subjects.

Definition. Let $A \subseteq \mathbb{N}_0$ and assume $0 \in A$. We say that A is a *(non-negative integer) basis* of order $h > 0$ if $hA = \mathbb{N}_0$ and $(h-1)A \neq \mathbb{N}_0$.

In words, A is a basis of order h if every positive integer can be expressed as a sum of at most h non-zero elements of A, and h is the least number for which this is the case.

Examples. Theorems 9.6, 10.1 and 10.2 together say that the squares form a basis of order 4. The Goldbach Conjecture asserts that the set $A = \{0, 1\} \cup \mathbb{P}$, where $\mathbb{P}$ denotes the set of primes, is a basis of order 3.