

# FOUNDATIONS AND UNIQUE FACTORIZATION

PETE L. CLARK

## 1. FOUNDATIONS

What is number theory?

This is a difficult question to answer: number theory is an area, or collection of areas, of pure mathematics that have been studied for well over two thousand years. As such, it means different things to different number theorists (of which I am one). Nevertheless the question is not nearly as subjective as “What is truth?” or “What is beauty?”: all of the things that various people call number theory are related, in fact deeply and increasingly related over time.

If you think about it, it is hard to give a satisfactory definition of any area of mathematics that would make much sense to someone who has not taken one or several courses in it. One might say that analysis is the study of limiting processes, especially summation, differentiation and integration; that algebra is the study of algebraic structures like groups, rings and fields; and that topology is the study of topological spaces and continuous maps between them. But these descriptions function more by way of *dramatis personae* than actual explanations; less pretentiously, they indicate (some of) the *objects* studied in each of these fields, but they do not really tell us which properties of these objects are of most interest and which questions we are trying to answer about them.<sup>1</sup> Such motivation is hard to provide in the abstract – much easier, and more fruitful, is to give examples of the types of problems that mathematicians in these areas are or were working on. For instance, in algebra one can point to the classification of finite simple groups, and in topology the Poincaré conjecture. Both of these are problems that had been open for long periods of time and have been solved relatively recently<sup>2</sup>, so one may glean that these topics have been central to their various subjects for some time.

What are the “objects” of number theory analogous to the above description? A good one sentence answer is that number theory is the study of the integers, i.e., the positive and negative whole numbers.

Of course this is not really satisfactory: astrology, accounting and computer science, for instance, could plausibly be described in the same way. What properties of the integers are we interested in?

---

<sup>1</sup>Actually it seems easier to describe analysis than algebra or topology.

<sup>2</sup>Whether or not the three-dimensional Poincaré conjecture has truly been solved is not really clear at the moment of writing – December 2006 – but there is no apparent reason to doubt that it has.

The most succinct response seems to be that we are interested in the integers *as a ring*: namely, as endowed with the two fundamental operations of addition  $+$  and multiplication  $\cdot$  and – especially – the interactions between these two operations.

Let us elaborate on this a bit. Consider first the non-negative integers – which, as is traditional, we will denote by  $\mathbb{N}$  – endowed with the operation  $+$ . This is a very simple structure: we start with 0, the additive identity, and get every positive integer by repeatedly adding 1.<sup>3</sup> In some sense the natural numbers under addition are the simplest nontrivial algebraic structure.

Note that subtraction is not in general defined on the natural numbers: we would like to define  $a - b = c$  in case  $a = b + c$ , but of course there is not always such a natural number  $c$  – e.g.  $3 - 5$ . There are two complementary responses to this: the first is to *formally extend* the natural numbers so that additive inverses always exist: in this way we get the integers  $\mathbb{Z}$  under addition.<sup>4</sup> The second response is to record the relation between two natural numbers  $a$  and  $b$  such that  $b - a$  exists as a natural number. Of course this relation is just that  $a \leq b$ . This is quite a simple relation on  $\mathbb{N}$ : indeed, for any pair of integers, we have either  $a \leq b$  or  $b \leq a$ , and we have both exactly when  $a = b$ .<sup>5</sup>

Now for comparison consider the positive integers

$$\mathbb{Z}^+ = 1, 2, 3, \dots$$

under the operation of multiplication. This is a richer structure: whereas additively, there is a single building block – 1 – the multiplicative building blocks are the prime numbers 2, 3, 5, 7,  $\dots$ . Of course the primes are familiar objects, but the precise analogy with the additive case may not be as familiar, so let us spell it out carefully: just as subtraction is not in general defined on  $\mathbb{N}$ , division is not in general defined on  $\mathbb{Z}^+$ . On the one hand we can “formally complete”  $\mathbb{Z}^+$  by adjoining multiplicative inverses, getting this time the positive rational numbers  $\mathbb{Q}^+$ . However, again one can view the fact that  $a/b$  is not always a positive integer as being intriguing rather than problematic, and we again consider the relation between two positive integers  $a$  and  $b$  that  $b/a$  be a positive integer: in other words, that there exist a positive integer  $c$  such that  $b = a \times c$ . In such a circumstance we say that  $a$  *divides*  $b$ , and write it as  $a|b$ .<sup>6</sup> It is easy to see that the relation of divisibility is more complicated than the relation  $\leq$  since divisibility is not a total ordering: e.g.  $2 \nmid 3$  and also  $3 \nmid 2$ . What are we to make of this divisibility relation?

First, on a case-by-case basis, we do indeed know how to determine whether or not  $a|b$ : we have the following fact which is truly one of the first pieces of mathematics we learn:

---

<sup>3</sup>Here I am alluding to the fact that in the natural numbers, addition can be defined in terms of the “successor” operation  $s(n) = n + 1$ , as was done by the 19th century mathematical logician Giuseppe Peano. No worries if you have never heard of the Peano axioms – their importance lies in the realm of mathematical logic rather than arithmetic itself.

<sup>4</sup>When regarded in proper generality, the process which gets us from  $\mathbb{N}$  to  $\mathbb{Z}$  can be viewed as the *group completion* of any commutative monoid. In this course, we shall only pursue algebraic formalism when it is essential to do so: in this case, it is not.

<sup>5</sup>That is to say, the relation  $\leq$  on  $\mathbb{N}$  is a linear, or total, ordering.

<sup>6</sup>Careful:  $a|b \iff \frac{b}{a}$  is an integer: sorry about that!

**Proposition 1.** (*Division Theorem*) For any positive integers  $n$  and  $d$ , there exist unique non-negative integers  $q$  and  $r$  with  $0 \leq r < d$  and  $n = qd + r$ .

This is a very useful tool, but it does not tell us the *structure* of  $\mathbb{Z}^+$  under the divisibility relation. To address this, the primes inevitably come into play: there is a unique minimal element of  $\mathbb{Z}^+$  under divisibility, namely 1 (in other words, 1 divides every positive integer and is the only positive integer with this property): it therefore plays the analogous role to 0 under  $\leq$  on  $\mathbb{N}$ . In  $\mathbb{N} \setminus 0$ , the unique smallest element is 1. In  $\mathbb{Z}^+ \setminus 1$  the smallest elements are the primes  $p$ . Given that the definition of a prime is precisely an integer greater than one divisible only by one and itself, this is clear. The analogue to repeatedly adding 1 is taking repeated powers of a single prime: e.g., 2,  $2^2$ ,  $2^3$ ,  $\dots$ . However, we certainly have more than one prime – in fact, as you probably know and we will recall soon enough, there are infinitely many primes – and this makes things more complicated. This suggests that maybe we should consider the divisibility relation one prime at a time.

So, for any prime  $p$ , let us define  $a|_p b$  to mean that  $\frac{b}{a}$  is a rational number which, when written in lowest terms, has denominator *not* divisible by  $p$ . For instance,  $3|_2 5$ , since  $\frac{5}{3}$ , while not an integer, doesn't have a 2 in the denominator. For that matter we see that  $3|_p 5$  for all primes  $p$  different from 3, and this suggests the following:

**Proposition 2.** For any  $a, b \in \mathbb{Z}^+$ ,  $a|b \iff a|_p b$  for all primes  $p$ .

Proof: Certainly if  $a|b$ , then  $a|_p b$  for all primes  $p$ . For the converse, write  $\frac{b}{a}$  in lowest terms, say as  $\frac{B}{A}$ . Then  $a|_p b$  iff  $A$  is not divisible by  $p$ . But the only positive integer which is not divisible by any primes is 1.

In summary, we find that the multiplicative structure of  $\mathbb{Z}^+$  is similar but infinitely more complicated than the additive structure of  $\mathbb{N}$ : instead of there being one “generator”, namely 1, such that every element can be obtained as some power of that generator, we have infinitely many generators – the primes – and every element can be obtained (uniquely!) by taking each prime a non-negative integer number of times (which must be zero for all but finitely many primes). Nevertheless this switch from one generator to infinitely many does not in itself cause much trouble: given

$$a = p_1^{a_1} \cdots p_n^{a_n} \cdots$$

and

$$b = p_1^{b_1} \cdots p_n^{b_n} \cdots$$

we find that  $a | b$  iff  $a |_p b$  for all  $p$  iff  $a_i \leq b_i$  for all  $i$ . Similarly, it is no problem to multiply the two integers: we just have

$$ab = p_1^{a_1+b_1} \cdots p_n^{a_n+b_n} \cdots$$

Thus we can treat positive integers under multiplication as vectors with infinitely many components, which are not fundamentally more complicated than vectors with a single component.

The real trouble begins when we attempt to *mix* the additive and multiplicative structures. If we write integers in standard decimal notation, it is easy to add them, and if we write integers in the above “vector” factored form, it is easy to multiply

them. But what is the prime factorization of  $2^{13} + 3^{12}$ ? It's not trivial to say: in practice, the problem of given an integer  $n$ , finding its prime power factorization (1) is extremely computationally difficult, to the extent that most present-day security rests on this difficulty.<sup>7</sup>

If we ask even the easiest questions which mix the additive and multiplicative structure, we find ourselves in trouble fast. For instance, although in the multiplicative structure, each of the primes just rests “on its own axis” as a generator, in the additive structure we can ask where the primes occur with respect to the relation  $\leq$ . We do not have anything approaching a formula for  $p_n$ , and the task of describing the distribution of the  $p_n$ 's inside  $\mathbb{N}$  is a branch of number theory in and of itself (we will see a taste of it later on). For instance, consider the quantity  $g(n) = p_{n+1} - p_n$ , the “ $n$ th prime gap.” For  $n > 1$ , the primes are all odd, so  $g(n) \geq 2$ . Computationally one finds lots of instances when  $g(n)$  is exactly 2, e.g. 5, 7, 11, 13, and so forth: an instance of  $g(n) = 2$  – equivalently, of a prime  $p$  such that  $p + 2$  is also a prime – is called a *twin prime pair*. The trouble is that knowing the factorization of  $p$  tells us exactly nothing about the factorization of  $p + 2$ . Whether or not there are infinitely many twin primes is a big open problem in number theory.

It goes on like this: suppose we ask to represent numbers as a sum of two odd primes. Then such a number must be even and at least 6, and experimenting, one soon is led to guess that every even number at least 6 is a sum of two odd primes: this is known as Goldbach's Conjecture, and is about 400 years old. It remains unsolved. There are many, many such easily stated unsolved problems which mix primes and addition: for instance, how many primes  $p$  are of the form  $n^2 + 1$ ? Again, it is a standard conjecture that there are infinitely many, and it is wide open. Note that if we asked instead how many primes were of the form  $n^2$ , we would have no trouble answering – the innocent addition of 1 gives us terrible problems.

Lest you think we are just torturing ourselves by asking such questions, let me mention three amazing positive results:

**Theorem 3.** (*Fermat's Two Squares Theorem*) *A prime  $p > 2$  is of the form  $x^2 + y^2$  iff it is of the form  $4k + 1$ .*

This is arguably the first beautiful theorem of number theory. It says that to check whether an odd prime satisfies the very complicated condition of being a sum of two (integer, of course!) squares, all we need to do is divide it by four: if its remainder is 1, then it is a sum of two squares; otherwise its remainder will be 3 and it will not be a sum of two squares.

**Theorem 4.** (*Lagrange's Four Squares Theorem*) *Every prime number – indeed, every positive integer – is of the form  $x^2 + y^2 + z^2 + w^2$ .*

**Theorem 5.** (*Dirichlet, 1837*) *Suppose  $a$  and  $b$  are coprime positive integers (i.e., they are not both divisible by any integer  $n > 1$ ). Then there are infinitely many primes of the form  $an + b$ .*

---

<sup>7</sup>A systematic study of the difficulty of factoring and its cryptographic implications is the topic of our “sister” course 4450, so I will say almost nothing about it here.

Remark: In particular, taking  $a = 4$ ,  $b = 1$ , see that there are infinitely many primes of the form  $4k + 1$ , so in particular there are infinitely many primes which are a sum of two squares.

We will see proofs of Theorems 3 and 4 in this course: indeed we will prove Theorem 3 several times and try to extract as much insight as possible from the different proofs. The proof of Theorem 5 is beyond our ambitions in this course: it requires more sophisticated techniques – both algebraic and analytic – than we shall introduce.

**Admission:** In fact there is a branch of number theory which studies only the addition operation on subsets of  $\mathbb{N}$ : if  $A$  and  $B$  are two subsets of natural numbers, then by  $A+B$  we mean the set of all numbers of the form  $a+b$  for  $a \in A$  and  $b \in B$ . For a positive integer  $h$ , by  $hA$  we mean the set of all  $h$ -fold sums  $a_1 + \dots + a_h$  of elements of  $A$  (repetitions allowed). There are plenty of interesting theorems concerning these operations, and this is a branch of mathematics called *additive number theory*. In truth, though, it is much more closely related to other branches of mathematics like combinatorics, Fourier analysis and ergodic theory than to the sort of number theory we will be exploring in this course.

## 2. THE FUNDAMENTAL THEOREM (IN $\mathbb{Z}$ )

We had better pay our debts by giving a proof of the uniqueness of the prime power factorization. This is justly called the *Fundamental Theorem of Arithmetic*. For completeness, we nail down the *existence* of a prime power factorization, although as mentioned above this is almost obvious:

**Proposition 6.** *Every positive integer  $n$  is a product of primes  $p_1^{a_1} \dots p_r^{a_r}$  (when  $n = 1$  this is the empty product).*

Proof: By induction on  $n$ , the case of  $n = 1$  being trivial. Assume  $n > 1$  and the result holds for all  $m < n$ . Among all divisors  $d > 1$  of  $n$ , the least is necessarily a prime, say  $p$ . So  $n = pm$  and apply the result inductively to  $m$ .

**Important Remark:** Note that the result seemed obvious, and we proved it by induction. Formally speaking, just about any statement about the integers contain an appeal to induction at some point, since induction – or equivalently, the well-ordering principle that any nonempty subset of integers has a smallest element – is (along with a few much more straightforward axioms) their characteristic property. But induction proofs can be straightforward, tedious, or both. Often I will let you fill in such induction proofs; I will either just say “by induction” or, according to taste, present the argument in less formal noninductive terms. To be sure, sometimes an induction argument is nontrivial, and those will be given in detail.

Let us say that a factorization  $n = p_1^{a_1} \dots p_r^{a_r}$  is in *standard form* if  $p_1 < \dots < p_r$ . Clearly any factorization can be put in standard form just by correctly ordering the prime divisors.

**Theorem 7.** *The standard form factorization of a positive integer is unique.*

The proof is, perhaps surprisingly, not trivial. Indeed it requires several steps. The key is the following, an important result in its own right:

**Theorem 8.** (*Euclid's Lemma*) Suppose  $p$  is prime and  $p \mid ab$ . Then  $p \mid a$  or  $p \mid b$ .

Remark: This result, and not the uniqueness of factorization, appears in Euclid's *Elements*. But, as we are about to see, the uniqueness of factorization follows readily enough from this result (which is itself less than easy to prove!) that it is traditional to credit Euclid with "essentially" proving the fundamental theorem. The first explicit statement and proof is due to Gauss.

Theorem 8  $\implies$  Theorem 7: Let us induct on the (minimal!) number  $r$  of factors in a prime factorization of  $n$ . The case of  $r = 0$  – i.e.,  $n = 1$  – is trivial. Suppose the result holds for numbers with  $< r$  factors, and consider

$$n = p_1^{a_1} \cdots p_r^{a_r} = q_1^{b_1} \cdots q_s^{b_s}.$$

Now  $p_1 \mid n$ , so by Theorem 8  $p_1$  divides some  $q_j^{b_j}$ , and this implies that  $p_1 \mid q_j$ . Therefore we can cancel a common prime factor of both sides, reducing to the case where  $n$  has a factorization with  $r - 1$  prime factors, and the induction hypothesis does the rest.

Now we prove Theorem 8 by introducing the following important notion:

An *ideal* of  $\mathbb{Z}$  is a nonempty subset  $I$  of  $\mathbb{Z}$  such that  $a, b \in I$  implies  $a + b \in I$  and  $a \in I, c \in \mathbb{Z}$  implies  $ca \in I$ .<sup>8</sup>

For any integer  $d$ , the set  $(d) = \{nd \mid n \in \mathbb{Z}\}$  of all multiples of  $d$  is an ideal.

**Proposition 9.** Any nonzero ideal  $I$  of  $\mathbb{Z}$  is of the form  $(d)$ , where  $d$  is the least positive element of  $I$ .

Proof: Suppose not: then there exists an element  $n$  which is not a multiple of  $d$ . Applying the Division Theorem (Proposition 1), we may write  $n = qd + r$  with  $0 < r < d$ . Since  $d \in I, qd \in I$  and hence  $r = n - qd \in I$ . But  $r$  is positive and smaller than  $d$ , a contradiction.

In the next couple of results we would like to refer to the greatest common divisor of two integers, but since we have two different orderings on the positive integers – an additive ordering ( $\leq$ ) and a multiplicative ordering ( $\mid$ ) – the terms is ambiguous, and the ambiguity is a key point! Watch for it in the following:

**Proposition 10.** For integers  $a$  and  $b$ , not both zero, the set  $I_{a,b} = \{xa + yb \mid x, y \in \mathbb{Z}\}$  is a nonzero ideal. Its positive generator  $d$  has the following property:

$$(1) \quad e \mid a \ \& \ e \mid b \iff e \mid d.$$

Proof: It is easy to see that the set  $I_{a,b}$  is closed under addition and under multiplication by all integers, so it is an ideal. By the previous result, it is generated by its smallest positive element, say  $d = Xa + Yb$ .

Now, suppose  $e \mid d$ . Then, since  $a \in (d)$ ,  $(a) \subset (d)$  and thus  $d \mid a$  (to contain is to divide) and by transitivity  $e \mid a$ ; similarly  $e \mid b$ . (In fact we made a bigger production of this than was necessary: we could have said that  $d$  is a multiple of  $e$ , and  $a$  and  $b$  are multiples of  $d$ , so of course  $a$  and  $b$  are multiples of  $e$ . This is the easy

<sup>8</sup>We hope that the reader recognizes this as a special case of an ideal in a commutative ring.

direction.) Conversely, suppose that  $e|a$  and  $e|b$  (so  $e$  is a common divisor of  $a$  and  $b$ ). Then  $e | Xa + Yb = d$ . (Since  $d$  could be smaller than  $a$  and  $b$  – e.g.  $a = 17$ ,  $b = 10^{10}$ ,  $d = 1$ , this is the nontrivial implication.)

We have therefore shown the existence of a common divisor  $d > 0$  of  $a$  and  $b$  which is divisible by all other common divisors: in other words, it is the unique maximal common divisor of  $a$  and  $b$  in the *multiplicative* sense. From this it follows immediately that it is also the *largest* common divisor in the additive sense: i.e.,  $d \geq e$  for any common divisor  $e$  of  $a$  and  $b$  (because, among positive integers,  $a|b \implies a \leq b$ ), but the multiplicative sense is nontrivial and more useful. Henceforth we will always use greatest common divisor in this multiplicative sense: a common divisor which is divisible by all other common divisors. We shall denote it either by  $\gcd(a, b)$ , or just  $(a, b)$  (the latter notation is suggestive of the ideal generated by  $a$  and  $b$ ).

**Corollary 11.** *If  $a$  and  $b$  are integers, not both zero, then for any integer  $m$  there exist integers  $x$  and  $y$  such that*

$$xa + yb = m \gcd(a, b).$$

Proof: This follows immediately from the equality of ideals  $I_{a,b} = (\gcd(a, b))$ : the left hand side is an arbitrary element of  $I_{a,b}$  and the right hand side is an arbitrary element of  $(\gcd(a, b))$ .

An important special case is when  $\gcd(a, b) = 1$  – we say  $a$  and  $b$  are **relatively prime**. The corollary then asserts that for any integer  $m$ , we can find integers  $x$  and  $y$  such that  $xa + yb = m$ .

Indeed we can use this to prove Euclid’s Lemma (Theorem 8): if  $p | ab$  and  $p$  does not divide  $a$ , then the greatest common divisor of  $p$  and  $a$  must be 1. Thus there are integers  $x$  and  $y$  such that  $xa + yp = 1$ . Multiplying through by  $b$  we get  $xab + ypb = b$ . Since  $p | xab$  and  $p | ypb$ , we conclude  $p | b$ . This completes the proof of the Fundamental Theorem of Arithmetic.

### 3. A RING WITHOUT UNIQUE FACTORIZATION

The train of thought involved in proving the fundamental theorem is quite subtle. The first time one sees it, it is hard to believe that such complications are necessary: is it not “obvious” that the factorization of integers into primes is unique?

It is not obvious, but rather familiar and true. The best way to perceive the non-obviousness is to consider new and different contexts. Consider the following example: let  $\mathbb{E}$  denote the set of even integers.<sup>9</sup> Because this is otherwise known as the ideal  $(2) = 2\mathbb{Z}$ , it has a lot of structure: it forms a group under addition, and there is a well-defined multiplication operation satisfying all the properties of a ring except one: namely, there is no 1, or multiplicative identity. (A ring without identity is sometimes wryly called a *rng*, so the title of this section is not a typo.)

<sup>9</sup>This example is taken from Silverman’s book. In turn Silverman took it, I think, from Harold Stark’s introductory number theory text. Maybe it is actually due to Stark (but probably not...)

Let us consider factorization in  $\mathbb{E}$ : in general, an element  $x$  of some structure should be prime if every factorization  $x = yz$  is “trivial” in some sense. However, in  $\mathbb{E}$ , since there is no 1, there are no trivial factorizations, and we can define an element  $x$  of  $\mathbb{E}$  to be prime if it cannot be written as the product of two other elements of  $\mathbb{E}$ . Of course this is a new notion of prime: 2 is a conventional prime and also a prime of  $\mathbb{E}$ , but clearly none of the other conventional primes are  $\mathbb{E}$ -prime. Moreover there are lots of  $\mathbb{E}$ -primes which are not prime in the usual sense: e.g., 6 is  $\mathbb{E}$ -prime. Indeed, it is not hard to see that an element of  $\mathbb{E}$  is an  $\mathbb{E}$ -prime iff it is divisible by 2 but not by 4, because then it is impossible to factor it as a product of two even numbers. (So, in fact, the  $\mathbb{E}$ -primes are much simpler in structure than the usual primes.)

Now consider

$$36 = 2 \cdot 18 = 6 \cdot 6.$$

Since 2, 18 and 6 are all divisible by 2 and not 4, they are  $\mathbb{E}$ -primes, so 36 has two different factorizations into  $\mathbb{E}$ -primes.

This example begins to arouse our skepticism about unique factorization: it is not, for instance, inherent in the nature of factorization that factorization into primes must be unique. On the other hand, the rng  $\mathbb{E}$  is quite artificial: it is an inconveniently small substructure of a better behaved ring  $\mathbb{Z}$ . Later we will see more distressing examples.

#### 4. CONSEQUENCES

Even if we were not seriously in doubt of unique factorization, the previous proof exposes quite a lot of other useful material. Let us look at some of it in more detail:

##### 4.1. Applications of the prime power factorization.

There are certain functions of  $n$  which are most easily defined in terms of the prime power factorization. This includes many so-called **arithmetic functions** that we will discuss a bit later in the course. But here let us give some basic examples. First, let us write the prime power factorization as

$$n = \prod_i p_i^{a_i},$$

where  $p_i$  denotes the  $i$ th prime in sequence, and  $a_i$  is a non-negative integer. This looks like an infinite product, but we impose the condition that  $a_i = 0$  for all but finitely many  $i$ ,<sup>10</sup> so that past a certain point we are just multiplying by 1. The convenience of this is that we do not need different notation for the primes dividing some other integer.

Now suppose we have two such factored positive integers

$$a = \prod_i p_i^{a_i},$$

---

<sup>10</sup>In fact, this representation is precisely analogous to the expression of  $(\mathbb{Z}, \cdot) = (\mathbb{N}, +)^\infty$  of problem G1).

$$b = \prod_i p_i^{b_i}.$$

Then we can give a simple and useful formula for the gcd and the lcm. Namely, the greatest common divisor of  $a$  and  $b$  is

$$\gcd(a, b) = \prod_i p_i^{\min(a_i, b_i)},$$

where  $\min(c, d)$  just gives the smaller of the two integers  $c$  and  $d$  (and, of course, the common value  $c = d$  when they are equal). More generally, we have that, writing out two integers  $a$  and  $b$  in factored form above, we have that  $a \mid b \iff a_i \leq b_i$  for all  $i$ . In fact this is exactly the statement that  $a \mid b \iff a \mid_p b$  for all  $p$  that we expressed earlier.

We often (e.g. now) find ourselves wanting to make reference to the  $a_i$  in the prime power factorization of an integer  $a$ . The  $a_i$  is the highest power of  $p_i$  that divides  $a$ . One often says that  $p_i^{a_i}$  *exactly divides*  $a$ , meaning that  $p_i^{a_i} \mid a$  and  $p_i^{a_i+1}$  does not. So let us define, for any prime  $p$ ,  $\text{ord}_p(a)$  to be the highest power of  $p$  that divides  $a$ : equivalently:

$$n = \prod_i p_i^{\text{ord}_{p_i}(n)}.$$

Notice that  $\text{ord}_p$  is reminiscent of a logarithm to the base  $p$ : in fact, that's exactly what it is when  $n = p^a$  is a power of  $p$  only:  $\text{ord}_p(p^a) = a$ . However, for integers  $n$  divisible by some prime  $q \neq p$ ,  $\log_p(n)$  is nothing nice – in fact, it is an irrational number – whereas  $\text{ord}_p(n)$  is by definition always a non-negative integer. In some sense, the beauty of the functions  $\text{ord}_p$  is that they allow us to “localize” our attention at one prime at a time: every integer  $n$  can be written as  $p^r \cdot m$  with  $\gcd(m, p) = 1$ , and the  $\text{ord}_p$  just politely ignores the  $m$ :  $\text{ord}_p(p^r \cdot m) = \text{ord}_p(p^r) = r$ .

This is really just notation, but it is quite useful: for instance, we can easily see that for all  $p$ ,

$$\text{ord}_p(\gcd(a, b)) = \min(\text{ord}_p(a), \text{ord}_p(b));$$

this just says that the power of  $p$  which divides the gcd of  $a$  and  $b$  should be the largest power of  $p$  which divides both  $a$  and  $b$ . And then a positive integer  $n$  is determined by all of its  $\text{ord}_p(n)$ 's via the above equation.

Similarly, define the least common multiple  $\text{lcm}(a, b)$  of positive integers  $a$  and  $b$  to be a positive integer  $m$  with the property that  $a \mid m$  &  $b \mid m \implies m \mid m$ . Then essentially the same reasoning gives us that

$$\text{ord}_p(\text{lcm}(a, b)) = \max(\text{ord}_p(a), \text{ord}_p(b)),$$

and then that

$$\text{lcm}(a, b) = \prod_p p^{\max(\text{ord}_p(a), \text{ord}_p(b))}.$$

We can equally well define  $\text{ord}_p$  on a negative integer  $n$ : it is again the largest power  $i$  of  $p$  such that  $p_i \mid n$ . Since multiplying by  $-1$  doesn't change divisibility in any way, we have that  $\text{ord}_p(n) = \text{ord}_p(-n)$ . Note however that  $\text{ord}_p(0)$  is slightly problematic – every  $p^i$  divides  $0$ :  $0 \cdot p^i = 0$  – so if we are going to define this at all it would make sense to put  $\text{ord}_p(0) = \infty$ .

We do lose a little something by extending the ord functions to negative integers: namely, since for all  $p$ ,  $\text{ord}_p(n) = \text{ord}_p(-n)$ , the ord functions do not allow us to distinguish between  $n$  and  $-n$ . From a more abstract algebraic perspective, this is because  $n$  and  $-n$  generate the same ideal (are **associates**; more on this later), and we make peace with the fact that different generators of the same ideal are more or less equivalent when it comes to divisibility. However, in  $\mathbb{Z}$  we do have a remedy: we could define a map  $\text{ord}_{-1} : \mathbb{Z} \setminus \{0\} \rightarrow \pm 1$  such that  $\text{ord}_{-1}(n) = +1$  if  $n > 0$  and  $-1$  if  $n < 0$ . Then  $-1$  acts as a “prime of order 2,” in contrast to the other “infinite order primes,” and we get a corresponding unique factorization statement.<sup>11</sup> But although there is some sense to this, we will not adopt it formally here.<sup>12</sup>

**Proposition 12.** *For  $p$  a prime and  $m$  and  $n$  integers, we have:*

- a)  $\text{ord}_p(mn) = \text{ord}_p(m) + \text{ord}_p(n)$ .
- b)  $\text{ord}_p(m+n) \geq \min(\text{ord}_p(m), \text{ord}_p(n))$ .
- c) If  $\text{ord}_p(m) \neq \text{ord}_p(n)$ ,  $\text{ord}_p(m+n) = \min(\text{ord}_p(m), \text{ord}_p(n))$ .

We leave these as exercises: suitably decoded, they are familiar facts about divisibility. Note that part a) says that  $\text{ord}_p$  is some sort of *homomorphism* from  $\mathbb{Z} \setminus \{0\}$  to  $\mathbb{Z}$ . However,  $\mathbb{Z} \setminus \{0\}$  under multiplication is not our favorite kind of algebraic structure: it lacks inverses, so is a monoid rather than a group. This perhaps suggests that we should try to extend it to a map on the nonzero rational numbers  $\mathbb{Q}^\times$  (which, if you did problem G1), you will recognize as the group completion of  $\mathbb{Z} \setminus \{0\}$ ; if not, no matter), and this is no sooner said than done:

For a nonzero rational number  $\frac{a}{b}$ , we define

$$\text{ord}_p\left(\frac{a}{b}\right) = \text{ord}_p(a) - \text{ord}_p(b).$$

In other words, powers of  $p$  dividing the numerator count positively; powers of  $p$  dividing the denominator count negatively. There is something to check here, namely that the definition does not depend upon the choice of representative of  $\frac{a}{b}$ . But it clearly doesn't:

$$\begin{aligned} \text{ord}_p\left(\frac{ac}{bc}\right) &= \text{ord}_p(ac) - \text{ord}_p(bc) \\ &= \text{ord}_p(a) + \text{ord}_p(c) - \text{ord}_p(b) - \text{ord}_p(c) = \text{ord}_p(a) - \text{ord}_p(b) = \text{ord}_p\left(\frac{a}{b}\right). \end{aligned}$$

So we get a map

$$\text{ord}_p : \mathbb{Q}^\times \rightarrow \mathbb{Z}$$

which has all sorts of uses: among other things, we can use it to recognize whether a rational number  $x$  is an integer: it will be iff  $\text{ord}_p(x) \geq 0$  for all primes  $p$ .

Example: Let us look at the partial sums  $S_i$  of the harmonic series  $\sum_{n=1}^{\infty} \frac{1}{n}$ . The first partial sum  $S_1 = 1$  – that's a whole number. The second one is  $S_2 = 1 + \frac{1}{2} = \frac{3}{2}$  which is not. Then  $S_3 = 1 + \frac{1}{2} + \frac{1}{3} = \frac{11}{6}$  is not an integer either; neither is  $S_4 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = \frac{25}{12}$ .

It is natural to ask whether *any* partial sum  $S_n$  for  $n \geq 1$  is an integer. Indeed,

<sup>11</sup>This perspective is apparently due to John Horton Conway, and was explained to me by Manjul Bhargava.

<sup>12</sup>By the way, Manjul never told me what  $\text{ord}_{-1}(0)$  should be...

this is a standard question in honors math classes because...well, frankly, because it's rather hard.<sup>13</sup> But using properties of the ord function we can give a simple proof. The first step is to look carefully at the data and see if we can find a pattern. (This is, of course, something to do whenever you are presented with a problem whose solution you do not immediately know. Modern presentations of mathematics – including, alas, these notes, to a large extent – often hide this experimentation and discovery process.) What we see in the small partial sums is that not only are they not integers, they are all not integers for “the same reason”: there is always a power of 2 in the denominator.

So what we'd like to show is that for all  $n \geq 1$ ,  $\text{ord}_2(S_n) < 0$ . It is true for  $n = 2$ ; moreover we don't have to do the calculation for  $n = 3$ : since  $\text{ord}_2(\frac{1}{3}) = 0 \neq \text{ord}_2(S_2)$ , we must have  $\text{ord}_2(S_2 + \frac{1}{3}) = \min(\text{ord}_2(S_2), \text{ord}_2(S_3)) = -1$ . And then we get  $\frac{1}{4}$ , which 2-order  $-2$ , which is different from  $\text{ord}_2(S_3)$ , so again, using that when we add two rational numbers with different 2-orders, the 2-order of the sum is the smaller of the 2 2-orders, we get that  $\text{ord}_2(S_4) = -2$ . Excitedly testing a few more values, we see that this pattern continues:  $\text{ord}_2(S_n)$  and  $\text{ord}_2(\frac{1}{n+1})$  are always different; if only we can show that this always holds, this will prove the result. In fact one can say even more: one can *precisely* what  $\text{ord}_2(S_n)$  is as a function of  $n$  and thus see in particular that it is always negative. I will leave the final observation and proof to you – why should I steal your fun? (Here is one hint: consider for each  $k$  the set of integers  $n$  such that  $\text{ord}_2(n) = -k$ .)

#### 4.2. Linear Diophantine equations.

Recall that one of the two main things we agreed that number theory is about was solving Diophantine equations, i.e., looking for solutions over  $\mathbb{Z}$  and/or over  $\mathbb{Q}$  to polynomial equations. Certainly we saw some primes in the previous section; now we solve the simplest class of Diophantine equations, namely the linear ones.

Historical remark: as I said before, nowadays when someone says Diophantine equation, they mean that we are interested either in solutions over  $\mathbb{Z}$  or solutions over  $\mathbb{Q}$ , or both. Diophantus himself considered positive rational solutions. Nowadays the restriction to positive numbers seems quite artificial (and I must wonder whether Diophantus massaged his equations so as to get positive rather than negative solutions); it also makes things quite a bit more difficult: it stands to reason that since equations become easier to solve if we allow ourselves to divide numbers, correspondingly they become more difficult if we do not allow subtraction!

This also means that the term “Linear Diophantine equation” is, strictly speaking, an anachronism. If you want to solve any number of linear equations with coefficients in  $\mathbb{Q}$ , then – since  $\mathbb{Q}$  is a field – you are just doing linear algebra, which works equally well over  $\mathbb{Q}$  as it does over  $\mathbb{R}$  or  $\mathbb{C}$ . For instance, suppose we want to solve the equation

$$ax + by = c$$

in rational numbers, where  $a$  and  $b$  are nonzero rational numbers and  $c$  is any rational number. Well, it's not much fun, is it? Let  $x$  be any rational number at

---

<sup>13</sup>When I first got assigned this problem (my very first semester at college), I found – or looked up? – some quite elaborate solution which used, in particular, **Bertrand's Postulate** that for  $n > 1$  there is always a prime  $p$  with  $n < p < 2n$ . (This was proven in the latter half of the 19th century by Chebyshev. One of Paul Erdős' early mathematical triumphs was an elegant new proof of this result.)

all, and solve for  $y$ :

$$y = \frac{c - ax}{b}.$$

Speaking more geometrically, any line  $y = mx + b$  in the plane passing through one rational point and with rational slope – roughly speaking, with  $m$  and  $b$  rational – will have lots of rational solutions: one for every rational choice of  $x$ .

Exercise: Consider the line  $y = mx + b$  for  $m, b \in \mathbb{R}$ . As above, if  $m$  and  $b$  are rational, there are clearly infinitely many rational points on the line.

a) Suppose  $m$  is rational and  $b$  is irrational. Show that there are no rational points on the line.

b) Suppose  $m$  is irrational and  $b$  is rational. Show that there is exactly one rational point on the line.

c) Now suppose that  $m$  and  $b$  are both irrational. What are the possibilities for the number of rational solutions?

So for Diophantus, the first interesting example was quadratic polynomial equations. Indeed, after this section, the quadratic case will occupy our interest for perhaps the majority of the course.

However, over  $\mathbb{Z}$  things are never so easy: for instance, the equation

$$3x + 3y = 1$$

clearly does not have an integer solution, since no matter what integers  $x$  and  $y$  we choose,  $3x + 3y$  will be divisible by  $3$ . More generally, if  $a$  and  $b$  have a common divisor  $d > 1$ , then it is hopeless to try to solve

$$ax + by = 1.$$

But this is the only restriction, and indeed we saw this before: en route to proving the fundamental theorem, we showed that for any integers  $a$  and  $b$ , not both zero, then  $\gcd(a, b)$  generates the ideal  $\{xa + yb \mid x, y \in \mathbb{Z}\}$ , meaning that for any integer  $m$ , the equation

$$ax + by = m \gcd(a, b)$$

has solutions in  $x$  and  $y$ . In other words, we can solve

$$ax + by = n$$

if  $n$  is a multiple of the gcd of  $a$  and  $b$ . By the above, it is also true that we can only solve the equation if  $n$  is a multiple of the gcd of  $x$  and  $y$  – the succinct statement is the *equality* of ideals  $I_{a,b} = (\gcd(a, b))$  – so we have (and already had, really) the following important result.

**Theorem 13.** *For fixed integers  $a$  and  $b$ , not both zero, and any integer  $m$ , the equation*

$$ax + by = m$$

*has a solution in integers  $(x, y)$  iff  $\gcd(a, b) \mid m$ .*

In particular, if  $\gcd(a, b) = 1$ , then we can solve the equation for any integer  $m$ . The fundamental case is to solve

$$ax + by = 1,$$

because if we can find such  $x$  and  $y$ , then just by multiplying through by  $m$  we can solve the general equation.

This is a nice result, but it raises two further questions. First, we found one solution. Now what can we say about *all* solutions?<sup>14</sup> Second, given that we know that solution(s?) exist, how do we actually find them?

Example: We are claiming that  $3x + 7y = 1$  has an integer solution. What could it be? Well, a little experimentation yields  $x = -2, y = 1$ . Is this the only solution? Indeed not: we could add 7 to  $x$  and the sum would increase by 21, and then subtract 3 from  $y$  and the sum would decrease by 21. This leads us to write down the family of solutions  $x_n = -2 + 7n, y_n = 1 - 3n$ . Are there any more? Well, we have found one integral solutions whose  $x$ -coordinates are evenly spaced, 7 units apart from each other. If there is any other solution  $3X + 7Y = 1$ , there must be some  $n$  such that  $0 < X - x_n < 7$ . This would give a solution  $3(X - x_n) = -7(Y - y_n)$  with  $0 < X - x_n < 7$ . But this is absurd: the left hand side would therefore be prime to 7, whereas the right hand side is divisible by 7. So we evidently found the general solution. A similar argument shows:

**Theorem 14.** *For  $a$  and  $b$  coprime positive integers, the general integral solution to  $ax + by = 1$  is  $x_n = x_0 + nb, y_n = y_0 - na$ , where  $x_0a + y_0b = 1$  is any particular solution guaranteed to exist by Theorem 13.*

We ask the reader to verify this as an informal exercise (i.e., not to be turned in).

Note that although one might have thought that in addressing the question of how to find all solutions we would necessarily have had to nail down how to find a particular solution, this turned out not to be the case: Theorem 14 cleverly evades the question of finding the first solution  $(x_0, y_0)$ .<sup>15</sup>

It should be said that the above theorem does suggest an algorithm (“is effective,” in the standard mathematical jargon): there will be exactly one solution  $(x_0, y_0)$  with  $0 \leq x_0 < |b|$ , so what we could do is check, each integer  $x$  in this range until we find the one for which  $1 - xa$  is a multiple of  $b$ . But this is a terribly inefficient algorithm, and in fact Euclid famously had a better one. In the exercises we describe Euclid’s algorithm for computing gcd’s and how to use it to obtain a solution much faster.

**4.3. The Fundamental Theorem in a PID.** Conceptually, our proof of the fundamental theorem can be broken down into several steps:

Step 1: We show that the integers  $\mathbb{Z}$  form a principal ideal domain (PID), i.e., a commutative ring without zero divisors in which every ideal is principal.

<sup>14</sup>Diophantus was for the most part content with finding a single solution. The more penetrating inquiry into the set of all solutions was apparently first made by Fermat.

<sup>15</sup>Those who have studied differential equations will find this situation familiar: the general solution of an inhomogeneous equation is obtained by finding the general solution of the associated homogeneous equation and adding to it some particular solution of the inhomogeneous equation. But how that particular solution is found is often left unclear.

Step 2: We show that greatest common divisors exist in a PID, and more precisely, that the equation  $xa + yb = \gcd(a, b)$  has a solution.

Step 3: We use Step 2 to show that the unfactorable elements  $p$  in a PID satisfy Euclid's Lemma.

Step 4: We deduce unique factorization from Euclid's Lemma.

Or, we can compress the argument and see that we are really proving two different kinds of statements: (i)  $\mathbb{Z}$  is a PID; and (ii) every PID has unique factorization. The second statement is of quite a general (and hence abstract algebraic) character: once we have the correct terminology and definitions to express *what it means* for an integral domain to have unique factorization, one sees that Steps 2-4 of the argument apply to prove (ii). On the other hand, not every integral domain has unique factorization (since an integral domain necessarily has a multiplicative identity, the earlier business with  $\mathbb{E}$  is *not* an example of this; we will see examples later on), so it follows that not every integral domain is a PID. (In fact being a PID is sufficient but not necessary for unique factorization; however, among the rings of most interest to us in this course – certain rings of integers in number fields – the two are equivalent.) Thus what seemed like the easiest part of the proof – namely that  $\mathbb{Z}$  is a PID – is in fact the part that we are most eager to see to what extent it can be generalized. It turns out that there is an idea to the proof of Step 1, the existence of a so-called **Euclidean norm**, which we will be able to generalize to prove that some other rings are PID's, most notably the Gaussian integers and the Eisenstein integers.

We could throw ourselves into this right now and prove that the ring  $\mathbb{Z}[\sqrt{-1}]$  (the Gaussian integers) has unique factorization. However, this would be treading rather far down the path in the algebraic direction. We would rather like to see a bit of the analytic, combinatorial and geometric sides of the subject before delving into any more technical matters, so we will come back to this later on in the course. In the meantime, I have prepared an algebra handout on the theory of factorization in integral domains which discusses the “general” part of the argument in a modern way. Those with a good algebra background / taking the course for graduate credit should look through these notes in the meantime. Those whose algebra backgrounds are more modest may prefer to wait and see these ideas used in the relatively concrete context of  $\mathbb{Z}[\sqrt{-1}]$  and closely related rings.

## 5. HOMEWORK

Here is an explanation of the strange letters and symbols which follow many of the problems:

(E) This denotes an **easier** problem. Students who find these problems *too* easy can write “OK” as the solution to the problem; but students with more modest backgrounds might appreciate having a supply of more straightforward problems.

(\*) This denotes a **harder** problem. Harder problems are almost optional: in 4400 one can get up to an  $A^-$  grade without doing any star problems; in 6400 one

can get up to a  $B$  grade without doing any star problems and up to an  $A^-$  by doing only a few starred problems. It should be said that the difficulty varies more widely in these problems than in any other. If I were to be honest about things, there may be a few problems which could be labelled:

(\*\*) It may not be possible to solve this problem without more advanced knowledge (and/or I might not quite remember how to solve it!), or indeed

(\*\*\*) This problem is to the best of my knowledge unsolved, and it is not at all clear to me how to solve it.

But I am not above leaving out the second and third stars to try to get you to think about problems you might otherwise skip over: welcome to the deep end of the pool.

(O) This problem is **open-ended**, meaning exactly what is being asked may not be quite clear, and several solutions (or in some cases, no solution) may be equally acceptable. These problems are all optional, and can be omitted by all students without penalty.

(G) This means graduate-level. In many cases it would have deserved a (\*), but in addition to being challenging it may also be more abstract and may call upon more background: in particular more abstract algebra. All (G) problems are optional at the 4400 level and if solved have the same benefits as (\*) problems. Not every 6400 student is expected to be able to solve every 6400 problem.

(H) This means a **historical** problem. Historical problems are also optional; however, students at the 4400 level may do (H) problems instead of (\*) problems and still get an A in the course.

“Can you...?” In multi-part problems, one of the parts might ask for a sharpening of the previous parts in an interrogative way. These are also optional, and in some cases they are quite unreasonable, e.g., can you write a computer program which plays Schuh’s divisor game better than humans do? Clearly this is not required.

## 6. PROBLEM SET 2

2.1)(E) Prove the Division Theorem (Proposition 1). Hint: It suffices to take  $q$  to be the largest non-negative integer such that  $n - qd \geq 0$ .

2.2)(E) Show that  $d|n \iff$  we have  $r = 0$  in the Division Theorem.

2.3) Prove the converse of Euclid’s Lemma: suppose  $d$  is a positive integer such that whenever  $d|ab$ ,  $d|a$  or  $d|b$ . Then  $d$  is prime.

Remark: Among other things, this allows us to generalize the notion of primes to not-necessarily principal ideals.

2.4) “To contain is to divide”: for integers  $a$  and  $b$ , we have  $a|b \iff (a) \supset (b)$ .

2.5) For any integers  $a$  and  $b$ , not both zero, there are exactly two integers  $d_1$  and  $d_2$  with the property that  $e|a$  &  $e|b \implies e|d$ , and  $d_2 = -d_1$ .

2.6) Show that if  $a = b = 0$ , there is no integer  $d$  such that  $e|a$  &  $e|b \implies e|d$ .

The next 2 exercises concern the rng  $\mathbb{E}$ .

2.7)(O) Should the factorizations  $6 = 2 \cdot 3$  and  $6 = (-2) \cdot (-3)$  be counted as “essentially different” or not? (I could go either way on this one.)

2.8) Give a necessary and sufficient condition on a positive element  $x \in \mathbb{E}$  to have two different factorizations into positive  $\mathbb{E}$ -primes.<sup>16</sup> Hint: pay attention to  $\text{ord}_2(x)$  and also to the number of odd primes dividing  $x$ .

2.9) Prove Proposition 12.

2.10) Complete the proof that  $S_n = 1 + \frac{1}{2} + \dots + \frac{1}{n}$  is never an integer by showing that for all  $n \geq 1$ ,  $\text{ord}_2(S_n) \neq \text{ord}_2(\frac{1}{n+1})$ .

2.11)\*\* Show that except for  $n = 1, 2, 6$ , the decimal expansion of  $S_n$  is non-terminating. (I.e., show that except for these values,  $\text{ord}_p(S_n) < 0$  for some prime  $p \neq 2, 5$ .)<sup>17</sup>

2.12) For any nonzero integers  $a$  and  $b$ , show that

$$\gcd(a, b) \cdot \text{lcm}(a, b) = ab.$$

2.13) For integers  $a$  and  $b$ , show that the intersection of the two ideals  $(a) \cap (b)$  is an ideal of  $\mathbb{Z}$ . In fact, if you know the definition of an ideal in a commutative ring, show that the intersection of any two (or more...) ideals is always an ideal.<sup>18</sup> Because  $\mathbb{Z}$  is a PID, we must have  $(a) \cap (b) = (c)$  for some  $c \in \mathbb{Z}$ , well-determined up to a sign. What is  $c$  in terms of  $a$  and  $b$ ?

2.14) a) Let  $a_1, \dots, a_n$  be a (finite) set of integers, not all zero. Define the *greatest common divisor*  $\gcd(a_1, \dots, a_n)$  of the set, and show that it exists and is unique up to a sign. In fact, show that the set

$$(a_1, \dots, a_n) = \{x_1 a_1 + \dots + x_n a_n \mid x_i \in \mathbb{Z}\}$$

is an ideal of  $\mathbb{Z}$ , and that the positive generator of this (necessarily principal!) ideal is the gcd we are looking for.

b)\* Define  $\text{lcm}(a_1, \dots, a_n)$ , show it exists, and explain how to get it from the ideals  $(a_1), \dots, (a_n)$ .

2.15) Show that  $\gcd(a_1, a_2, a_3) = \gcd(\gcd(a_1, a_2), a_3)$ .

<sup>16</sup>Note that the “positives” are to avoid the problems of 2.7.

<sup>17</sup>Note the double-star: this is quite difficult.

<sup>18</sup>It is a metatheorem of algebra that if  $H_1$  and  $H_2$  are some substructures of a structure  $G$ , then  $H_1 \cap H_2$  is also a substructure. Unions do not work nearly as nicely.

2.16)\* Find an identity relating  $\gcd(a_1, a_2, a_3)$ ,  $\text{lcm}(a_1, a_2, a_3)$  and  $a_1 a_2 a_3$ . Can you extend this to more than three numbers? (Hint: inclusion/exclusion.)

2.17) One says that a set of integers  $a_1, \dots, a_n$  is **relatively prime in pairs** if for all  $i \neq j$ ,  $\gcd(a_i, a_j) = 1$ .

a) Show that if  $a_1, \dots, a_n$  are relatively prime in pairs, then  $\gcd(a_1, \dots, a_n) = 1$ .

b) Show that the converse does not hold when  $n \geq 3$ : indeed, find the smallest example of three integers which are not simultaneously divisible by any  $d > 1$  but for which any two have a nontrivial common divisor.<sup>19</sup>

Remark: The phrase “let  $a_1, \dots, a_n$  be relatively prime integers” is therefore ambiguous when  $n \geq 3$ . Probably it ought to mean the weaker condition that  $\gcd(a_1, \dots, a_n) = 1$  but careful authors rephrase to avoid the ambiguity. If you hear someone say it, stop and ask them which one they mean!

G2) Let  $F$  be a field and let  $v : F^\times \rightarrow \mathbb{Z}$  be a surjective map satisfying properties a) and b) of Proposition 12;  $v$  is said to be a **discrete valuation** of  $F$ .

a) Let  $R_v := \{x \in F^\times \mid v(x) \geq 0\} \cup \{0\}$ . Show that  $R_v$  is a subring of  $F$ , the **valuation ring**. (It is common to formally set  $v(0) = \infty$  to avoid having to keep “manually inserting 0.”)

c) Since  $v$  is surjective, there is an element  $\pi \in R_v$  with  $v(\pi) = 1$ .<sup>20</sup> Show that for any  $n \geq 1$ ,

$$\{x \in F^\times \mid v(x) \geq n\} \cup \{0\} = (\pi^n),$$

the principal ideal of  $R_v$  generated by  $\pi^n$  (of course  $\pi^0 = 1$ ).

d) Show that every ideal of  $R_v$  is of the form  $(\pi^n)$  for a suitable  $n \in \mathbb{N}$ . In particular, every ideal of  $R_v$  is principal, and there is a unique maximal ideal,  $(\pi)$ .

e) When  $F = \mathbb{Q}$ ,  $v = \text{ord}_p$ , what is the valuation ring  $R_v$ ?

f) Suppose  $k$  is a field, and consider  $F = k(t)$ , the quotient field of the ring of polynomials  $k[t]$  with coefficients in  $k$ . Show that the map  $v$  which takes a rational function  $\frac{p(x)}{q(x)}$  to  $\deg(q(x)) - \deg(p(x))$  is a discrete valuation of  $k(t)$ . Note that this is consistent with our previous convention that the degree of the zero polynomial is  $-\infty$ !

<sup>19</sup>This is reminiscent of the fact that a set of vectors can be linearly dependent even when any two of them are linearly independent from each other, a fact that gives linear algebra students no end of trouble.

<sup>20</sup>Denoting this element by  $\pi$  is traditional. Needless to say (?) it has nothing to do with 3.1415926535897...