FOUNDATIONS AND THE FUNDAMENTAL THEOREM

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 mathematicians. 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. 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, so one may reasonably infer that these topics have been central to their respective 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?

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. 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.[1] 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 – consider e.g. $3 - 5$.

As you well know, there are two different responses to this: the first is to *formally extend* the natural numbers so that additive inverses always exist. In other words, for every positive integer n, we formally introduce a corresponding "number" $-n$ with the property that $n + (-n) = 0$. Although it is not *a priori* obvious that such a construction works – rather, the details and meaning of this construction were a point of confusion even among leading mathematicians for a few thousand years – nowadays we understand that it works to give a consistent structure: the integers $\mathbb{Z}$, endowed with an associative addition operation $+$, which has an identity 0 and for which each integer n has a unique additive inverse $-n$.

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 $a = b$.[2]

Now for comparison consider the positive integers

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

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, ... 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 $a|b$.[3] 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. $2 \not| 3$ and also $3 \not| 2$. What are we to make of this divisibility relation?

First, on a case-by-case basis, we *do* know how to determine whether $a \mid b$.

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 $n = qd + r$.*

[1]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.

[2]That is to say, the relation $\leq$ on $\mathbb{N}$ is a linear, or total, ordering.

[3]Careful: $a|b \iff \frac{b}{a}$ is an integer.

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, ... 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 $a \mid_p b$ to mean that $\frac{b}{a}$ is a rational number which, when written in lowest terms, has denominator *not* divisible by p. For instance, $3 \mid_2 5$, since $\frac{5}{3}$, while not an integer, doesn't have a 2 in the denominator. For that matter, $3 \mid_p 5$ for all primes p different from 3, and this suggests the following:

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

Proof: Certainly if $a|b$, then $a \mid_p b$ for all primes p. For the converse, write $\frac{b}{a}$ in lowest terms, say as $\frac{B}{A}$. Then $a \mid_p b$ 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 to the additive structure of $\mathbb{N}$, except that 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, as we shall see!) by taking each prime a non-negative integer number of times (which must be zero for all but finitely many primes). 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 $a \mid b$ iff $a \mid_p b$ 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 "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.[4]

Itis remarkable how quickly we can find ourselves in very deep waters by asking apparently innocuous questions that mix additive and multiplicative structure. 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 "nth 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 nothing[5] 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, 12/25/1640) A prime $p > 2$ is of the form $x^2 + y^2$ iff it is of the form $4k + 1$.*

This is, to my mind, 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, 1770) 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 $an + b$.*

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.

[4]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.

[5]Well, nothing except that $p + 2$ is not divisible by 2 for all $p > 2$.

We will see proofs of Theorems 3 and 4 in this course. To be more precise, we will give two different proofs of Theorem 3. The first theorem uses the observation that x^2+y^2 can be factored in the ring $\mathbb{Z}[i]$ of Gaussian integers as $(x+iy)(x-iy)$ and will be our jumping off point to the use of algebraic methods. There is an analogous proof of Theorem 4 using a noncommutative ring of "integral quaternions". This proof however has some technical complications which make it less appealing for in-class presentation, so we do not discuss it in these notes.[6] On the other hand, we will give parallel proofs of Theorems 3 and 4 using geometric methods. The proof of Theorem 5 is of a different degree of sophistication than any other proofs in this course. We do present a complete proof at the end of these notes, but I have not managed to persuade myself that our treatment is appropriate for a one-semester undergraduate course in the subject.

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 $a_1 + \ldots + a_h$ 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 $\mathbb{Z}$)

2.1. **Existence of prime factorizations.**

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*.

Let us first 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 $p_1^{a_1} \cdots p_r^{a_r}$ (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

[6] It was, in fact, the subject of a student project in the 2007 course.

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.

A factorization $n = p_1^{a_1} \cdots p_r^{a_r}$ is in **standard form** if $p_1 < \ldots < p_r$. Any factorization can be put in standard form by correctly ordering the prime divisors.

2.2. **The fundamental theorem and Euclid's Lemma.**

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

Note that this is just a mildly laundered version of the more common statement: the factorization of a positive integer into primes is unique up to the order of the factors.

Theorem 7 was first stated and proved by Gauss in his *Disquisitiones Arithmeticae*. However, it is generally agreed that the result is "essentially" due to the ancient (circa 300 BC) Greek mathematician Euclid of Alexandria. Euclid proved:

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

The point is that, assuming the very easy Proposition 6, Theorems 7 and 8 are equivalent. From a strictly logical point of view two assertions are equivalent if they are both true or both false – or, if they range over a set of possible parameters then they are true for exactly the same values of those parameters. Since a theorem in mathematics is a true assertion, strictly speaking any two theorems are equivalent. But in common use the statement "Theorem A is equivalent to Theorem B" carries the connotation that it is much easier to deduce the truth of each theorem from the other than to prove either theorem. This is the case here.

Theorem 7 $\implies$ Theorem 8: Suppose for a contradiction that $p \mid ab$ but p does not divide either a or b. Writing out $a = \prod_i p_i^{a_i}$ and $b = \prod_j q_j^{b_j}$, our assumptions are equivalent to $p_i \neq p \neq q_j$ for all i, j. But then $ab = \prod p_i^{a_i} q_j^{a_j}$, and collecting this into standard form we get that no positive power of the prime p appears in the standard form factorization of ab. On the other hand, by assumption $p \mid ab$ so $ab = p \cdot m$, and then factoring m into primes we will get a standard form factorization of ab in which p does apear to some positive power, contradicting the uniqueness of the standard form prime factorization.

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, reducing to the case where n has a factorization with $r-1$ prime factors, and the induction hypothesis does the rest.

Therefore one way to prove Theorem 7 is to give Euclid's proof of Theorem 8. Euclid's proof goes by way of giving an explicit – and efficient – algorithm for finding the greatest common divisor of a pair of positive integers. This **Euclidean algorithm** can be put to a variety of uses in elementary number theory, so Euclid's proof is generally the one given in introductory courses. By making use of algebraic

ideas it is possible to streamline Euclid's proof of Theorem 8 in a way which bypasses the algorithm: the idea is to show that the ring of integers has the property of being a **Principal Ideal Domain**, which is for a general ring a stronger result than the uniqueness of factorization into primes. In fact there is a third strategy, which directly proves Theorem 7. This proof, due to Hasse, Lindemann and Zermelo, is not sufficiently widely known. It is an archetypical instance of bypassing seemingly "necessary" machinery by sheer cleverness.

2.3. **The HLZ proof of the uniqueness of factorization.**

We claim that the standard form factorization of a positive integer is unique. Assume not; then the set of positive integers which have at least two different standard form factorizations is nonempty, so has a least elment, say n, where:

$$n = p_1 \cdots p_r = q_1 \cdots q_s. \tag{1}$$

Here the p_i's and q_j's are prime numbers, not necessarily distinct from each other. However,we must have $p_1 \neq q_j$ for any j. Indeed, if we had such an equality, then after relabelling the q_j's we could assume $p_1 = q_1$ and then divide through by $p_1 = q_1$ to get a smaller positive integer $\frac{n}{p_1}$. By the assumed minimality of n, the prime factorization of $\frac{n}{p_1}$ must be unique: i.e., $r - 1 = s - 1$ and $p_i = q_i$ for all $2 \leq i \leq r$. But then multiplying back by $p_1 = q_1$ we see that we didn't have two different factorizations after all. (In fact this shows that for all i, j, $p_i \neq q_j$.)

In particular $p_1 \neq q_1$. Without loss of generality, assume $p_1 < q_1$. Then, if we subtract $p_1 q_2 \cdots q_s$ from both sides of (1), we get

$$m := n - p_1 q_2 \cdots q_s = p_1(p_2 \cdots p_r - q_2 \cdots q_s) = (q_1 - p_1)(q_2 \cdots q_s). \tag{2}$$

Evidently $0 < m < n$, so by minimality of n, the prime factorization of m must be unique. However, (2) gives two different factorizations of m, and we can use these to get a contradiction. Specifically, $m = p_1(p_2 \cdots p_r - q_2 \cdots q_s)$ shows that $p_1 \mid m$. Therefore, when we factor $m = (q_1 - p_1)(q_2 \cdots q_s)$ into primes, at least one of the prime factors must be p_1. But $q_2, \ldots, q_j$ are already primes which are different from p_1, so the only way we could get a p_1 factor is if $p_1 \mid (q_1 - p_1)$. But this implies $p_1 \mid q_1$, and since q_1 is also prime this implies $p_1 = q_1$. Contradiction!

2.4. **Proof using ideals.**

Now we turn things around by giving a direct proof of Euclid's Lemma. We (still!) do not follow Euclid's original proof, which employs the **Euclidean algorithm**, but rather a modernized version using ideals.

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 $ca \in I$.[7]

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.*

[7]We hope that the reader recognizes this as a special case of an ideal in a commutative ring.

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.

Existence of gcd's: Let a and b be two nonzero integers. An integer d is said to be a **greatest common divisor** of a and b if

(GCD1) $d \mid a$ and $d \mid b$.
(GCD2) If $e \mid a$ and $e \mid b$ then $e \mid d$.

Note well that this is (at least apparently) different from the definition of greatest common divisor one learns in school: in the set of all common divisors of a and b, d is defined to be a divisor which is divisible by every other divisor, not a divisor which is numerically largest. In particular, unlike the school definition, it is not obvious that greatest common divisors exist! However:

Proposition 10. *For $a, b \in \mathbb{Z}$, 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:*

$$e|a \ \& \ e|b \iff e|d, \tag{3}$$

and is therefore a greatest common divisor of a and b.

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 $e|d$. Then, since $a \in (d)$, $(a) \subset (d)$ and thus $d|a$ (to contain is to divide) and by transitivity $e|a$; similarly $e|b$. (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 direction.) Conversely, suppose that $e|a$ and $e|b$ (so e is a common divisor of a and b). Then $e \mid Xa + Yb = d$. (Since d could be smaller than a and b – e.g. $a = 17$, $b = 10^{10}$, $d = 1$, this is the nontrivial implication.)

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 $p \mid ab$ 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 $p \mid xab$ and $p \mid ypb$, we conclude $p \mid b$. This completes the proof of the Fundamental Theorem of Arithmetic.

3. Some examples of failure of 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.

Example: let $\mathbb{E}$ denote the set of even integers.[8] 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.)

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 $\mathbb{E}$-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.

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.

Example 2: Let $R_\circ = \mathbb{R}[\cos\theta, \sin\theta]$ be the ring of real trigonometric polynomials: i.e., the ring whose elements are polynomial expressions in $\sin\theta$ and $\cos\theta$ with real coefficients. We view the elements as functions from $\mathbb{R}$ to $\mathbb{R}$ and add and multiply them pointwise.

Of course this ring is not isomorphic to the polynomial ring $\mathbb{R}[x, y]$, since we have the Pythagorean identity $\cos^2\theta + \sin^2\theta = 1$. It is certainly plausible – and can be shown to be true – that all polynomial relations between the sine and cosine are consequences of this one relation, in the sense that $R_\circ$ is isomorphic to the quotient ring $\mathbb{R}[x, y]/(x^2 + y^2 - 1)$.

Now consider the basic trigonometric identity

$$(\cos\theta)(\cos\theta) = (1 + \sin\theta)(1 - \sin\theta). \tag{4}$$

[8]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...)

It turns out that $\cos\theta$, $1+\sin\theta$ and $1-\sin\theta$ are all irreducible elements in the ring $R_\circ$. Moreover, the only units in $R_\circ$ are the nonzero real numbers, so all three of these elements are nonassociate, and therefore (4) exhibits two different factorizations into irreducible elements! Thus, in a sense, the failure of unique factorization in $R_\circ$ is the explanation for the subject of trigonometric identities!

To see how subtle the issue of unique factorization can be, consider now the ring

$$C_\circ = \mathbb{C}[\cos\theta, \sin\theta]$$

of trigonometric polynomials with complex coefficients. But the classic "Euler identity"

$$e^{i\theta} = \cos\theta + i\sin\theta$$

shows that $e^{i\theta}$ is an element of $C_\circ$, and conversely, both the sine and cosine functions are expressible in terms of $e^{i\theta}$:

$$\cos\theta = \frac{1}{2}\left(e^{i\theta} + \frac{1}{e^{i\theta}}\right),$$

$$\sin\theta = \frac{1}{2i}\left(e^{i\theta} - \frac{1}{e^{i\theta}}\right).$$

Thus $C_\circ = \mathbb{C}[e^{i\theta}, \frac{1}{e^{i\theta}}]$. Now the ring $\mathbb{C}[e^{it}]$ is isomorphic to the polynomial ring $\mathbb{C}[T]$, so $C_\circ$ is, up to isomorphism, obtained from $\mathbb{C}[T]$ by adjoining T^{-1}. Recall that $\mathbb{C}[t]$ is a principal ideal domain (PID). Finally, if R is any PID with fraction field K, and S is any ring such that $R \subset S \subset K$ – i.e., any ring obtained by adjoining to R the multiplicative inverses of each of some set of nonzero elements of R – then it can be shown that S is also a PID, hence in particular a unique factorization domain.

The foregoing discussion has been quite brief, with no pretense of presenting a complete argument. A nice writeup, with all details provided, is to be found in H.F. Trotter, *An overlooked example of nonunique factorization*, American Mathematical Monthly 95 (1988), 339-342. It would make a nice final project to read and understand this article.[9]

4. Consequences of the fundamental theorem

The second proof of the fundamental theorem develops material which is very useful in its own right. 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

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

[9]Alternately, in my preprint *Elliptic Dedekind domains revisited* – see http://math.uga.edu/∼pete/ellipticded.pdf – there is a discussion of how this result can be immediately deduced from a much more general theorem of M. Rosen. But I honestly think you will find Trotter's discussion much easier to understand.

where p_i denotes the ith 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,[10] 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

$$a = \prod_i p_i^{a_i},$$
$$b = \prod_i p_i^{b_i}.$$

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 $a|b \iff a|_p b$ 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 $p_i^{a_i}|a$ and $p_i^{a_i+1}$ does not. So let us define, for any prime p, $\mathrm{ord}_p(a)$ to be the highest power of p that divides a: equivalently:

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

Notice that 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: $\mathrm{ord}_p(p^a) = a$. 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 $\mathrm{ord}_p(n)$ is by definition always a non-negative integer. In some sense, the beauty of the functions 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 ord_p just politely ignores the m: $\mathrm{ord}_p(p^r \cdot m) = \mathrm{ord}_p(p^r) = r$.

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

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

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 $\mathrm{ord}_p(n)$'s via the above equation.

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

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

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

and then that

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

We can equally well define ord_p on a negative integer n: it is again the largest power i of p such that $p_i|n$. Since multiplying by -1 doesn't change divisibility in any way, we have that $\mathrm{ord}_p(n) = \mathrm{ord}_p(-n)$. Note however that $\mathrm{ord}_p(0)$ 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 $\mathrm{ord}_p(0) = \infty$.

We do lose a little something by extending the ord functions to negative integers: namely, since for all p, $\mathrm{ord}_p(n) = \mathrm{ord}_p(-n)$, 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 $\mathrm{ord}_{-1} : \mathbb{Z} \setminus \{0\} \to \pm 1$ such that $\mathrm{ord}_{-1}(n) = +1$ 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.[11] But although there is some sense to this, we will not adopt it formally here.[12]

Proposition 12. *For p a prime and m and n integers, we have:*
a) $\mathrm{ord}_p(mn) = \mathrm{ord}_p(m) + \mathrm{ord}_p(n)$.
b) $\mathrm{ord}_p(m+n) \geq \min(\mathrm{ord}_p(m), \mathrm{ord}_p(n))$.
c) If $\mathrm{ord}_p(m) \neq \mathrm{ord}_p(n)$, $\mathrm{ord}_p(m+n) = \min(\mathrm{ord}_p(m), \mathrm{ord}_p(n))$.

We leave these as exercises: suitably decoded, they are familiar facts about divisibility. Note that part a) says that 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

$$\mathrm{ord}_p(\frac{a}{b}) = \mathrm{ord}_p(a) - \mathrm{ord}_p(b).$$

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,

[11] This perspective is apparently due to John Horton Conway, and was explained to me by Manjul Bhargava.

[12] By the way, Manjul never told me what $\mathrm{ord}_{-1}(0)$ should be...

namely that the definition does not depend upon the choice of representative of $\frac{a}{b}$. But it clearly doesn't:

$$\mathrm{ord}_p(\frac{ac}{bc}) = \mathrm{ord}_p(ac) - \mathrm{ord}_p(bc)$$

$$= \mathrm{ord}_p(a) + \mathrm{ord}_p(c) - \mathrm{ord}_p(b) - \mathrm{ord}_p(c) = \mathrm{ord}_p(a) - \mathrm{ord}_p(b) = \mathrm{ord}_p(\frac{a}{b}).$$

So we get a map

$$\mathrm{ord}_p : \mathbb{Q}^\times \to \mathbb{Z}$$

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 $\mathrm{ord}_p(x) \geq 0$ 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, this is a standard question in honors math classes because...well, frankly, because it's rather hard.[13] 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 $n \geq 1$, $\mathrm{ord}_2(S_n) < 0$. It is true for $n = 2$; moreover we don't have to do the calculation for $n = 3$: since $\mathrm{ord}_2(\frac{1}{3}) = 0 \neq \mathrm{ord}_2(S_2)$, we must have $\mathrm{ord}_2(S_2 + \frac{1}{3}) = \min(\mathrm{ord}_2(S_2), \mathrm{ord}_2(S_3)) = -1$. And then we get $\frac{1}{4}$, which 2-order -2, which is different from $\mathrm{ord}_2(S_3)$, 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 $\mathrm{ord}_2(S_4) = -2$. Excitedly testing a few more values, we see that this pattern continues: $\mathrm{ord}_2(S_n)$ and $\mathrm{ord}_2(\frac{1}{n+1})$ 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 $\mathrm{ord}_2(S_n)$ 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?

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.

[13] 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 Cebyshev. One of Paul Erdös' early mathematical triumphs was an elegant new proof of this result.)

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 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.

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 y. 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 $a, b \in \mathbb{Z}$, not both zero, and any $m \in \mathbb{Z}$, 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?[14] Second, given that we know that solutions 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.

The above argument does not, of course, use any special properties of 3 and 7: with purely notational changes it carries over to a proof of the following result.

Theorem 14. *For a and b coprime positive integers, the general integral solution to $xa + yb = 1$ 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.*

However, let us take the oppotunity to give a slightly different reformulation and reproof of Theorem 14. We will work in slightly more generality: for fixed, relatively prime nonzero integers a and b and a variable integer N, consider all integral solutions of the equation

$$ax + by = N \tag{5}$$

To borrow terminology from other areas of mathematics,[15] (5) is **linear and inhomogeneous** in x and y. What this means is that the left hand side is an expression which is linear in x and y but the right-hand side is nonzero. There is an associated **homogeneous linear equation**:

$$ax + by = 0 \tag{6}$$

[14]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.

[15]Especially, from the elementary theory of differential equations.

Here we are saying something quite basic in a fancy way: the real solutions of (6) form a line through the origin in $\mathbb{R}^2$, with slope $m = \frac{-a}{b}$. But the set of integer solutions to (6) also has a nice algebraic structure: if (x_1, y_1), (x_2, y_2) are any two integer solutions and C is any integer, then since

$$a(x_1 + x_2) + b(y_1 + y_2) = (ax_1 + by_1) + (ax_2 + by_2) = 0 + 0 = 0,$$

$$a(Cx_1) + b(Cy_2) = C(ax_1 + by_1) = C \cdot 0 = 0,$$

both the sum $(x_1, y_1) + (x_2, y_2)$ and the integer multiple $C(x_1, y_1)$ are solutions. To be algebraically precise about it, the set of integer solutions to (6) forms a subgroup of the additive group of the one-dimensional $\mathbb{R}$-vector space of all real solutions.

Now we claim that it is easy to solve the homogeneous equation directly. The $\mathbb{Q}$-solutions are clearly $\{(x, \frac{-a}{b}x) \mid x \in \mathbb{Q}\}$. And, since a and b are relatively prime, in order for x and $\frac{-a}{b}x$ to both be integers, it is necessary and sufficient that x itself be an integer and that it moreover be divisible by b. Therefore the general integral solution to the homogeneous equation is $\{(nb, -na) \mid n \in \mathbb{Z}\}$.

Now we make the fundamental observation about solving inhomogeneous linear equations in terms of the associated homogeneous linear equation. We claim that if (x_0, y_0) is any one solution to the inhomogeneous equation (5) and $(x_n, y_n) = (nb, -na)$ is the general solution to the associated homogeneous equation (6), then the general solution to the inhomogeneous equation is $(x_0, y_0) + (x_n, y_n)$. Let's check this. On the one hand, we have

$$a(x_0 + x_n) + b(y_0 + y_n) = (ax_0 + by_0) + (ax_n + by_n) = N + 0 = N,$$

so these are indeed solutions to the inhomogeneous equation. On the other hand, if (x, y) and (x', y') are any two solutions to the inhomogeneous equation, then, by a very similar computation, their difference $(x - x', y - y')$ is a solution to the homogeneous equation.

In other words the set of all solutions to the inhomogeneous equation is simply a **translate** of the abelian group of all solutions to the homogeneous equation. Thus, since the solutions to the homogeneous equation are simply a set of points along the line with distance $\sqrt{a^2 + b^2}$ between consecutive solutions, the same holds for **all** the inhomogeneous equations, independent of N.

Remark aside: At the cost of introducing some further fancy terminology, the discussion can be summarized by saying that the solution set to the inhomogeneous equation is a **principal homogeneous space** for the commutative group of solutions to the homogeneous equation. The general meaning of this is in terms of group actions on sets: let G be a group, X a set, and $\bullet : G \times X \to X$ an action of G on X. (We are assuming familiarity with this algebraic concept only to make the present digression. It will not be needed in the rest of the course.) Then we say that X is a principal homogeneous space for G if the action is simply transitivie: for all $x, y \in X$, there exists a unique element g of G such that $g \cdot x = y$.

To look back this homogeneous/inohomogeneous argument, what it *doesn't* give

us is any particular solution to the inhomogeneous equation.[16] To get this in any given case we can use Euclid's algorithm, but in thinking about things in general it is useful to acknowledge a certain amount of fuzziness in the picture: we can only say where any particular solution will be located on the line to within an accuracy of $d = \sqrt{a^2 + b^2}$.

What is interesting is that we can use these seemingly very primitive geometric ideas to extract useful information about a more difficult problem. Namely, let us now suppose that a, b, N are all positive integers, and we seek to solve the linear Diophantine equation

$$ax + by = N$$

in positive integers (x, y). Then the geometric picture shows right away that we are interested in the intersection of the infinite family of all integral solutions with the first quadrant of $\mathbb{R}^2$. More precisely, we have a line segment L_N which joins $(0, \frac{N}{b})$ to $(\frac{N}{a}, 0)$, and we are asking whether there are integer solutions on L_N.

Notice that the length of L_N is

$$\ell_N = \sqrt{\left(\frac{N}{a}\right)^2 + \left(\frac{N}{b}\right)^2} = N\sqrt{\frac{1}{a^2} + \frac{1}{b^2}} = N\frac{\sqrt{a^2 + b^2}}{ab} = \left(\frac{d}{ab}\right) N.$$

Thus when N is small, L_N is a very small line segment, and since consecutive integral solutions on the line are spaced d units apart, it is by no means guaranteed that there are any integral solutions on L_N. For instance, since $ax+by \geq a+b \geq 2$, there is no positive integral solution to $ax + by = 1$. But since L_N grows linearly with N and d is independent of N, when N is sufficiently large we must have some integral points on L_N. In fact this must happen as soon as $\ell_N > d$.[17] By similar reasoning, the number of solutions must be extremely close to $\frac{\ell_n}{d} = \frac{N}{ab}$. Precisely:

Theorem 15. *Let $a, b \in \mathbb{Z}^+$ be relatively prime, and let $N \in \mathbb{Z}^+$.*
a) If $N > ab$, then there exist positive integers x, y such that $ax + by = N$.
b) Let $\mathcal{N}_N$ be the number of positive integral solutions (x, y) to $ax + by = N$. Then

$$\lfloor \frac{N}{ab} \rfloor - 1 \leq \mathcal{N}_N \leq \lfloor \frac{N}{ab} \rfloor + 1.$$

We leave the details of the proof of Theorem 15 to the interested reader.

It turns out that the lower bound on N in part a) is of the right order of magnitude, but is never sharp: for instance, if $a = 2$, $b = 3$, then the theorem asserts $2x + 3y = N$ has a positive integral solution if $N > 6$, whereas pure thought shows that it suffices to take $N \geq 2$. The sharp lower bound is known (in terms of a and b, of course) and is a result of J.J. Sylvester: c.f. Problem Set XX.

[16] Indeed, so far as this abtract reasoning goes, such a solution might not exist: according to the definition we gave for a principal homogeneous space, taking $X = \varnothing$ gives a principal homogeneous space under any group G.

[17] To understand the reasoning here, imagine that you know that a certain bus comes once every hour at a fixed time – i.e., at a certain number of minutes past each hour – but you don't know exactly what that fixed time is. Nevertheless, if you wait for any full hour, you will be able to catch the bus.