1. Introduction

1.1. Organisation

1.2. Literature

1.3. The plan

1.4. Motivation: \( n = x^2 + y^2 \)

1.5. Motivation: Algebraic numbers

1.6. Motivation: Shamir’s Secret Sharing Scheme

1.6.1. The problem

1.6.2. The \( k = 1 \) case

1.6.3. The \( k = n \) case: what doesn’t work

1.6.4. The XOR operations

1.6.5. The \( k = n \) case: an answer

1.6.6. The \( k = 2 \) case

1.6.7. The \( k = 3 \) case

2. Elementary number theory

2.1. Notations

2.2. Divisibility

2.3. Congruence modulo \( n \)

2.4. Chains of congruences

2.5. Division with remainder

2.6. Even and odd numbers

2.7. The floor function

2.8. Common divisors, the Euclidean algorithm and the Bezout theorem

2.9. Substitutivity for congruences
1. Introduction

This file will contain the notes from the Math 4281 class (“Introduction to Modern Algebra”) I am teaching at UMN in Spring 2019. I will type the first draft directly in the classroom, and subsequently expand it into proper writing. Occasionally, I will also add extra sections not covered in class.

The website of the class is http://www-users.math.umn.edu/~dgrinber/19s/index.html; you will find homework sets there.

1.1. Organisation

See the syllabus for the organization of this class and for the homework.

1.2. Literature

Many books have been written about abstract algebra. I have only a passing familiarity with most of them. Some of the “bibles” of the subject (bulky texts covering lots of material) are Dummit/Foote [DumFoo04], Knapp [Knapp16a] and [Knapp16b] (both freely available), van der Waerden [Waerde91a] and [Waerde91b] (one of the oldest texts on modern algebra, thus rather dated, but still as readable as ever).

Of course, any book longer than 200 pages likely goes further than our course will (unless it is full of details or solved exercises or printed in really large letters). Thus, let me recommend some more introductory sources. Siksek’s lecture notes [Siksek15] are a readable introduction that is a lot more amusing than I had ever expected an algebra text to be. Goodman’s free book [Goodma16] combines introductory material with geometric motivation and applications, such as the classification of regular polyhedra and 2-dimensional crystals. In a sense, it is a great complement to our ungeometric course. Pinter’s [Pinter10] often gets used in classes like ours. Armstrong’s notes [Armstr18] cover a significant part of what we do (and he will likely have notes for a second course written up by the end of this semester).

Keith Conrad’s blurbs [Conrad*] are not a book, as they only cover selected topics. But at pretty much every topic they cover, they are one of the best sources (clear, full of examples, and often going fairly deep). We shall follow one of them particularly closely: the one on Gaussian integers [ConradG].

We will use some basic linear algebra, all of which can be found in Hefferon’s book [Heffer17] (but we won’t need all of this book). As far as determinants are concerned, we will briefly build up their theory; we refer to [Strick13, Section 12 & Appendix B] for proofs (and to [Grinbe15, Chapter 6] for a really detailed and formal treatment).

This course will begin (after some motivating questions) with a survey of elementary number theory. This is in itself a deep subject (despite the name) with a long history (perhaps as old as mathematics), and of course we will just scratch the surface. Books like [NiZuMo91], [Burton10] and [UspHea39] cover a lot more than
we can do. The Gallier/Quaintance survey [GalQua17] covers a good amount of basics and more.

We assume that the reader is familiar with the commonplaces of mathematical argumentation, such as induction (including strong induction), “WLOG” arguments, proof by contradiction, summation signs ($\sum$) and polynomials (a vague notion of polynomials will suffice; we will give a precise definition when it becomes necessary). If not, several texts can be helpful in achieving such familiarity: e.g., [LeLeMe18, particularly Chapters 1–5], [Hammac18], [Day16].

2019-01-23 lecture

1.3. The plan

The material I am going to cover is mostly standard. However, the order in which I will go through it is somewhat unusual: I will spend a lot of time studying the basic examples before defining abstract notions such as “group”, “monoid”, “ring” and “field”. This way, once I come to these notions, you’ll already have many examples to work with. (Don’t be fooled by the word “example”: We will prove a lot about them, much of which is neither straightforward nor easy.)

First, I will show some motivating questions that are easy to state yet require abstract algebra to prove. We will hopefully see their answers by the end of this class. (Some of them can also be answered elementarily, without using abstract algebra, but such answers usually take more work and are harder to find.)

1.4. Motivation: $n = x^2 + y^2$

A perfect square means the square of an integer. Thus, the perfect squares are

$$0^2 = 0, \quad 1^2 = 1, \quad 2^2 = 4, \quad 3^2 = 9, \quad 4^2 = 16, \quad \ldots$$

Here is an old problem (first solved by Pierre de Fermat in 1640, but apparently already studied by Diophantus in the 3rd Century):

**Question 1.1.** What integers can be written as sums of two perfect squares?

For example, 5 can be written in this way, since $5 = 2^2 + 1^2$.

So can 4, since $4 = 2^2 + 0^2$. (Keep in mind that 0 is a perfect square.)

However, 7 cannot be written in this way. In fact, if we had $7 = a^2 + b^2$ for two integers $a$ and $b$, then $a^2$ and $b^2$ would have to be $\leq 7$ (since $a^2$ and $b^2$ are always $\geq 0$, no matter what sign $a$ and $b$ have); but the only perfect squares that are $\leq 7$ are $0, 1, 4$, and there is no way to write 7 as a sum of two of these perfect squares (just check all the possibilities).

For a similar but simpler reason, no negative number can be written as a sum of two perfect squares.
We can of course approach Question 1.1 using a computer: It is very easy to check, for a given integer $n$, whether $n$ is a sum of two perfect squares. (Just check all possibilities for $a$ and $b$ for the validity of the equation $n = a^2 + b^2$. You only need to try $a$ and $b$ belonging to \( \{0, 1, \ldots, \lfloor \sqrt{n} \rfloor \} \), where $\lfloor y \rfloor$ (for a real number $y$) denotes the smallest integer that is less or equal than $y$ (also known as “$y$ rounded down”).) If you do this, you will see that among the first 101 nonnegative integers, the ones that can be written as sums of two perfect squares are precisely

\[
0, 1, 2, 4, 5, 8, 9, 10, 13, 16, 17, 18, 20, 25, 26, 29, \\
32, 34, 36, 37, 40, 41, 45, 49, 50, 52, 53, 58, 61, 64, \\
65, 68, 72, 73, 74, 80, 81, 82, 85, 89, 90, 97, 98, 100.
\]

Having this data, you can look up the sequence in the Online Encyclopedia of Integer Sequences (short OEIS) and see that the sequence of these integers is known as OEIS Sequence A001481. In the “Comments” field, you can read a lot of what is known about it (albeit in telegraphic style).

For example, one of the comments says “Closed under multiplication”. This is short for “if you multiply two entries of the sequence, then the product will again be an entry of the sequence”. In other words, if you multiply two integers that are sums of two perfect squares, then you get another sum of two perfect squares. Why is this so?

It turns out that there is a “simple” reason for this: the identity

\[
(a^2 + b^2)(c^2 + d^2) = (ad + bc)^2 + (ac - bd)^2,
\]

which holds for arbitrary reals $a, b, c, d$ (and thus, in particular, for integers). This is known as the Brahmagupta-Fibonacci identity, and of course can easily be proven by expanding both sides. But how would you come up with such an identity?

If you stare at the above sequence long enough, you may also discover another pattern: An integer of the form $4k + 3$ with integer $k$ (that is, an integer that is larger by 3 than a multiple of 4) can never be written as a sum of two perfect squares. (Thus, 3, 7, 11, 15, 19, 23, \ldots cannot be written in this way.) This does not account for all integers that cannot be written in this way, but it does provide some clues to the answer that we will later see. In order to prove this observation, we shall need basic modular arithmetic (or at least division with remainder); we will see this proof very soon (see Exercise 2.11 (c)).

Further questions can be asked. One of them is: Given an integer $n$, how many ways are there to represent $n$ as a sum of two perfect squares? This is actually several questions masquerading as one, since it is not so clear what a “way” is. Do $5 = 1^2 + 2^2$ and $5 = 2^2 + 1^2$ count as two different ways? What about $5 = 1^2 + 2^2$ versus $5 = (-1)^2 + 2^2$ (here, the perfect squares are the same, but do we really want to count the squares or rather the numbers we are squaring?).

Let me formalize the question as follows:
**Question 1.2.** Let \( n \) be an integer.

(a) How many pairs \((a, b)\) \(\in\mathbb{N}^2\) are there that satisfy \( n = a^2 + b^2 \)? Here, and in the following, \( \mathbb{N} \) denotes the set \( \{0, 1, 2, \ldots\} \) of all nonnegative integers.

(b) How many pairs \((a, b)\) \(\in\mathbb{Z}^2\) are there that satisfy \( n = a^2 + b^2 \)? Here, and in the following, \( \mathbb{Z} \) denotes the set \( \{\ldots, -2, -1, 0, 1, 2, \ldots\} \) of all integers.

(c) How do these counts change if we count unordered pairs instead (i.e., count \((a, b)\) and \((b, a)\) as one only)?

Note that when I say “pair”, I always mean “ordered pair” by default, unless I explicitly say “unordered pair”.

Again, a little bit of programming easily yields answers to all three parts of this question for small values of \( n \), and the resulting data can be plugged into the OEIS and yields lots of information.

**First steps toward answering Question 1.2.** (a) I claim that the number of such pairs is even unless \( n \) is twice a perfect square (i.e., unless \( n = 2m^2 \) for some integer \( m \)); in the latter case, this number is odd instead.

Why? Let me define a solution to be a pair \((a, b)\) such that \( n = a^2 + b^2 \). So I want to know whether the number of solutions is even or odd. But we have \( a^2 + b^2 = b^2 + a^2 \) for all \( a \) and \( b \). Thus, if \((a, b)\) is a solution, then so is \((b, a)\). Hence, the solutions themselves “come in pairs”, with each solution \((a, b)\) being matched to the solution \((b, a)\), unless there is a solution \((a, b)\) with \( a = b \) (because such a solution would be matched to itself, and thus not form an actual pair). But solutions \((a, b)\) with \( a = b \) are easy to classify: If \( n \) is twice a perfect square, then there is exactly one such solution (namely, \((\sqrt{n/2}, \sqrt{n/2})\)); otherwise there is none (because \( n = a^2 + b^2 \) with \( a = b \) leads to \( n = b^2 + b^2 = 2b^2 \)). Since we know that all the other solutions “come in pairs”, we thus conclude that the number of solutions is odd if \( n \) is twice a perfect square and even otherwise. This proves our claim.

Of course, we have not made much headway into Question 1.2; knowing whether a number is even or odd is far from knowing the number itself. But I think the argument above was worth showing; similar reasoning is used a lot in algebra.

(b) By reasoning analogous to the one we used in part (a), we can see that the number of such pairs will be divisible by 8 whenever \( n \) is neither a perfect square nor twice a perfect square. Indeed, this relies on the fact that

\[
\begin{align*}
a^2 + b^2 &= b^2 + a^2 = (-a)^2 + b^2 = b^2 + (-a)^2 = a^2 + (-b)^2 = (-b)^2 + a^2 \\
&= (-a)^2 + (-b)^2 = (-b)^2 + (-a)^2
\end{align*}
\]

for all \( a \) and \( b \). Thus the pairs \((a, b)\) \(\in\mathbb{Z}^2\) that satisfy \( n = a^2 + b^2 \) don’t just come in pairs; they come in sets of 8 (namely, each \((a, b)\) comes in a set with \((b, a)\), \((-a, b)\), \((b, -a)\), \((a, -b)\), \((-b, a)\), \((-a, -b)\) and \((-b, -a)\)). These sets of 8 can “degenerate” to smaller sets when some of their elements coincide, but this can only happen when \( n \) is a perfect square (in which case we can have \((a, b) = (-a, b)\) for example) or twice a perfect square (in which case we can have \((a, b) = (b, a)\) or \((a, b) = (-b, -a)\) or other such coincidences). (Check this!)
(c) We can reduce this to parts (a) and (b). Indeed:

- When $n$ is not twice a perfect square, the number of unordered pairs will be half the number of ordered pairs, since each unordered pair $(u, v)$ unordered corresponds to precisely two ordered pairs $(u, v)$ and $(v, u)$.

- When $n$ is twice a perfect square, we have

\[
\frac{\text{(the number of unordered pairs)}}{2} = \frac{\text{(the number of ordered pairs)}}{2} + \text{(the number of pairs with } a = b).}
\]

Indeed, each unordered pair $(u, v)$ unordered corresponds to precisely two ordered pairs $(u, v)$ and $(v, u)$ unless $u = v$, in which case it corresponds to only one ordered pair. Thus, if we multiply the number of unordered pairs by 2, then we overcount the number of ordered pairs, because we are counting the pairs $(u, v)$ with $u = v$ (that is, the pairs with $a = b$) twice. So we get (the number of ordered pairs) + (the number of pairs with $a = b$). This proves our above formula.

What is the number of pairs with $a = b$? If $n = 0$, then it is 1 (and the only such pair is $(0, 0)$). Otherwise, it is 1 if we are counting pairs in $\mathbb{N}^2$ (and the only such pair is $(\sqrt{n/2}, \sqrt{n/2})$), and is 2 if we are counting pairs in $\mathbb{Z}^2$ (and the only two such pairs are $(\sqrt{n/2}, \sqrt{n/2})$ and $(-\sqrt{n/2}, -\sqrt{n/2})$).

Note that sums of squares have a geometric meaning (going back to Pythagoras): Two real numbers $a$ and $b$ satisfy $a^2 + b^2 = n$ (for a given integer $n \geq 0$) if and only if the point with Cartesian coordinates $(a, b)$ lies on the circle with center 0 and radius $\sqrt{n}$. This will actually prove a valuable insight that will lead us to the answers to the above questions.

Just as a teaser: There are formulas for all three parts of Question 1.2 in terms of divisors of $n$ of the forms $4k + 1$ and $4k + 3$. We will see these formulas after we have properly understood the concept of Gaussian integers.

1.5. Motivation: Algebraic numbers

A real number $z$ is said to be algebraic if there exists a nonzero polynomial $P$ with rational coefficients such that $P(z) = 0$. In other words, a real number $z$ is algebraic if and only if it is a root of a nonzero polynomial with rational coefficients.

(If you know the complex numbers, you can replace “real” by “complex” in this definition; but we shall only see real numbers in this little motivational subsection.)

Examples:

- Each rational number $a$ is algebraic (being a root of the nonzero polynomial $x - a$ with rational coefficients).

\[\text{In the rest of this argument, “pair” will always mean “pair } (a, b) \text{ satisfying } n = a^2 + b^2.\]
• The number $\sqrt{2}$ is algebraic (being a root of the nonzero polynomial $x^2 - 2$).

• The number $\sqrt[3]{5}$ is algebraic (being a root of $x^3 - 5$).

• All the roots of the polynomial $f(x) := \frac{3}{2}x^4 + 17x^3 - 12x + \frac{9}{4}$ (whatever they are) are algebraic.
  Speaking of these roots, what are they? Using a computer, one can show that this polynomial $f(x)$ has 4 real roots $(-11.269\ldots, -0.960\ldots, 0.198\ldots, 0.697\ldots)$, which can be written as complicated expressions with radicals (i.e., $\sqrt[\ldots]$ signs), though complex numbers appear in these expressions (despite the roots being real!). All this does not matter to the fact that they are algebraic :)

• All the roots of the polynomial $g(x) := x^7 - x^5 + 1$ are algebraic.
  This polynomial has only one real root. This root cannot be written as an expression with radicals (as can be proven using Galois theory – indeed, the discovery of this theory greatly motivated the development of abstract algebra). Nevertheless, it is algebraic, by definition. (The same holds for the remaining 6 complex roots of $g$ – we are working with real numbers here only for the sake of familiarity.)

• The most famous number that is not algebraic is $\pi$. This is a famous result of Lindemann, but it belongs to analysis, not to algebra, because $\pi$ is not defined algebraically in the first place (it is defined as the length of a curve or as an area of a curved region – but either of these definitions boils down to a limit of a sequence).

• The second most famous number that is not algebraic is Euler’s number $e$ (the basis of the natural logarithm). Again, analysis is needed to define $e$, and thus also to prove its non-algebraicity.

Numbers that are not algebraic are called transcendental. We shall not study them much, since most of them do not come from algebra. Instead, we shall try our hands at the following question:

**Question 1.3.**  
(a) Is the sum of two (or, more generally, finitely many) algebraic numbers always algebraic? 
(b) What if we replace “sum” by “difference” or “product”?

Let me motivate why this is a natural question to ask. The sum of two integers is still an integer; the sum of two rational numbers is still a rational number. These facts are fundamental; without them we could hardly work with integers and rational numbers. If a similar fact would not hold for algebraic numbers, it would mean that the algebraic numbers are not a good “number system” to work in; on a practical level, it would mean that (e.g.) if we defined a function on the set of all algebraic numbers, then we could not plug a sum of algebraic numbers into it.
Attempts at answering Question 1.3 (a). Let us try a particularly simple example of a sum of two algebraic numbers: Let \( w = \sqrt{2} + \sqrt{3} \). Is \( w \) algebraic?

To answer this question affirmatively, we need to find a nonzero polynomial \( f(x) \) with rational coefficients that has \( w \) as a root.

Just looking at the equality \( w = \sqrt{2} + \sqrt{3} \), we cannot directly eyeball such an \( f \). The problem, in a sense, is that there are too many (namely, two) square roots in this equality.

However, if we square this equality, then we obtain

\[
w^2 = (\sqrt{2} + \sqrt{3})^2 = 2 + 2\sqrt{2}\cdot\sqrt{3} + 3 = 5 + 2\sqrt{6},
\]

which is an equality with only one square root (a sign of progress). Subtracting 5 from this equality (in order to “isolate” this remaining square root), we obtain \( w^2 - 5 = 2\sqrt{6} \). If we now square this equality, then we obtain \( (w^2 - 5)^2 = (2\sqrt{6})^2 = 24 \).

At this point all square roots are gone, and we are left with an equality that contains rational numbers and \( w \) only! We can further rewrite it as \( (w^2 - 5)^2 - 24 = 0 \). Thus, \( w \) is a root of the polynomial \( f(x) := (x^2 - 5)^2 - 24 = x^4 - 10x^2 + 1 \). This means that \( w \) is algebraic (since \( f \) is nonzero).

Let us try a more complicated example: Let \( z = \sqrt{2} + 3\sqrt{2} \). Is \( z \) algebraic? The squaring trick no longer works, since squaring \( \sqrt{2} + 3\sqrt{2} \) does not reduce the number of radicals (= root signs). Let’s instead try rewriting \( z = \sqrt{2} + 3\sqrt{2} \) as \( z - \sqrt{2} = 3\sqrt{2} \). Cubing this equality, we obtain \( (z - \sqrt{2})^3 = 2 \). In view of

\[
(z - \sqrt{2})^3 = z^3 - 3z^2\sqrt{2} + 3z\left(\sqrt{2}\right)^2 - \left(\sqrt{2}\right)^3
\]

(this is a particular case of the identity \((a - b)^3 = a^3 - 3a^2b + 3ab^2 - b^3\), which is one form of the Binomial Theorem for exponent 3), this becomes

\[
z^3 - 3z^2\sqrt{2} + 3z\left(\sqrt{2}\right)^2 - \left(\sqrt{2}\right)^3 = 2.
\]

This simplifies to

\[
z^3 - 3\sqrt{2}z^2 + 6z - 2\sqrt{2} = 2.
\]

Let us transform this inequality in such a way that all terms with a \( \sqrt{2} \) in them end up on the right hand side while all the remaining terms end up on the left. We thus obtain

\[
z^3 + 6z - 2\left(3z^2 + 2\right)\sqrt{2}.
\]

Now, squaring this equality yields

\[
(z^3 + 6z - 2)^2 = (3z^2 + 2)^2 2.
\]
Hence, $z$ is a root of the polynomial

$$g(x) := \left(x^3 + 6x - 2\right)^2 - 2\left(3x^2 + 2\right)^2 = x^6 - 6x^4 - 4x^3 + 12x^2 - 24x - 4.$$  

This is a nonzero polynomial with rational coefficients; hence, $z$ is algebraic.

We thus have verified that the sum of two algebraic numbers is algebraic in two cases. What about more complicated cases, such as 

$$\sqrt{2} + \sqrt{3} + \sqrt{11} ?$$

This is a sum of two algebraic numbers (since we already know that $\sqrt{2} + \sqrt{3} = w$ is algebraic). Is it algebraic? Neither of our above two methods properly works here; do we have to come up with new ad-hoc tricks?

\[ \square \]

### 1.6. Motivation: Shamir's Secret Sharing Scheme

#### 1.6.1. The problem

Adi Shamir is one of the founders of modern mathematical cryptography (famous in particular for the RSA cryptosystem, see later).

Shamir’s Secret Sharing Scheme is a way in which a secret $a$ (a piece of data – e.g., nuclear launch codes) can be distributed among $n$ people in such a way that

- any $k$ of them can (if they come together) reconstruct it uniquely, but
- any $k - 1$ of them (if they come together) cannot gain any insight about it (i.e., not only cannot they reconstruct it, but they cannot even tell that some values are more likely than others to be $a$).

Here $n$ and $k$ are fixed positive integers.

Understanding this scheme completely will require some abstract algebra, but we can already start thinking about the problem and get reasonably far.

So we have $n$ people $1, 2, \ldots, n$, a positive integer $k \in \{1, 2, \ldots, n\}$ and a secret piece of data $a$. We assume that this data $a$ is encoded as a bitstring – i.e., a finite sequence of bits. A bit is an element of the set $\{0, 1\}$. Thus, examples of bitstrings are $(0, 1, 1, 0)$ and $(1)$ and $(1, 1, 0, 1, 0, 0, 0)$ as well as the empty sequence $(\ )$. When writing bitstring, we shall usually omit both the commas and the parentheses; thus, e.g., the bitstring $(1, 1, 0, 1, 0, 0, 0)$ will become 1101000. Make sure you don’t mistake it for a number. Our goal is to give each of the $n$ people $1, 2, \ldots, n$ some bitstring in such a way that:

- **Requirement 1:** Any $k$ of the $n$ people can (if they come together) reconstruct $a$ uniquely.
• *Requirement 2*: Any $k - 1$ of the $n$ people are unable to gain any insight about $a$ (even if they collaborate).

We denote the bitstrings given to the people $1, 2, \ldots, n$ by $a_1, a_2, \ldots, a_n$, respectively.

We assume that the length of our secret bitstring $a$ is known in advance to all parties; i.e., it is not a secret. Thus, when we say "$k - 1$ persons cannot gain any insight about $a$", we do not mean that they don’t know the length; and when we say "some values are more likely than others to be $a$", we only mean values that fit this length.

1.6.2. The $k = 1$ case

One simple special case of our problem is when $k = 1$. In this case, it suffices to give each of the $n$ people the full secret $a$ (that is, we set $a_i = a$ for all $i$). Then, Requirement 1 is satisfied (since any 1 of the $n$ people already knows $a$), while Requirement 2 is satisfied as well (0 people know nothing).

1.6.3. The $k = n$ case: what doesn’t work

Let us now consider the case when $k = n$. This case will not help us solve the general problem, but it will show some ideas that we will encounter again and again in abstract algebra.

We want to ensure that all $n$ people needed to reconstruct the secret $a$, while any $n - 1$ of them will be completely clueless.

It sounds reasonable to split $a$ into $n$ parts, and give each person one of these parts$^2$ (i.e., we let $a_i$ be the $i$-th part of $a$ for each $i \in \{1, 2, \ldots, n\}$). This method satisfies Requirement 1 (indeed, all $n$ people together can reconstruct $a$ simply by fusing the $n$ parts back together), but fails Requirement 2 (indeed, any $n - 1$ people know $n - 1$ parts of the secret $a$, which is a far from being clueless about $a$). So this method doesn’t work. It is not that easy.

1.6.4. The XOR operations

One way to solve the $k = n$ case is using the XOR operation.

Let us first define some basic language. A *binary operation* on a set $S$ is (informally speaking) a function that takes two elements of $S$ and assigns a new element of $S$ to them. More formally:

**Definition 1.4.** A binary operation on a set $S$ is a map $f$ from $S \times S$ to $S$. When $f$ is a binary operation on $S$ and $a$ and $b$ are two elements of $S$, we shall write $afb$ for the value $f(a, b)$.

$^2$assuming that $a$ is long enough for that
Example 1.5. Addition, subtraction and multiplication of integers are three binary operations on the set $\mathbb{Q}$ (the set of all rational numbers). For example, addition is the map from $\mathbb{Q} \times \mathbb{Q}$ to $\mathbb{Q}$ that sends each pair $(a, b) \in \mathbb{Q} \times \mathbb{Q}$ to $a + b$.

Division is not a binary operation on the set $\mathbb{Q}$. Indeed, if it was, then it would send the pair $(1, 0)$ to some integer called $1/0$; but there is no such integer.

There are myriad more complicated binary operations around waiting for someone to name them. For example, you could define a binary operation $\otimes$ on the set $\mathbb{Q}$ by $a \otimes b = \frac{a - b}{1 + a^2 + b^2}$. Indeed, you can do this because $1 + a^2 + b^2$ is always nonzero when $a, b \in \mathbb{Q}$ (after all, squares are nonnegative, so that $\sum a^2 \geq 0$). I am not saying that you should...

Now, we define some specific binary operations on the set $\{0, 1\}$ of all bits, and on the set $\{0, 1\}^n$ of all length-$n$ bitstrings (for a given $n$).

**Definition 1.6.** We define a binary operation XOR on the set $\{0, 1\}$ by setting

$$0 \text{ XOR } 0 = 0, \quad 0 \text{ XOR } 1 = 1, \quad 1 \text{ XOR } 0 = 1, \quad 1 \text{ XOR } 1 = 0.$$  

This is a valid definition, because there are only four pairs $(a, b) \in \{0, 1\} \times \{0, 1\}$, and we have just defined $a \text{ XOR } b$ for each of these four options. We can also rewrite this definition as follows:

$$a \text{ XOR } b = \begin{cases} 1, & \text{if } a \neq b; \\ 0, & \text{if } a = b \end{cases} = \begin{cases} 1, & \text{if exactly one of } a \text{ and } b \text{ is 1;} \\ 0, & \text{otherwise.} \end{cases}$$

For lack of a better name, we refer to $a \text{ XOR } b$ as the “XOR of $a$ and $b$”.

The name “XOR” is short for “exclusive or”. In fact, if you identify bits with boolean truth values (so the bit 0 stands for “False” and the bit 1 stands for “True”), then $a \text{ XOR } b$ is precisely the truth value for “exactly one of $a$ and $b$ is True”, which is also known as “$a$ exclusive-or $b$”.

**Definition 1.7.** Let $m$ be a nonnegative integer. We define a binary operation XOR on the set $\{0, 1\}^m$ (this is the set of all length-$m$ bitstrings) by

$$(a_1, a_2, \ldots, a_m) \text{ XOR } (b_1, b_2, \ldots, b_m) = (a_1 \text{ XOR } b_1, a_2 \text{ XOR } b_2, \ldots, a_m \text{ XOR } b_m).$$

In other words, if $a$ and $b$ are two length-$m$ bitstrings, then $a \text{ XOR } b$ is obtained by taking the XOR of each entry of $a$ with the corresponding entry of $b$, and packing these $m$ XORs into a new length-$m$ bitstring.
For example,

\[(1001) \text{ XOR } (1100) = 0101;\]
\[(11011) \text{ XOR } (10101) = 01110;\]
\[(11010) \text{ XOR } (01011) = 10001;\]
\[(1) \text{ XOR } (0) = 1;\]
\[](0) \text{ XOR } (0) = (0).

Note that if \(a\) and \(b\) are two length-\(m\) bitstrings, then the 0's in the bitstring \(a \text{ XOR } b\) are at the positions where \(a\) and \(b\) have equal entries, and the 1's in \(a \text{ XOR } b\) are at the positions where \(a\) and \(b\) have different entries. Thus, the operation XOR on bitstring essentially pinpoints the differences between \(a\) and \(b\).

We observe the following simple properties of these operations XOR on bits and on bitstrings:

- We have \(a \text{ XOR } 0 = a\) for any bit \(a\). (This can be trivially checked by considering both possibilities for \(a\).)
- Thus, \(a \text{ XOR } 0 = a\) for any bitstring \(a\), where 0 denotes the bitstring \(00\cdots0 = (0,0,\ldots,0)\) (of appropriate length – i.e., of the same length as \(a\)).
- We have \(a \text{ XOR } a = 0\) for any bit \(a\). (This can be trivially checked by considering both possibilities for \(a\).)
- Thus, \(a \text{ XOR } a = 0\) for any bitstring \(a\). We shall refer to this as the self-cancellation law.
- We have \(a \text{ XOR } b = b \text{ XOR } a\) for any bits \(a, b\). (Again, this is easy to check by going through all four options for \(a\) and \(b\).)
- Thus, \(a \text{ XOR } b = b \text{ XOR } a\) for any bitstrings \(a, b\).
- We have \(a \text{ XOR } (b \text{ XOR } c) = (a \text{ XOR } b) \text{ XOR } c\) for any bits \(a, b, c\). (Again, this is easy to check by going through all eight options for \(a, b, c\).)
- Thus, \(a \text{ XOR } (b \text{ XOR } c) = (a \text{ XOR } b) \text{ XOR } c\) for any bitstrings \(a, b, c\).
- Thus, for any bitstrings \(a\) and \(b\), we have

\[
(a \text{ XOR } b) \text{ XOR } b = a \text{ XOR } \underbrace{(b \text{ XOR } b)}_{=0} = a \text{ XOR } 0 = a.
\]
(by the self-cancellation law)

This observation gives rise to a primitive cryptosystem (known as a one-time pad): If you have a secret bitstring \(a\) that you want to encrypt, and another

\[\text{As a mnemonic, we shall try to use boldfaced letters like } a \text{ and } b \text{ for bitstrings and regular italic letters like } a \text{ and } b \text{ for single bits.}\]
secret bitstring \( b \) that can be used as a key, then you can encrypt \( a \) by XORing it with \( b \) (that is, you transform it into \( a \text{ XOR } b \)). Then, you can decrypt it again by XORing it with \( b \) again; indeed, if you do this, you will obtain \((a \text{ XOR } b) \text{ XOR } b = a\). This is a highly safe cryptosystem as long as you can safely communicate the key \( b \) to whomever needs to be able to decrypt (or encrypt) your secrets, and as long as you are able to generate uniformly random keys \( b \) of sufficient length. Its only weakness is its impracticality (in many situations): If the secret you want to encrypt is long (say, a whole book), your key will need to be equally long. Even storing such keys can become difficult.

We shall refer to the properties \( a \text{ XOR } b = b \text{ XOR } a \) and \( a \text{ XOR } b = b \text{ XOR } a \) as laws of commutativity, and we shall refer to the properties \( a \text{ XOR } (b \text{ XOR } c) = (a \text{ XOR } b) \text{ XOR } c \) and \( a \text{ XOR } (b \text{ XOR } c) = (a \text{ XOR } b) \text{ XOR } c \) as laws of associativity. These are, of course, similar to well-known facts like \( \alpha + \beta = \beta + \alpha \) and \( \alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma \) for numbers \( \alpha, \beta, \gamma \) (which is why we are giving them the same name). This similarity is not coincidental. Just as for addition or multiplication of numbers, these laws lead to a notion of “XOR-products”:

**Proposition 1.8.** Let \( m \) be a positive integer. Let \( a_1, a_2, \ldots, a_m \) be \( m \) bitstrings. Then, the “XOR-product” expression

\[
a_1 \text{ XOR } a_2 \text{ XOR } a_3 \text{ XOR } \cdots \text{ XOR } a_m
\]

is well-defined, in the sense that it does not depend on the parenthesization.

What do we mean by “parenthesization”? To clarify things, let us set \( m = 4 \). In this case, we want to make sense of the expression \( a_1 \text{ XOR } a_2 \text{ XOR } a_3 \text{ XOR } a_4 \). This expression does not make sense a priori, since it is a XOR of four bitstrings, whereas we have defined only the XOR of two bitstrings. But there are five ways to put parentheses around some of its sub-expressions such that the expression becomes meaningful:

\[
\begin{align*}
(a_1 \text{ XOR } a_2) \text{ XOR } (a_3 \text{ XOR } a_4), \\
((a_1 \text{ XOR } a_2) \text{ XOR } a_3) \text{ XOR } a_4, \\
(a_1 \text{ XOR } (a_2 \text{ XOR } a_3)) \text{ XOR } a_4, \\
a_1 \text{ XOR } (a_2 \text{ XOR } (a_3 \text{ XOR } a_4)), \\
(a_1 \text{ XOR } (a_2 \text{ XOR } a_3)) \text{ XOR } a_4.
\end{align*}
\]

Each of these five parenthesizations (= placements of parentheses) turns our expression \( a_1 \text{ XOR } a_2 \text{ XOR } a_3 \text{ XOR } a_4 \) into a combination of XOR’s of two bitstrings each, and thus gives it meaning. The question is: Do these five parenthesizations give it the same meaning?
Well, let us calculate:

\[
(a_1 \text{ XOR } a_2) \text{ XOR } (a_3 \text{ XOR } a_4) \\
= a_1 \text{ XOR } (a_2 \text{ XOR } (a_3 \text{ XOR } a_4)) \\
= (a_2 \text{ XOR } a_3) \text{ XOR } a_4 \\
= a_1 \text{ XOR } ((a_2 \text{ XOR } a_3) \text{ XOR } a_4) \\
= (a_1 \text{ XOR } (a_2 \text{ XOR } a_3)) \text{ XOR } a_4 \\
= ((a_1 \text{ XOR } a_2) \text{ XOR } a_3) \text{ XOR } a_4,
\]

where we used the law of associativity in each step. This shows that our five parenthesizations yield the same result. Thus, they all give our “XOR-product” expression \(a_1 \text{ XOR } a_2 \text{ XOR } a_3 \text{ XOR } a_4\) the same meaning; so we can say that this expression is well-defined. This confirms Proposition 1.8 for \(m = 4\).

Of course, proving Proposition 1.8 is less simple. Such a proof will appear in Exercise 4 on homework set #0.

1.6.5. The \(k = n\) case: an answer

Let us now return to our problem. We have \(n\) persons 1, 2, \(\ldots\), \(n\) and a secret \(a\) (encoded as a bitstring). We want to give each person \(i\) some bitstring \(a_i\) such that only all \(n\) of them can recover \(a\) but any \(n - 1\) of them cannot gain any insight about \(a\).

We let \(a_1, a_2, \ldots, a_{n - 1}\) be \(n - 1\) uniformly random bitstrings of the same length as \(a\). (Think of them as random gibberish.) Set

\[
a_n = a \text{ XOR } a_1 \text{ XOR } a_2 \text{ XOR } \cdots \text{ XOR } a_{n - 1}.
\]

(This expression makes sense because of Proposition 1.8.)

Then,

\[
a_n \text{ XOR } a_{n - 1} \text{ XOR } a_{n - 2} \text{ XOR } \cdots \text{ XOR } a_1 \\
= (a \text{ XOR } a_1 \text{ XOR } a_2 \text{ XOR } \cdots \text{ XOR } a_{n - 1}) \text{ XOR } a_{n - 1} \text{ XOR } a_{n - 2} \text{ XOR } \cdots \text{ XOR } a_1 \\
= a \text{ XOR } a_1 \text{ XOR } a_2 \text{ XOR } \cdots \text{ XOR } a_{n - 1} \text{ XOR } a_{n - 1} \text{ XOR } a_{n - 2} \text{ XOR } \cdots \text{ XOR } a_1 \\
= a \text{ XOR } a_1 \text{ XOR } a_2 \text{ XOR } \cdots \text{ XOR } a_{n - 2} \text{ XOR } 0 \text{ XOR } a_{n - 2} \text{ XOR } \cdots \text{ XOR } a_1 \\
= a \text{ XOR } a_1 \text{ XOR } a_2 \text{ XOR } \cdots \text{ XOR } a_{n - 2} \text{ XOR } a_{n - 2} \text{ XOR } \cdots \text{ XOR } a_1 \\
= 0
\]

\[
= a
\]
(here, we have been unravelling the big XOR-product from the middle on, by cancelling equal bitstrings using the self-cancellation law and then removing the resulting 0 using the \( a \) XOR 0 = \( a \) law). Hence, the \( n \) people together can decrypt the secret \( a \).

Can \( n - 1 \) people gain any insight about it? The \( n - 1 \) people 1, 2, \ldots, \( n - 1 \) certainly cannot, since all they know are the random bitstrings \( a_1, a_2, \ldots, a_{n-1} \). But the \( n - 1 \) people 2, 3, \ldots, \( n \) cannot gain any insight about \( a \) either: In fact, all they know are the random bitstrings \( a_2, a_3, \ldots, a_{n-1} \) and the bitstring
\[
a_n = a \text{ XOR } a_1 \text{ XOR } a_2 \text{ XOR } \cdots \text{ XOR } a_{n-1};
\]
therefore, all the information they have about \( a \) and \( a_1 \) comes to them through \( a \) XOR \( a_1 \), which says nothing about \( a \) as long as they know nothing about \( a_1 \).

(We used a bit of handwaving in this argument, but then again we never formally defined what it means to "gain no insight"; this is done in courses on cryptography and information theory.) Similar arguments show that any other choice of \( n - 1 \) persons remains equally clueless about \( a \). So we have solved the problem in the case \( k = n \).

1.6.6. The \( k = 2 \) case

The next simple case is when \( k = 2 \). So we want to ensure that any 2 of our \( n \) people can together recover the secret, but no 1 person can learn anything about it alone.

A really nice approach was suggested by Nathan in class: We pick \( n \) random bitstrings \( x_1, x_2, \ldots, x_{n-1} \) of the same length as \( a \). Set
\[
x_n = a \text{ XOR } x_1 \text{ XOR } x_2 \text{ XOR } \cdots \text{ XOR } x_{n-1};
\]
thus, as in the \( k = n \) case, we have
\[
x_n \text{ XOR } x_{n-1} \text{ XOR } x_{n-2} \text{ XOR } \cdots \text{ XOR } x_1 = a. \tag{2}
\]

Each person \( i \) now receives the bitstring
\[
a_i = x_1 x_2 \cdots x_{i-1} x_{i+1} x_{i+2} \cdots x_n,
\]
where the product stands for concatenation (i.e., the bitstring \( a_i \) is formed by writing down all of the bitstring \( x_1, x_2, \ldots, x_n \) one after the other but skipping \( x_i \)). Thus, each person \( i \) can recover all the \( n - 1 \) bitstrings \( x_1, x_2, \ldots, x_{i-1}, x_{i+1}, x_{i+2}, \ldots, x_n \) (because their lengths are the length of \( a \), which is known), but knows nothing about \( x_i \) (his "blind spot"). Hence, 2 people together can recover all the \( n \) bitstrings \( x_1, x_2, \ldots, x_n \) and therefore recover the secret \( a \) (by (2)). On the other hand, each single person has no insight about \( a \) (this is proven similarly to the \( k = n \) case). So again, the problem is solved in this case.
1.6.7. The \( k = 3 \) case

Now, let us come to the case when \( k = 3 \). Now I think the usefulness of the XOR approach has come to its end: at least I don’t know how to make it work here. Instead, out of the blue, I will invoke something completely different: polynomials (let’s say with rational coefficients).

Recall a fact you might have heard in high school: A polynomial \( p(x) = cx^2 + bx + a \) of degree \( \leq 2 \) is uniquely determined by any three of its values. More precisely: If \( u, v, w \) are three fixed distinct numbers, then a polynomial \( p(x) = cx^2 + bx + a \) of degree \( \leq 2 \) is uniquely determined by the values \( p(u), p(v), p(w) \).

We will put this to use now, and sort-of solve the problem.

Also recall that any bitstring of given length \( N \) can be encoded as an integer in \( \{0, 1, \ldots, 2^{N-1}\} \); just read it as a number in binary. More precisely, any bitstring \( a_{N-1}a_{N-2} \cdots a_0 \) of length \( N \) becomes the integer \( a_{N-1} \cdot 2^{N-1} + a_{N-2} \cdot 2^{N-2} + \cdots + a_0 \cdot 2^0 \in \{0, 1, \ldots, 2^N - 1\} \). For example, the bitstring 010110 of length 6 becomes the integer \( 0 \cdot 2^5 + 1 \cdot 2^4 + 0 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = 22 \in \{0, 1, \ldots, 2^6 - 1\} \).

Choose two uniformly random bitstrings \( c \) and \( b \) (of the same length as \( a \)) and encode them as numbers \( c \) and \( b \) (as just explained). Encode the secret \( a \) as a number \( a \) as well (in the same way). Define the polynomial \( p(x) = cx^2 + bx + a \). Reveal to each person \( i \in \{1, 2, \ldots, n\} \) the value \( p(i) \) – or, rather, a bitstring that encodes it in binary – as \( a_i \).

As we know, any three of the values \( p(i) \) uniquely determine the polynomial \( p \). Thus, any three people can use their bitstrings \( a_i \) to recover three values \( p(i) \) and therefore \( p \) and therefore \( a \) (as the constant term of \( p \) and therefore \( a \)) (by decoding \( a \)). So our method satisfies Requirement 1.

Now, let us see whether it satisfies Requirement 2. Any 2 people can recover two values \( p(i) \), which generally do not determine \( p \) uniquely. It is not hard to show that they do not even determine \( a \) uniquely; thus, they do not determine \( a \) uniquely. What’s better: If you know just two values of \( p \), there are infinitely many possible choices for \( p \), and all of them have distinct constant terms (unless one of the two values you know is \( p(0) \), which of course pins down the constant term). So we get infinitely many possible values for \( a \), and thus infinitely many possible values for \( a \). This means that our 2 people don’t gain any insight about \( a \), right?

Not so fast! We cannot really have “infinitely many possible values for \( a \)”, since \( a \) is bound to be a bitstring of a given length – there are only finitely many of those! You can only get infinitely many possible values for \( p \) if you forget how \( p \) was constructed (from \( c \), \( b \) and \( a \)) and pretend that \( p \) is just a “uniformly random” polynomial (whatever this means). But no one can force the 2 people to do this; it is certainly not in their interest! Here are some things they might do with this knowledge:

- Let \( N \) be the length of \( a \) (which, as we said, is known). Thus, \( c \) and \( b \) are
bitstrings of length $N$, so that $c$ and $b$ are integers in $\{0, 1, \ldots, 2^N - 1\}$. Assume that one of the 2 people is person 2. Now, person 2 knows $p(2) = c2^2 + b2 + a = 4c + 2b + a$, and thus knows whether $a$ is even or odd (because $a$ is even resp. odd if and only if $4c + 2b + a$ is even resp. odd). This means she knows the last bit of the secret $a$. This is not “clueless”.

- You might try to fix this by picking $c$ and $b$ to be uniformly random rational numbers instead (rather than using uniformly random bitstrings $c$ and $b$).

Unfortunately, there is no such thing as a “uniformly random rational number” (in the sense that, e.g., larger numbers aren’t less likely to be picked than smaller numbers). Any probability distribution will make some numbers more likely than others, and this will usually cause information about $a$ to “leak”. For example, if $c$ and $b$ are chosen from the interval $[0, 2^N - 1]$, then person 1’s knowledge of $p(1) = c1^2 + b1 + a = c + b + a$ will sometimes reveal to person 1 that $a \geq 0.5 \cdot (2^N - 1)$ (namely, this will happen when $p(1) \geq 2.5 \cdot (2^N - 1)$, which occasionally happens). This, again, is nontrivial information about the secret $a$, which a single person (or even two people) should not be having.

So we cannot make Requirement 2 hold, and the culprit is that there are too many numbers (namely, infinitely many). What would help is a finite “number system” in which we can add, subtract, multiply and divide (so that we can define polynomials over it, and a polynomial of degree $\leq 2$ is still uniquely determined by any 3 values). Assuming that this “number system” is large enough that we can encode bitstrings using “numbers” of this system (instead of integers), we can then play the above game using this “number system” and obtain actually uniformly random numbers.

It turns out that such “number systems” exist. They are called finite fields, and we will construct them later in this course.

Assuming that they can be constructed, we thus obtain a method of solving the problem for $k = 3$. A similar method works for arbitrary $k$, using polynomials of degree $\leq k - 1$. This is called Shamir’s secret sharing scheme.

2019-01-30 lecture (virtual)

2. Elementary number theory

Let us now begin a systematic introduction to algebra. We start with studying integers and their divisibility properties – the beginnings of number theory. Part of these will be used directly in what will follow; part of these will inspire more general results and proofs.
2.1. Notations

**Definition 2.1.** Let \( \mathbb{N} = \{0, 1, 2, \ldots\} \) be the set of **nonnegative** integers.

Let \( \mathbb{P} = \{1, 2, 3, \ldots\} \) be the set of **positive** integers.

Let \( \mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\} \) be the set of integers.

Let \( \mathbb{Q} \) be the set of rational numbers.

Let \( \mathbb{R} \) be the set of real numbers.

Be careful with the notation \( \mathbb{N} \): While I use it for \( \{0, 1, 2, \ldots\} \), various other authors use it for \( \{1, 2, 3, \ldots\} \) instead. There is no consensus in sight on what \( \mathbb{N} \) should mean.

Same holds for the word “natural number” (which I will avoid): It means “element of \( \mathbb{N} \)”, so again its ultimate meaning depends on the author.

2.2. Divisibility

We now go through the basics of divisibility of integers.

**Definition 2.2.** Let \( a \) and \( b \) be two integers. We say that \( a \mid b \) (or “\( a \) divides \( b \)” or “\( b \) is divisible by \( a \)” or “\( b \) is a multiple of \( a \)”) if there exists an integer \( c \) such that \( b = ac \).

Some authors define the “divisibility” relation a bit differently, in that they forbid \( a = 0 \). From the viewpoint of abstract algebra, this feels like an unnecessary exception, so we don’t follow them.

**Example 2.3.**

(a) We have \( 4 \mid 12 \), since \( 12 = 4 \cdot 3 \).

(b) We have \( a \mid 0 \) for any \( a \in \mathbb{Z} \), since \( 0 = a \cdot 0 \).

(c) An integer \( b \) satisfies \( 0 \mid b \) only when \( b = 0 \), since \( 0 \mid b \) implies \( b = 0c = 0 \) (for some \( c \in \mathbb{Z} \)).

(d) We have \( a \mid a \) for any \( a \in \mathbb{Z} \), since \( a = a \cdot 1 \).

(e) We have \( 1 \mid b \) for each \( b \in \mathbb{Z} \), since \( b = 1 \cdot b \).

**Proposition 2.4.**

Let \( a \) and \( b \) be two integers.

(a) We have \( a \mid b \) if and only if \( |a| \mid |b| \).

(b) If \( a \mid b \) and \( b \neq 0 \), then \( |a| \leq |b| \).

(c) Assume that \( a \neq 0 \). Then, \( a \mid b \) if and only if \( \frac{b}{a} \in \mathbb{Z} \).

Before we prove this proposition, let us recall a well-known fact: We have

\[
|xy| = |x| \cdot |y|
\]  

(3)

for any two integers \( x \) and \( y \). (This can be easily proven by case distinction: \( x \) is either nonnegative or negative, and so is \( y \).)

\(^4\)or real numbers
Proof of Proposition 2.4 (a) \[\implies\]
Assume that \(a \mid b\). Thus, there exists an integer \(d\) such that \(b = ad\) (by Definition 2.2). Consider this \(d\). We have \(b = ad\) and thus \(|b| = |ad| = |a| \cdot |d|\) (by (3)). Thus, there exists an integer \(c\) such that \(|b| = |a| \cdot c\) (namely, \(c = |d|\)). In other words, \(|a| \mid |b|\). This proves the “\(\implies\)” direction of Proposition 2.4 (a).

\[\iff\]
Assume that \(|a| \mid |b|\). Thus, there exists an integer \(f\) such that \(|b| = |a| \cdot f\) (by Definition 2.2). Consider this \(f\).

The definition of \(|b|\) shows that \(|b|\) equals either \(b\) or \(−b\). Hence, \(b\) equals either \(|b|\) or \(−|b|\). In other words, \(b\) equals either 1 \(|b|\) or \((−1) \mid |b|\). In other words, \(b = q \mid b\) for some \(q \in \{1, −1\}\). Similarly, \(a = r \mid a\) for some \(r \in \{1, −1\}\). Consider these \(q\) and \(r\).

From \(r \in \{1, −1\}\), we obtain \(r^2 = 1\). Now, \(\frac{r \cdot a}{\frac{r}{|r|}} = \frac{r}{\frac{r}{|r|}} \cdot |a| = |a|\).

Now, \(b = a \cdot qfr\) (since \(a \cdot qfr = qfr \cdot ra = q \cdot \frac{a}{|a|} = q \frac{b}{b} = b\)). Hence, there exists an integer \(c\) such that \(b = ac\) (namely, \(c = qfr\)). In other words, \(a \mid b\). This proves the “\(\iff\)” direction of Proposition 2.4 (a).

Thus, the proof of Proposition 2.4 (a) is complete.

(b) Assume that \(a \mid b\) and \(b \neq 0\).

From \(a \mid b\), we conclude that there exists an integer \(c\) such that \(b = ac\). Consider this \(c\). We have \(ac = b \neq 0\), thus \(c \neq 0\). Hence, \(|c| > 0\), and thus \(|c| \geq 1\) (since \(|c|\) is an integer). We can multiply this inequality by \(|a|\) (since \(|a| \geq 0\), and obtain \(|a| \cdot |c| \geq |a| \cdot 1 = |a|\).

From \(b = ac\), we obtain \(|b| = |ac| = |a| \cdot |c|\) (by (3)). Hence, \(|b| = |a| \cdot |c| \geq |a|\). This proves Proposition 2.4 (b).

(c) \[\implies\]
Assume that \(a \mid b\). Thus, there exists an integer \(d\) such that \(b = ad\). Consider this \(d\). We can divide the equality \(b = ad\) by \(a\) (since \(a \neq 0\)), and thus obtain \(\frac{b}{a} = d \in \mathbb{Z}\). This proves the \(\implies\) direction of Proposition 2.4 (c).

\[\iff\]
Assume that \(\frac{b}{a} \in \mathbb{Z}\). Thus, there exists an integer \(c\) such that \(b = ac\) (namely, \(c = \frac{b}{a}\)). In other words, \(a \mid b\). This proves the \(\iff\) direction of Proposition 2.4 (c).

Hence, the proof of Proposition 2.4 (c) is complete.

Proposition 2.4 (a) shows that both \(a\) and \(b\) in “the statement \(a \mid b\)” can be

\[\frac{b}{a}\]
This is complete.

\[\text{If you are unfamiliar with the shorthand notation “\(\implies\)”}, \text{ let me explain it. Our goal is to prove that \(a \mid b\) if and only if \(|a| \mid |b|\). In other words, we need to prove the equivalence (a \(\mid b\) \(\iff\) (|a| \(\mid |b|\)). In order to prove this equivalence, it suffices to prove the two implications (a \(\mid b\) \(\implies\) (|a| \(\mid |b|\)) (called the “forward implication” or the “\(\implies\)” direction of the equivalence) and (a \(\mid b\) \(\iff\) (|a| \(\mid |b|\)) (called the “backward implication” or the “\(\iff\)” direction”). The shorthand “\(\implies\)” simply marks the beginning of the proof of the forward implication; similarly, the symbol “\(\iff\)” heralds the end of the proof of the backward implication.}

\[\text{Me saying “Consider this \(d’\)” means that I am picking some integer \(d\) such that \(b = ad\) (this can be done, since we have just proven that such a \(d\) exists), and will be referring to it as \(d\) from now on.}\]
replaced by their absolute values. Thus, when we talk about divisibility of integers, the sign of the integers does not really matter – it usually suffices to work with nonnegative integers. We will often use this (tacitly, after a couple times) in proofs.

The next proposition shows some basic properties of the divisibility relation:

\textbf{Proposition 2.5. (a)} We have \(a \mid a\) for every \(a \in \mathbb{Z}\). (This is called the \textit{reflexivity} of divisibility.)

\(\text{(b)}\) If \(a, b, c \in \mathbb{Z}\) satisfy \(a \mid b\) and \(b \mid c\), then \(a \mid c\). (This is called the \textit{transitivity} of divisibility.)

\(\text{(c)}\) If \(a_1, a_2, b_1, b_2 \in \mathbb{Z}\) satisfy \(a_1 \mid b_1\) and \(a_2 \mid b_2\), then \(a_1a_2 \mid b_1b_2\).

\textbf{Proof.} (a) Let \(a \in \mathbb{Z}\). Then, there exists an integer \(c\) such that \(a = ac\) (namely, \(c = 1\)). In other words, \(a \mid a\) (by Definition 2.2). This proves Proposition 2.5 (a).

\(\text{(b)}\) Let \(a, b, c \in \mathbb{Z}\) satisfy \(a \mid b\) and \(b \mid c\).

From \(a \mid b\), we conclude that there exists an integer \(d\) such that \(b = ad\). Consider this \(d\).

From \(b \mid c\), we conclude that there exists an integer \(e\) such that \(c = be\). Consider this \(e\).

We have \(c = b e = ade\). Hence, there exists an integer \(f\) such that \(c = af\) (namely, \(f = de\)). In other words, \(a \mid c\) (by Definition 2.2). This proves Proposition 2.5 (b).

\(\text{(c)}\) Let \(a_1, a_2, b_1, b_2 \in \mathbb{Z}\) satisfy \(a_1 \mid b_1\) and \(a_2 \mid b_2\).

From \(a_1 \mid b_1\), we conclude that there exists an integer \(d\) such that \(b_1 = a_1d\). Consider this \(d\).

From \(a_2 \mid b_2\), we conclude that there exists an integer \(e\) such that \(b_2 = a_2e\). Consider this \(e\).

We have \(b_1 = a_1d\) \(a_2 e = a_1 a_2 d\). Hence, there exists an integer \(f\) such that \(b_1 b_2 = a_1a_2f\) (namely, \(f = de\)). In other words, \(a_1a_2 \mid b_1b_2\) (by Definition 2.2). This proves Proposition 2.5 (c). \(\square\)

\textbf{Exercise 2.1.} Let \(a\) and \(b\) be two integers such that \(a \mid b\) and \(b \mid a\). Prove that \(|a| = |b|\).

\textbf{Exercise 2.2.} Let \(a, b, c\) be three integers such that \(c \neq 0\). Prove that \(a \mid b\) holds if and only if \(ac \mid bc\).

\textit{Solution to Exercise 2.2} \(\implies\): Assume that \(a \mid b\) holds. We must prove that \(ac \mid bc\).

It is easy to do this straight from the definition of divisibility, but here is a shorter argument: Proposition 2.5 (a) (applied to \(c\) instead of \(a\)) yields \(c \mid c\). Also, \(a \mid b\). Hence, Proposition 2.5 (c) (applied to \(a_1 = a, b_1 = b, a_2 = c\) and \(b_2 = c\)) yields \(ac \mid bc\). This proves the “\(\implies\)” direction of Exercise 2.2.

\(\impliedby\): Assume that \(ac \mid bc\) holds. We must prove that \(a \mid b\).
We have $ac \mid bc$. In other words, there exists an integer $d$ such that $bc = (ac)d$ (by Definition 2.2). Consider this $d$. We have $bc = (ac)d = ad$. We can divide both sides of this equality by $c$ (since $c \neq 0$), and thus obtain $b = ad$. Thus, there exists an integer $e$ such that $b = ae$ (namely, $e = d$). In other words, $a \mid b$ (by Definition 2.2). This proves the “$\iff$” direction of Exercise 2.2.

\[\square\]

### 2.3. Congruence modulo $n$

The next definition is simple but crucial:

**Definition 2.6.** Let $n, a, b \in \mathbb{Z}$. We say that $a$ is congruent to $b$ modulo $n$ (or, short, “$a \equiv b \text{ mod } n$”) if and only if $n \mid a - b$.

**Example 2.7.** (a) Is $3 \equiv 7 \text{ mod } 2$? Yes, since $2 \mid 3 - 7 = -4$.

(b) Is $3 \equiv 6 \text{ mod } 2$? No, since $2 \nmid 3 - 6 = -3$.

Now, let $a$ and $b$ be two integers.

(c) We have $a \equiv b \text{ mod } 0$ if and only if $a = b$. (Indeed, $a \equiv b \text{ mod } 0$ is defined to mean $0 \mid a - b$, but the latter divisibility happens only when $a - b = 0$, which is tantamount to saying $a = b$.)

(d) We have $a \equiv b \text{ mod } 1$ always, since $1 \mid a - b$ always holds (remember: $1$ divides everything).

Note that being congruent modulo 2 means having the same parity: i.e., two even numbers will be congruent modulo 2, and two odd numbers will be, but an even number will never be congruent to an odd number modulo 2. (To be rigorous: This is not quite obvious at this point yet; but it will be easy once we have properly introduced division with remainder. See Exercise 2.10 (i) below for the proof.)

The word “modulo” in the phrase “$a$ is congruent to $b$ modulo $n$” is due to Gauss and means something like “with respect to”. You should think of “$a$ is congruent to $b$ modulo $n$” as a relation between all three of the numbers $a$, $b$ and $n$, but $a$ and $b$ are the “main characters” and $n$ sets the scenery.

**Exercise 2.3.** Let $a, b \in \mathbb{Z}$. Prove that $a + b \equiv a - b \text{ mod } 2$.

**Solution to Exercise 2.3** According to Definition 2.6, we have $a + b \equiv a - b \text{ mod } 2$ if and only if $2 \mid (a + b) - (a - b)$. Thus, it remains to prove that $2 \mid (a + b) - (a - b)$. But this follows immediately from $(a + b) - (a - b) = 2b$. Thus Exercise 2.3 is solved. 

**Proposition 2.8.** Let $n \in \mathbb{Z}$.

(a) We have $a \equiv a \text{ mod } n$ for every $a \in \mathbb{Z}$.

(b) If $a, b, c \in \mathbb{Z}$ satisfy $a \equiv b \text{ mod } n$ and $b \equiv c \text{ mod } n$, then $a \equiv c \text{ mod } n$.

(c) If $a, b \in \mathbb{Z}$ satisfy $a \equiv b \text{ mod } n$, then $b \equiv a \text{ mod } n$.

(d) If $a_1, a_2, b_1, b_2 \in \mathbb{Z}$ satisfy $a_1 \equiv b_1 \text{ mod } n$ and $a_2 \equiv b_2 \text{ mod } n$, then
\[
\begin{align*}
  a_1 + a_2 &\equiv b_1 + b_2 \text{ mod } n; \\
  a_1 - a_2 &\equiv b_1 - b_2 \text{ mod } n; \\
  a_1 a_2 &\equiv b_1 b_2 \text{ mod } n.
\end{align*}
\]

\[\square\]
(e) Let \( m \in \mathbb{Z} \) be such that \( m \mid n \). If \( a, b \in \mathbb{Z} \) satisfy \( a \equiv b \mod n \), then \( a \equiv b \mod m \).

Proof. (a) Let \( a \in \mathbb{Z} \). Recall that \( a \equiv a \mod n \) is defined to mean \( n \mid a - a \). Since \( n \mid a - a \) holds (because \( a - a = 0 = n \cdot 0 \)), we thus see that \( a \equiv a \mod n \) holds. This proves Proposition 2.8 (a).

(b) Let \( a, b, c \in \mathbb{Z} \) satisfy \( a \equiv b \mod n \) and \( b \equiv c \mod n \).

We have \( a \equiv b \mod n \). In other words, \( n \mid a - b \) (by Definition 2.6). In other words, there exists an integer \( p \) such that \( a - b = np \) (by Definition 1.2). Consider this \( p \).

We have \( b \equiv c \mod n \). In other words, \( n \mid b - c \) (by Definition 2.6). In other words, there exists an integer \( q \) such that \( b - c = nq \) (by Definition 1.2). Consider this \( q \).

Now,
\[
(a - c) = (a - b) + (b - c) = n(p + q).
\]
Hence, there exists an integer \( r \) such that \( a - c = nr \) (namely, \( r = p + q \)). In other words, \( n \mid a - c \) (by Definition 1.2). In other words, \( a \equiv c \mod n \) (by Definition 2.6). This proves Proposition 2.8 (b).

(c) Let \( a, b \in \mathbb{Z} \) satisfy \( a \equiv b \mod n \).

We have \( a \equiv b \mod n \). In other words, \( n \mid a - b \) (by Definition 2.6). In other words, there exists an integer \( p \) such that \( a - b = np \) (by Definition 1.2). Consider this \( p \). Now,
\[
(b - a) = -(a - b) = -np = n(-p).
\]
Hence, there exists an integer \( c \) such that \( b - a = nc \) (namely, \( c = -p \)). In other words, \( n \mid b - a \) (by Definition 1.2). In other words, \( b \equiv a \mod n \) (by Definition 2.6). This proves Proposition 2.8 (c).

(d) Let \( a_1, a_2, b_1, b_2 \in \mathbb{Z} \) satisfy \( a_1 \equiv b_1 \mod n \) and \( a_2 \equiv b_2 \mod n \).

We have \( a_1 \equiv b_1 \mod n \). In other words, \( n \mid a_1 - b_1 \) (by Definition 2.6). In other words, there exists an integer \( p \) such that \( a_1 - b_1 = np \) (by Definition 1.2). Consider this \( p \).

We have \( a_2 \equiv b_2 \mod n \). In other words, \( n \mid a_2 - b_2 \) (by Definition 2.6). In other words, there exists an integer \( q \) such that \( a_2 - b_2 = nq \) (by Definition 1.2). Consider this \( q \).

We have
\[
(a_1 + a_2) - (b_1 + b_2) = (a_1 - b_1) + (a_2 - b_2) = np + nq = n(p + q).
\]
Hence, there exists an integer \( c \) such that \( (a_1 + a_2) - (b_1 + b_2) = nc \) (namely, \( c = p + q \)). In other words, \( n \mid (a_1 + a_2) - (b_1 + b_2) \) (by Definition 1.2). In other words, \( a_1 + a_2 \equiv b_1 + b_2 \mod n \) (by Definition 2.6). A similar argument (using \( p - q \) instead
of \( p + q \) shows that \( a_1 - a_2 \equiv b_1 - b_2 \mod n \). It thus remains to show that \( a_1 a_2 \equiv b_1 b_2 \mod n \).

Let us first show that \( a_1 a_2 \equiv a_1 b_2 \mod n \). Indeed, \( a_1 a_2 - a_1 b_2 = a_1 (a_2 - b_2) = a_1 n q = n (a_1 q) \). Hence, there exists an integer \( c \) such that \( a_1 a_2 - a_1 b_2 = nc \) (namely, \( c = a_1 q \)). In other words, \( n | a_1 a_2 - a_1 b_2 \) (by Definition 2.2). In other words, \( a_1 a_2 \equiv a_1 b_2 \mod n \) (by Definition 2.6).

Next, let us show that \( a_1 b_2 \equiv b_1 b_2 \mod n \). Indeed, \( a_1 b_2 - b_1 b_2 = b_2 (a_1 - a_2) = b_2 n p = n (b_2 p) \). Hence, there exists an integer \( c \) such that \( a_1 b_2 - b_1 b_2 = nc \) (namely, \( c = b_2 p \)). In other words, \( n | a_1 b_2 - b_1 b_2 \) (by Definition 2.2). In other words, \( a_1 b_2 \equiv b_1 b_2 \mod n \) (by Definition 2.6).

From \( a_1 a_2 \equiv a_1 b_2 \mod n \) and \( a_1 b_2 \equiv b_1 b_2 \mod n \), we now conclude that \( a_1 a_2 \equiv b_1 b_2 \mod n \) (by Proposition 2.8(c), applied to \( a = a_1 a_2 \), \( b = a_1 b_2 \) and \( c = b_1 b_2 \)). This completes the proof of Proposition 2.8(d).

(e) Let \( a, b \in \mathbb{Z} \) satisfy \( a \equiv b \mod n \). We have \( a \equiv b \mod n \). In other words, \( n | a - b \) (by Definition 2.6). From \( m | n \) and \( n | a - b \), we obtain \( m | a - b \) (by Proposition 2.5(b), applied to \( m, n \) and \( a - b \) instead of \( a, b \) and \( c \)). In other words, \( a \equiv b \mod m \) (by Definition 2.6). This proves Proposition 2.8(e).

In the above proof, we took care to explicitly cite Definition 2.2 and Definition 2.6 whenever we used them; in the following, we will not be this detailed.

Proposition 2.8(d) is saying that congruences modulo \( n \) (for a fixed integer \( n \)) can be added, subtracted and multiplied together. This does not mean that you can do everything with them that you can do with equalities. The next exercise shows that dividing congruences and taking a congruence to the power of another does not generally work:

**Exercise 2.4.** Let \( n, a_1, a_2, b_1, b_2 \in \mathbb{Z} \) satisfy \( a_1 \equiv b_1 \mod n \) and \( a_2 \equiv b_2 \mod n \). Then, in general, neither \( a_1 / a_2 \equiv b_1 / b_2 \mod n \) nor \( a_1^{a_2} \equiv b_1^{b_2} \mod n \) is necessarily true. Of course, this is partly due to the fact that \( a_1 / a_2, b_1 / b_2 \) and \( a_1^{a_2} \) and \( b_1^{b_2} \) are not always integers in the first place (and being congruent modulo \( n \) only makes sense for integers, at least for now). But even when \( a_1 / a_2, b_1 / b_2 \) and \( a_1^{a_2} \) and \( b_1^{b_2} \) are integers, the congruences \( a_1 / a_2 \equiv b_1 / b_2 \mod n \) nor \( a_1^{a_2} \equiv b_1^{b_2} \mod n \) are often false. Find examples of \( n, a_1, a_2, b_1, b_2 \) such that \( a_1 / a_2, b_1 / b_2 \) and \( a_1^{a_2} \) and \( b_1^{b_2} \) are integers but the congruences \( a_1 / a_2 \equiv b_1 / b_2 \mod n \) nor \( a_1^{a_2} \equiv b_1^{b_2} \mod n \) are false.

**Solution to Exercise 2.4** There are many such examples. Here is one:

\[
\begin{align*}
n &= 8, & a_1 &= 10, & a_2 &= 2, & b_1 &= 10, & b_2 &= 10.
\end{align*}
\]

These satisfy \( a_1 \equiv b_1 \mod n \) and \( a_2 \equiv b_2 \mod n \), but neither \( a_1 / a_2 \equiv b_1 / b_2 \mod n \) nor \( a_1^{a_2} \equiv b_1^{b_2} \mod n \).
It is much easier to find examples which fail only one of the two congruences \( a_1/a_2 \equiv b_1/b_2 \mod n \) and \( a_1^{b_2} \equiv b_1^{b_2} \mod n \).

However, we can divide a congruence \( a \equiv b \mod n \) by a nonzero integer \( d \) when all of \( a, b, n \) are divisible by \( d \):

**Exercise 2.5.** Let \( n, d, a, b \in \mathbb{Z} \), and assume that \( d \neq 0 \). Assume that \( d \) divides each of \( a, b, n \), and assume that \( a \equiv b \mod n \). Prove that \( a/d \equiv b/d \mod n/d \).

*Solution to Exercise 2.5* We have \( a \equiv b \mod n \). In other words, \( n \mid a - b \) (by the definition of congruence). Note that all of \( a/d, b/d \) and \( n/d \) are integers (since \( d \) divides each of \( a, b, n \)). Hence, \((a - b)/d = a/d - b/d\) is an integer as well. Hence, Exercise 2.2 (applied to \( n/d \), \((a - b)/d \) and \( d \) instead of \( a, b \) and \( c \)) shows that \( n/d \mid (a - b)/d \) holds if and only if \((n/d)d \mid ((a - b)/d)d \). Since \((n/d)d \mid ((a - b)/d)d \) holds (indeed, this is just a complicated way to say \( n \mid a - b \)), we thus conclude that \( n/d \mid (a - b)/d \) holds. In other words, \( n/d \mid a/d - b/d \) (since \((a - b)/d = a/d - b/d \)). In other words, \( a/d \equiv b/d \mod n/d \) (by the definition of congruence). This solves Exercise 2.5.

We can also take a congruence to the \( k \)-th power when \( k \in \mathbb{N} \):

**Exercise 2.6.** Let \( n, a, b \in \mathbb{Z} \) be such that \( a \equiv b \mod n \). Prove that \( a^k \equiv b^k \mod n \) for each \( k \in \mathbb{N} \).

(Note that the “\( n \)” is not being taken to the \( k \)-th power here.)

*First solution to Exercise 2.6* We want to prove that

\[
a^k \equiv b^k \mod n \quad \text{for each } k \in \mathbb{N}.
\]  

(7)

We shall prove this by induction on \( k \):

*Induction base:* Proposition 2.8 (a) yields \( 1 \equiv 1 \mod n \). In view of \( a^0 = 1 \) and \( b^0 = 1 \), this rewrites as \( a^0 \equiv b^0 \mod n \). In other words, (7) holds for \( k = 0 \). This completes the induction base.

*Induction step:* Let \( \ell \in \mathbb{N} \). Assume that (7) holds for \( k = \ell \). We must prove that (7) holds for \( k = \ell + 1 \).

We have assumed that (7) holds for \( k = \ell \). In other words, we have \( a^\ell \equiv b^\ell \mod n \). Also, recall that \( a \equiv b \mod n \). Hence, (6) (applied to \( c = a^\ell \) and \( d = b^\ell \)) yields \( a\ell = b\ell \mod n \). In other words, \( a\ell + 1 \equiv b\ell + 1 \mod n \) (since \( a\ell = a\ell + 1 \) and \( b\ell = b\ell + 1 \)). In other words, (7) holds for \( k = \ell + 1 \). This completes the induction step. Thus, (7) is proven by induction. Therefore, Exercise 2.6 is solved.

*Second solution to Exercise 2.6* Recall that

\[
(a - b) \left( a^{k-1} + a^{k-2}b + a^{k-3}b^2 + \cdots + ab^{k-2} + b^{k-1} \right) = a^k - b^k
\]  

(8)

for every \( k \in \mathbb{N} \). (This is a well-known identity, and it appears (with \( k \) renamed as \( n \)) as the first half of Exercise 1 on homework set #0.)
Now, let \( k \in \mathbb{N} \). We have assumed that \( a \equiv b \mod n \). In other words, \( n \mid a - b \). In other words, there exists an integer \( c \) such that \( a - b = nc \). Consider this \( c \). Now, (8) yields

\[
a^k - b^k = (a - b) \left( a^{k-1} + a^{k-2}b + a^{k-3}b^2 + \cdots + ab^{k-2} + b^{k-1} \right)
\]

The right hand side of this equality is clearly divisible by \( n \). Hence, so is the left hand side. In other words, \( n \mid a^k - b^k \). In other words, \( a^k \equiv b^k \mod n \). Hence, Exercise 2.6 is solved again. □

We can add not just two, but any number of congruences (where “number” means “finite number”):

**Exercise 2.7.** Let \( n \) be an integer. Let \( S \) be a finite set. For each \( s \in S \), let \( a_s \) and \( b_s \) be two integers. Assume that

\[
a_s \equiv b_s \mod n \quad \text{for each} \quad s \in S. \tag{9}
\]

(a) Prove that

\[
\sum_{s \in S} a_s \equiv \sum_{s \in S} b_s \mod n. \tag{10}
\]

(b) Prove that

\[
\prod_{s \in S} a_s \equiv \prod_{s \in S} b_s \mod n. \tag{11}
\]

(Keep in mind that if the set \( S \) is empty, then \( \sum_{s \in S} a_s = \sum_{s \in S} b_s = 0 \) and \( \prod_{s \in S} a_s = \prod_{s \in S} b_s = 1 \); this holds by the definition of empty sums and of empty products.)

**Solution to Exercise 2.7** (a) We shall solve Exercise 2.7(a) by induction on \( |S| \):

*Induction base:* Exercise 2.7(a) holds whenever \( |S| = 0 \). This completes the induction base.

*Induction step:* Fix \( k \in \mathbb{N} \). Assume that Exercise 2.7(a) holds whenever \( |S| = k \). We must prove that Exercise 2.7(a) holds whenever \( |S| = k + 1 \).

We have assumed that Exercise 2.7(a) holds whenever \( |S| = k \). In other words, the following statement is true:

Statement 1: Let \( n, S, a_s \) and \( b_s \) be as in Exercise 2.7. Assume that \( |S| = k \). Then,

\[
\sum_{s \in S} a_s \equiv \sum_{s \in S} b_s \mod n.
\]

Proof. Let \( n, S, a_s \) and \( b_s \) be as in Exercise 2.7 and assume that \( |S| = 0 \). Then, the set \( S \) is empty (since \( |S| = 0 \)), and thus we have \( \sum_{s \in S} a_s = (\text{empty sum}) = 0 \). Similarly, \( \sum_{s \in S} b_s = 0 \). Now, Proposition 2.8(a) yields \( 0 \equiv 0 \mod n \). In view of \( \sum_{s \in S} a_s = 0 \) and \( \sum_{s \in S} b_s = 0 \), this rewrites as \( \sum_{s \in S} a_s \equiv \sum_{s \in S} b_s \mod n \). Thus, Exercise 2.7(a) holds in our case.

So we have shown that Exercise 2.7(a) holds whenever \( |S| = 0 \).
Now, we must prove that Exercise 2.7 (a) holds whenever $|S| = k + 1$. In other words, we must prove the following statement:

**Statement 2:** Let $n, S, a_s$ and $b_s$ be as in Exercise 2.7. Assume that $|S| = k + 1$.
Then, $\sum_{s \in S} a_s \equiv \sum_{s \in S} b_s \mod n$.

([Proof of Statement 2:] We have $|S| = k + 1 > k \geq 0$; thus, the set $S$ is nonempty. Hence, there exists some $t \in S$. Pick such a $t$. Thus, $|S \setminus \{t\}| = |S| - 1 = k$ (since $|S| = k + 1$). Moreover, from (9), we immediately obtain that $a_s \equiv b_s \mod n$ (since each $s \in S \setminus \{t\}$ belongs to $S$). Hence, we can apply Statement 1 to $S \setminus \{t\}$ instead of $S$. We thus obtain $\sum_{s \in S \setminus \{t\}} a_s \equiv \sum_{s \in S \setminus \{t\}} b_s \mod n$.

Also, we have $a_t \equiv b_t \mod n$ (by (9), applied to $s = t$). Adding these two congruences together, we obtain $\sum_{s \in S \setminus \{t\}} a_s + a_t \equiv \sum_{s \in S \setminus \{t\}} b_s + b_t \mod n$.

In view of $\sum_{s \in S} a_s = \sum_{s \in S \setminus \{t\}} a_s + a_t$ (here, we have split off the addend for $s = t$ from the sum) and $\sum_{s \in S} b_s = \sum_{s \in S \setminus \{t\}} b_s + b_t$ (here, we have split off the addend for $s = t$ from the sum), this can be rewritten as $\sum_{s \in S} a_s \equiv \sum_{s \in S} b_s \mod n$.

This proves Statement 2.)

We have now proven Statement 2; this means that Exercise 2.7 (a) holds whenever $|S| = k + 1$. This completes the induction step; thus, Exercise 2.7 (a) is solved.

(b) The solution to Exercise 2.7 (b) is analogous to the one we gave above for Exercise 2.7 (a); the main difference is that we have to replace sums by products (and 0 by 1).

---

**Exercise 2.8.** Is it true that if $a_1, a_2, b_1, b_2, n_1, n_2 \in \mathbb{Z}$ satisfy $a_1 \equiv b_1 \mod n_1$ and $a_2 \equiv b_2 \mod n_2$, then $a_1 a_2 \equiv b_1 b_2 \mod n_1 n_2$?

**Solution to Exercise 2.8** No, it is not true. For example, $a_1 = 1$, $a_2 = 1$, $b_1 = 1$, $b_2 = 0$, $n_1 = 0$ and $n_2 = 1$ yield a counterexample.
2.4. Chains of congruences

For this whole Section 2.4, we fix an integer $n$.

Chains of equalities are a fundamental piece of notation used throughout mathematics. For example, here is a chain of equalities:

\[
\begin{align*}
(ad + bc)^2 + (ac - bd)^2 &= (ad)^2 + 2ad \cdot bc + (bc)^2 + (ac)^2 - 2ac \cdot bd + (bd)^2 \\
&= a^2d^2 + 2abcd + b^2c^2 + a^2c^2 - 2abcd + b^2d^2 \\
&= a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2 \\
&= (a^2 + b^2)(c^2 + d^2)
\end{align*}
\]

(where $a, b, c, d$ are arbitrary numbers). This chain proves the equality (1). But why does it really? If we look closely at this chain of equalities, we see that it has the form 

\[
A = B = C = D = E
\]

and so on. This kind of statement is called a “chain of equalities”, and, a priori, it simply means that any two adjacent numbers in this chain are equal: $A = B$ and $B = C$ and $C = D$ and $D = E$. Without as much as noticing it, we have concluded that any two numbers in this chain are equal; thus, in particular, $A = E$, which is precisely the equality (1) we wanted to prove.

That this kind of “chaining” is possible is one of the most basic facts in mathematics. Let us define a chain of equalities formally:

**Definition 2.9.** If $a_1, a_2, \ldots, a_k$ are $k$ objects, then the statement 

\[
A = B = C = \cdots = E
\]

shall mean that $a_i = a_{i+1}$ holds for each \(i \in \{1, 2, \ldots, k-1\} \).

(In other words, it shall mean that $a_1 = a_2$ and $a_2 = a_3$ and $a_3 = a_4$ and $a_4 = a_5$ and $a_5 = a_6$. This is vacuously true when $k \leq 1$. If $k = 2$, then it simply means that $a_1 = a_2$.)

Such a statement will be called a chain of equalities.

**Proposition 2.10.** Let $a_1, a_2, \ldots, a_k$ be $k$ objects such that $a_1 = a_2 = \cdots = a_k$. Let $u$ and $v$ be two elements of \{1, 2, ..., k\}. Then, $a_u = a_v$.

So we have defined a chain of equalities to be true if and only if any two adjacent terms in this chain are equal (i.e., if “each equality sign in the chain is satisfied”). Proposition 2.10 shows that in such a chain, any two terms are equal. This is

---

8"Objects" can be numbers, sets, tuples or any other well-defined things in mathematics.
intuitively rather clear, but can also be formally proven by induction using the basic properties of equality (transitivity\(^9\), reflexivity\(^{10}\) and symmetry\(^{11}\)).

But our goal is to understand basic number theory, not to scrutinize the foundations of mathematics. So let us recall that we have fixed an integer \(n\), and consider congruences modulo \(n\). We claim that these can be chained just as equalities:

**Definition 2.11.** If \(a_1, a_2, \ldots, a_k\) are \(k\) integers, then the statement “\(a_1 \equiv a_2 \equiv \cdots \equiv a_k \mod n\)” shall mean that

\[
    a_i \equiv a_{i+1} \mod n \text{ holds for each } i \in \{1, 2, \ldots, k-1\}.
\]

(In other words, it shall mean that \(a_1 \equiv a_2 \mod n\) and \(a_2 \equiv a_3 \mod n\) and \(a_3 \equiv a_4 \mod n\) and \(\cdots\) and \(a_{k-1} \equiv a_k \mod n\). This is vacuously true when \(k \leq 1\). If \(k = 2\), then it simply means that \(a_1 \equiv a_2 \mod n\).)

Such a statement will be called a chain of congruences modulo \(n\).

**Proposition 2.12.** Let \(a_1, a_2, \ldots, a_k\) be \(k\) integers such that \(a_1 \equiv a_2 \equiv \cdots \equiv a_k \mod n\). Let \(u\) and \(v\) be two elements of \(\{1, 2, \ldots, k\}\). Then, \(a_u \equiv a_v \mod n\).

Proposition 2.12 shows that any two terms in a chain of congruences modulo \(n\) must be congruent to each other modulo \(n\). Again, this can be formally proven by induction; see [Grinbe15, proof of Proposition 2.16]. The ingredients of the proof are basic properties of congruence modulo \(n\): transitivity, reflexivity and symmetry. These are fancy names for parts (b), (a) and (c) of Proposition 2.8.

We will use Proposition 2.12 tacitly (just as you would use Proposition 2.10): i.e., every time we prove a chain of congruences like \(a_1 \equiv a_2 \equiv \cdots \equiv a_k \mod n\), we assume that the reader will automatically conclude that any two of its terms are congruent to each other modulo \(n\) (and will remember this conclusion). For instance, if we show that \(1 \equiv 4 \equiv 34 \equiv 334 \equiv 304 \mod 3\), then we automatically get the congruences \(1 \equiv 304 \mod 3\) and \(334 \equiv 1 \mod 3\) and \(4 \equiv 334 \mod 3\) and several others out of this chain.

Chains of congruences can also include equality signs. For example, if \(a, b, c, d\) are integers, then “\(a \equiv b \equiv c \equiv d \mod n\)” means that \(a \equiv b \mod n\) and \(b = c\) and \(c \equiv d \mod n\). Such a chain is still a chain of congruences, because \(b = c\) implies \(b \equiv c \mod n\) (by Proposition 2.8 (a)).

Just as there are chains of equalities and chains of congruences, there are chains of divisibilities:

**Definition 2.13.** If \(a_1, a_2, \ldots, a_k\) are \(k\) integers, then the statement “\(a_1 | a_2 | \cdots | a_k\)” shall mean that

\[
    a_i | a_{i+1} \text{ holds for each } i \in \{1, 2, \ldots, k-1\}.
\]

\(^9\)Transitivity of equality says that if \(a, b, c\) are three objects satisfying \(a = b\) and \(b = c\), then \(a = c\).

\(^{10}\)Reflexivity of equality says that every object \(a\) satisfies \(a = a\).

\(^{11}\)Symmetry of equality says that if \(a, b\) are two objects satisfying \(a = b\), then \(b = a\).
(In other words, it shall mean that \( a_1 \mid a_2 \) and \( a_2 \mid a_3 \) and \( a_3 \mid a_4 \) and \( \cdots \) and \( a_{k-1} \mid a_k \). This is vacuously true when \( k \leq 1 \). If \( k = 2 \), then it simply means that \( a_1 \mid a_2 \).)

Such a statement will be called a chain of divisibilities.

**Proposition 2.14.** Let \( a_1, a_2, \ldots, a_k \) be \( k \) integers such that \( a_1 \mid a_2 \mid \cdots \mid a_k \). Let \( u \) and \( v \) be two elements of \( \{1, 2, \ldots, k\} \) such that \( u \leq v \). Then, \( a_u \mid a_v \).

Note that we had to require \( u \leq v \) in this proposition, unlike the analogous propositions for chains of equalities and chains of congruences, because there is no “symmetry of divisibility” (i.e., if \( a \mid b \), then we don’t generally have \( b \mid a \)). The proof of Proposition 2.14 relies on the reflexivity of divisibility (Proposition 2.5 (a)) and on the transitivity of divisibility (Proposition 2.5 (b)).

Again, chains of divisibilities can include equality signs.

---

### 2.5. Division with remainder

The following fact you likely remember from high school:

**Theorem 2.15.** Let \( n \) be a positive integer. Let \( u \in \mathbb{Z} \). Then, there exists a unique pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\} \) such that \( u = qn + r \).

Before we prove this theorem, let us introduce the notations that it justifies:

**Definition 2.16.** Let \( n \) be a positive integer. Let \( u \in \mathbb{Z} \). Theorem 2.15 shows that there exists a unique pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\} \) such that \( u = qn + r \). Consider this pair.

(a) We denote the integer \( q \) by \( u \div n \), and refer to it as the quotient of the division of \( u \) by \( n \).

(b) We denote the integer \( r \) by \( u \mod n \), and refer to it as the remainder of the division of \( u \) by \( n \).

The words “quotient” and “remainder” are standard, but the notations “\( u \div n \)” and “\( u \mod n \)” are not (I have taken them from the Python programming language); be prepared to see other notations in the literature (e.g., the notations “\( \text{quo}(u, n) \)” and “\( \text{rem}(u, n) \)” for \( u \div n \) and \( u \mod n \), respectively).

**Example 2.17.** (a) We have \( 14 \div 3 = 4 \) and \( 14 \mod 3 = 2 \), because \( (4, 2) \) is the unique pair \((q, r) \in \mathbb{Z} \times \{0, 1, 2\} \) satisfying \( 14 = q \cdot 3 + r \).

(b) We have \( 18 \div 3 = 6 \) and \( 18 \mod 3 = 0 \), because \( (6, 0) \) is the unique pair \((q, r) \in \mathbb{Z} \times \{0, 1, 2\} \) satisfying \( 18 = q \cdot 3 + r \).

(c) We have \( (-2) \div 3 = -1 \) and \( (-2) \mod 3 = 1 \), because \( (-1, 1) \) is the unique pair \((q, r) \in \mathbb{Z} \times \{0, 1, 2\} \) satisfying \( -2 = q \cdot 3 + r \).

(d) For each \( u \in \mathbb{Z} \), we have \( u \div 1 = u \) and \( u \mod 1 = 0 \), because \( (u, 0) \) is the unique pair \((q, r) \in \mathbb{Z} \times \{0\} \) satisfying \( u = q \cdot 1 + r \).
But we have gotten ahead of ourselves: We need to prove Theorem 2.15 before we can use the notations “u / v” and “u % v”.

Let us split Theorem 2.15 into two parts: existence and uniqueness:

**Lemma 2.18.** Let $n$ be a positive integer. Let $u \in \mathbb{Z}$. Then, there exists at least one pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}$ such that $u = qn + r$.

**Lemma 2.19.** Let $n$ be a positive integer. Let $u \in \mathbb{Z}$. Then, there exists at most one pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}$ such that $u = qn + r$.

We begin by proving Lemma 2.19 (which is the easier one):

**Proof of Lemma 2.19** Let $(q_1, r_1)$ and $(q_2, r_2)$ be two pairs $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}$ such that $u = qn + r$. We shall show that $(q_1, r_1) = (q_2, r_2)$.

We know that $(q_1, r_1)$ is a pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}$ such that $u = qn + r$. In other words, $(q_1, r_1) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}$ and $u = q_1n + r_1$. Similarly, $(q_2, r_2) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}$ and $u = q_2n + r_2$.

From $(q_1, r_1) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}$, we obtain $q_1 \in \mathbb{Z}$ and $r_1 \in \{0, 1, \ldots, n - 1\}$. Similarly, $q_2 \in \mathbb{Z}$ and $r_2 \in \{0, 1, \ldots, n - 1\}$. Thus, in particular, $q_1, q_2, r_1, r_2$ are integers.

From $r_1 \in \{0, 1, \ldots, n - 1\}$ and $r_2 \in \{0, 1, \ldots, n - 1\}$, we can easily derive

$$|r_2 - r_1| \leq n - 1.$$  \hspace{1cm} (12)

[Proof of (12): Intuitively, this should be clear: Both $r_1$ and $r_2$ belong to the integer interval $\{0, 1, \ldots, n - 1\}$, and thus the unsigned distance between $r_1$ and $r_2$ is at most $n - 1$ (with the worst case being when $r_1$ and $r_2$ are at opposite ends of this interval).

Here is a formal restatement of this argument: We have $r_1 \in \{0, 1, \ldots, n - 1\}$, thus $r_1 \geq 0$.

Also, $r_2 \in \{0, 1, \ldots, n - 1\}$, hence $r_2 \leq n - 1$. Hence, $r_2 - r_1 \leq (n - 1) - 0 = n - 1$. Similarly, $r_1 - r_2 \leq n - 1$. But recall that $|x| \in \{x, -x\}$ for each $x \in \mathbb{Z}$. Applying this to $x = r_2 - r_1$, we obtain

$$|r_2 - r_1| \in \begin{cases} r_2 - r_1, & \text{if } r_2 - r_1 \geq 0 \\ r_1 - r_2, & \text{if } r_2 - r_1 < 0 \end{cases} = \{r_2 - r_1, r_1 - r_2\}.$$

In other words, $|r_2 - r_1|$ is one of the two numbers $r_2 - r_1$ and $r_1 - r_2$. Since both of these numbers $r_2 - r_1$ and $r_1 - r_2$ are $\leq n - 1$ (as we have just shown), we thus conclude that $|r_2 - r_1| \leq n - 1$. This proves (12).]

We have $q_1n + r_1 = u = q_2n + r_2$, thus $q_1n - q_2n = r_2 - r_1$. Hence,

$$r_2 - r_1 = q_1n - q_2n = (q_1 - q_2)n.$$  \hspace{1cm} (13)

Assume (for the sake of contradiction) that $q_1 \neq q_2$. Thus, $q_1 - q_2 \neq 0$, so that $|q_1 - q_2| > 0$ and therefore $|q_1 - q_2| \geq 1$ (since $|q_1 - q_2|$ is an integer). We can
multiply this inequality by $n$ (since $n$ is positive) and thus obtain $|q_1 - q_2| n \geq 1n = n$. But from (13), we obtain

$$|r_2 - r_1| = |(q_1 - q_2) n| = |q_1 - q_2| \cdot \frac{|n|}{n} \quad \text{(by (3))}$$

(since $n$ is positive)

$$= |q_1 - q_2| n \geq n > n - 1.$$ This contradicts (12). This contradiction shows that our assumption (that $q_1 \neq q_2$) was false. Hence, we have $q_1 = q_2$. Thus, $q_1 - q_2 = 0$, so that (13) becomes $r_2 - r_1 = (q_1 - q_2) n = 0$ and thus $r_2 = r_1$, so that $r_1 = r_2$. Combining this with $q_1 = q_2$, we obtain $(q_1, r_1) = (q_2, r_2)$.

Now, forget that we have fixed $(q_1, r_1)$ and $(q_2, r_2)$. We thus have proven that if $(q_1, r_1)$ and $(q_2, r_2)$ are two pairs $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n-1\}$ such that $u = qn + r$, then $(q_1, r_1) = (q_2, r_2)$. In other words, any two pairs $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n-1\}$ such that $u = qn + r$ must be equal. In other words, there exists at most one such pair. This proves Lemma 2.19.

But we also need to prove Lemma 2.18. This lemma can be proven by induction on $u$, but not without some complications: Since it is stated for all integers $u$ (rather than just for nonnegative or positive integers), the classical induction principle (with an induction base and a “$u$ to $u + 1$” step) cannot prove it directly. Instead, we have to either add a “$u$ to $u - 1$” step to our induction (resulting in a “two-sided induction” or “up- and down-induction” argument), or to treat the case of negative $u$ separately. A proof using the first of these two methods can be found in [Grinbe15] (where $n$ and $u$ are denoted by $N$ and $n$). We shall instead give a proof using the second method; thus, we first state the particular case of Lemma 2.18 when $u$ is nonnegative:

**Lemma 2.20.** Let $n$ be a positive integer. Let $u \in \mathbb{N}$. Then, there exists at least one pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n-1\}$ such that $u = qn + r$.

This lemma can be proven by induction on $u$ as in [Grinbe15] [proof of Proposition 2.150]. Let us instead prove this by strong induction on $u$. See [Grinbe15, §2.8] for an introduction to strong induction; in particular, recall that a strong induction needs no induction base (but often contains a case distinction in its “induction step” that, in some way, does give the first few values a special treatment). The proof of Lemma 2.20 that we give below follows a stupid but valid method of finding the pair $(q, r)$: Keep subtracting $n$ from $u$ until $u$ becomes $< n$; then $r$ will be the resulting number, whereas $q$ will be the number of times you have subtracted $n$.

**Proof of Lemma 2.20.** We proceed by strong induction on $u$.

Let $U \in \mathbb{N}$. Assume (as the induction hypothesis) that Lemma 2.20 holds for every $u \in \mathbb{N}$ satisfying $u < U$. We must prove that Lemma 2.20 also holds for
\( u = U \). In other words, we must prove that there exists at least one pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) such that \( U = qn + r \).

We are in one of the following two cases:

Case 1: We have \( U < n \).

Case 2: We have \( U \geq n \).

Let us first consider Case 1. In this case, we have \( U < n \). Thus, \( U \leq n - 1 \) (since \( U \) and \( n \) are integers), so that \( U \in \{0, 1, \ldots, n - 1\} \) (since \( U \in \mathbb{N} \)). Combining this with \( 0 \in \mathbb{Z} \), we obtain \((0, U) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\). Hence, \((0, U)\) is a pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) such that \( U = qn + r \) (since \( U = 0n + U \)). Thus, there exists at least one pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) such that \( U = qn + r \) (namely, \((q, r) = (0, U)\)).

Let us now consider Case 2. In this case, we have \( U \geq n \). Hence, \( U - n \geq 0 \), so that \( U - n \in \mathbb{N} \) (remember that \( \mathbb{N} = \{0, 1, 2, \ldots\} \)). Also, \( U - n < U \) (since \( n \) is positive). But our induction hypothesis said that Lemma 2.20 holds for every \( u \in \mathbb{N} \) satisfying \( u < U \). Hence, in particular, Lemma 2.20 holds for \( u = U - n \) (since \( U - n \in \mathbb{N} \) and \( U - n < U \)). In other words, there exists at least one pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) such that \( U - n = qn + r \). Fix such a pair and denote it by \((q_0, r_0)\). Thus, \((q_0, r_0) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) and \( U - n = q_0n + r_0 \).

From \( U - n = q_0n + r_0 \), we obtain \( U = n + (q_0n + r_0) = (q_0 + 1)n + r_0 \). Also, from \((q_0, r_0) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\), we obtain \( q_0 \in \mathbb{Z} \) and \( r_0 \in \{0, 1, \ldots, n - 1\}\), and thus \((q_0 + 1, r_0) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\). Thus, the pair \((q_0 + 1, r_0)\) is a pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) such that \( U = qn + r \) (since \( U = (q_0 + 1)n + r_0 \)). Therefore, there exists at least one pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) such that \( U = qn + r \) (namely, \((q, r) = (q_0 + 1, r_0)\)).

Now, in each of the two Cases 1 and 2, we have shown that there exists at least one pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) such that \( U = qn + r \). This completes the induction step; thus, Lemma 2.20 is proven by strong induction.

In order to derive Lemma 2.18 from Lemma 2.20 (that is, to extend Lemma 2.20 to the case of negative \( u \)), we shall need a simple but important trick:

**Lemma 2.21.** Let \( n \) be a positive integer. Let \( u \in \mathbb{Z} \). Then, there exists a \( v \in \mathbb{N} \) such that \( u \equiv v \mod n \).

**Proof of Lemma 2.21.** We are in one of the following two cases:

Case 1: We have \( u \geq 0 \).

Case 2: We have \( u < 0 \).

Let us first consider Case 1. In this case, we have \( u \geq 0 \). Thus, \( u \in \mathbb{N} \). Also, \( u \equiv u \mod n \) (by Proposition 2.3(a)). Thus, there exists a \( v \in \mathbb{N} \) such that \( u \equiv v \mod n \) (namely, \( v = u \)). This proves Lemma 2.21 in Case 1.

Let us now consider Case 2. In this case, we have \( u < 0 \). Hence, \( -u > 0 \). Now, \( u - (n - 1)(-u) = nu \) is divisible by \( n \) (since \( u \in \mathbb{Z} \)). In other words, \( n \mid u - (n - 1)(-u) \). In other words, \( u \equiv (n - 1)(-u) \mod n \). Moreover, \( n \geq 1 \) (since \( n \) is a positive integer), so that \( n - 1 \geq 0 \). We can multiply this inequality
with \(-u\) (since \(-u > 0\)), and thus obtain \((n - 1)(-u) \geq 0(-u) = 0\). In other words, \((n - 1)(-u) \in \mathbb{N}\). Thus, there exists a \(v \in \mathbb{N}\) such that \(u \equiv v \mod n\) (namely, \(v = (n - 1)(-u)\)). This proves Lemma 2.21 in Case 2.

We have now proven Lemma 2.21 in both Cases 1 and 2; hence, Lemma 2.21 always holds.

**Proof of Theorem 2.15.** Theorem 2.15 follows by combining Lemma 2.18 with Lemma 2.21.

**Proof of Lemma 2.18.** Lemma 2.21 shows that there exists a \(v \in \mathbb{N}\) such that \(u \equiv v \mod n\). Consider this \(v\).

Note that \(n \mid u - v\) (since \(u \equiv v \mod n\)). In other words, there exists an integer \(c\) such that \(u - v = nc\). Consider this \(c\). From \(u - v = nc\), we obtain \(u = v + nc\).

Lemma 2.20 (applied to \(v\) instead of \(u\)) yields that there exists at least one pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) such that \(v = qn + r\). Fix such a pair, and denote it by \((q_0, r_0)\). Thus, \((q_0, r_0) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) and \(v = q_0n + r_0\). Now,

\[
\frac{u}{n} = \frac{v}{n} + nc = (q_0n + r_0) + nc = (q_0 + c)n + r_0.
\]

Also, from \((q_0, r_0) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\), we obtain \(q_0 \in \mathbb{Z}\) and \(r_0 \in \{0, 1, \ldots, n - 1\}\), and thus \((q_0 + c, r_0) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\). Thus, the pair \((q_0 + c, r_0)\) is a pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) such that \(u = qn + r\) (since \(u = (q_0 + c)n + r_0\)). Therefore, there exists at least one pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) such that \(u = qn + r\) (namely, \((q, r) = (q_0 + c, r_0)\)). This proves Lemma 2.18.

**Proof of Theorem 2.15.** Theorem 2.15 follows by combining Lemma 2.18 with Lemma 2.19.

The following properties of the quotient and the remainder are simple but will be used all the time:

**Corollary 2.22.** Let \(n\) be a positive integer. Let \(u \in \mathbb{Z}\).

(a) Then, \(\frac{u}{n} \in \{0, 1, \ldots, n - 1\}\) and \(u \equiv \frac{u}{n} \mod n\).

(b) We have \(n \mid u\) if and only if \(u \equiv 0 \mod n\).

(c) If \(c \in \{0, 1, \ldots, n - 1\}\) is such that \(c \equiv u \mod n\), then \(c = \frac{u}{n}\).

(d) We have \(u = (\frac{u}{n})n + \frac{u}{n}\).

Before we prove this corollary, let us explain its purpose. Corollary 2.22 (a) says that \(\frac{u}{n}\) is a number in the set \(\{0, 1, \ldots, n - 1\}\) that is congruent to \(u\) modulo \(n\). Corollary 2.22 (c) says that \(\frac{u}{n}\) is the only such number (as it says that any further such number \(c\) must be equal to \(\frac{u}{n}\)). Corollary 2.22 (b) gives an algorithm to check whether \(n \mid u\) holds (namely, compute \(\frac{u}{n}\) and check whether \(\frac{u}{n} = 0\)). Corollary 2.22 (d) is a trivial consequence of the definition of quotient and remainder.

**Proof of Corollary 2.22** Theorem 2.15 says that there is a unique pair \((q, r) \in \mathbb{Z} \times \{0, 1, \ldots, n - 1\}\) such that \(u = qn + r\). Consider this pair \((q, r)\). The uniqueness of
Recall that \( u \equiv n \) was defined to be \( r \) (in Definition 2.16(b)). Thus, \( u \equiv n \) \( r \). Now, \( n \mid qn = u - r \) (since \( u = qn + r \)). In other words, \( u \equiv r \mod n \). Hence, \( r \equiv u \mod n \) (by Proposition 2.8(c)). This rewrites as \( u \equiv u \mod n \) (since \( r = u \mod n \)). Furthermore, \( u \equiv n = r \in \{0,1,\ldots,n-1\} \) (since \( (q,r) \in \mathbb{Z} \times \{0,1,\ldots,n-1\} \)). This completes the proof of Corollary 2.22(a).

Also, \( u \equiv n \) was defined to be \( q \) (in Definition 2.16(a)). Hence, \( u \equiv n = q \). Now,

\[
(14) \quad u = \frac{q}{n} + \frac{r}{u \equiv n} = (u \equiv n) n + (u \equiv n).
\]

This proves Corollary 2.22(d).

(b) \implies: Assume that \( n \mid u \). We must prove that \( u \equiv n = 0 \).

We have \( n \mid u \). In other words, there exists some integer \( w \) such that \( u = nw \). Consider this \( w \).

We have \( n - 1 \in \mathbb{N} \) (since \( n \) is a positive integer), thus \( 0 \in \{0,1,\ldots,n-1\} \).
Hence, \((w,0) \in \mathbb{Z} \times \{0,1,\ldots,n-1\} \) (since \( w \in \mathbb{Z} \)). Also, \( u = nw = wn = wn + 0 \).
Hence, \((14) \) (applied to \( (q',r') = (w,0) \)) yields \((w,0) = (q,r) \). In other words, \( w = q \) and \( 0 = r \). Hence, \( r = 0 \), so that \( u \equiv n = 0 \). This proves the \( \implies \) implication of Corollary 2.22(b).

\( \Longleftarrow: \) Assume that \( u \equiv n = 0 \). We must prove that \( n \mid u \).

We have \( u = qn + \frac{r}{u \equiv n} = qn = nq \). Thus, \( n \mid u \). This proves the \( \Longleftarrow \) implication of Corollary 2.22(b).

(c) Let \( c \in \{0,1,\ldots,n-1\} \) be such that \( c \equiv u \mod n \).

We have \( c \equiv u \mod n \). In other words, \( n \mid c - u \). In other words, there exists some integer \( w \) such that \( c - u = nw \). Consider this \( w \).

From \(-w \in \mathbb{Z} \) and \( c \in \{0,1,\ldots,n-1\} \), we obtain \((c - u) \in \mathbb{Z} \times \{0,1,\ldots,n-1\} \).
Also, from \( c - u = nw \), we obtain \( u = c - nw = (-w) n + c \). Hence, \((14) \) (applied to \( (q',r') = (-w,c) \)) yields \((-w,c) = (q,r) \). In other words, \(-w = q \) and \( c = r \).
Hence, \( c = r = u \equiv n \). This proves Corollary 2.22(c). \( \square \)

**Exercise 2.9.** Let \( n \) be a positive integer. Let \( u \) and \( v \) be integers. Prove that \( u \equiv v \mod n \) if and only if \( u \equiv v \mod n \).

### 2.6. Even and odd numbers

Recall the following:
Definition 2.23. Let \( u \) be an integer.

(a) We say that \( u \) is **even** if \( u \) is divisible by 2.

(b) We say that \( u \) is **odd** if \( u \) is not divisible by 2.

So an integer is either even or odd (but not both at the same time). The following exercise collects various properties of even and odd integers:

**Exercise 2.10.** Let \( u \) be an integer.

(a) Prove that \( u \) is even if and only if \( u \% 2 = 0 \).

(b) Prove that \( u \) is odd if and only if \( u \% 2 = 1 \).

(c) Prove that \( u \) is even if and only if \( u \equiv 0 \mod 2 \).

(d) Prove that \( u \) is odd if and only if \( u \equiv 1 \mod 2 \).

(e) Prove that \( u \) is odd if and only if \( u + 1 \) is even.

(f) Prove that exactly one of the two numbers \( u \) and \( u + 1 \) is even.

(g) Prove that \( u \equiv v \mod 2 \) holds if and only if \( u \) and \( v \) are either both odd or both even.

**Solution to Exercise 2.10.** TODO.

**Exercise 2.11.** (a) Prove that each even integer \( u \) satisfies \( u^2 \equiv 0 \mod 4 \).

(b) Prove that each odd integer \( u \) satisfies \( u^2 \equiv 1 \mod 4 \).

(c) Prove that no two integers \( x \) and \( y \) satisfy \( x^2 + y^2 \equiv 3 \mod 4 \).

(d) Prove that if \( x \) and \( y \) are two integers satisfying \( x^2 + y^2 \equiv 2 \mod 4 \), then \( x \) and \( y \) are both odd.

**Solution to Exercise 2.11.** TODO.

Exercise 2.11(c) establishes our previous experimental observation that an integer of the form \( 4k + 3 \) with integer \( k \) (that is, an integer that is larger by 3 than a multiple of 4) can never be written as a sum of two perfect squares.

2019-02-04 lecture

2.7. The floor function

Definition 2.24. Let \( x \) be a real number. Then, \( \lfloor x \rfloor \) is defined to be the unique integer \( n \) satisfying \( n \leq x < n + 1 \). This integer \( \lfloor x \rfloor \) is called the **floor** of \( x \), or the **integer part** of \( x \).

**Remark 2.25.** (a) Why is \( \lfloor x \rfloor \) well-defined? I mean, why does the unique integer \( n \) in Definition 2.24 exist, and why is it unique? This question is trickier than it sounds and relies on the construction of real numbers. However, in the case
when \( x \) is rational, the well-definedness of \( \lfloor x \rfloor \) follows from Proposition 2.26 below.

(b) What we call \( \lfloor x \rfloor \) is typically called \( [x] \) in older books (such as [NiZuMo91]). I suggest avoiding the notation \( [x] \) wherever possible; it has too many different meanings (whereas \( \lfloor x \rfloor \) almost always means the floor of \( x \)).

(c) The map \( \mathbb{R} \to \mathbb{Z}, \, x \mapsto \lfloor x \rfloor \) is called the floor function or the greatest integer function.

There is also a ceiling function, which sends each \( x \in \mathbb{R} \) to the unique integer \( n \) satisfying \( n - 1 < x \leq n \); this latter integer is called \( \lceil x \rceil \). The two functions are connected by the rule \( \lceil x \rceil = -\lfloor -x \rfloor \) (for all \( x \in \mathbb{R} \)).

The floor and the ceiling functions are some of the simplest examples of discontinuous functions.

(d) Here are some examples of floors:

\[
\begin{align*}
\lfloor n \rfloor &= n & \text{for every } n \in \mathbb{Z}; \\
\lfloor 1.32 \rfloor &= 1; & \lfloor \pi \rfloor = 3; & \lfloor 0.98 \rfloor = 0; \\
\lfloor -2.3 \rfloor &= -3; & \lfloor -0.4 \rfloor &= -1.
\end{align*}
\]

(e) You might have the impression that \( \lfloor x \rfloor \) is “what remains from \( x \) if the digits behind the comma are removed”. This impression is highly imprecise. For one, it is completely broken for negative \( x \) (for example, \( \lfloor -2.3 \rfloor \) is \(-3\), not \(-2\)). But more importantly, the operation of “removing the digits behind the comma” from a number is not well-defined; the periodic decimal representations 0.999\ldots and 1.000\ldots belong to the same real number (1), but removing their digits behind the comma leaves us with different integers.

(f) A related map is the map \( \mathbb{R} \to \mathbb{Z}, \, x \mapsto \lfloor x + \frac{1}{2} \rfloor \). It sends each real \( x \) to the integer that is closest to \( x \), choosing the larger one in the case of a tie. This is one of the many things that are commonly known as “rounding” a number.

---

**Proposition 2.26.** Let \( a \) and \( b \) be integers such that \( b > 0 \). Then, \( \lfloor a / b \rfloor \) is well-defined and equals \( a / \lceil b \rceil \).

*Proof of Proposition 2.26* TODO (exercise).

---

**2.8. Common divisors, the Euclidean algorithm and the Bezout theorem**

**Definition 2.27.** Let \( b \in \mathbb{Z} \). The divisors of \( b \) are defined as the integers that divide \( b \).

Be aware that some books have a mildly different definition of “divisors”; namely, they additionally require them to be positive. We don’t make such a requirement.
For example, the divisors of 6 are \(-6, -3, -2, -1, 1, 2, 3, 6\). Of course, the negative divisors of an integer \(b\) are merely the reflections of the positive divisors through the origin\(^{12}\) (this follows easily from Proposition 2.4(a)); thus, the positive divisors are usually the only thing one is interested in.

Here are some basic properties of divisors:

**Proposition 2.28.** (a) If \(b \in \mathbb{Z}\), then 1 and \(b\) are divisors of \(b\).

(b) The divisors of 0 are all the integers.

(c) Let \(b \in \mathbb{Z}\) be nonzero. Then, all divisors of \(b\) belong to the set \(\{−|b|, −|b| + 1, \ldots, |b|\} \setminus \{0\}\).

**Proof of Proposition 2.28.** (a) Clearly, \(1 \mid b\) (since \(b = 1b\)), so that 1 is a divisor of \(b\). Also, \(b \mid b\) (since \(b = b \cdot 1\)), so that \(b\) is a divisor of \(b\).

(b) Each integer \(a\) divides 0 (since \(0 = a \cdot 0\)) and thus is a divisor of 0. This proves Proposition 2.28(b).

(c) Let \(a\) be a divisor of \(b\). Then, \(a\) divides \(b\). In other words, \(a \mid b\). Hence, Proposition 2.4(b) yields \(|a| \leq |b|\) (since \(b \neq 0\)). But \(|a| \geq a\) (since \(|x| \geq x\) for each \(x \in \mathbb{R}\)), so that \(a \leq |a| \leq |b|\). Also, \(|a| \geq −a\) (since \(|x| \geq −x\) for each \(x \in \mathbb{R}\)) and thus \(−a \leq |a| \leq |b|\), so that \(a \geq −|b|\). Combining this with \(a \leq |b|\), we obtain \(−|b| \leq a \leq |b|\) and thus \(a \in \{−|b|, −|b| + 1, \ldots, |b|\}\) (since \(a\) is an integer).

From Example 2.3(c), we know that \(0 \mid b\) only when \(b = 0\). Thus, we don’t have \(0 \mid b\) (since \(b \neq 0\)).

If we had \(a = 0\), then we would have \(0 \mid b\), which would contradict the fact that we don’t have \(0 \mid b\). Thus, we cannot have \(a = 0\). Hence, \(a \neq 0\). Combining \(a \in \{−|b|, −|b| + 1, \ldots, |b|\}\) with \(a \neq 0\), we obtain \(a \in \{−|b|, −|b| + 1, \ldots, |b|\} \setminus \{0\}\).

We have proven this for each divisor \(a\) of \(b\). Thus, we conclude that all divisors of \(b\) belong to the set \(\{−|b|, −|b| + 1, \ldots, |b|\} \setminus \{0\}\). This proves Proposition 2.28(c).

Thanks to Proposition 2.28, we have a method to find all divisors of an integer \(b\): If \(b = 0\), then Proposition 2.28(b) directly yields the result; otherwise, Proposition 2.28(c) shows that there is only a finite set of numbers we have to check. When \(b\) is large, this is slow, but to some extent that is because the problem is computationally hard (or at least suspected to be hard).

It is somewhat more interesting to consider the common divisors of two or more integers:

**Definition 2.29.** Let \(b_1, b_2, \ldots, b_k\) be finitely many integers. Then, the common divisors of \(b_1, b_2, \ldots, b_k\) are defined to be the integers \(a\) that satisfy

\[
(a \mid b_i \text{ for all } i \in \{1, 2, \ldots, k\})
\]

(in other words, that divide all of the integers \(b_1, b_2, \ldots, b_k\)). We let \(\text{Div}(b_1, b_2, \ldots, b_k)\) denote the set of these common divisors.

\(^{12}\)Reflection through the origin” is just a poetic way to say “negative”; i.e., the reflection of a number \(a\) through the origin is \(-a\).
Note that the concept of common divisors encompasses the concept of divisors: The common divisors of a single integers $b$ are merely the divisors of $b$. Thus, \( \text{Div}(b) \) is the set of all divisors of $b$ whenever $b \in \mathbb{Z}$.

(Also, the common divisors of an empty list of integers are all integers, because the requirement (15) is vacuously true for $k = 0$. Thus, \( \text{Div}() = \mathbb{Z} \).)

**Example 2.30.** The common divisors of 6 and 8 are $-2, -1, 1, 2$. (In order to see this, just observe that the divisors of 6 are $-6, -3, -2, -1, 1, 2, 3, 6$, whereas the divisors of 8 are $-8, -4, -2, -1, 1, 2, 4, 8$; now you can find the common divisors of 6 and 8 by taking the numbers common to these two lists.) Thus,

\[
\text{Div}(6, 8) = \{-2, -1, 1, 2\}.
\]

**Proposition 2.31.** Let $b_1, b_2, \ldots, b_k$ be finitely many integers that are not all 0. Then, the set \( \text{Div}(b_1, b_2, \ldots, b_k) \) has a largest element, and this largest element is a positive integer.

**Proof of Proposition 2.31.** The integer 1 satisfies \( (1 \mid b_i \text{ for all } i \in \{1, 2, \ldots, k\}) \). Thus, 1 is a common divisor of $b_1, b_2, \ldots, b_k$ (by the definition of a “common divisor”). In other words, \( 1 \in \text{Div}(b_1, b_2, \ldots, b_k) \) (by the definition of \( \text{Div}(b_1, b_2, \ldots, b_k) \)). Hence, the set \( \text{Div}(b_1, b_2, \ldots, b_k) \) is nonempty.

We have assumed that $b_1, b_2, \ldots, b_k$ are not all 0. In other words, there exists an $i \in \{1, 2, \ldots, k\}$ such that $b_i \neq 0$. Consider such an $i$.

TODO.

- Greatest common divisor. (Also set \( \gcd(0, 0) = 0 \).)
- \( \gcd(a, ua + b) = \gcd(a, b) \), or equivalently \( \gcd(a, b) = \gcd(a, c) \) if \( b \equiv c \mod a \).
- \( \gcd(a, b) = \gcd(b, a) \).
- Euclidean algorithm.
- Bezout for $a, b \geq 0$, proven by strong ind $/a + b$.
- Bezout for arbitrary $a, b$.
- “\( \mathbb{Z} \)-linear combination”.
- corollary \( m \mid a, m \mid b \implies m \mid \gcd(a, b) \).
- coprimality.
- notation $a \perp b$.
- corollary $ab \mid c, a \perp c \implies b \mid c$. 

}(Also set gcd(0, 0) = 0.)

- gcd(a, ua + b) = gcd(a, b), or equivalently gcd(a, b) = gcd(a, c) if b \equiv c mod a.
- gcd(a, b) = gcd(b, a).
- Euclidean algorithm.
- Bezout for a, b ≥ 0, proven by strong ind/a + b.
- Bezout for arbitrary a, b.
- “Z-linear combination”.
- corollary m | a, m | b \implies m | gcd(a, b).
- coprimality.
- notation a ⊥ b.
- corollary ab | c, a ⊥ c \implies b | c.
• corollary: existence of modular inverses.
• \( s \gcd(a, b) = \gcd(sa, sb) \).
• \( a \mid bc \), \( b \perp c \implies a = uv \) with \( u \mid b \) and \( v \mid c \).
• lcm.
• gcd*lcm.
• gcd of multiple numbers.

2.9. Substitutivity for congruences

In Section 2.4, we have learnt that congruences modulo an integer \( n \) can be chained together like equalities. A further important feature of congruences is the principle of substitutivity for congruences. This is yet another way in which congruences behave like equalities. We are not going to state it fully formally (as it is a metamathematical principle), but merely explain its meaning. Later on, once we understand what the rings \( \mathbb{Z}/n \) (for integer \( n \)) are, we will no longer need this principle, since it will just boil down to “equal things can be substituted for one another” (the whole point of \( \mathbb{Z}/n \) is to “make congruent numbers equal”); but for now, we cannot treat “congruent modulo \( n \)” as “equal”, so we have to state it.

You are probably used to making computations like these:

\[
(a + b)^2 = a^2 + 2ab + b^2
\]

(for any two numbers \( a \) and \( b \)). What is going on in these underbraces (like \( (a + b)^2 \))? Something pretty simple is going on: You are replacing a number (in this case, \( (a + b)^2 \)) by an equal number (in this case, \( a^2 + 2ab + b^2 \)). This relies on a fundamental principle of mathematics (called the principle of substitutivity for equalities), which says that an object in an expression can indeed be replaced by any object equal to it (without changing the value of the expression). (This is also known as Leibniz’s equality law.) To be precise, we are using this principle twice in some of our equality signs above, since we are making several replacements at the same time; but this is fine (we can just do the replacement one by one instead).

We would like to have a similar principle for congruences modulo \( n \): We would like to be able to replace any integer by an integer congruent to it modulo \( n \). For example, we would like to be able to say that if seven integers \( a, a', b, b', c, c', n \) satisfy 
\[
a \equiv a' \pmod{n} \quad b \equiv b' \pmod{n} \quad c \equiv c' \pmod{n}
\]

then
\[
\underbrace{b}_{\equiv b' \pmod{n}} + \underbrace{c}_{\equiv c' \pmod{n}} + \underbrace{a}_{\equiv a' \pmod{n}} + \underbrace{a}_{\equiv a' \pmod{n}} + \underbrace{b}_{\equiv b' \pmod{n}} \equiv b'c' + c'a' + a'b' \pmod{n}.
\]
We have to be careful with this: For example, we run into troubles if division is involved in our expressions. For example, we have $6 \equiv 9 \mod 3$, but we do not have $\frac{6}{3} \equiv \frac{9}{3} \mod 3$. Similarly, exponentiation can be problematic. So we need to state the principle we are using here in clearer terms, so that we know what we can do.

For this whole Section 2.9, we fix an integer $n$.

The principle of substitutivity for equalities says the following:

**Principle of substitutivity for equalities (PSE):** If two objects $x$ and $x'$ are equal, and if we have any expression $A$ that involves the object $x$, then we can replace this $x$ (or, more precisely, any arbitrary appearance of $x$ in $A$) in $A$ by $x'$; the resulting expression $A'$ will be equal to $A$.

Here are two examples of how this principle can be used:

- If $a, b, c, d, e, c'$ are numbers such that $c = c'$, then the PSE says that we can replace $c$ by $c'$ in the expression $a \left( b - (c + d) e \right)$, and the resulting expression $a \left( b - (c' + d) e \right)$ will be equal to $a \left( b - (c + d) e \right)$; that is, we have

\[
a \left( b - (c + d) e \right) = a \left( b - (c' + d) e \right).
\]

(16)

- If $a, b, c, a'$ are numbers such that $a = a'$, then

\[
(a - b) \left( a + b \right) = (a' - b) \left( a + b \right),
\]

(17)

because the PSE allows us to replace the first $a$ appearing in the expression $(a - b) \left( a + b \right)$ by an $a'$. (We can also replace the second $a$ by $a'$, of course.)

More generally, we can make several such replacements at the same time. The PSE is one of the headstones of mathematical logic; it is the essence of what it means for two objects to be equal.

The principle of substitutivity for congruences is similar, but far less fundamental; it says the following:

**Principle of substitutivity for congruences (PSC):** If two numbers $x$ and $x'$ are congruent to each other modulo $n$ (that is, $x \equiv x' \mod n$), and if we have any expression $A$ that involves only integers, addition, subtraction and multiplication, and involves the object $x$, then we can replace this $x$ (or, more precisely, any arbitrary appearance of $x$ in $A$) in $A$ by $x'$; the resulting expression $A'$ will be congruent to $A$ modulo $n$.

This principle is less general than the PSE, since it only applies to expressions that are built from integers and certain operations (note that division is not one of these operations). But it still lets us prove analogues of our above examples (16) and (17):
• If $a, b, c, d, e, c'$ are integers such that $c \equiv c' \mod n$, then the PSC says that we can replace $c$ by $c'$ in the expression $a(b - (c + d)e)$, and the resulting expression $a(b - (c' + d)e)$ will be congruent to $a(b - (c + d)e)$ modulo $n$; that is, we have
\[ a(b - (c + d)e) \equiv a(b - (c' + d)e) \mod n. \tag{18} \]

• If $a, b, c, a'$ are integers such that $a \equiv a' \mod n$, then
\[ (a - b)(a + b) \equiv (a' - b)(a + b) \mod n, \tag{19} \]

because the PSC allows us to replace the first $a$ appearing in the expression $(a - b)(a + b)$ by an $a'$. (We can also replace the second $a$ by $a'$, of course.)

We shall not prove the PSC, since we have not formalized it (after all, we have not defined what an “expression” is). But we shall prove the specific congruences (18) and (19) using Proposition 2.8, the way in which we prove these congruences is symptomatic: Every congruence obtained from the PSC can be proven in a manner like these. Thus, the proofs of (18) and (19) given below can serve as templates which can easily be adapted to any other situation in which an application of the PSC needs to be justified.

Proof of (18). Let $n$ be any integer, and let $a, b, c, d, e, c'$ be integers such that $c \equiv c' \mod n$.

Adding the congruence $c \equiv c' \mod n$ with the congruence $d \equiv d \mod n$ (which follows from Proposition 2.8 (a)), we obtain $c + d \equiv c' + d \mod n$. Multiplying this congruence with the congruence $e \equiv e \mod n$ (which follows from Proposition 2.8 (a)), we obtain $(c + d)e \equiv (c' + d)e \mod n$. Subtracting this congruence from the congruence $b \equiv b \mod n$ (which, again, follows from Proposition 2.8 (a)), we obtain $b - (c + d)e \equiv b - (c' + d)e \mod n$. Multiplying the congruence $a \equiv a \mod n$ (which follows from Proposition 2.8 (a)) with this congruence, we obtain $a(b - (c + d)e) \equiv a(b - (c' + d)e) \mod n$. This proves (18). \qed

Proof of (19). Let $n$ be any integer, and let $a, b, c, a'$ be integers such that $a \equiv a' \mod n$.

Subtracting the congruence $b \equiv b \mod n$ (which follows from Proposition 2.8 (a)) from the congruence $a \equiv a' \mod n$, we obtain $a - b \equiv a' - b \mod n$. Multiplying this congruence with the congruence $a + b \equiv a + b \mod n$ (which follows from Proposition 2.8 (a)), we obtain $(a - b)(a + b) \equiv (a' - b)(a + b) \mod n$. This proves (19). \qed

As we said, these two proofs are exemplary: Any congruence obtained from the PSC can be proven in such a way (starting with the congruence $x \equiv x' \mod n$, and then “wrapping” it up in the expression $A$ by repeatedly adding, multiplying and subtracting congruences that follow from Proposition 2.8 (a)).

\footnote{Proposition 2.8 (d) shows that we can add, subtract and multiply congruences modulo $n$ at will. We are using this freedom here and will use it many times below.}
When we apply the PSC, we shall use underbraces to point out which integers we are replacing. For example, when deriving (18) from this principle, we shall write
\[
 a \left( b - \left( c \equiv c' \mod n \right) + d \right) e \equiv a \left( b - (c' + d) e \right) \mod n,
\]
in order to stress that we are replacing \( c \) by \( c' \). Likewise, when deriving (19) from the PSC, we shall write
\[
 \left( a \equiv a' \mod n \right) - b \right) (a + b) \equiv (a' - b) (a + b) \mod n,
\]
in order to stress that we are replacing the first \( a \) (but not the second \( a \)) by \( a' \).

The PSC allows us to replace a single integer \( x \) appearing in an expression by another integer \( x' \) that is congruent to \( x \) modulo \( n \). Applying this principle many times, we thus conclude that we can also replace several integers at the same time (because we can get to the same result by performing these replacements one at a time, and Proposition 2.12 shows that the final result will be congruent to the original result). For example, if seven integers \( a, a', b, b', c, c', n \) satisfy
\[
 a \equiv a' \mod n \quad \text{and} \quad b \equiv b' \mod n \quad \text{and} \quad c \equiv c' \mod n,
\]
then
\[
 bc + ca + ab \equiv b'c' + c'a' + a'b' \mod n,
\]
because we can replace all the six integers \( b, c, c, a, a, b \) in the expression \( bc + ca + ab \) (listed in the order of their appearance in this expression) by \( b', c', c', a', a', b' \), respectively. If we want to derive this from the principle of substitutivity for congruences, we must perform the replacements one at a time, e.g., as follows:
\[
 b \equiv b' \mod n \quad c + ca + ab \equiv b'c' \equiv c' \mod n \quad +ca + ab \equiv b'c' + c' \equiv c \mod n \quad a + ab \\
 \equiv b'c' + c' \equiv c' \mod n \quad a \equiv a' \mod n \quad +ab \equiv b'c' + c'a' + a \equiv a' \mod n \quad b \\
 \equiv b'c' + c'a' + a' \equiv b'c' + c'a' + a'b' \mod n.
\]

Of course, we shall always just show the replacements as a single step:
\[
 b \equiv b' \mod n \quad c \equiv c' \mod n \quad +a \equiv a' \mod n \quad b \equiv b' \mod n \equiv b'c' + c'a' + a'b' \mod n.
\]

References


http://www.math.miami.edu/~armstrong/561fa18.php


[Conrad*] Keith Conrad, Expository notes (“blurbs”).  
https://kconrad.math.uconn.edu/blurbs/


https://www.math.vt.edu/people/day/ProofsBook/IPaMV.pdf.


https://www.cis.upenn.edu/~jean/RSA-primality-testing.pdf

http://homepage.divms.uiowa.edu/~goodman/algebrabook.dir/algebrabook.html

The numbering of theorems and formulas in this link might shift when the project gets updated; for a “frozen” version whose numbering is guaranteed to match that in the citations above, see https://github.com/darijgr/detnotes/releases/tag/2019-01-10.

http://www.people.vcu.edu/~rhammack/BookOfProof/

http://joshua.smcvt.edu/linearalgebra/