

# **Discrete Mathematics**

*Solutions with Commentary*

---

Marcelo Fiore

Ohad Kammar

Dima Szamozvancev

2020


# Contents

1.	On proofs . . . . .	1
1.1.	Basic exercises . . . . .	1
1.2.	Core exercises . . . . .	4
1.3.	Optional exercises . . . . .	12
2.	On numbers . . . . .	15
2.1.	Basic exercises . . . . .	15
2.2.	Core exercises . . . . .	20
2.3.	Optional exercises . . . . .	25
3.	More on numbers . . . . .	28
3.1.	Basic exercises . . . . .	28
3.2.	Core exercises . . . . .	29
3.3.	Optional exercises . . . . .	40
4.	On induction . . . . .	42
4.1.	Basic exercises . . . . .	42
4.2.	Core exercises . . . . .	45
4.3.	Optional exercises . . . . .	56

# 1. On proofs

## 1.1. Basic exercises

*The main aim is to practice the analysis and understanding of mathematical statements (e.g. by isolating the different components of composite statements), and exercise the art of presenting a logical argument in the form of a clear proof (e.g. by following proof strategies and patterns).*


The solutions will consist of the proper formal proof, showing the required level of detail and precision – you should try to present your answers in a similar form. Of course, a formal written proof is just a polished facade of the (usually more difficult) process of finding the correct sequence of reasoning steps and proof techniques, which often constitutes the “scratchwork”. Therefore, most of the formal proofs will be accompanied with notes (marked ) on how the problem was approached, what guided the reasoning process and what mistakes should be avoided. Mastering the art of formal proof may take some practice, but it is a very important skill to acquire both for this course and your whole scientific education.


Prove or disprove the following statements.

Some fairly simple statements, but they showcase a wide range of proposition types and proof techniques. Accordingly, the proof notes can apply to most of the statements you will encounter, no matter how complicated.

1. Suppose  $n$  is a natural number larger than 2, and  $n$  is not a prime number. Then  $2 \cdot n + 13$  is not a prime number.

The statement is false. Choose  $n = 9$ . Then  $n = 3 \cdot 3$  isn't prime, yet  $2 \cdot n + 13 = 31$  is prime, and we disproved the statement by a counterexample.

 “Prove or disprove” questions should usually start with a sanity check: try a few numbers, and if things seem to work, try a formal proof. Unfortunately, this is not a sure-fire technique, as you may need to backtrack after realising that the statement is false after all. If you realise this in the middle of writing the formal proof (rather than just the scratchwork), you need to cross everything out and start again: there is no space for “plot twists” in a proof attempt, and you should state if the statement is true or false right away.

 To disprove a statement, all we need to present is a counterexample which falsifies it. There is no need to explain the general situation where the statement doesn't work, or try to prove the negation of the statement. The counterexample doesn't have to be very elaborate, often edge cases like 0 or the empty set do the job perfectly. However, we have to make sure that our counterexample falls under the consideration of the statement: 0 will not falsify a proposition that starts with “for every positive integer”.

2. If  $x^2 + y = 13$  and  $y \neq 4$  then  $x \neq 3$ .

We equivalently prove that if  $x^2 + y = 13$  then  $y \neq 4$  implies  $x \neq 3$ . Assume that  $x^2 + y = 13$ . We establish the contrapositive of the goal, i.e. if  $x = 3$  then  $y = 4$ . Indeed,

assume  $x = 3$ . Then,  $y = 13 - x^2 = 13 - 9 = 4$ , as required.

♪ The statement is of the form  $(P \wedge Q) \Rightarrow R$ , so our proof algorithm dictates that we should start by assuming  $P$  and  $Q$ , and prove  $R$ . However, in this case,  $Q$  and  $R$  are *negative* assertions: knowing that  $y$  could be anything other than 4 is less useful than knowing that it is equal to something. Since the consequent  $R$  is also negated, we may instead consider proving a contrapositive. However, the contrapositive of  $(P \wedge Q) \Rightarrow R$  is  $\neg R \Rightarrow \neg(P \wedge Q)$ , which will turn our useful assumption  $P$  into a negated goal. Instead, we perform a *partial assumption*: we assume  $P$  only, and prove  $Q \Rightarrow R$ ; this can now be done easily, since the contrapositive will give us the additional assumption  $x = 3$  and the goal will be  $y = 4$ . Logically, this technique follows from the equivalence  $(P \wedge Q) \Rightarrow R \simeq P \Rightarrow (Q \Rightarrow R)$ , which can be seen as “currying” the proposition  $Q$ .

3. For an integer  $n$ ,  $n^2$  is even if and only if  $n$  is even.

( $\Rightarrow$ ) We prove the following more general result: The product of an even integer with any integer is an even integer. The required proposition follows as a corollary.

Consider any two integers  $m, n$  and assume that  $m$  is even. By definition of even integers,  $m = 2k$  for some integer  $k$ . Therefore,  $m \cdot n = 2k \cdot n = 2(k \cdot n)$  and thus, by definition,  $m \cdot n$  is an even integer.

( $\Leftarrow$ ) We prove the contrapositive; i.e.,  $n$  odd implies  $n^2$  odd. Assume  $n$  is odd, then by [Proposition 8](#) (product of odd numbers is odd) of the notes,  $n \cdot n = n^2$  is odd.

♪ As soon as we see the phrase “if and only if”, we need to remember not to move on to the next question halfway through the proof. Most of the time such proofs will consist of two parts, so stating which direction is being proved is helpful for the reader (and a good reminder to you to complete both directions).

♪ Both directions follow as *corollaries* (logical consequences which are important in their own right) of more general statements about multiplying two different even/odd numbers. The advantage of making theorem statements as general as possible is that they can be applied in many contexts and give rise to useful corollaries; we could have proved that the square of an odd number is odd directly, but the underlying proof approach would have been exactly the same as Proposition 8 so we might as well use it!

4. For all real numbers  $x$  and  $y$  there is a real number  $z$  such that  $x + z = y - z$ .

Consider arbitrary real numbers  $x, y$ , and choose  $z = \frac{y-x}{2}$ . Then,  $z$  is a real number satisfying  $x + z = x + \frac{y-x}{2} = \frac{y+x}{2} = y + \frac{y-x}{2} = y - z$ . Therefore, there exists a real number  $z$  satisfying  $x + z = y - z$ .

♪ This is an example of an existence proof, which tend to look a bit backwards when written formally: rather than deriving the witness as part of the proof, we “give away” the answer right away, then show that it satisfies the required property. Of course, we don’t just pluck the witness out of thin air, or resort to a lucky guess. We find what it should be from

the required properties via some sort of calculation, and once we found an answer, we present it as a witness at the very beginning of the formal proof. Showing that it satisfies the properties is of course straightforward (but can't be omitted), since that's how we found the witness in the first place. In this specific example, the answer  $z = \frac{y-x}{2}$  can be found by simple rearrangement of the condition  $x + z = y - z$ , but, to present this as a formal existence proof, we need to state the witness and rigorously demonstrate that it satisfies the requirement.

5. For all integers  $x$  and  $y$  there is an integer  $z$  such that  $x + z = y - z$ .

The statement is false. Indeed, for the integers  $x = 0$  and  $y = 1$  we will prove that there does not exist an integer  $z$  satisfying  $x + z = y - z$ ; i.e., equivalently, such that  $z = 1 - z$ . Assume to the contrary that such an integer, say  $z_0$ , existed. Then, we would have  $2 \cdot z_0 = 1$  and hence  $z_0 = \frac{1}{2}$ ; which is absurd as  $\frac{1}{2}$  is not an integer. Therefore, there are integers  $x$  and  $y$  for which there is no integer  $z$  such that  $x + z = y - z$ .

🎵 This proof may seem more verbose than it needs to be: surely we can just say “let  $x = 0$  and  $y = 1$ , then  $z = \frac{y-x}{2} = \frac{1}{2}$  which is not an integer”. The problem with this reasoning is that it is not a nonexistence proof: we showed that the specific  $z$  that can be computed with the method above is not an integer, but that doesn't mean there cannot be any other  $z$  that works. To be completely rigorous, we need to show that the existence of any  $z$  that satisfies the property is a logical absurdity – from this follows the lengthy but airtight proof of the answer.

🎵 Note how in two lines we got to a statement that is *obviously* false: that there exists an integer  $z$  such that  $z = 1 - z$ . We need to resist the temptation to take logical shortcuts and appeal to the intuition of the reader to fill in the holes of our argument: anyone with familiarity of basic arithmetic will recognise this as false (just rearrange the equation to get  $z = \frac{1}{2}$ , which is not an integer), but this will not convince someone who reasons purely by logic (and, unfortunately, supervisors and examiners are such people). As explained in the previous point, the easiest logically rigorous way to show that such an integer  $z$  does not exist is by contradiction.

6. The addition of two rational numbers is a rational number.

Consider any two rational numbers  $r, s$ . By definition, there exist some integers  $a, c$  and some positive integers  $b, d$  such that  $r = \frac{a}{b}$  and  $s = \frac{c}{d}$ . Then,  $r + s = \frac{a \cdot d + b \cdot c}{b \cdot d}$  is a quotient of an integer (namely  $a \cdot d + b \cdot c$ ) by a positive integer (namely  $b \cdot d$ ), and hence a rational number.

🎵 A large part of writing formal proofs is just expanding definitions: rather than trying to reason about rational numbers, we use their formal definition to transition into a proof about integers. The more abstract the statement (quite common in set theory), the more layers of definitions we may need to unwrap. However, this can allow us to prove some rather difficult-looking propositions with a very simple, low-level reasoning step.

7. For every real number  $x$ , if  $x \neq 2$  then there is a unique real number  $y$  such that  $2 \cdot y / (y + 1) = x$ .

We need to show that for every real number  $x$ , if  $x \neq 2$  then there exists a real number  $y$  satisfying: ①  $\frac{2y}{y+1} = x$  and ② for all real numbers  $z$ , if  $\frac{2z}{z+1} = x$  then  $y = z$ .

Consider an arbitrary real number  $x$ , and assume  $x \neq 2$ . Then,  $y = \frac{x}{2-x}$  is a real number satisfying ①, and if  $z$  is any real number satisfying  $\frac{2z}{z+1} = x$  then  $2 \cdot z = z \cdot x + x$ . Hence,  $(2 - x) \cdot z = x$ . As  $x \neq 2$ ,  $z = \frac{x}{2-x} = y$ .

♪ This is a unique existence proof, so requires two separate arguments: existence and uniqueness. The standard way of proving uniqueness is to assume another value with the same property, and show that it must be equal to the existing witness. Uniqueness may seem like a relatively unimportant result, but in fact, it forms the basis of powerful proof techniques which we'll see later on.

8. For all integers  $m$  and  $n$ , if  $m \cdot n$  is even, then either  $m$  is even or  $n$  is even.

One may prove the contrapositive of the statement; i.e. that if  $m$  and  $n$  are odd then  $m \cdot n$  is odd. But this is nothing but [Proposition 8](#) of the notes.

♪ Negation-based proof techniques (contradiction or contraposition) are often used to avoid awkward proof patterns, usually involving existence or disjunction. Rather than have a disjunctive goal (which requires some sort of case-splitting), we negate it to turn (via the de Morgan laws) into a conjunctive assumption.

## 1.2. Core exercises

*Having practised how to analyse and understand basic mathematical statements and clearly present their proofs, the aim is to get familiar with the basics of divisibility.*

1. Characterise those integers  $d$  and  $n$  such that:

a)  $0 \mid n$

We prove that an integer  $n$  satisfies  $0 \mid n$  iff  $n = 0$ .

( $\Rightarrow$ ) Assume  $0 \mid n$ . By definition, for some integer  $l$ ,  $n = l \cdot 0 = 0$ .

( $\Leftarrow$ ) Assume  $n = 0$ . Then,  $n = 0 \cdot 0$  and, by definition,  $0 \mid n$ .

♪ A good example of the need to be precise when applying definitions. We may intuitively interpret  $d \mid n$  (" $d$  divides  $n$ ") as " $\frac{n}{d}$  is an integer", and conclude that  $0 \mid n$  is impossible because  $\frac{n}{0}$  is undefined. However, the formal definition of  $d \mid n$  makes no mention of the division operator: it is an algebraically more fundamental concept which only requires multiplication to express. Strictly speaking, we haven't yet formally defined division in the course – sure, you *know* what division is from school, but giving a precise and rigorous definition is more difficult than it may seem! If we use the proper definition of divisibility for this exercise, we do find an appropriate value for  $n$ , namely 0: zero divides zero because there exists an integer  $l$  (any integer

will work) such that  $0 = l \cdot 0$ .

b)  $d \mid 0$

We prove that  $d \mid 0$  for all integers  $d$ . Indeed, let  $d$  be an arbitrary integer. Then,  $0 = 0 \cdot d$  and hence  $d \mid 0$ .

2. Let  $k, m, n$  be integers with  $k$  positive. Show that:

$$(k \cdot m) \mid (k \cdot n) \iff m \mid n$$

Consider any positive integer  $k$  and any two integers  $m, n$ .

( $\Rightarrow$ ) Assume  $(k \cdot m) \mid (k \cdot n)$ . Then,  $k \cdot n = l \cdot (k \cdot m)$ . As  $k > 0$ , we can cancel  $k$  and deduce  $n = l \cdot m$ . Hence,  $m \mid n$ .

( $\Leftarrow$ ) Assume  $m \mid n$ . Then,  $n = a \cdot m$  for some integer  $a$ ; and multiplying by  $k$ , we have  $k \cdot n = a \cdot (k \cdot m)$ . Hence,  $(k \cdot m) \mid (k \cdot n)$ .

🎵 “Cancelling things” on both sides of an equation is a very standard process in elementary (“high-school”) algebra. While in many cases it is still allowed in this course, you should pay extra attention to any side-conditions required for the cancellation, or if cancellation is even possible for the algebraic structure you’re working with! It does hold for addition and multiplication (with a side-condition), but, for example, an equation between function composites  $f \circ g = f \circ e$  cannot be simplified to  $g = e$  in general (only if  $f$  is an *injection* – see later). Cancellability may be a property of the structure, or particular elements in a structure, rather than something you can just do arbitrarily.

🎵 The ( $\Rightarrow$ ) direction of this proof relied on the fact that  $k$  is positive, and in particular, nonzero – otherwise we wouldn’t be able to cancel the  $k$ s (and the property wouldn’t actually hold). We did not require any assumptions on  $k$  in the ( $\Leftarrow$ ) direction, so we could extract a weaker form of the theorem, stating that for every integer  $k, m$  and  $n$ ,

$$m \mid n \implies (k \cdot m) \mid (k \cdot n)$$

In some cases you may not have the assumption that  $k$  is positive but may still be able to apply this weaker form to make progress. However, this is technically not a corollary of the stronger statement, because that requires an unneeded assumption on  $k$ .

3. Prove or disprove that: For all natural numbers  $n$ ,  $2 \mid 2^n$ .

This is false, as  $2 \nmid 2^0$ .

🎵 This is just a gentle reminder that 0 is a natural number!

4. Show that for all integers  $l, m, n$ ,

$$l \mid m \wedge m \mid n \implies l \mid n$$

Consider any integers  $l, m, n$ , and assume  $l \mid m \wedge m \mid k$ . As  $l \mid m$ ,  $m = a \cdot l$  for some integer  $a$ . As  $m \mid n$ ,  $n = b \cdot m$  for some integer  $b$ . But then:  $n = b \cdot m = b \cdot (a \cdot l) = (b \cdot a) \cdot l$  and, as  $b \cdot a$  is an integer, we have  $l \mid n$ .

♪ An example of a proof which is not particularly difficult or illuminating, but it's still presented in a clear, structured, formal manner. It should take about a line of scratchwork to convince yourself that the statement is true, but that is only the first step: next, you need to convince the reader of the proof, who may not find your sketch clear or rigorous enough. Learning how to present even the simplest arguments in a formal, systematic manner will massively aid you in tackling more difficult propositions which may seem very daunting at first, but are actually much easier to digest connective-by-connective, definition-by-definition.

5. Find a counterexample to the statement: For all positive integers  $k, m, n$ ,

$$(m \mid k \wedge n \mid k) \implies (m \cdot n) \mid k$$

Choose  $k = m = n = 2$ . Then,  $k, m, n$  are positive integers. As  $2 \mid 2$ , we have  $m \mid k \wedge n \mid k$  yet  $(2 \cdot 2) \nmid 2$ .

♪ While questions like this don't explicitly ask for it, you need to find a counterexample and also show that it is a counterexample, i.e. that it contradicts the statement. Only writing  $k = m = n = 2$  is not enough; you need to justify your answer.

6. Prove that for all integers  $d, k, l, m, n$ ,

a)  $d \mid m \wedge d \mid n \implies d \mid (m + n)$

Assume  $d \mid m \wedge d \mid n$ . As  $d \mid m$ ,  $m = a \cdot d$  for some integer  $a$ . As  $d \mid n$ ,  $n = b \cdot d$  for some integer  $b$ . Therefore,  $m + n = a \cdot d + b \cdot d = (a + b) \cdot d$ . As  $a + b$  is an integer, we have  $d \mid (m + n)$  as required.

b)  $d \mid m \implies d \mid k \cdot m$

Assume  $d \mid m$ ; i.e.  $m = a \cdot d$  for some integer  $a$ . Then,  $k \cdot m = k \cdot (a \cdot d) = (k \cdot a) \cdot d$ . As  $k \cdot a$  is an integer,  $d \mid (k \cdot m)$ .

c)  $d \mid m \wedge d \mid n \implies d \mid (k \cdot m + l \cdot n)$

Assume  $d \mid m \wedge d \mid n$ . As  $d \mid m$ , by 6(b) above,  $d \mid (k \cdot m)$ . Analogously, from  $d \mid n$  we have  $d \mid (l \cdot n)$ . Thus,  $d \mid (k \cdot m) \wedge d \mid (l \cdot n)$  so that applying 6(a) we conclude  $d \mid (k \cdot m + l \cdot n)$  as required.

♪ Science is about building on the shoulders of giants – even if that giant is us, ten minutes ago. After proving two useful properties of divisibility in parts 6(a) and 6(b), they are now part of our “knowledge base” and we can refer back to them freely, without having to reprove them again.



Mathematics and computer science are all about decomposition and composition (also known as divide-and-conquer). Faced with a complicated proposition/problem, we break it up into smaller components which are much easier to reason about. Then, we find ways to solve the subproblems: prove lemmas and sub-theorems or write functions, classes and methods to perform well-defined tasks. Finally, we combine the sub-solutions and reap the rewards. In practice, however, the challenge is not always in solving the subproblems from scratch, but figuring out which existing elements of the knowledge base/programming library can be glued together to give the desired results: after all, if we or someone else has solved some difficult problem already, we shouldn't need to do it again! There may be a striking one-liner proof/program that does the job, but finding it may take significantly more effort than just solving the problem manually. But, seeing how programmers can spend hours finding the shortest, simplest, fastest, most space-efficient algorithms, there is a lot of enjoyment to be had in crafting concise and elegant proofs that combine clever reasoning techniques with existing propositions in satisfying ways. We will hopefully see examples of this in the course so you can appreciate proof-writing not as a chore, but something intellectually stimulating and often quite addictive!

7. Prove that for all integers  $n$ ,


$$30 \mid n \iff (2 \mid n \wedge 3 \mid n \wedge 5 \mid n)$$

( $\Rightarrow$ ) Assume  $30 \mid n$ . Then,  $n = 30 \cdot a$  for some integer  $a$ . Thus,  $n = 2 \cdot (15 \cdot a)$  and so  $2 \mid n$ . Similarly,  $n = 3 \cdot (10 \cdot a)$  and therefore  $3 \mid n$ . And, as  $n = 5 \cdot (6 \cdot a)$ , we also deduce  $5 \mid n$ . Therefore  $2 \mid n \wedge 3 \mid n \wedge 5 \mid n$ .

( $\Leftarrow$ ) Assume  $2 \mid n \wedge 3 \mid n \wedge 5 \mid n$ . As  $2 \mid n$  and  $3 \mid n$  and  $5 \mid n$ , we have  $n = 2 \cdot a$  and  $n = 3 \cdot b$  and  $n = 5 \cdot c$  for some integers  $a, b, c$ . Moreover, we have:

$$30 \cdot (a + b - 4 \cdot c) = 15 \cdot 2 \cdot a + 10 \cdot 3 \cdot b - 4 \cdot 6 \cdot 5 \cdot c = 15 \cdot n + 10 \cdot n - 24 \cdot n = n$$

Thus,  $n = 30 \cdot k$  for the integer  $k = a + b - 4 \cdot c$ , as required.

 The ( $\Leftarrow$ ) direction of this proof is more subtle than it may look – we can't just multiply 2, 3 and 5 together (see §1.2.5 above). Instead, we know that  $30 \mid 30a$ , so  $30 \mid 15n$ ; similarly,  $30 \mid 10n$  and  $30 \mid 6n$ . We need to put these together to get  $30 \mid n$ , for which we make use of §1.2.6(a) above to find a linear combination of  $15n$ ,  $10n$  and  $6n$  that adds up to  $n$ . After some thinking, we find that  $15 + 10 + (-4) \times 6$  works, giving us the desired coefficients of  $a$ ,  $b$  and  $c$ .

8. Show that for all integers  $m$  and  $n$ ,

$$(m \mid n \wedge n \mid m) \implies (m = n \vee m = -n)$$

Consider any pair of integers  $m, n$ , and assume that  $m \mid n$  and that  $n \mid m$ . If  $m = 0$  then, by §1.2.1(a) above,  $n = 0$  and we have  $m = n$ .

Consider henceforth the case  $m \neq 0$ . As  $m \mid n$  and  $n \mid m$ , there are integers  $a, b$  such that  $n = a \cdot m$  and  $m = b \cdot n$ . Thus,  $m = b \cdot a \cdot m$  and, as  $m \neq 0$ , we have  $b \cdot a = 1$ . Then, since  $a$  and  $b$  are integers, either  $a = b = 1$  or  $a = b = -1$  (otherwise, one would have  $a \cdot b \geq 2$  or  $a \cdot b \leq -2$ ). Finally, if  $a = b = 1$  then  $m = n$ , and if  $a = b = -1$  then  $m = -n$ . Either way,  $m = n$  or  $m = -n$  as required.

♪ You may have started the proof from the second paragraph, without assuming that  $m \neq 0$ . Then, at the step  $m = b \cdot a \cdot m$ , you would be stuck (if you're being careful): you can't divide by  $m$  because it may be 0. In such cases a common solution is to handle the problematic case ( $m = 0$ ) separately, then have the desired extra assumption  $m \neq 0$  in the main proof and continue from there.

9. Prove or disprove that: For all positive integers  $k, m, n$ ,

$$k \mid (m \cdot n) \implies k \mid m \vee k \mid n$$

We disprove it by means of a counterexample. Choose  $m = n = 2$  and  $k = 4$ . Then  $k \mid m \cdot n$ , yet neither  $k \mid m$  nor  $k \mid n$ .

♪ It may sometimes be easier to disprove the contrapositive statement, since an implication holds if and only if its contrapositive holds.

10. Let  $P(m)$  be a statement for  $m$  ranging over the natural numbers, and consider the following derived statement (with  $n$  also ranging over the natural numbers):

$$P^\#(n) \triangleq \forall k \in \mathbb{N}. 0 \leq k \leq n \implies P(k)$$

- a) Show that, for all natural numbers  $\ell$ ,  $P^\#(\ell) \implies P(\ell)$ .

Let  $\ell$  be a natural number, and assume that

$$P^\#(\ell) = (\forall \text{ natural number } k. 0 \leq k \leq \ell \implies P(k))$$

holds.

Since  $\ell$  is a natural number, it follows by instantiation that

$$0 \leq \ell \leq \ell \implies P(\ell)$$

and, since  $0 \leq \ell \leq \ell$  is true by reflexivity of  $\leq$ , it follows by Modus Ponens that  $P(\ell)$  holds as required.

♪ This last exercise starts to trip some students up, understandably: so far we've been proving properties about numbers and divisibility, while now we're proving things about seemingly nothing in particular. Such abstract proofs are very common


in mathematics, for the very simple reason that they can be applied in a huge number of ways – in this case,  $P(m)$  can be *any* logical statement about natural numbers, and the propositions will hold no matter how simple or complicated the definition of  $P$  is! You may think of this as a “polymorphic” theorem, since we are proving something about an arbitrary predicate  $P$  (any first-class function `nat -> bool`, if you will), but as a consequence, we cannot assume anything about how it’s defined.

Thinking abstractly takes some getting used to, as you may feel like there isn’t anything to go on or any familiar notion to grasp in order to build intuition. However, abstractness has the major benefit of avoiding any distracting details and low-level “fluff” that could lead the proof attempt astray. If the above proposition was specialised to  $P(m)$  meaning “ $m$  is even”, you might start by unwrapping the definition of evenness and incorporate it into the proof somehow, despite the property holding no matter what  $P(m)$  actually is. Abstract proofs like this often involve purely logical reasoning, without invoking any number theory or algebra – and logical reasoning is often easier, since we essentially have an algorithm for proving logical statements. Thus, when you are faced with an incomprehensible jumble of logical symbols, the task may well be easier than proving a simple statement about natural numbers!

- b) Exhibit a concrete statement  $P(m)$  and a specific natural number  $n$  for which the following statement *does not* hold:

$$P(n) \implies P^\#(n)$$

Let  $P(m) \triangleq (m = 1)$  and  $n = 1$ . Then  $P(1)$  is the true proposition  $(1 = 1)$ , but  $P^\#(1) \iff P(0) \wedge P(1)$  is equivalent to  $(0 = 1) \wedge (1 = 1)$  which is false.

 Here we actually needed to “decode” the definition of  $P^\#$  in order to find a way to falsify the above statement. Fortunately this is not too difficult in this case:  $P^\#(n)$  holds if  $P(k)$  holds for all naturals less than or equal to  $n$ , essentially turning a predicate  $P$  about naturals into a predicate  $P^\#$  about a finite collection of naturals (similarly to how `map` turns a function on values into a function on lists of values). Then, we need to find a predicate  $P$  and  $n \in \mathbb{N}$  that does not satisfy  $P(n) \implies P^\#(n)$ . This is trickier than just finding a number, since there are lots of ways we could define  $P$ . But, once again, we try something very simple ( $P(m)$  holds for  $m = 1$  only) and find that it can easily be turned into a counterexample. There are lots of other options for  $P$  of course, but there’s no need to try something convoluted or interesting to get a contradiction (and equally, there’s no need to spend time finding the *simplest* counterexample if you’ve already found a more complicated one).

- c) Prove the following:

- $P^\#(0) \iff P(0)$

$(\Rightarrow)$  Assume  $P^\#(0)$ ; that is, for all  $0 \leq k \leq 0$ ,  $P(k)$ . As  $0 \leq 0 \leq 0$ ,  $P(0)$  holds.  
 $(\Leftarrow)$  Assume  $P(0)$ . Consider any  $k$ , and assume  $0 \leq k \leq 0$ . Then,  $k = 0$  and  $P(k)$  holds by assumption.

$$\bullet \forall n \in \mathbb{N}. (P^\#(n) \Rightarrow P^\#(n+1)) \iff (P^\#(n) \Rightarrow P(n+1))$$

$(\Rightarrow)$  Assume that  $(P^\#(n) \Rightarrow P^\#(n+1))$ , and further assume that  $P^\#(n)$  holds. Then, it follows that also  $P^\#(n+1)$  holds; i.e. that

$$\forall \text{ natural number } k. 0 \leq k \leq n+1 \Rightarrow P(k).$$

In particular, by instantiation, we have that

$$0 \leq n+1 \leq n+1 \Rightarrow P(n+1)$$

and since the antecedent of this implication is true, we deduce that  $P(n+1)$  holds, as required.


$(\Leftarrow)$  Assume that ①  $(P^\#(n) \Rightarrow P(n+1))$ , and further assume that ②  $P^\#(n)$  holds. We need show that  $P^\#(n+1)$  also holds; i.e. that

$$\forall \text{ natural number } k. 0 \leq k \leq n+1 \Rightarrow P(k).$$

or, equivalently, that

$$P^\#(n) \wedge P(n+1)$$

hold, which is indeed the case because  $P^\#(n)$  holds by assumption ② and  $P(n+1)$  follows by Modus Ponens from assumptions ① and ②.

 This is the most complicated statement in this exercise sheet, so do not worry if had difficulties with it. We have universal quantification, bi-implication, nested implication, and unwrapping the definition of  $P^\#$  gives another layer of quantification and implication. You may take a minute to get a feel for what the statement is saying, but the nice thing about purely logical proofs is that you can often dive in head-first without really thinking about what you're proving!

Look at the top-level construct (universal quantification, bi-implication, etc.), apply the proof pattern for that construct (often giving you some assumptions), and continue until your goal becomes some atomic statement like  $P(n+1)$  (which you can't unwrap further, since you don't know what  $P$  is). After "digesting" the proof goal, you should have a bunch of assumptions that you can work with: unwrapping some definitions, instantiating universals, applying Modus Ponens. Eventually you should end up with an assumption that matches the atomic proof goal, and that's enough to conclude the proof.

At first, doing the "assume | prove"-style scratchwork is very helpful for practicing proof patterns and keeping track of goals and assumptions. It should mostly

feel like an algorithmic process: with a few rule applications you can turn a very scary-looking formula into a primitive goal and a lot of assumptions to work with, and the task usually boils down to finding a way of combining assumptions on the LHS to get something that matches the RHS. Occasionally there is a small bit of actual “thinking” required to make progress, such as finding an appropriate value to instantiate a  $\forall$  with, or transforming an assumption in some useful way: in the  $(\Leftarrow)$  direction above, really the only clever bit was figuring out that

$$\forall \text{ natural number } k. 0 \leq k \leq n+1 \implies P(k).$$

is equivalent to

$$P^\#(n) \wedge P(n+1)$$

but even this step was not made in a vacuum, since we already had the assumptions  $P^\#(n)$  and  $P(n+1)$ . With some practice these methodical proofs should become second nature, and you will be able to keep track of things in your head, directly writing down the formal proof without any prior scratchwork.

$$\bullet (\forall m \in \mathbb{N}. P^\#(m)) \iff (\forall m \in \mathbb{N}. P(m))$$

$(\Rightarrow)$  Assume that  $\forall$  natural number  $m. P^\#(m)$ , and let  $n$  be an arbitrary natural number. Then, by assumption,  $P^\#(n)$  holds; that is

$$\forall \text{ natural number } k. 0 \leq k \leq n \implies P(k)$$


and, by instantiation,  $0 \leq n \leq n \implies P(n)$  so that  $P(n)$  holds. Thus, we have shown

$$\forall \text{ natural number } M. P(M)$$

$(\Leftarrow)$  Assume that ①  $\forall$  natural number  $M. P(M)$ . We need show that for all natural numbers  $m$  and  $k$ ,

$$0 \leq k \leq m \implies P(k)$$

To this end, let  $m$  and  $k$  be arbitrary natural numbers, and assume  $0 \leq k \leq m$ . Since  $k$  is a natural number, we may instantiate assumption ① with it yielding  $P(k)$  as required.

 This theorem may seem both surprising and unsurprising. Even though  $P^\#(m)$  is definitely more general than  $P(m)$  (since  $P^\#(m)$  implies  $P(m)$  but not vice versa), in the “limit” of quantifying over *all* natural numbers, they become equivalent. Then again, if  $P(m)$  holds for all natural numbers  $m$ , of course it would hold for all natural numbers smaller than any  $n$ ! This theorem (and the properties proved as part of this exercise) form the basis of an important proof technique which will be discussed later in the course.

### 1.3. Optional exercises

1. A series of questions about the properties and relationship of triangular and square numbers (adapted from David Burton).

- a) A natural number is said to be *triangular* if it is of the form  $\sum_{i=0}^k i = 0 + 1 + \cdots + k$ , for some natural  $k$ . For example, the first three triangular numbers are  $t_0 = 0$ ,  $t_1 = 1$  and  $t_2 = 3$ .

Find the next three triangular numbers  $t_3$ ,  $t_4$  and  $t_5$ .

$$t_3 = 6, t_4 = 10, t_5 = 15.$$

- b) Find a formula for the  $k^{\text{th}}$  triangular number  $t_k$ .

**Geometric approach.**

$$2 \cdot t_k = \begin{array}{ccc} \circ & & \bullet \quad \cdots \quad \bullet \\ \circ & \circ & \\ \vdots & & \vdots \\ \circ & \cdots & \circ \end{array} + \begin{array}{ccc} & \bullet & \\ & \vdots & \\ & \bullet & \end{array} = \begin{array}{ccc} \circ & \bullet & \cdots \quad \bullet \\ \circ & \circ & \bullet \quad \bullet \\ \vdots & & \vdots \\ \circ & \cdots & \circ \quad \bullet \end{array} = k \cdot (k+1)$$

**Algebraic approach.**

Note that, on the one hand,

$$\begin{aligned} \sum_{i=0}^k (i+1)^2 - \sum_{i=0}^k i^2 &= (k+1)^2 + \left( \sum_{i=0}^{k-1} (i+1)^2 \right) - \left( \sum_{i=1}^k i^2 \right) - 0^2 \\ &= (k+1)^2 \end{aligned}$$

and that, on the other,

$$\begin{aligned} \sum_{i=0}^k (i+1)^2 - \sum_{i=0}^k i^2 &= \sum_{i=0}^k ((i+1)^2 - i^2) \\ &= \sum_{i=0}^k (2 \cdot i + 1) \\ &= \left( 2 \cdot \sum_{i=0}^k i \right) + \sum_{i=0}^k 1 \\ &= 2 \cdot t_k + k + 1 \end{aligned}$$

$$\text{so that } t_k = \frac{k^2+k}{2}.$$

- c) A natural number is said to be *square* if it is of the form  $k^2$  for some natural number  $k$ .

Show that  $n$  is triangular iff  $8 \cdot n + 1$  is a square. (Plutarch, circ. 100BC)

( $\Rightarrow$ ) Assume  $n$  is triangular; i.e.  $n = t_k$  for some natural number  $k$ . By the previous item,  $n = \frac{k \cdot (k+1)}{2}$  and one has that  $8 \cdot n + 1 = (2 \cdot k + 1)^2$  is a square number.

( $\Leftarrow$ ) Assume that  $8 \cdot n + 1$  is a square number; i.e.  $8 \cdot n + 1 = a^2$  for some natural number  $a$ . Then  $a^2$  is odd and, by [Proposition 12](#) of the notes, thus so is  $a$ . Therefore,  $a = 2 \cdot k + 1$

for some natural number  $k$ . Finally, since  $8 \cdot n + 1 = a^2 = (2 \cdot k + 1)^2 = 4 \cdot k^2 + 4 \cdot k + 1$  one has  $n = \frac{k^2+k}{2} = t_k$  as required.

- d) Show that the sum of every two consecutive triangular numbers is square. (Nicomachus, circ. 100BC)

Consider any two consecutive triangular numbers  $t_k$  and  $t_{k+1}$ . Then, a simple calculation shows that the sum  $t_k + t_{k+1}$  equals  $(k+1)^2$  and hence is square:

$$\frac{k^2+k}{2} + \frac{(k+1)^2+k+1}{2} = \frac{2k^2+4k+2}{2} = k^2+2k+1 = (k+1)^2$$

- e) Show that, for all natural numbers  $n$ , if  $n$  is triangular, then so are  $9 \cdot n + 1$ ,  $25 \cdot n + 3$ ,  $49 \cdot n + 6$  and  $81 \cdot n + 10$ . (Euler, 1775)

Consider any natural number  $n$ , and assume that  $n$  is triangular; i.e.  $n = \frac{k \cdot (k+1)}{2}$  for some natural number  $k$ . Then, calculate that  $9 \cdot n + 1 = t_{3k+1}$ :

$$9 \frac{k^2+k}{2} + 1 = \frac{9k^2+9k+2}{2} = \frac{9k^2+6k+1+3k+1}{2} = \frac{(3k+1)^2+3k+1}{2} = t_{3k+1}$$

Similarly, by completing the square, we can show that  $25 \cdot n + 3 = t_{5k+2}$ ,  $49 \cdot n + 6 = t_{7k+3}$ , and  $81n + 10 = t_{9k+4}$ .

- f) Prove the generalisation: For all  $n$  and  $k$  natural numbers, there exists a natural number  $q$  such that  $(2n+1)^2 \cdot t_k + t_n = t_q$ . (Jordan, 1991, attributed to Euler)

Here's a proof by a 2014/15 student (who wished to remain anonymous). Let  $n$  and  $k$  be arbitrary natural numbers. We know that:

$$t_k = \frac{k(k+1)}{2} \quad \text{and} \quad t_n = \frac{n(n+1)}{2}$$

Choose  $q = 2nk + n + k$ , and calculate:

$$\begin{aligned} t_q &= \frac{q(q+1)}{2} = \frac{(2nk+n+k) \cdot (2nk+n+k+1)}{2} \\ &= \frac{4n^2k^2+4n^2k+4nk^2+4nk+k^2+k+n^2+n}{2} \\ &= \frac{(4n^2+4n+1)(k^2+k)+n^2+n}{2} \\ &= (2n+1)^2 \cdot \frac{k(k+1)}{2} + \frac{n(n+1)}{2} \\ &= (2n+1)^2 t_k + t_n \end{aligned}$$

Therefore we are done.

2. Let  $P(x)$  be a predicate on a variable  $x$  and let  $Q$  be a statement not mentioning  $x$ . Show that the following equivalence holds:

$$((\exists x. P(x)) \implies Q) \iff (\forall x. (P(x) \implies Q))$$

$(\Rightarrow)$  Assume  $(\exists x. P(x)) \Rightarrow Q$ . We need show  $\forall x. (P(x) \Rightarrow Q)$ . We do this by considering an arbitrary  $a$  and showing that  $P(a) \Rightarrow Q$ , for which in turn we further assume  $P(a)$  and finally show  $Q$ .

To recap, then, we are in the following situation:

Assumptions	Goal
$(\exists x. P(x)) \Rightarrow Q$	$Q$
for arbitrary $a$	
$P(a)$	


Then, by the last assumption,  $\exists x. P(x)$  and from this and the first assumption, by Modus Ponens, we deduce  $Q$  as required.

$(\Leftarrow)$  Assume  $\forall x. (P(x) \Rightarrow Q)$ . We need show  $(\exists x. P(x)) \Rightarrow Q$ . For which we further assume  $\exists x. P(x)$  and show  $Q$

To recap, then, we are in the following situation:

Assumptions	Goal
$\forall x. (P(x) \Rightarrow Q)$	$Q$
$\exists x. P(x)$	

From the second assumption, there is an  $a$  for which ①  $P(a)$  holds and, by instantiation from the first assumption, ②  $P(a) \Rightarrow Q$ . By Modus Ponens from ② and ①,  $Q$  follows as required.

 This is a very important duality that crops up in many different forms in mathematics and computer science (and you will certainly encounter variants of it in future courses). Despite this, it may seem quite unintuitive: it almost seems to say that we can convert existential quantification into universal! Of course, we can't ignore the shifting of the parentheses: it's certainly *not* the case that

$$(\exists x. P(x) \Rightarrow Q) \Longleftrightarrow (\forall x. P(x) \Rightarrow Q)$$

A good way to get an intuition for this property is as a generalisation of case analysis. If a property  $Q$  depends on the existence of a witness  $x$  satisfying  $P(x)$ , but not  $x$  itself, we need to prove  $Q$  no matter what  $x$  is. That is, our proof must hold for any actual value of the witness, so we can instead look at what possible values can  $x$  take, and show  $Q$  by assuming  $P(x)$  for *all* values  $x$ . We are case analysing the potential values of the witness, and proving  $Q$  no matter what it is.

Alternatively, we can look at the contrapositives of both sides:

$$(\neg Q \Rightarrow \neg(\exists x. P(x))) \Longleftrightarrow (\forall x. (\neg Q \Rightarrow \neg P(x)))$$

The LHS becomes  $\neg Q \Rightarrow (\forall x. \neg P(x))$  using the de Morgan rule for quantifiers; but now



the universal can be extended over the whole implication, since assuming  $\neg Q$  first and then taking an arbitrary  $x$  is the same as taking an arbitrary  $x$  and then assuming  $\neg Q$  (which doesn't say anything about  $x$ ).

## 2. On numbers

### 2.1. Basic exercises

1. Let  $i, j$  be integers and let  $m, n$  be positive integers. Show that:

a)  $i \equiv i \pmod{m}$

By §1.2.1(b), every number divides  $i - i = 0$ , so  $m \mid i - i$ .

b)  $i \equiv j \pmod{m} \implies j \equiv i \pmod{m}$

Assume  $i \equiv j \pmod{m}$ . Then  $m \mid i - j$ ; i.e.  $i - j = k \cdot m$  for some integer  $k$ . Thus,  $j - i = (-k) \cdot m$ , and as  $-k$  is an integer  $m \mid j - i$ ; i.e.  $j \equiv i \pmod{m}$ .

c)  $i \equiv j \pmod{m} \wedge j \equiv k \pmod{m} \implies i \equiv k \pmod{m}$

Assume  $i \equiv j \pmod{m} \wedge j \equiv k \pmod{m}$ . Then,  $m \mid i - j$  and  $m \mid j - k$ . Hence, by §1.2.6(a),  $m \mid (i - j) + j - k = i - k$  and thus  $i \equiv k \pmod{m}$ .

🎵 When working with congruence, we have three layers of definitions:  $i \equiv j \pmod{m}$ , defined as  $m \mid i - j$ , defined as  $\exists k \in \mathbb{Z}. i - j = k \cdot m$ . To prove fundamental properties about congruence (symmetry or transitivity), we usually need to go “down a level” and reason about divisibility. At this level, we may be able to use known properties of divisibility, such as in part (c); other times it may be easier to go further down, and talk about the primitive definition of divisibility, such as in part (b). In the second case we are essentially proving a lemma about divisibility “inline”: that  $d \mid m$  implies  $d \mid -m$ . Alternatively, we may notice that this property follows as a direct corollary of §1.2.6(b), with the multiplicative constant  $k = -1$ . The statement we prove is valid either way, but in some cases writing a quick inline proof may be easier or harder than finding if it is an instance of some existing property.

2. Prove that for all integers  $i, j, k, l, m, n$  with  $m$  positive and  $n$  nonnegative,

a)  $i \equiv j \pmod{m} \wedge k \equiv l \pmod{m} \implies i + k \equiv j + l \pmod{m}$

Assume  $i \equiv j \pmod{m} \wedge k \equiv l \pmod{m}$ . Then,  $m \mid i - j$  and  $m \mid k - l$ . Hence, by §1.2.6(a),  $m \mid (i - j) + (k - l) = (i + k) - (j + l)$  and  $i + k \equiv j + l \pmod{m}$ .

b)  $i \equiv j \pmod{m} \wedge k \equiv l \pmod{m} \implies i \cdot k \equiv j \cdot l \pmod{m}$

Assume  $i \equiv j \pmod{m} \wedge k \equiv l \pmod{m}$ . Then,  $m \mid (i - j)$  and  $m \mid (k - l)$ . By §1.2.6(b),  $m \mid i \cdot (k - l)$  and  $m \mid l \cdot (i - j)$ ; and, by §1.2.6(a),  $m \mid i \cdot (k - l) + l \cdot (i - j) = i \cdot k - j \cdot l$ . Hence,  $i \cdot k \equiv j \cdot l \pmod{m}$ .

c)  $i \equiv j \pmod{m} \implies i^n \equiv j^n \pmod{m}$

For  $n = 0$ ,  $i^n \equiv j^n \pmod{m}$  always. Assume now ①  $i \equiv j \pmod{m}$ . Then, for  $n = 1$ , we are done by assumption. For  $n = 2$ , by the previous item, we have ②  $i^2 \equiv j^2 \pmod{m}$ . From ① and ②, again by the previous item, we have  $i^3 \equiv j^3 \pmod{m}$ . Iterating this process we get  $i^n \equiv j^n \pmod{m}$  for every value of  $n$ .

🎵 If you're familiar with it, you may be screaming "induction!" – indeed, a formal proof requires the mathematical Principle of Induction, which will be studied later in the course.

🎵 These properties of congruence are fairly simple to state and prove, but combined with the previous exercise they form the basis of equational proofs about congruence. They allow us to extend a congruence between two integers into a congruence between two algebraic (polynomial) expressions of arbitrary nesting which differ in those two integers. For example, if we know that  $i \equiv j \pmod{m}$ , we also know  $(3i^2 + 5i - 7)^4 \equiv (3j^2 + 5j - 7)^4 \pmod{m}$  by repeatedly applying the properties proved in this exercise:  $i \equiv j \pmod{m}$  implies  $i^2 \equiv j^2 \pmod{m}$  implies  $3i^2 \equiv 3j^2 \pmod{m}$  and so on. This is really helpful in equational proofs in modular arithmetic, because we can rewrite parts of an expression not only if they are equal, but also when they are merely congruent. We will see examples of this shortly.

3. Prove that for all natural numbers  $k, l$  and positive integers  $m$ ,

a)  $\text{rem}(k \cdot m + l, m) = \text{rem}(l, m)$

By the Division Theorem,

$$l = \text{quo}(l, m) \cdot m + \text{rem}(l, m)$$

and hence

$$k \cdot m + l = (k + \text{quo}(l, m)) \cdot m + \text{rem}(l, m)$$

from which it follows by the Division Theorem that

$$\text{quo}(k \cdot m + l, m) = k + \text{quo}(l, m) \quad \text{and} \quad \text{rem}(k \cdot m + l, m) = \text{rem}(l, m).$$

🎵 The [Division Theorem](#) may seem like a dramatic name for a fairly obvious and unremarkable statement: that numbers can be divided with a remainder. But, in fact, the theorem is quite powerful and allows one to prove properties surprisingly easily. Let's remind ourselves of the full statement:

For every natural number  $m$  and positive natural number  $n$ , there exists a unique pair of integers  $q$  and  $r$  such that  $0 \leq q$ ,  $0 \leq r < n$ , and  $m = q \cdot n + r$ .

This is a *unique existence* statement, a form very common in mathematics. The associated proof technique relies both on the existence and uniqueness components. To highlight the former, consider the following alternative statement of the Division

**Theorem:**

Given any natural number  $m$  and for any choice of positive integer  $n$ , we can write  $m$  as  $m = q \cdot n + r$  where  $q$  and  $r$  are unique integers satisfying  $0 \leq q$  and  $0 \leq r < n$ .

This form emphasises the fact that if we have a natural number  $m$ , we can choose *any* natural number  $n$ , and the theorem guarantees that it's possible to write  $m$  in terms of  $n$  in the specific form  $m = q \cdot n + r$  for two unique naturals satisfying  $0 \leq q$  and  $0 \leq r < n$ . In essence, we get immediate “access” to two naturals  $q$  and  $r$  and two new assumptions about these naturals, as well as their uniqueness proofs.

Since  $q$  and  $r$  are uniquely determined by  $m$  and  $n$ , we can write them as  $\text{quo}(m, n)$  and  $\text{rem}(m, n)$  as if  $\text{quo}$  and  $\text{rem}$  were functions. In reality, they are just shorthands for “the natural  $q$  (resp.  $r$ ) determined from  $m$  and  $n$  by the Division Theorem”. With these, you can succinctly state the Division Theorem as

Any natural number  $m$  can be expressed as  $m = \text{quo}(m, n) \cdot n + \text{rem}(m, n)$  for any choice of positive integer  $n$ , with  $\text{quo}(m, n), \text{rem}(m, n) \in \mathbb{N}$  and  $\text{rem}(m, n) < n$ .

You may well ask “why go through all this when we have the integer division and remainder operators”? Well, we haven't formally defined them yet (and one way to define them formally is precisely via  $\text{quo}$  and  $\text{rem}$ !), but even ignoring that, proofs using uniqueness wouldn't really work if we just treated  $\text{rem}$  and  $\text{quo}$  as operators. To see how this works, let's expand the solution to the question above.

We are required to show that for all natural numbers  $k, l$  and positive integers  $m$ ,  $\text{rem}(k \cdot m + l, m) = \text{rem}(l, m)$  – any multiple of  $m$  can be cancelled out in a remainder by  $m$ . If we think of  $\text{rem}$  as the remainder operator (e.g. `%` in Java), this seems obvious – but other than spelling out the details of division as repeated subtraction (the Division Algorithm), it's quite tricky to prove! Instead, as we said above,  $\text{rem}(l, m)$  should be treated as “the unique  $r$  determined by  $l$  and  $m$  by the Division Theorem”. This is where uniqueness comes in: we know that for any other expansion  $l = \text{quo}(l, m) \cdot m + r'$  with  $r' < m$ ,  $r'$  must be equal to  $\text{rem}(l, m)$ . Thus, equality of remainders can be derived from showing that they satisfy the same property: that they can appear in the same expansion of  $l$  (via  $m$ ) and are both strictly less than  $m$ .

The question is exactly a proof of equality of two remainders:  $\text{rem}(k \cdot m + l, m)$  and  $\text{rem}(l, m)$ . If we show that they appear in two “different” expansions of the same natural number, they must be equal. What expansion would  $\text{rem}(l, m)$  appear in? Easy: the Division Theorem tells us that  $l$  can always be rewritten in terms of  $m$  as

$$l = \text{quo}(l, m) \cdot m + \text{rem}(l, m)$$

Similarly,  $\text{rem}(k \cdot m + l, m)$  appears in the expansion

$$k \cdot m + l = \text{quo}(k \cdot m + l, m) \cdot m + \text{rem}(k \cdot m + l, m)$$

All we did is apply the streamlined form of the Division Theorem, expanding both  $l$  and  $k \cdot m + l$  in terms of  $m$ . They can't directly be compared yet, because they are expansions of different naturals. To resolve that, we just add  $k \cdot m$  to the first equation and factorise to get:

$$k \cdot m + l = (k + \text{quo}(l, m)) \cdot m + \text{rem}(l, m)$$

And with that, we are done! How? Well, we have two different expansions of the number  $k \cdot m + l$ : it's equal both to

$$\text{quo}(k \cdot m + l, m) \cdot m + \text{rem}(k \cdot m + l, m) \quad \text{and} \quad (k + \text{quo}(l, m)) \cdot m + \text{rem}(l, m)$$

where both  $\text{rem}(k \cdot m + l, m)$  and  $\text{rem}(l, m)$  are less than  $m$ . But the Division Theorem tells us that there is exactly one such expansion of  $k \cdot m + l$  possible, so these two remainders *cannot* be different! That is to say,

$$\text{rem}(k \cdot m + l, m) = \text{rem}(l, m)$$

which was precisely our proof goal.


Such surprising and abrupt conclusions are very much characteristic of proofs by *universal properties*: rather than proving equality directly, we show that both remainders satisfy the universal property (specified by the Division Theorem) of the same number  $k \cdot m + l$  and therefore must be equal. We will see a lot of examples of this in the course and the exercises: while many statements can be proved by alternative means, proofs by universal properties are often remarkably compact and elegant, achieving the same goal with only a few clever reasoning steps.

b)  $\text{rem}(k + l, m) = \text{rem}(\text{rem}(k, m) + l, m)$

Because

$$\begin{aligned} \text{rem}(k + l, m) &= \text{rem}(\text{quo}(k, m) \cdot m + \text{rem}(k, m) + l, m) && \text{(by DT on } k \text{ with } m) \\ &= \text{rem}(\text{rem}(k, m) + l, m) && \text{(by §2.1.3(a))} \end{aligned}$$

Note that, as a corollary,  $\text{rem}(k + l, m) = \text{rem}(\text{rem}(k, m) + \text{rem}(l, m), m)$ .

 The previous property of remainders is quite useful, especially in combination with the Division Theorem: since we have a choice of expanding  $k$  in terms of any positive integer, we can choose  $m$  to then ensure that the term  $\text{quo}(k, m) \cdot m$  – being a multiple of  $m$  – can be cancelled out.

c)  $\text{rem}(k \cdot l, m) = \text{rem}(k \cdot \text{rem}(l, m), m)$

Because

$$\begin{aligned} \text{rem}(k \cdot l, m) &= \text{rem}(k \cdot \text{quo}(l, m) \cdot m + k \cdot \text{rem}(l, m), m) && \text{(by DT on } l \text{ with } m) \\ &= \text{rem}(k \cdot \text{rem}(l, m), m) && \text{(by §2.1.3(a))} \end{aligned}$$

Note that, as a corollary,  $\text{rem}(k \cdot l, m) = \text{rem}(\text{rem}(k, m) \cdot \text{rem}(l, m), m)$ .

🎵 Once again, we start by expanding a natural in terms of  $m$ , then use part §2.1.3(a) to cancel the whole term. In this case, we choose  $l$ : this was guided by the need to end up with a  $\text{rem}(l, m)$ , which we wouldn't get by expanding  $k$ .

4. Let  $m$  be a positive integer.

a) Prove the associativity of the addition and multiplication operations in  $\mathbb{Z}_m$ ; that is:

$$\forall i, j, k \in \mathbb{Z}_m. (i +_m j) +_m k = i +_m (j +_m k) \quad \text{and} \quad (i \cdot_m j) \cdot_m k = i \cdot_m (j \cdot_m k)$$

Consider arbitrary  $i, j, k$  in  $\mathbb{Z}_m$ , and calculate as follows:

$$\begin{aligned} (i +_m j) +_m k &= \llbracket [i + j]_m + k \rrbracket_m && \text{(by definition of } +_m) \\ &= \text{rem}(\text{rem}(i + j, m) + k, m) && \text{(by definition of } \llbracket \cdot \rrbracket_m) \\ &= \text{rem}((i + j) + k, m) && \text{(by §2.1.3(b))} \\ &= \text{rem}(i + (j + k), m) && \text{(by associativity of addition)} \\ &= \text{rem}(i + \text{rem}(j + k, m), m) && \text{(by §2.1.3(b))} \\ &= \llbracket i + [j + k]_m \rrbracket_m && \text{(by definition of } \llbracket \cdot \rrbracket_m) \\ &= i +_m (j +_m k) && \text{(by definition of } +_m) \end{aligned}$$

Similarly, consider arbitrary  $i, j, k$  in  $\mathbb{Z}_m$ , and calculate as follows:

$$\begin{aligned} (i \cdot_m j) \cdot_m k &= \llbracket [i \cdot j]_m \cdot k \rrbracket_m && \text{(by definition of } \cdot_m) \\ &= \text{rem}(\text{rem}(i \cdot j, m) \cdot k, m) && \text{(by definition of } \llbracket \cdot \rrbracket_m) \\ &= \text{rem}((i \cdot j) \cdot k, m) && \text{(by §2.1.3(c))} \\ &= \text{rem}(i \cdot (j \cdot k), m) && \text{(by associativity of multiplication)} \\ &= \text{rem}(i \cdot \text{rem}(j \cdot k, m), m) && \text{(by §2.1.3(c))} \\ &= \llbracket i \cdot [j \cdot k]_m \rrbracket_m && \text{(by definition of } \llbracket \cdot \rrbracket_m) \\ &= i \cdot_m (j \cdot_m k) && \text{(by definition of } \cdot_m) \end{aligned}$$

🎵 When defining something in terms of an existing construction, its properties will often directly follow from the known properties of the underlying definition. In this case, associativity of  $+_m$  relies on the associativity of  $+$  in terms of which  $+_m$  is defined. However, we needed a lemma about addition and remainders to simplify the expressions until we can directly appeal to the associativity of  $+$ .

🎵 These are examples of *equational proofs*, a very common and useful technique for mathematical reasoning, generalising the algebraic calculations you are familiar with from school. Whenever we need to prove equality or equivalence of two mathematical objects (numbers, sets, functions, etc.), we can build it up as a chain of equalities, each rewriting some part of the expression via some known property, definition, or

lemma. There's often a symmetry in the proofs, nicely showcased in this exercise: the first half unwraps several layers of definitions and simplifies the resulting expressions; the second half does the same in reverse. Indeed, it's often helpful to write equational proofs starting from both ends, until they meet in the middle.


b) Prove that the additive inverse of  $k$  in  $\mathbb{Z}_m$  is  $[-k]_m$ .

We need show that  $k +_m [-k]_m \equiv 0 \pmod{m}$ ; and indeed, since

$$l \equiv [l]_m \pmod{m} \text{ for all } l \in \mathbb{Z}$$

one has that

$$k +_m [-k]_m = [k + [-k]_m]_m \equiv k + [-k]_m \equiv k + (-k) = 0 \pmod{m}$$

 This is an example of a *congruence proof*: a weaker form of an equational proof where some of the steps are not strict equalities, but congruences modulo  $m \in \mathbb{Z}^+$ . Since congruence is a so-called *equivalence relation* (it's reflexive, symmetric, and transitive, all proved in §2.1.1), a chain of congruences establishes a congruence between the endpoints. Reflexivity allows us to strengthen some of the congruences into equalities: in the example above,  $k +_m [-k]_m = [k + [-k]_m]_m$  is a strict equality, since it is the definition of  $+_m$ . Importantly, all congruences must be modulo the same  $m \in \mathbb{Z}^+$ , which is denoted at the end of the proof, ranging over the entire chain of congruences.

## 2.2. Core exercises

- Find an integer  $i$ , natural numbers  $k, l$  and a positive integer  $m$  for which  $k \equiv l \pmod{m}$  holds while  $i^k \equiv i^l \pmod{m}$  does not.

Take  $i = 2, k = 0, l = 3$ , and  $m = 3$ . Then,  $k = 0 \equiv 3 = l \pmod{3}$ , yet  $2^0 = 1 \not\equiv 8 = 2^3 \pmod{3}$ .

- Formalise and prove the following statement: A natural number is a multiple of 3 iff so is the number obtained by summing its digits. Do the same for the analogous criterion for multiples of 9 and a similar condition for multiples of 11.

For all natural numbers  $n$  and digits  $a_1, \dots, a_n$ ,

- $\left( \sum_{i=0}^n a_i \cdot 10^i \right) \equiv 0 \pmod{3} \iff \left( \sum_{i=0}^n a_i \right) \equiv 0 \pmod{3}$
- $\left( \sum_{i=0}^n a_i \cdot 10^i \right) \equiv 0 \pmod{9} \iff \left( \sum_{i=0}^n a_i \right) \equiv 0 \pmod{9}$
- $\left( \sum_{i=0}^n a_i \cdot 10^i \right) \equiv 0 \pmod{11} \iff \left( \sum_{i=0}^n (-1)^i \cdot a_i \right) \equiv 0 \pmod{11}$

The above follow from the following stronger statements

- $\left( \sum_{i=0}^n a_i \cdot 10^i \right) \equiv \left( \sum_{i=0}^n a_i \right) \pmod{3}$
- $\left( \sum_{i=0}^n a_i \cdot 10^i \right) \equiv \left( \sum_{i=0}^n a_i \right) \pmod{9}$

$$\cdot \left( \sum_{i=0}^n a_i \cdot 10^i \right) \equiv \left( \sum_{i=0}^n (-1)^i \cdot a_i \right) \pmod{11}$$

The rule for 3 uses the fact that  $10 \equiv 1 \pmod{3}$ , which, by the exponentiation property shown in §2.1.2(c), implies  $10^l \equiv 1 \pmod{3}$  for all  $l \in \mathbb{Z}^+$ . This can be applied in every term of the sum (since congruences can be applied within sums and products as shown in §2.1.2, reducing the  $10^l$  coefficients to 1. The technique works the same for divisibility by 9, since  $10 \equiv 1 \pmod{9}$ ; for 11, we notice that  $10^{2n} \equiv 1 \pmod{11}$ , but  $10^{2n+1} \equiv 10 \equiv -1 \pmod{11}$  for all  $n \in \mathbb{N}$ .

There are also other proofs. Below is one based on the Binomial Theorem, rather than on the theory of divisibility and/or congruences for the case of divisibility by 11. Please study it and re-adapt it to the cases of divisibility by 3 and by 9.

First we calculate that

$$\begin{aligned} \sum_{i=0}^n a_i \cdot 10^i &= \sum_{i=0}^n a_i \cdot (11 - 1)^i \\ &= \sum_{i=0}^n a_i \cdot \sum_{j=0}^i \binom{i}{j} \cdot 11^j \cdot (-1)^{i-j} \\ &= \sum_{i=0}^n a_i \cdot \left[ (-1)^i + 11 \cdot \sum_{j=1}^i \binom{i}{j} \cdot 11^{j-1} \cdot (-1)^{i-j} \right] \\ &= \left( \sum_{i=0}^n (-1)^i \cdot a_i \right) + 11 \cdot \left[ \sum_{i=1}^n a_i \cdot \sum_{j=1}^i \binom{i}{j} \cdot 11^{j-1} \cdot (-1)^{i-j} \right] \end{aligned}$$

and then argue as follows:

( $\Rightarrow$ ) Assume  $11 \mid \left( \sum_{i=0}^n a_i \cdot 10^i \right)$ ; so that  $\sum_{i=0}^n a_i \cdot 10^i = 11 \cdot k$  for some integer  $k$ . Then,

$$\sum_{i=0}^n (-1)^i \cdot a_i = 11 \cdot \left( k - \left[ \sum_{i=1}^n a_i \cdot \sum_{j=1}^i \binom{i}{j} \cdot 11^{j-1} \cdot (-1)^{i-j} \right] \right)$$

showing that  $11 \mid \left( \sum_{i=0}^n (-1)^i \cdot a_i \right)$ .


( $\Leftarrow$ ) Assume  $11 \mid \left( \sum_{i=0}^n (-1)^i \cdot a_i \right)$ ; so that  $\sum_{i=0}^n (-1)^i \cdot a_i = 11 \cdot l$  for some integer  $l$ . Then,

$$\sum_{i=0}^n a_i \cdot 10^i = 11 \cdot \left( l + \left[ \sum_{i=1}^n a_i \cdot \sum_{j=1}^i \binom{i}{j} \cdot 11^{j-1} \cdot (-1)^{i-j} \right] \right)$$

showing that  $11 \mid \sum_{i=0}^n a_i \cdot 10^i$ .

3. Show that for every integer  $n$ , the remainder when  $n^2$  is divided by 4 is either 0 or 1.

This is [Lemma 26](#) of the notes.

 The question here refers to the “intuitive” notions of division and remainder, but by recognising their connection to congruence we can refer to the known number-theoretic properties of modular arithmetic.

4. What are  $\text{rem}(55^2, 79)$ ,  $\text{rem}(23^2, 79)$ ,  $\text{rem}(23 \cdot 55, 79)$  and  $\text{rem}(55^{78}, 79)$ ?

$$\text{rem}(55^2, 79) = 23, \text{rem}(23^2, 79) = 55, \text{rem}(23 \cdot 55, 79) = 1, \text{ and}$$

$$\begin{aligned} \text{rem}((55^2)^{39}, 79) &= \text{rem}(23 \cdot (23^2)^{19}, 79) = \text{rem}(23 \cdot 55 \cdot (55^2)^9, 79) \\ &= \text{rem}(23 \cdot (23^2)^4, 79) = \text{rem}(23 \cdot (55^2)^2, 79) \\ &= \text{rem}(23 \cdot 23^2, 79) = \text{rem}(23 \cdot 55, 79) \\ &= 1 \end{aligned}$$


Of course, since we know the last one from [Fermat's Little Theorem](#), there was really no need to calculate it!

5. Calculate that  $2^{153} \equiv 53 \pmod{153}$ . At first sight this seems to contradict Fermat's Little Theorem, why isn't this the case though? *Hint:* Simplify the problem by applying known congruences to subexpressions using the properties in §2.1.2.

One possible sequence of steps, using the fact that  $153 = 2^7 + 25$ :

$$\begin{aligned} 2^{153} &= 2^6 \cdot (2^7)^{21} = 2^6 \cdot 2^7 \cdot (2^7)^{20} \\ &\equiv 2^6 \cdot (-25) \cdot (-25)^{20} = 2^6 \cdot (-25) \cdot (25^2)^{10} = 2^6 \cdot (-25) \cdot 625^{10} \\ &\equiv 2^6 \cdot (-25) \cdot (13^2)^5 = 2^6 \cdot (-25) \cdot 169^5 \\ &\equiv (-25) \cdot 2^6 \cdot 16^5 = (-25) \cdot 2^6 \cdot (2^4)^5 = (-25) \cdot 2^5 \cdot (2^7)^3 \\ &\equiv (-25) \cdot 2^5 \cdot (-25) \cdot 25^2 = 2^5 \cdot (25^2)^2 \\ &\equiv 2^5 \cdot 13^2 \equiv 2^5 \cdot 16 = 2^2 \cdot 2^7 \equiv 4 \cdot (-25) \\ &\equiv 53 \pmod{153} \end{aligned}$$

This doesn't contradict Fermat's Little Theorem, since  $153 = 3^2 \cdot 17$  is composite.

 This may seem like a daunting exercise, but we actually didn't need to do anything more complicated than squaring and addition. The key is being able to make impactful simplifications using congruence: as soon as we have a number greater than 153, we can replace it with the remainder after dividing by 153.

6. Calculate the addition and multiplication tables, and the additive and multiplicative inverses tables for  $\mathbb{Z}_3$ ,  $\mathbb{Z}_6$  and  $\mathbb{Z}_7$ .

•  $\mathbb{Z}_3$

+	0	1	2	·	0	1	2	−(·)	(·) <sup>−1</sup>
0	0	1	2	0	0	0	0	0	
1	1	2	0	1	0	1	2	1	1
2	2	0	1	2	0	2	1	2	2




•  $\mathbb{Z}_6$

+	0	1	2	3	4	5	·	0	1	2	3	4	5		$-(\cdot)$		$(\cdot)^{-1}$
0	0	1	2	3	4	5	0	0	0	0	0	0	0	0	0	0	
1	1	2	3	4	5	0	1	0	1	2	3	4	5	1	5	1	1
2	2	3	4	5	0	1	2	0	2	4	0	2	4	2	4	2	
3	3	4	5	0	1	2	3	0	3	0	3	0	3	3	3	3	
4	4	5	0	1	2	3	4	0	4	2	0	4	2	4	2	4	
5	5	0	1	2	3	4	5	0	5	4	3	2	1	5	1	5	5

•  $\mathbb{Z}_7$


+	0	1	2	3	4	5	6	·	0	1	2	3	4	5	6		$-(\cdot)$		$(\cdot)^{-1}$
0	0	1	2	3	4	5	6	0	0	0	0	0	0	0	0	0	0	0	
1	1	2	3	4	5	6	0	1	0	1	2	3	4	5	6	1	6	1	1
2	2	3	4	5	6	0	1	2	0	2	4	6	1	3	5	2	5	2	4
3	3	4	5	6	0	1	2	3	0	3	6	2	5	1	4	3	4	3	5
4	4	5	6	0	1	2	3	4	0	4	1	5	2	6	3	4	3	4	2
5	5	6	0	1	2	3	4	5	0	5	3	1	6	4	2	5	2	5	3
6	6	0	1	2	3	4	5	6	0	6	5	4	3	2	1	6	1	6	6

 Great demonstration of the property that every element of  $\mathbb{Z}_p$  has a multiplicative inverse if  $p$  is a prime. Algebraically, this makes  $\mathbb{Z}_p$  a *field*: a place where you can do division.

7. Let  $i$  and  $n$  be positive integers and let  $p$  be a prime. Show that if  $n \equiv 1 \pmod{p-1}$  then  $i^n \equiv i \pmod{p}$  for all  $i$  not multiple of  $p$ .

Assume that  $i$  and  $n$  are positive integers and that  $p$  is a prime. Assume further that  $n \equiv 1 \pmod{p-1}$ ; so that  $n-1 = k \cdot (p-1)$  for some *natural number*  $k$ . Then, for  $i$  not a multiple of  $p$ , we have that

$$\begin{aligned}
 i^n &= i \cdot (i^{p-1})^k \\
 &\equiv i \cdot 1^k \pmod{p} && \text{(by Fermat's Little Theorem)} \\
 &= i
 \end{aligned}$$

 When the question involves prime numbers, you should expect to require properties and theorems specific to primes. In this course – which is only an introduction to number theory – this will quite often be Fermat's Little Theorem.

8. Prove that  $n^3 \equiv n \pmod{6}$  for all integers  $n$ .

We can proceed by case analysis: since either  $n \equiv 0 \pmod{6}$ , or  $n \equiv 1 \pmod{6}$ , or  $n \equiv 2 \pmod{6}$ , or  $n \equiv 3 \pmod{6}$ , or  $n \equiv 4 \pmod{6}$ , or  $n \equiv 5 \pmod{6}$ , we check that  $n^3 \equiv n \pmod{6}$  in each case.

- Case  $n \equiv 0 \pmod{6}$ :  $n^3 \equiv 0^3 = 0 \equiv n \pmod{6}$ .
- Case  $n \equiv 1 \pmod{6}$ :  $n^3 \equiv 1^3 = 1 \equiv n \pmod{6}$ .
- Case  $n \equiv 2 \pmod{6}$ :  $n^3 \equiv 2^3 = 8 \equiv 2 \equiv n \pmod{6}$ .
- Case  $n \equiv 3 \pmod{6}$ :  $n^3 \equiv 3^3 = 27 \equiv 3 \equiv n \pmod{6}$ .
- Case  $n \equiv 4 \pmod{6}$ :  $n^3 \equiv 4^3 = 64 \equiv 4 \equiv n \pmod{6}$ .
- Case  $n \equiv 5 \pmod{6}$ :  $n^3 \equiv 5^3 = 125 \equiv 5 \equiv n \pmod{6}$ .

Of course, this wouldn't really work for larger moduli – see next question. A more elegant solution is proving  $6 \mid n^3 - n$ , which, by the well-known divisibility rule for 6, follows from showing  $3 \mid n^3 - n$  and  $2 \mid n^3 - n$ . Now,

$$n^3 - n = n \cdot (n^2 - 1) = (n - 1) \cdot n \cdot (n + 1);$$


but this is a product of three consecutive integers, so at least one of them must be even and one must be divisible by 3. That is,  $n^3 - n = 2 \cdot 3 \cdot k$  for some  $k \in \mathbb{Z}$ , so  $n^3 \equiv n \pmod{6}$ .

Yet another approach is formally establishing the lemma (which can be seen as the generalisation of the divisibility rule of 6):

$$(a \equiv b \pmod{2} \wedge a \equiv b \pmod{3}) \iff a \equiv b \pmod{6}$$

In one direction, we have that  $a = 2k + b = 3l + b$ , so  $2k = 3l$  for integers  $k$  and  $l$ ; since  $3l$  must be even and 3 is odd,  $l$  must itself be even:  $l = 2m$  for some  $m \in \mathbb{Z}$ . Substituting back, we have  $a = 3 \cdot 2m + b$ , so  $a - b = 6m$ . In the opposite direction,  $a - b = 6k = 2 \cdot 3 \cdot k$ , which immediately implies  $2 \mid a - b$  and  $3 \mid a - b$ .

Now, it is sufficient to prove that  $n^3 \equiv n \pmod{2}$  and  $n^3 \equiv n \pmod{3}$ . The latter is a direct instance of Fermat's Little Theorem for the prime 3; the former holds by the congruence chain  $n^3 \equiv n^2 \equiv n \pmod{2}$ , with both steps using Fermat's Little Theorem  $n^2 \equiv n \pmod{2}$ , multiplied by  $n$  in the first step using the product property of §2.1.2.

 There are usually many ways of approaching a proof, ranging from “brute force” methods to elegant and concise number-theoretic arguments. It doesn't technically matter what you do as long as the proof is correct – but just like how “working” code doesn't always mean “neat and readable” code, you should strive to make your proofs as streamlined as possible. It's also very useful to practice recognising patterns and realising where some known lemma or property can be applied, since they often end up doing the bulk of the work: you shouldn't need to reprove a specific case of a known, more general statement.

## 9. Prove that $n^7 \equiv n \pmod{42}$ for all integers $n$ .

An exhaustive case analysis would be impractical in this case. Instead, we adapt our more conceptual solutions above.

First, we use a very similar proof as above for the lemma

$$(a \equiv b \pmod{6} \wedge a \equiv b \pmod{7}) \iff a \equiv b \pmod{42}$$

(notice how the crucial step is  $6k = 7l$  implying that  $6 \mid l$ , because  $6 \nmid 7$  – the lemma wouldn't hold for non-coprime numbers (see §1.2.5). Another trick in this case is recognising that  $a - b = 7(a - b) - 6(a - b)$ , and, since by assumption  $a - b = 6k = 7l$ , we have  $a - b = 7 \cdot 6k - 6 \cdot 7l = 42 \cdot (k - l)$ ).

Now,  $n^7 \equiv n \pmod{7}$  holds by Fermat's Little Theorem. To show  $n^7 \equiv n \pmod{6}$ , we can equivalently show  $n^7 \equiv n \pmod{2}$  and  $n^7 \equiv n \pmod{3}$ ; both follow by repeated applications of Fermat's Little Theorem.

## 2.3. Optional exercises

1. Prove that for all integers  $n$ , there exist natural numbers  $i$  and  $j$  such that  $n = i^2 - j^2$  iff either  $n \equiv 0 \pmod{4}$  or  $n \equiv 1 \pmod{4}$  or  $n \equiv 3 \pmod{4}$ .

Consider an arbitrary integer  $n$ .

( $\Rightarrow$ ) Assume there exist natural numbers  $i$  and  $j$  such that  $n = i^2 - j^2$ . By [Proposition 25](#) of the notes, we have that

$$\text{either } i^2 \equiv 0 \pmod{4} \text{ or } i^2 \equiv 1 \pmod{4}$$

and

$$\text{either } j^2 \equiv 0 \pmod{4} \text{ or } j^2 \equiv 1 \pmod{4}$$

We therefore have four cases:

- $i^2 \equiv 0 \pmod{4}$  and  $j^2 \equiv 0 \pmod{4}$ , in which case  $n \equiv 0 \pmod{4}$ ;
- $i^2 \equiv 0 \pmod{4}$  and  $j^2 \equiv 1 \pmod{4}$ , in which case  $n \equiv -1 \equiv 3 \pmod{4}$ ;
- $i^2 \equiv 1 \pmod{4}$  and  $j^2 \equiv 0 \pmod{4}$ , in which case  $n \equiv 1 \pmod{4}$ ;
- $i^2 \equiv 1 \pmod{4}$  and  $j^2 \equiv 1 \pmod{4}$ , in which case  $n \equiv 0 \pmod{4}$ ;

Hence, either  $n \equiv 0 \pmod{4}$ , or  $n \equiv 1 \pmod{4}$ , or  $n \equiv 3 \pmod{4}$  as required.

( $\Leftarrow$ ) Assume that either  $n \equiv 0 \pmod{4}$ , or  $n \equiv 1 \pmod{4}$ , or  $n \equiv 3 \pmod{4}$ . We need to find natural numbers  $i$  and  $j$  such that  $n = i^2 - j^2$ .

Graphically, we want to show that one can distribute any number of balls (as long as it's congruent to 0, 1 or 3 modulo 4) in a square grid leaving an empty square sub-grid, for instance as follows (for  $i = 7$ ,  $j = 3$ , and  $n = 40$ ):

•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•				•
•	•	•				•
•	•	•				•
•	•	•	•	•	•	•

We split our analysis in three cases.

- Case  $n$  is zero.

There are natural numbers  $i = j = 0$  such that  $n = i^2 - j^2$ , and we are done.

- Case  $n$  is a non-zero even integer.

As  $\text{rem}(n, 4) = n - \text{quo}(n, 4) \cdot 4$  (by the Division Theorem), it follows that  $\text{rem}(n, 4)$  is even and since hence it is necessarily 0. Thus,  $n$  is in fact a non-zero multiple of 4; say of the form  $4 \cdot k$  for some non-zero integer  $k$ . Then,

$$n = (k+1)^2 - (k-1)^2 = (-k-1)^2 - (1-k)^2$$

and since either

$$k+1 \text{ and } k-1 \text{ are natural numbers}$$

or

$$-k-1 \text{ and } 1-k \text{ are natural numbers}$$

there are natural numbers  $i, j$  such that  $n = i^2 - j^2$ . (Note that this argument slightly generalises that of Proposition 22 of the notes.)

Graphically, we are in the following kind of situation:

• <sub>2</sub>	• <sub>2</sub>	• <sub>2</sub>	• <sub>2</sub>	• <sub>2</sub>	• <sub>2</sub>	• <sub>2</sub>	• <sub>3</sub>
• <sub>1</sub>							• <sub>3</sub>
• <sub>1</sub>							• <sub>3</sub>
• <sub>1</sub>							• <sub>3</sub>
• <sub>1</sub>							• <sub>3</sub>
• <sub>1</sub>							• <sub>3</sub>
• <sub>1</sub>							• <sub>3</sub>
• <sub>1</sub>	• <sub>4</sub>	• <sub>4</sub>	• <sub>4</sub>	• <sub>4</sub>	• <sub>4</sub>	• <sub>4</sub>	• <sub>4</sub>

- Case  $n$  is odd.

Then  $n = 2 \cdot k + 1$  for some integer  $k$ , and

$$n = (k+1)^2 - k^2 = (-k-1)^2 - (-k)^2.$$

Since either

$$k+1 \text{ and } k \text{ are natural numbers}$$

or

$$-k-1 \text{ and } -k \text{ are natural numbers}$$

there are natural numbers  $i, j$  such that  $n = i^2 - j^2$ .

Graphically, we are in the following kind of situation:

•	• <sub>2</sub>	• <sub>2</sub>	• <sub>2</sub>	• <sub>2</sub>	• <sub>2</sub>	• <sub>2</sub>
• <sub>1</sub>						
• <sub>1</sub>						
• <sub>1</sub>						
• <sub>1</sub>						
• <sub>1</sub>						
• <sub>1</sub>						

🎵 Graphical proofs are great for intuition: so called “proofs without words” are often as illuminating as they are beautiful. However, they are not (usually) a substitute for a formal proof by logical reasoning, especially if the proposition to be shown is more general than what could be encoded graphically. In this case, the statement is about all *integers*  $n$ , while the graphical proof can only work for a *natural number*  $n$ .

2. A *decimal* (respectively *binary*) *repunit* is a natural number whose decimal (respectively binary) representation consists solely of 1's.

a) What are the first three decimal repunits? And the first three binary ones?

The first three decimal repunits are 1, 11, and 111; while the first three binary repunits are 1, 3, and 7.

b) Show that no decimal repunit strictly greater than 1 is a square, and that the same holds for binary repunits. Is this the case for every base? *Hint:* Use [Lemma 26](#) of the notes.

Let  $n$  be a decimal repunit greater than 1; that is,  $n = \sum_{i=0}^l 10^i$  for some  $l \geq 1$ . Then,

$$n \equiv \sum_{i=0}^l 2^i \equiv 1 + 2 = 3 \pmod{4}$$

and, by [Proposition 25](#) of the notes, we deduce that  $n$  is not square.

Incidentally, the calculation above already contains the proof of the property for binary repunits, since they are of the form  $n = \sum_{i=0}^l 2^i$

The statement:

For every base  $r$ , there are no  $r$ -ary repunits greater than 1 that are square.


is false. As a counterexample, take the base  $r = 3$  and the 3-ary repunit 4 consisting of two 1's.

### 3. More on numbers

#### 3.1. Basic exercises

1. Calculate the set  $\text{CD}(666, 330)$  of common divisors of 666 and 330.

We have that  $666 = 2 \cdot 3^2 \cdot 37$  and  $330 = 2 \cdot 3 \cdot 5 \cdot 11$ . Hence,  $\text{CD}(666, 330) = \{1, 2, 3, 2 \cdot 3\} = \{1, 2, 3, 6\}$ .

 You may be familiar with this method of computing the common divisors of two numbers using their prime factorisation – this of course relies on the Fundamental Theorem of Arithmetic, introduced later in the course.

2. Find the gcd of 21212121 and 12121212.


We run [Euclid's Algorithm](#):

$$\begin{aligned}\gcd(21212121, 12121212) &= \gcd(12121212, 9090909) \\ &= \gcd(9090909, 3030303) \\ &= 3030303\end{aligned}$$

3. Prove that for all positive integers  $m$  and  $n$ , and integers  $k$  and  $l$ ,

$$\gcd(m, n) \mid (k \cdot m + l \cdot n)$$

Let  $m, n$  be positive integers and  $k, l$  be integers. As  $\gcd(m, n) \mid m$  and  $\gcd(m, n) \mid n$  it follows from §1.2.6(a) that  $\gcd(m, n) \mid k \cdot m$  and  $\gcd(m, n) \mid l \cdot n$ ; from which it further follows by §1.2.6(b) that  $\gcd(m, n) \mid (k \cdot m + l \cdot n)$ .

 Like  $\text{rem}$ , we can treat  $\gcd(m, n)$  as a function of two positive integers  $m$  and  $n$ , or as a symbol for the greatest common divisor of  $m$  and  $n$  defined using the universal property of gcds. For example, we make use of the fact that  $\gcd(m, n)$  is a common divisor of  $m$  and  $n$ , so we “automatically” get  $\gcd(m, n) \mid m$  and  $\gcd(m, n) \mid n$ . We will see more examples of this in the upcoming exercises.

4. Find integers  $x$  and  $y$  such that  $x \cdot 30 + y \cdot 22 = \gcd(30, 22)$ . Now find integers  $x'$  and  $y'$  with  $0 \leq y' < 30$  such that  $x' \cdot 30 + y' \cdot 22 = \gcd(30, 22)$ .

Run the [Extended Euclid's Algorithm](#) to find that  $\gcd(30, 22) = 2$  and  $x \cdot 30 + y \cdot 22 = 2$  for  $x = 3$  and  $y = -4$ . To get a  $y'$  between the range  $0 \leq y' < 30$ , we notice that

$$(x + 11 \cdot l) \cdot 30 + (y - 15 \cdot l) \cdot 22 = 2$$

for all integers  $l$  ([Slide 219](#)), and find a value  $l_0$  such that  $0 \leq y - 15 \cdot l_0 < 30$  setting  $x' = x + 11 \cdot l_0$  and  $y' = y - 15 \cdot l_0$ . The two options are  $l_0 = -1$  for  $(-8) \cdot 30 + 11 \cdot 22 = 2$ , and  $l_0 = -2$  for  $(-19) \cdot 30 + 26 \cdot 22 = 2$ .

5. Prove that for all positive integers  $m$  and  $n$ , there exists integers  $k$  and  $l$  such that  $k \cdot m + l \cdot n = 1$  iff  $\gcd(m, n) = 1$ .

( $\Rightarrow$ ) By [Corollary 62](#) of the notes: if 1 can be expressed as a linear combination of  $m$  and  $n$ , and  $\gcd(m, n)$  must divide any linear combination of  $m$  and  $n$ , we must have  $\gcd(m, n) = 1$ .

( $\Leftarrow$ ) By [Theorem 70](#) of the notes:  $\gcd(m, n)$  is a linear combination of  $m$  and  $n$ .

6. Prove that for all integers  $n$  and primes  $p$ , if  $n^2 \equiv 1 \pmod{p}$  then either  $n \equiv 1 \pmod{p}$  or  $n \equiv -1 \pmod{p}$ .

Assume  $n^2 \equiv 1 \pmod{p}$ . Then  $p$  divides  $n^2 - 1 = (n - 1) \cdot (n + 1)$ . By [Euclid's Theorem](#),  $p \mid (n - 1)$  or  $p \mid (n + 1)$ ; that is, either  $n \equiv 1 \pmod{p}$  or  $n \equiv -1 \pmod{p}$ .

### 3.2. Core exercises

1. Prove that for all positive integers  $m$  and  $n$ ,  $\gcd(m, n) = m$  iff  $m \mid n$ .

Let  $m$  and  $n$  be arbitrary positive integers.

( $\Rightarrow$ ) Assume that  $\gcd(m, n) = m$ . Then  $m$  is the greatest common divisor of both  $m$  and  $n$ , and in particular a divisor of  $n$ .

( $\Leftarrow$ ) Assume  $m \mid n$ .

Here are two arguments.

- a) We have that  $n = k \cdot m$  for some positive integer  $k$ , and hence that


$$\gcd(m, n) = \gcd(m, k \cdot m) = m \cdot \gcd(1, k) = m$$

where the second equality is a consequence of the linearity property ([Lemma 63\(3\)](#) of the notes) of  $\gcd$ .

- b) By [Theorem 61](#) of the notes, it suffices to prove that

- $m \mid m$  and  $m \mid n$ , and
- for all positive integers  $d$  such that  $d \mid m$  and  $d \mid n$  it necessarily follows that  $d \mid m$ ;

all of which hold trivially.

 It's worth analysing the second approach, as it's quite characteristic of proofs by universal properties: the proof just "pops out" without us having to do a whole lot of work, similar to our use of the Division Theorem in [§2.1.3\(a\)](#).

As mentioned in [§3.1.3](#), there are several equivalent ways of thinking about gcds. One is as a function of two positive integers  $m$  and  $n$ , computed via [Euclid's Algorithm](#); another is as a label for a unique number characterised by the universal property of being the greatest common divisor of  $m$  and  $n$ . The difference may seem insignificant, but that is precisely because of [Theorem 61](#), which states that the value computed by Euclid's Algorithm coincides with the greatest common divisor. The universal property of gcds (which we'll get to shortly) is the *specification* of what it is to be a greatest common divisor;

Theorem 61 states that Euclid's Algorithm satisfies the specification. We don't *define* the greatest common divisor of  $m$  and  $n$  as "the number returned by Euclid's Algorithm"; just as how we don't define a sorted list as "the list returned by the quicksort algorithm" or a lasagna as "the dish you get by following this specific recipe in this specific cookbook". We already know what a gcd/sorted list/lasagna is supposed to be, and we can then ask whether some algorithm computes the gcd or some recipe makes a lasagna, or it doesn't. Of course, what makes a lasagna and what is the *best* lasagna is entirely subjective, while mathematical concepts can be unambiguously characterised using universal properties.

Universal properties have two parts: the *property* and the *universality*. The former characterises the set of candidates for the concept we are considering; the latter selects a specific candidate which is "better" than all the other ones. In the case of the greatest common divisor of  $m$  and  $n$ , the property is that of being a common divisor of  $m$  and  $n$ : the set of candidates that satisfy this property is  $\text{CD}(m, n)$ . The "best" such candidate that we are looking for is the one which is greater than all the other ones, and since  $\text{CD}(m, n)$  is a finite non-empty set of natural numbers, it must have a unique greatest element  $\max(\text{CD}(m, n))$ . We can denote this element (which depends entirely on  $m$  and  $n$ ) as  $\text{gcd}(m, n)$  and call it the greatest common divisor of  $m$  and  $n$ .

From this description (or, really, definition) of  $\text{gcd}(m, n)$  as the greatest element of the set of common divisors, we can directly extract two "axioms":  $\text{gcd}(m, n) \in \text{CD}(m, n)$  (since it is a common divisor), and for all  $d \in \text{CD}(m, n)$ ,  $d \leq \text{gcd}(m, n)$  (since it is the greatest common divisor). In fact, we can state something stronger: not only are all other common divisors numerically greater than  $\text{gcd}(m, n)$ , they also all divide it:  $\forall d \in \text{CD}(m, n). d \mid \text{gcd}(m, n)$ . Expanding these, we universally characterise  $\text{gcd}(m, n)$  as the unique natural number  $g$  satisfying the properties of being a common divisor and a multiple of all common divisors:

$$\textcircled{1} g \mid m \wedge g \mid n \quad \textcircled{2} \forall d \in \mathbb{Z}^+. (d \mid m \wedge d \mid n) \implies d \mid g$$

Using the transitivity of divisibility (§1.2.4), we can combine these into the concise specification of the universal property of greatest common divisors:

$$\textcircled{3} \forall d \in \mathbb{Z}^+. (d \mid m \wedge d \mid n) \iff d \mid \text{gcd}(m, n)$$

It's easy to show that gcds are unique: if we had two gcds, both would have to satisfy  $\textcircled{2}$  and, in particular, they must divide each other; but divisibility (on positive integers) is antisymmetric (§1.2.8), so the two gcds must be equal. Uniqueness in turn gives rise to the following important proof principle:

To prove that a number  $g \in \mathbb{Z}^+$  is equal to  $\text{gcd}(m, n)$ ,  
it is sufficient to show that  $g$  satisfies  $\textcircled{1}$  and  $\textcircled{2}$ .

This is similar to the approach we used with the Division Theorem: to prove that a number  $r$  is equal to  $\text{rem}(m, n)$ , it was sufficient to show that it is less than  $n$  and it can appear in an expansion  $m = q \cdot k + r$  with  $q \in \mathbb{N}$ . Adapting this technique to the combined form  $\textcircled{3}$ ,



we get a useful and particularly simple variation:

To prove that a number  $d$  divides  $\gcd(m, n)$ , it's sufficient to show that  $d \mid m$  and  $d \mid n$ .

This, combined with the antisymmetry of divisibility (on positive integers), allows us to prove equality of gcds, as shown in the example proofs of [Lemma 63](#) in the notes. In essence, the first step in proving something about  $\gcd(m, n)$  or  $\text{rem}(n, m)$  is “forgetting” about the gcd or rem and approach the proof via the universal property; it may seem like a very roundabout technique (as opposed to, for example, a direct chain of equalities ending in  $\gcd(m, n)$ ), but it often leads to short and straightforward proofs. However, it's definitely not the case that *all* proofs about gcds have to be done this way, and we'll see more examples later!

To conclude the discussion, let us expand on proof (b) of this exercise, which uses the UP of gcds. To recap, in the ( $\Leftarrow$ ) direction we need to show:

$$\forall m, n \in \mathbb{Z}^+. m \mid n \implies \gcd(m, n) = m$$

As always, assume  $m, n \in \mathbb{Z}^+$  and  $m \mid n$ . The proof goal  $\gcd(m, n) = m$  asks us to show that  $m$  is equal to  $\gcd(m, n)$ ; but, by the proof principle above, it is sufficient to show that  $m$  satisfies ① and ②. That is,

$$\text{① } m \mid m \wedge m \mid n \quad \text{② } \forall d \in \mathbb{Z}^+. (d \mid m \wedge d \mid n) \implies d \mid m$$

① holds by reflexivity of  $\mid$  and our assumption  $m \mid n$ ; ② is a direct implication. And that's it! The proof (a) wasn't exactly complicated either, but (b) was rightly labelled as “trivial”.

The beautiful thing about this characterisation of gcds is that it is an instance of a much more general mathematical notion called a *greatest lower bound* (with the dual *least upper bound* being the least common multiple). These concepts appear all over mathematics and computer science, and you will encounter many examples in this course as well; accordingly, the proof technique described above can be (and will be, and has already been!) applied in several seemingly different contexts. As a teaser, see if you can spot the similarity between statement ③ above, and the pattern for proving a conjunction of two statements  $P$  and  $Q$  given any set  $A$  of assumptions:

$$\forall A. (A \Rightarrow P) \wedge (A \Rightarrow Q) \iff A \Rightarrow (P \wedge Q)$$

2. Let  $m$  and  $n$  be positive integers with  $\gcd(m, n) = 1$ . Prove that for every natural number  $k$ ,

$$m \mid k \wedge n \mid k \iff m \cdot n \mid k$$

Let  $m$  and  $n$  be arbitrary positive integers, and assume that ①  $\gcd(m, n) = 1$ . Further, let  $k$  be a natural number.

( $\Rightarrow$ ) Assume that ②  $m \mid k$  and ③  $n \mid k$ .

It follows from ① that

$$m \cdot i + n \cdot j = 1 \quad \text{④}$$

for some integers  $i, j$ ; and it follows from ② and ③ that

$$k = a \cdot m = b \cdot n \quad \text{⑤}$$


for some natural numbers  $a, b$ .

Multiplying ④ by  $k$  on both sides and using ⑤, we therefore have

$$k = b \cdot n \cdot m \cdot i + a \cdot m \cdot n \cdot j = (b \cdot i + a \cdot j) \cdot (m \cdot n)$$

showing that  $(m \cdot n) \mid k$ .

( $\Leftarrow$ ) Assume that  $(m \cdot n) \mid k$ . Then, since both  $m \mid (m \cdot n)$  and  $n \mid (m \cdot n)$ , by the transitivity of divisibility, we are done.

 The ( $\Rightarrow$ ) direction of this proof used another characterisation of  $\gcd(m, n)$  as the *least positive linear combination of  $m$  and  $n$* . (NB: “Least” here means “lowest”, not the superlative of “less positive”.) Now that we are more familiar with universal properties, we can decode this description as ①  $\gcd(m, n)$  is a linear combination of  $m$  and  $n$ , and ②  $\gcd(m, n)$  divides all linear combinations of  $m$  and  $n$ :

$$\text{① } \exists k_0, l_0 \in \mathbb{Z}. k_0 \cdot m + l_0 \cdot n = \gcd(m, n) \quad \text{② } \forall k, l \in \mathbb{Z}. \gcd(m, n) \mid k \cdot m + l \cdot n$$

This characterisation is especially useful if we are able to express 1 as a linear combination of  $m$  and  $n$ , since ② means they must be *coprime*, i.e.  $\gcd(m, n) = 1$ . Another common use of an assumption of coprimality  $\gcd(m, n) = 1$  is that multiplication by  $\gcd(m, n)$  is a no-op, so we can freely introduce  $\gcd(m, n)$  or  $k_0 \cdot m + l_0 \cdot n$  for some  $k_0, l_0 \in \mathbb{Z}$  into any expression. This is what we make use of in the question when multiplying ④ and ⑤.

3. Prove that for all positive integers  $a, b, c$ , if  $\gcd(a, c) = 1$  then  $\gcd(a \cdot b, c) = \gcd(b, c)$ .

Below are three different proofs of the property.

#### Proof by equational reasoning

For  $a, b, c$  positive integers such that  $\gcd(a, c) = 1$ , we have

$$\begin{aligned} \gcd(b, c) &= \gcd(\gcd(a, c) \cdot b, c) && \text{(since } \gcd(a, c) = 1\text{)} \\ &= \gcd(\gcd(a \cdot b, c \cdot b), c) && \text{(by linearity)} \\ &= \gcd(a \cdot b, \gcd(c \cdot b, c)) && \text{(by associativity)} \\ &= \gcd(a \cdot b, c) && \text{(by §3.2.1)} \end{aligned}$$

#### Proof by universality

Let  $a, b, c$  positive integers such that  $\gcd(a, c) = 1$ . We need to prove that  $\gcd(a \cdot b, c) = \gcd(b, c)$ , or equivalently, that  $\gcd(a \cdot b, c) \mid \gcd(b, c)$  and  $\gcd(b, c) \mid \gcd(a \cdot b, c)$ . By the

universal property of gcds, it is sufficient to show the following two properties:

- $\gcd(a \cdot b, c) \mid b$  and  $\gcd(a \cdot b, c) \mid c$ . The latter holds since  $\gcd(a \cdot b, c)$  is a divisor of  $c$ . To establish the former, we note that  $b = \gcd(a, c) \cdot b$  (since  $a$  and  $c$  are coprime), and by distributivity,  $\gcd(a \cdot b, c \cdot b)$ . Thus, we can show that  $\gcd(a \cdot b, c) \mid \gcd(a \cdot b, c \cdot b)$ , or equivalently,  $\gcd(a \cdot b, c) \mid a \cdot b$  and  $\gcd(a \cdot b, c) \mid c \cdot b$ , both of which follow from  $\gcd(a \cdot b, c)$  being a common divisor of  $a \cdot b$  and  $c$ .
- $\gcd(b, c) \mid a \cdot b$  and  $\gcd(b, c) \mid c$ . Both follow from  $\gcd(b, c)$  being a divisor of  $b$  and  $c$ .

#### Proof using the Fundamental Theorem of Arithmetic

The [Fundamental Theorem of Arithmetic](#) states that every positive integer is expressible as the product of a unique finite sequence of ordered primes. If two integers are coprime, their unique prime factorisations must be disjoint: that is, there is no prime  $p$  that appears in the factorisation of both  $a$  and  $c$ . For any  $b \in \mathbb{Z}^+$ , the prime factorisation of  $a \cdot b$  will be the product of those of  $a$  and  $b$ . Therefore the common prime factors of  $a \cdot b$  and  $c$  must be the common factors of  $b$  and  $c$ , since there are no common factors of  $a$  and  $c$  by assumption. Since the greatest common divisor is the product of the common prime factors, we must have  $\gcd(a \cdot b, c) = \gcd(b, c)$ .

🎵 These are three fairly different proofs of the same (relatively simple) theorem: one uses equational reasoning and some properties of gcds, the second makes use of universality, while the third relies on a powerful and general theorem rather than gcd properties. The first is probably the most concise form, but of course it relies on us having established all the required properties of gcds already.

4. Prove that for all positive integers  $m$  and  $n$ , and integers  $i$  and  $j$ :

$$n \cdot i \equiv n \cdot j \pmod{m} \iff i \equiv j \pmod{\frac{m}{\gcd(m, n)}}$$

We have:

$$\begin{aligned} n \cdot i \equiv n \cdot j \pmod{m} &\iff k \cdot m = n(i - j) \\ &\iff k \cdot \frac{m}{\gcd(m, n)} = \frac{n}{\gcd(m, n)} \cdot (i - j) \\ &\iff \frac{m}{\gcd(m, n)} \mid \frac{n}{\gcd(m, n)} \cdot (i - j) \end{aligned}$$

Now we show that

$$\frac{m}{\gcd(m, n)} \mid \frac{n}{\gcd(m, n)} \cdot (i - j) \iff i \equiv j \pmod{\frac{m}{\gcd(m, n)}}$$

( $\Leftarrow$ ) We have  $\frac{m}{\gcd(m, n)} \mid i - j$  by assumption, and from the multiplication property of divisibility (§1.2.6(b)), we have  $\frac{m}{\gcd(m, n)} \mid \frac{n}{\gcd(m, n)} \cdot (i - j)$ .

( $\Rightarrow$ ) We first establish that  $\frac{m}{\gcd(m,n)}$  and  $\frac{n}{\gcd(m,n)}$  are coprime using linearity:


$$\gcd(m, n) = \gcd\left(\frac{m \cdot \gcd(m, n)}{\gcd(m, n)}, \frac{n \cdot \gcd(m, n)}{\gcd(m, n)}\right) = \gcd(m, n) \cdot \gcd\left(\frac{m}{\gcd(m, n)}, \frac{n}{\gcd(m, n)}\right)$$


Since  $\gcd(m, n)$  is a positive integer, this equality can only hold if  $\gcd\left(\frac{m}{\gcd(m, n)}, \frac{n}{\gcd(m, n)}\right) = 1$ .

This assumption of coprimality can then be used in [Euclid's Theorem](#) to conclude

$$\frac{m}{\gcd(m, n)} \mid \frac{n}{\gcd(m, n)} \cdot (i - j) \Rightarrow \frac{m}{\gcd(m, n)} \mid (i - j)$$

as required.

 The inspiration for the first “creative” step (dividing both sides by  $\gcd(m, n)$ ) comes from seeing the term  $\frac{m}{\gcd(m, n)}$  in the proof goal.

 A very useful corollary of this theorem is that we can always divide both sides of a congruence by a positive integer that is coprime with the modulus. Similarly, we can divide both sides of the congruence *and* the modulus with any positive integer that divides all three. The general theorem handles the case “in between”, when a positive integer divides both sides of the congruence, but not the modulus.

5. Prove that for all positive integers  $m, n, p, q$  such that  $\gcd(m, n) = \gcd(p, q) = 1$ , if  $q \cdot m = p \cdot n$  then  $m = p$  and  $n = q$ .

Let  $m, n, p, q$  be positive integers. Assume that  $\gcd(m, n) = \gcd(p, q) = 1$  and further that

$$\textcircled{1} \quad q \cdot m = p \cdot n.$$

Multiplying both sides of the identity  $1 = \gcd(m, n)$  by  $p$  and using the linearity property of  $\gcd$  we have that

$$p = p \cdot \gcd(m, n) = \gcd(p \cdot m, p \cdot n) \quad \textcircled{2}$$

Now, from  $\textcircled{1}$  and the linearity property of  $\gcd$ , we also have that

$$\gcd(p \cdot m, p \cdot n) = \gcd(p \cdot m, q \cdot m) = \gcd(p, q) \cdot m \quad \textcircled{3}$$

Finally, since  $\gcd(p, q) = 1$ , one has  $p = m$  from  $\textcircled{2}$  and  $\textcircled{3}$ .

We can show with an analogous argument that  $n = q$  as well.

6. Prove that for all positive integers  $a$  and  $b$ ,  $\gcd(13 \cdot a + 8 \cdot b, 5 \cdot a + 3 \cdot b) = \gcd(a, b)$ .

### Computational proof

For all positive integers  $a$  and  $b$ , one has

$$\begin{aligned} \gcd(13 \cdot a + 8 \cdot b, 5 \cdot a + 3 \cdot b) &= \gcd((13 \cdot a + 8 \cdot b) - (5 \cdot a + 3 \cdot b), 5 \cdot a + 3 \cdot b) \\ &= \gcd(8 \cdot a + 5 \cdot b, 5 \cdot a + 3 \cdot b) \\ &= \gcd((8 \cdot a + 5 \cdot b) - (5 \cdot a + 3 \cdot b), 5 \cdot a + 3 \cdot b) \end{aligned}$$

$$\begin{aligned}
&= \gcd(3 \cdot a + 2 \cdot b, 5 \cdot a + 3 \cdot b) \\
&= \gcd(3 \cdot a + 2 \cdot b, (5 \cdot a + 3 \cdot b) - (3 \cdot a + 2 \cdot b)) \\
&= \gcd(3 \cdot a + 2 \cdot b, 2 \cdot a + b) \\
&= \gcd((3 \cdot a + 2 \cdot b) - (2 \cdot a + b), 2 \cdot a + b) \\
&= \gcd(a + b, 2 \cdot a + b) \\
&= \gcd(a + b, (2 \cdot a + b) - (a + b)) \\
&= \gcd(a + b, a) \\
&= \gcd((a + b) - a, a) \\
&= \gcd(b, a) \\
&= \gcd(a, b)
\end{aligned}$$

### Conceptual proof (advanced)

We prove following general statement (see [2018/P8/Q9 exam question](#)):

$$\forall n \in \mathbb{N}. \gcd(a \cdot F_{n+3} + b \cdot F_{n+2}, a \cdot F_{n+1} + b \cdot F_n) = \gcd(a, b)$$

where  $F_n$  is the  $n^{\text{th}}$  Fibonacci number, defined recursively as

$$F_0 = 0 \quad F_1 = 1 \quad F_{n+2} = F_{n+1} + F_n$$

For  $n \in \mathbb{N}$ , we prove the following two properties, which, by the universal property of gcds, will imply the required equality.

- Both  $\gcd(a, b) \mid (aF_{n+3} + bF_{n+2})$  and  $\gcd(a, b) \mid (aF_{n+1} + bF_n)$ .

$\gcd(a, b)$  divides both  $a$  and  $b$ , so it divides every integer linear combination of them (§1.2.6(c)).

- For all positive integers  $d$ ,

$$\text{if } d \mid (aF_{n+3} + bF_{n+2}) \text{ and } d \mid (aF_{n+1} + bF_n) \text{ then } d \mid \gcd(a, b).$$

Let  $d$  be a positive integer such that  $d \mid (aF_{n+3} + bF_{n+2})$  and  $d \mid (aF_{n+1} + bF_n)$ ; so that  $di = aF_{n+3} + bF_{n+2}$  and  $dj = aF_{n+1} + bF_n$  for (positive) integers  $i$  and  $j$ .

It follows that

$$\begin{aligned}
d \cdot (iF_n - jF_{n+2}) &= (F_n \cdot F_{n+3} - F_{n+2} \cdot F_{n+1}) \cdot a \\
&= (F_n \cdot F_{n+2} + F_n \cdot F_{n+1} - F_n \cdot F_{n+1} - F_{n+1} \cdot F_{n+1}) \cdot a \\
&= (F_n \cdot F_{n+2} - F_{n+1}^2) \cdot a \\
&= (-1)^{n+1} a \qquad \qquad \qquad (\text{Cassini's Identity})
\end{aligned}$$

so that  $d \mid a$ ; and, analogously,

$$\begin{aligned}
 d \cdot (iF_{n+1} - jF_{n+3}) &= (F_{n+1} \cdot F_{n+2} - F_{n+3} \cdot F_n) \cdot a \\
 &= (F_{n+1} \cdot F_{n+1} + F_n \cdot F_{n+1} - F_n \cdot F_{n+1} - F_n \cdot F_{n+2}) \cdot b \\
 &= (F_{n+1}^2 - F_n \cdot F_{n+2}) \cdot b \\
 &= (-1)^n b \quad \text{(Cassini's Identity)}
 \end{aligned}$$

so that  $d \mid b$ . Thus,  $d \mid \gcd(a, b)$  as required.

 You will learn more about Fibonacci numbers in the next set of exercises.

7. Let  $n$  be an integer.

a) Prove that if  $n$  is not divisible by 3, then  $n^2 \equiv 1 \pmod{3}$ .

This is an instance of [Fermat's Little Theorem](#).

b) Show that if  $n$  is odd, then  $n^2 \equiv 1 \pmod{8}$ .

Let  $n$  be an odd integer, and thereby let  $k$  be an integer such that  $n = 2 \cdot k + 1$ .

We consider two cases.

- Case  $k$  is even.

Then,  $k = 2 \cdot l$  for some integer  $l$ , and  $n^2 = 8 \cdot l \cdot (2 \cdot l + 1) \equiv 1 \pmod{8}$ .

- Case  $k$  is odd.

Then,  $k = 2 \cdot l + 1$  for some integer  $l$ , and  $n^2 = 8 \cdot (2 \cdot l + 1) \cdot (l + 2) + 1 \equiv 1 \pmod{8}$ .

Either way  $n^2 \equiv 1 \pmod{8}$ , as required.

c) Conclude that if  $p$  is a prime number greater than 3, then  $p^2 - 1$  is divisible by 24.

Let  $p$  be a prime greater than 3. Then,  $p$  is an odd integer not divisible by 3 and it follows from part (a) that: ①  $3 \mid (p^2 - 1)$ . Moreover, as  $p$  is odd, we have from part (b) that: ②  $8 \mid (p^2 - 1)$ .

Finally, since  $\gcd(3, 8) = 1$ , by [§3.2.2](#) one has that ① and ② imply  $24 \mid (p^2 - 1)$  as required.

8. Prove that  $n^{13} \equiv n \pmod{10}$  for all integers  $n$ .

To show  $n^{13} \equiv n \pmod{10}$ , by the direct corollary of [§3.2.2](#) it is sufficient to show  $n^{13} \equiv n \pmod{2}$  and  $n^{13} \equiv n \pmod{5}$ . Both hold by successive applications of Fermat's Little Theorem, repeatedly reducing  $n^2$  or  $n^5$  to  $n$  until we reach  $n$ . For example:

$$n^{13} = n^5 \cdot n^5 \cdot n^3 \equiv n \cdot n \cdot n^3 = n^5 \equiv n \pmod{5}$$

9. Prove that for all positive integers  $l$ ,  $m$  and  $n$ , if  $\gcd(l, m \cdot n) = 1$  then  $\gcd(l, m) = 1$  and  $\gcd(l, n) = 1$ .

Let  $l$ ,  $m$ , and  $n$  be arbitrary positive integers, and assume that  $\gcd(l, m \cdot n) = 1$ .

By §3.1.5( $\Leftarrow$ ), there exist integers  $i$  and  $j$  such that  $i \cdot l + j \cdot m \cdot n = 1$ . Thus, we have that

$$\text{there exist integers } i \text{ and } a \text{ such that } i \cdot l + a \cdot m = 1$$

and

$$\text{there exist integers } i \text{ and } b \text{ such that } i \cdot l + b \cdot n = 1.$$

Therefore, by §3.1.5( $\Rightarrow$ ) one has that  $\gcd(l, m) = 1$  and  $\gcd(l, n) = 1$ .

10. Solve the following congruences:

a)  $77 \cdot x \equiv 11 \pmod{40}$

By §3.2.4, a solution will satisfy the congruence iff it satisfies  $\textcircled{+} 7 \cdot x \equiv 1 \pmod{40}$  ( $\gcd(40, 11) = 1$  so the modulus does not change). As 7 and 40 are coprime, this amounts to finding the multiplicative inverse of 7 in  $\mathbb{Z}_{40}$  (Corollary 75), which is the second coefficient in the expression of 1 as a linear combination of 40 and 7. We run the Extended Euclid's Algorithm to find that  $40 \cdot 3 + 7 \cdot (-17) = 1$ . Thus,  $x_0 = -17$  is a solution to  $\textcircled{+}$ , and therefore to  $77 \cdot x_0 \equiv 11 \pmod{40}$ . To find the general form of solutions, we note that the linear combination of 40 and 7 is not unique (Slide 219), so  $x$  can have the general form  $x = -17 + 40n \equiv 23 + 40n$  for any integer  $n$ .

b)  $12 \cdot y \equiv 30 \pmod{54}$

By §3.2.4, a solution will satisfy the congruence iff it satisfies  $\textcircled{+} 2 \cdot y \equiv 5 \pmod{9}$ , that is,  $2 \cdot y + 9 \cdot k = 5$  for some  $k \in \mathbb{Z}$ . Now, since 2 and 9 are coprime, we can express 1 as their linear combination, computing the coefficients using the Extended Euclid's Algorithm:  $2 \cdot (-4) + 9 \cdot 1 = 1$ . Multiplying both sides by 5 gives us  $2 \cdot (-20) + 9 \cdot 5 = 5$ , which is a solution to  $\textcircled{+}$  with  $y_0 = -20$ . To generate all the solutions, we note that  $\textcircled{+}$  is satisfied by  $y_0 + 9n$  for any  $n$ , so  $y$  can have the general form  $y = -20 + 9n \equiv 7 + 9n$  for any integer  $n$ .

c) 
$$\begin{cases} 13 \equiv z \pmod{21} \\ 3 \cdot z \equiv 2 \pmod{17} \end{cases}$$

To solve a system of congruences, we find the general form of solutions for the congruences individually, then find the ones that satisfy both.

All solutions to the first congruence are of the form  $z_1 = 13 + 21k$  for  $k \in \mathbb{Z}$ .

Solutions of the congruence  $3 \cdot z \equiv 2 \pmod{17}$  satisfy  $\textcircled{+} 3 \cdot z + 17 \cdot n = 2$ . Since 3 and 17 are coprime, we can express 1 as their linear combination using EEA:  $3 \cdot 6 + 17 \cdot (-1) = 1$ . Multiplying by 2 on both sides gives a solution to  $\textcircled{+}$ , and from there, we get the general form of solutions as  $z_2 = 12 + 17l$  for  $l \in \mathbb{Z}$ .

The solutions for the congruence system will be those which are both of the form  $z_1$

and  $z_2$  simultaneously:

$$13 + 21 \cdot k = 12 + 17 \cdot l$$

Albeit this looks like one equation with two unknowns, we can rearrange it to the form

$$21 \cdot (-k) + 17 \cdot l = 1 \quad (\oplus)$$


which we can solve using EEA, since 21 and 17 are coprime:

$$21 \cdot (-4) + 17 \cdot 5 = 1$$

Thus,  $(\oplus)$  has general solutions  $k = 4 + 17i$  and  $l = 5 + 21j$  for  $i, j \in \mathbb{Z}$ ; at these specific values of  $k$ , the general solution  $z_1 = 13 + 21 \cdot k$  for the first congruence also satisfies the second congruence (and similarly for  $z_2$ ). Substituting  $k$  into  $z_1$  or  $l$  into  $z_2$  gives

$$z = 97 + 357i \quad \forall i \in \mathbb{Z}.$$

which is the general form of solutions that satisfy the system of congruences.

 This question shows the usefulness of the characterisation of gcds via linear combinations: it allows us to solve one equation with two unknowns, as long as the RHS is a multiple of the gcd of the coefficients (so if the coefficients are coprime, the RHS can be any positive integer). Solving a congruence  $ax \equiv b \pmod{m}$  amounts to characterising the integer solutions of the equation  $ax - my = b$  (known as a linear Diophantine equation), which exist only if  $\gcd(a, m) \mid b$ .

If a congruence  $ax \equiv b \pmod{m}$  has one solution  $x_0$  (i.e. if  $\gcd(a, m) \mid b$ ), it has an infinite number of solutions of the form  $x = x_0 + pk$  for  $k \in \mathbb{Z}$ , all separated by a “period”  $p$ . In some cases (such as part (a)), the period coincides with the modulus, so all possible solutions can be derived from a single integer  $x_0 \in \mathbb{Z}_m$ . In other cases (such as part (b)) the solutions may be more “frequent” due to the period being a fraction of the modulus:  $m = dp$ . Then, the solutions  $x_0 + pk$  can be split into  $d$  classes, all with the period  $m$ , but different initial values  $x_0, x_1, \dots, x_{d-1} \in \mathbb{Z}_m$ . One such class  $\{\dots, x - 2m, x - m, x, x + m, x + 2m, \dots\}$  is often called the *congruence class of  $x$  modulo  $m$*  (denoted  $\overline{x}_m$  or sometimes  $[x]_m$ , although this course uses the latter notation to refer to the least positive element of  $\overline{x}_m$  in  $\mathbb{Z}_m$ ), so in essence, an infinite number of integer solutions to a congruence can be characterised by a finite number of congruence classes. With this interpretation, part (a) had only one solution  $\overline{23}_{40}$ , while part (b) had six:

$$\overline{7}_{54} \quad \overline{16}_{54} \quad \overline{25}_{54} \quad \overline{34}_{54} \quad \overline{43}_{54} \quad \overline{52}_{54}$$

By considering a solution to be a congruence class modulo  $m$ , we can show that a congruence  $ax \equiv b \pmod{m}$  has exactly  $\gcd(a, m)$  solutions if  $\gcd(a, m) \mid b$ , and 0 otherwise. Of course, the  $d = \gcd(a, m)$  congruence classes modulo  $m$  can be combined into one congruence class modulo  $m/d$  – the two representations are equivalent, but one may be



more useful in some contexts than the other. As an example, compare the phrases “every 8 hours starting at 1am” and “every day at 1am, 9am, and 5pm”, and how we must use the latter form to refer to events repeating regularly several times a week because 7 prime.

Since integer solutions of a congruence are not unique, we can ask which solutions of one congruence also satisfy another – that is, solve a *system of congruences*. These are quite different from the systems of equations you are familiar with, which involve  $n$  unknowns and  $n$  independent equations, and the solution is found by expressing one variable in terms of the others and performing substitutions. Congruence systems involve only one unknown, and the individual congruences are independent constraints on this one unknown. Rather than trying to combine the congruences via substitution, we solve each of them independently, getting sets of congruence classes for each individual congruence. Then, the task is finding the common elements of the congruence classes (their intersection), which therefore must satisfy the whole system of congruences simultaneously. If the individual solutions have the form  $x + pk$  and  $y + ql$ , the congruence classes  $\overline{x}_p$  and  $\overline{y}_q$  will intersect when  $x + pk = y + ql$ ; this now becomes another linear Diophantine equation of the form  $pk - ql = y - x$  that can be solved if  $\gcd(p, q) \mid y - x$ . The resulting integer values for  $k$  and  $l$  tell us the number of periods one needs to offset  $x$  and  $y$  by until they coincide, and since all solutions are uniformly periodic,  $k$  and  $l$  will themselves be periodic congruence classes. The general expressions can then be substituted back into either  $x + pk$  or  $y + ql$  to find an initial value and a larger period for the solutions that satisfy both parts of the congruence system.

As a simple example, consider the congruence classes  $\overline{1}_2$ ,  $\overline{2}_3$  and  $\overline{2}_4$ . The classes  $\overline{1}_2$  and  $\overline{2}_3$  will intersect whenever  $1 + 2n = 2 + 3k$ , and the linear Diophantine equation  $2n - 3k = 1$  has solutions  $n = 3m + 2$  and  $k = 2m + 1$ . What this means is that every 3<sup>rd</sup> ● starting from the second one (using 0-indexed counting) will coincide with every 2<sup>nd</sup> ■ starting from the first one, as can be seen below at step 5 (when  $m = 0$ ) and 11 (when  $m = 1$ ). To figure out what “every 3<sup>rd</sup> ● starting from the second one” means on the resolution of the integers, we substitute the solution for  $n$  back into  $1 + 2n$ , which combines the periods of “there is a solution at every 3<sup>rd</sup> circle” and “there is a circle every 2 steps” into “there is a solution every 6 steps” and similarly for the offset. Thus, the intersection of  $\overline{1}_2$  and  $\overline{2}_3$  will be  $\overline{5}_6$ . We can do a similar procedure to find the intersection of  $\overline{2}_3$  and  $\overline{2}_4$  to be  $\overline{12}_{12}$ . However,  $\overline{1}_2$  and  $\overline{2}_4$  will never intersect, since the Diophantine equation  $2n - 4l = 1$  has no solutions –  $\gcd(2, 4) = 2 \nmid 1$ . Congruence systems often arise from the interaction of periodic events: examples are scheduling, polyrhythms, predator-prey life cycles, etc.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$1 + 2n$	●		●		●		●		●		●		●	
$2 + 3k$		■			■			■			■			■
$2 + 4l$		▲				▲				▲				▲

11. What is the multiplicative inverse of: (a) 2 in  $\mathbb{Z}_7$ , (b) 7 in  $\mathbb{Z}_{40}$ , and (c) 13 in  $\mathbb{Z}_{23}$ ?

We apply [Corollary 75](#) of the notes, which states that if  $\gcd(m, n) = 1$ , the multiplicative inverse of  $[n]_m$  is  $[lc_2(m, n)]_m$ , where  $lc_2(m, n)$  is the second coefficient of the expression of 1 as a linear combination of  $m$  and  $n$  using EEC. With this, we get that:

- a)  $1 \cdot 7 + (-3) \cdot 2 = 1$ , so  $2^{-1} \equiv 4 \pmod{7}$
- b)  $3 \cdot 40 + (-17) \cdot 7 = 1$ , so  $7^{-1} \equiv 23 \pmod{40}$
- c)  $4 \cdot 23 + (-7) \cdot 13 = 1$ , so  $13^{-1} \equiv 16 \pmod{23}$

12. Prove that  $[22^{12001}]_{175}$  has a multiplicative inverse in  $\mathbb{Z}_{175}$ .

We first establish the following lemma:

For every pair of positive integers  $m$  and  $n$ , we have that  $[n]_m$  has a multiplicative inverse in  $\mathbb{Z}_m$  iff  $\gcd(m, n) = 1$ .

( $\Rightarrow$ ) Let  $m$  and  $n$  be arbitrary positive integers, and assume that  $[n]_m$  has a multiplicative inverse in  $\mathbb{Z}_m$ , say  $l$ . Then,

$$n \cdot l \equiv [n \cdot l]_m = [n]_m \cdot l = 1 \pmod{m}$$

and thus there exists an integer  $k$  such that  $n \cdot l + m \cdot k = 1$ . Thus, from [§3.1.5\( \$\Rightarrow\$ \)](#),  $\gcd(m, n) = 1$ .

( $\Leftarrow$ ) By [Corollary 75\(2\)](#) of the notes.

Now,  $\gcd(22^{12001}, 175) = \gcd(2^{12001} \cdot 11^{12001}, 5^2 \cdot 7)$ , and since the two numbers have no prime factors in common, they must be coprime. By the above lemma,  $\gcd(22^{12001}, 175) = 1$  implies that  $[22^{12001}]_{175}$  has a multiplicative inverse, as required.

### 3.3. Optional exercises

1. Let  $a$  and  $b$  be natural numbers such that  $a^2 \mid b \cdot (b + a)$ . Prove that  $a \mid b$ .

*Hint:* For positive  $a$  and  $b$ , consider  $a_0 = \frac{a}{\gcd(a, b)}$  and  $b_0 = \frac{b}{\gcd(a, b)}$  so that  $\gcd(a_0, b_0) = 1$ , and show that  $a^2 \mid b(b + a)$  implies  $a_0 = 1$ .

If either  $a$  or  $b$  are 0 the result is straightforward. Consider thus the case in which both  $a$  and  $b$  are positive integers, and assume that  $a^2 \mid b(b + a)$ .

Then, for  $a_0 = \frac{a}{\gcd(a, b)}$  and  $b_0 = \frac{b}{\gcd(a, b)}$ , we have that  $a_0 \mid b_0(b_0 + a_0)$  and, since  $\gcd(a_0, b_0) = 1$ , that  $a_0 \mid (b_0 + a_0)$  so that  $a_0 \mid b_0$  and thus  $a_0 = \gcd(a_0, b_0) = 1$ . Therefore,  $\gcd(a, b) = a$  and we are done.

2. Prove the converse of [§1.3.1\(f\)](#): For all natural numbers  $n$  and  $s$ , if there exists a natural number  $q$  such that  $(2n + 1)^2 \cdot s + t_n = t_q$ , then  $s$  is a triangular number. (49<sup>th</sup> Putnam, 1988)

*Hint:* Recall that if  $(\oplus) q = 2nk + n + k$  then  $(2n + 1)^2 t_k + t_n = t_q$ . Solving for  $k$  in  $(\oplus)$ , we get that  $k = \frac{q-n}{2n+1}$ ; so it would be enough to show that the fraction  $\frac{q-n}{2n+1}$  is a natural number.

Suggested by a 2014/15 student (who wished to remain anonymous).

Assume  $(2n+1)^2s + t_n = t_q$ . Then,  $t_n \equiv t_q \pmod{(2n+1)^2}$ ; so that  $n(n+1) \equiv q(q+1) \pmod{(2n+1)^2}$  and hence  $(q-n)(q-n+2n+1) \equiv 0 \pmod{(2n+1)^2}$ .

Therefore  $(2n+1)^2 \mid (q-n)(q-n+2n+1)$ , and it follows from the previous item that  $(2n+1) \mid (q-n)$ .

As  $t_q \geq t_n$ , we have that  $q \geq n$ , and therefore that  $k = \frac{q-n}{2n+1}$  is a natural number. By assumption and the solution to §1.3.1(f), we then have:

$$(2n+1)^2s + t_n = t_q = (2n+1)^2t_k + t_n$$

and so that  $s = t_k$ , as required.

3. Informally justify the correctness of the following alternative algorithm for computing the gcd of two positive integers:

```
let rec gcd0(m, n) = if m = n then m
                      else let p = min m n
                           and q = max m n
                           in gcd0(p, q - p)
```

The partial correctness of the algorithm follows from [Corollary 58\(2\)](#) of the notes. To establish the termination of `gcd0` on a pair of positive integers  $(m, n)$  we consider and analyse the computations arising from the call `gcd0(m, n)`. We consider three cases:

- Case  $m = n$ .

The computation of `gcd0(m, n)` reduces in one step to  $m$ , and therefore terminates.

- Case  $m \neq n$ .

The computation of `gcd0(m, n)` reduces in one step to that of `gcd0(p, q - p)`, where  $p = \min(m, n)$  and  $q = \max(m, n)$ . Thus, the passage of computing `gcd0(m, n)` by means of computing `gcd0(p, q - p)` maintains the invariant of having both components of the pair being positive integers; but, crucially, strictly decreases the sum of the pairs in each recursive call (as  $m + n > \max(m, n) = p + (q - p)$  because both  $m$  and  $n$  are positive). As this process cannot go on forever (the sum is of two strictly positive integers but decreases at every step, so the lowest it can go is  $1 + 1 = 2$ , at which point  $m = n$ ), the recursive calls must eventually stop and the overall computation terminate (in fact, in a number of steps necessarily less than or equal the sum of the input pair).

 We can use induction to make this argument formal; see §4.3.1.

## 4. On induction

### 4.1. Basic exercises

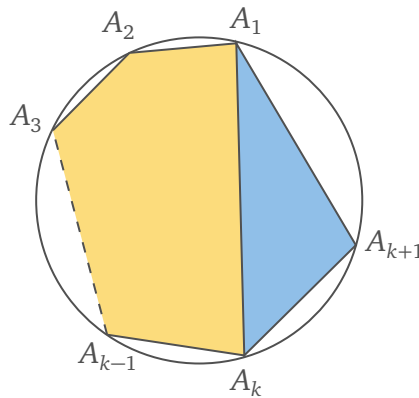
1. Prove that for all natural numbers  $n \geq 3$ , if  $n$  distinct points on a circle are joined in consecutive order by straight lines, then the interior angles of the resulting polygon add up to  $180 \cdot (n - 2)$  degrees.


We prove this property  $P(n)$  of all  $n \geq 3$  by mathematical induction from basis 3.


**Base case:**  $n = 3$ . Three connected points on a circle must form a triangle: since they are distinct, they cannot be collinear. The sum of internal angles of a triangle is  $180^\circ$ , which is  $180 \cdot (3 - 2)$  degrees.

**Inductive case:**  $n = k + 1$ . Assume that  $\textcircled{\text{IH}} P(k)$  holds and take an arbitrary polygon constructed from  $k + 1$  points  $A_1, \dots, A_{k+1}$  on a circle. The  $(k + 1)$ -gon can be separated into a  $k$ -gon and a triangle with a line segment connecting  $A_1$  and  $A_k$ . By the induction hypothesis  $\textcircled{\text{IH}}$ , the interior angles of the  $k$ -gon add up to  $S_k = 180 \cdot (k - 2)$  degrees. The sum of angles of the whole polygon is  $S_{k+1} = S_k + \angle A_k A_1 A_{k+1} + \angle A_1 A_{k+1} A_k + \angle A_1 A_k A_{k+1}$ , where the angle terms belong to the triangle  $\triangle A_1 A_k A_{k+1}$ . Since its interior angles must add up to  $180^\circ$ , we have the expression for the sum of internal angles of the  $(k + 1)$ -gon:

$$S_{k+1} = S_k + 180^\circ = 180 \cdot (k - 2) + 180^\circ = 180 \cdot ((k + 1) - 2)$$



 While the formula holds for any polygon, working with points on a circle makes the inductive proof easier, since we never need to worry about three points being on the same line and only making up one side.

 It may be tempting to approach the inductive step by starting with a  $k$ -gon, then *adding* a new point to turn it into a  $(k + 1)$ -gon and increasing the sum of internal angles by  $180^\circ$ . The problem with this is that we are *given* a  $(k + 1)$ -gon to start with, and its vertices are predetermined: we need to split it up into a triangle and a  $k$ -gon, no matter what the points are. This distinction is fairly minor in this case and doesn't cause any difficulties (any line segment connecting two vertices one point apart will split do the job), but remembering what parameters we have control over vs. what we are given (that is, what we need to assume as being arbitrary) is very important in proofs, especially inductive ones. We will

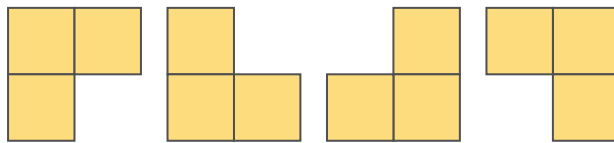
see examples of this throughout this sheet.

2. Prove that, for any positive integer  $n$ , a  $2^n \times 2^n$  square grid with any one square removed can be tiled with L-shaped pieces consisting of 3 squares.

We prove the property  $P(n)$  of all  $n \geq 1$  by mathematical induction from basis 1:

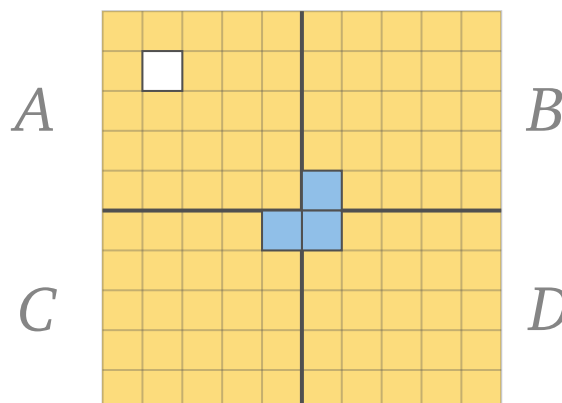
$$P(n) = \forall 0 \leq i, j \leq n. \text{ a } 2^n \times 2^n \text{ grid } A \text{ with square } A_{i,j} \text{ missing can be tiled}$$


**Base case:**  $n = 1$ . Take a  $2^1 \times 2^1 = 2 \times 2$  grid and assume one of the squares is missing. This must be one of the following four situations, depending on which one of the 4 squares was removed:



All resulting shapes can be tiled with one L-shaped piece consisting of three squares.

**Inductive case:**  $n = k + 1$ . Assume  $(IH) P(k)$ : a  $2^k \times 2^k$  grid with any square missing can be tiled with L-shaped pieces. Take a  $2^{k+1} \times 2^{k+1}$  grid with any one square missing. The grid can be split into four  $2^k \times 2^k$  quarters which we label by  $A, B, C$  and  $D$ ; assume, without loss of generality, that the missing square is in quarter  $A$  at position  $A_{i,j}$ . By the  $(IH)$  applied to  $i$  and  $j$ , the quarter  $A$  can be tiled with  $A_{i,j}$  missing. Next, we use the  $(IH)$  applied to  $i = k$  and  $j = 1$  to tile quarter  $B$  with the bottom left square missing. Similarly, we tile  $C$  and  $D$  with two applications of the induction hypothesis ( $(IH)(1, k)$  and  $(IH)(1, 1)$ , respectively) with the top right and left corners missing. The three missing corners form an L-shaped hole of 3 squares in the middle of the  $2^{k+1} \times 2^{k+1}$  grid, which can be filled in with one additional tile. This leaves only one missing square  $A_{i,j}$  with the rest of the grid tiled with L-shaped pieces, so we are done.



 This is an example of an inductive proof where the proposition  $P(n)$  is itself a universally quantified statement: we state property for all grid size parameters  $n$ , and within a particular grid of size  $2^n \times 2^n$ , for all possible grid cells that could be missing. Thus, after case-splitting on  $n$ , we still have a universally quantified proof obligation; however, in the inductive case,

we also have a universally quantified inductive assumption.

While the general pattern for proofs like this is just an instance of the standard induction principle, it is worth analysing nevertheless:

To prove a property of the form

$$\forall n \in \mathbb{N}. \forall x \in A. P(n, x)$$

it is sufficient to prove

$$\forall x \in A. P(0, x) \quad \text{and} \quad \forall k \in \mathbb{N}. (\forall y \in A. P(k, y)) \implies (\forall x \in A. P(k+1, x))$$

The base case – which is usually seen as the “trivial” step – is now itself a universally quantified statement which may not necessarily be easy to establish. Indeed, if the inner quantification is over natural numbers as well, we may end up having to do *another* inductive proof of  $\forall m \in \mathbb{N}. P(0, m)$  if a direct proof (“Let  $m$  be an arbitrary natural number and prove  $P(0, m)$ ...”) is not possible.

The inductive step highlights the interplay between the two quantifications. Unwrapping the formula, we get three assumptions: an arbitrary natural number  $k$ , an arbitrary element  $x \in A$ , and a *proof* that  $P(k, y)$  holds for any choice of  $y \in A$ . In the process of the proof, this induction hypothesis can be applied to any element  $y \in A$ , be it  $x \in A$ , a value computed from  $x$ , or any other value arbitrarily chosen by us. There is a significant difference between the inductive step above, and a formula such as


$$\forall k \in \mathbb{N}. \forall x \in A. P(k, x) \implies P(k+1, x)$$

which leaves us no flexibility in “tailoring” the IH to our needs by choosing an appropriate value for  $x$ .

The question above had an inner universal quantification over the position of the missing cell, so the proof cannot depend on any particular choice of position in the  $2^{k+1} \times 2^{k+1}$  grid. However, we do have control over the position of the missing cell when applying the induction hypothesis to the  $2^k \times 2^k$  quarter grids: we can essentially think of the ⑩ as a “function” from coordinates  $(i, j)$  to the proof of “tileability”. To complete the inductive step, we first apply the IH to the coordinates of the actual hole in the  $2^{k+1} \times 2^{k+1}$  within the  $A$  quarter, then select the appropriate locations for the holes in the quarters  $B$ ,  $C$  and  $D$  to leave an L-shaped hole in the middle. We apply the ⑩ both to the unknown values  $(i, j)$  given to us by the universal quantifier on the LHS of the implication, as well as values that we select deliberately to create space for an extra L-shaped tile.

🎵 The phrase “without loss of generality” is often used to reduce repetition or make simplifying assumptions that do not change the strength of the result. It is usually understood that if the assumption is violated, it can be altered in an obvious way to make the rest of the proof go through. It is important to ensure that the assumption really doesn’t affect

the generality of the statement: saying things like “w.l.o.g., assume  $n$  is even/nonzero/a power of two” is sometimes tempting, but it’s rarely clear how the proof could be extended to numbers which are odd/zero/not a power of two, and proving these cases may require entirely different approaches to the one considered. Above, we assumed w.l.o.g. that the hole is in quarter  $A$  so we don’t need to repeat the proof for all four quarters. The proofs would not be exactly the same (e.g. if the hole was in quarter  $B$ , the IH would need to be applied to the  $(i - k, j)$  coordinates of the  $2^k \times 2^k$  grid), but it’s clear that the general idea would work in each case.

 The proof above doesn’t just show that a tiling is possible, it gives a concrete algorithm for constructing it. Proofs like this are – unsurprisingly – called *constructive* proofs (also known as *effective* proofs to avoid confusion with constructive mathematics), as opposed to *nonconstructive* or *pure existence* proofs which show that a mathematical object exists, but doesn’t give a concrete example or way of computing one. Constructive proofs by induction naturally give rise to recursive algorithms, where the application of the  $\textcircled{\text{IH}}$  corresponds to a recursive call. Of course, when implementing the recursive algorithm, we don’t have the luxury of saying that “without loss of generality, assume the user will never call the function with the hole outside of quarter  $A$ ” – we have to explicitly handle all four possibilities and slightly different recursive calls to cover any possible input.

## 4.2. Core exercises

1. Establish the following:

(a) For all positive integers  $m$  and  $n$ ,

$$(2^n - 1) \cdot \sum_{i=0}^{m-1} 2^{i \cdot n} = 2^{m \cdot n} - 1$$

The first thing to note is that an inductive proof is not really necessary. Indeed, for arbitrary positive integers  $m$  and  $n$ , one can calculate that

$$\begin{aligned} (2^n - 1) \cdot \sum_{i=0}^{m-1} 2^{i \cdot n} &= \sum_{i=0}^{m-1} 2^{(i+1) \cdot n} - \sum_{i=0}^{m-1} 2^{i \cdot n} \\ &= \sum_{i=1}^{m-1} 2^{i \cdot n} + 2^{((m-1)+1) \cdot n} - 2^{0 \cdot n} - \sum_{i=1}^{m-1} 2^{i \cdot n} \\ &= 2^{m \cdot n} - 1 \end{aligned}$$

However, as it is very instructive, two inductive proofs follow. Note the different, though subtle, ways in which the inductive hypothesis is used in each proof.

For the *first proof*, we show

$$\forall m \in \mathbb{Z}^+. P(m)$$

for  $P(m)$  the statement

$$\forall n \in \mathbb{Z}^+. (2^n - 1) \cdot \sum_{i=0}^{m-1} 2^{i \cdot n} = 2^{m \cdot n} - 1$$

by the Principle of Induction.

**Base case:**  $m = 1$ . The statement  $P(1)$  amounts to

$$\forall n \in \mathbb{Z}^+. (2^n - 1) \cdot 2^{1 \cdot n} = 2^{1 \cdot n} - 1$$

which is vacuously true.

**Inductive step:**  $m = k + 1$ . Let  $k$  be an arbitrary positive integer, and assume that the Inductive Hypothesis  $P(k)$  holds for it; i.e. that

$$\forall n \in \mathbb{Z}^+. (2^n - 1) \cdot \sum_{i=0}^{k-1} 2^{i \cdot n} = 2^{k \cdot n} - 1 \quad (\text{IH})_1$$

We need show that  $P(k + 1)$  follows; i.e. that

$$\forall n \in \mathbb{Z}^+. (2^n - 1) \cdot \sum_{i=0}^{(k+1)-1} 2^{i \cdot n} = 2^{(k+1) \cdot n} - 1$$

To this end, we let  $l$  be an arbitrary positive integer and proceed to show that

$$(2^l - 1) \cdot \sum_{i=0}^k 2^{i \cdot l} = 2^{(k+1) \cdot l} - 1 \quad (1)$$

Indeed, instantiating the  $(\text{IH})_1$ , we have that

$$(2^l - 1) \cdot \sum_{i=0}^{k-1} 2^{i \cdot l} = 2^{k \cdot l} - 1 \quad (2)$$

and so that

$$\begin{aligned} (2^l - 1) \cdot \sum_{i=0}^k 2^{i \cdot l} &= \left( (2^l - 1) \cdot \sum_{i=0}^{k-1} 2^{i \cdot l} \right) + (2^l - 1) \cdot 2^{k \cdot l} \\ &= 2^{k \cdot l} - 1 + (2^l - 1) \cdot 2^{k \cdot l} \quad (\text{by } (2)) \\ &= 2^{(k+1) \cdot l} - 1 \end{aligned}$$

establishing (1) as required. □

For the *second proof*, to show

$$\forall n \in \mathbb{Z}^+. \forall m \in \mathbb{Z}^+. (2^n - 1) \cdot \sum_{i=0}^{m-1} 2^{i \cdot n} = 2^{m \cdot n} - 1$$



we let  $l$  be an arbitrary positive integer and prove

$$\forall m \in \mathbb{Z}^+. Q(l, m)$$

for  $Q(l, m)$  the statement

$$(2^l - 1) \cdot \sum_{i=0}^{m-1} 2^{i \cdot l} = 2^{m \cdot l} - 1$$

by the Principle of Induction.

**Base case:**  $m = 1$ . The statement  $Q(l, 1)$  amounts to

$$(2^l - 1) \cdot 2^{0 \cdot l} = 2^{1 \cdot l} - 1$$

which is vacuously true.

**Inductive step:**  $m = k + 1$ . Let  $k$  be an arbitrary positive integer, and assume that the Inductive Hypothesis  $Q(l, k)$  holds for it; i.e. that

$$(2^l - 1) \cdot \sum_{i=0}^{k-1} 2^{i \cdot l} = 2^{k \cdot l} - 1 \quad (\text{IH})_2$$


We need show that  $Q(l, k + 1)$  follows; i.e. that

$$\mathbb{Z}^+ (2^l - 1) \cdot \sum_{i=0}^{(k+1)-1} 2^{i \cdot l} = 2^{(k+1) \cdot l} - 1 \quad (1)$$

Indeed,

$$\begin{aligned} (2^l - 1) \cdot \sum_{i=0}^k 2^{i \cdot l} &= \left( (2^l - 1) \cdot \sum_{i=0}^{k-1} 2^{i \cdot l} \right) + (2^l - 1) \cdot 2^{k \cdot l} \\ &= 2^{k \cdot l} - 1 + (2^l - 1) \cdot 2^{k \cdot l} \quad (\text{by } (\text{IH})_2) \\ &= 2^{(k+1) \cdot l} - 1 \end{aligned}$$

establishing (1) as required.

 The core of the proof is the same in both cases; the difference is how they set up the induction hypothesis. The first proof includes the quantification over  $n$  in the  $(\text{IH})_1$  and applies it to the arbitrary  $l$  in the proof to get a specific instance (2). The second proof fixes this  $l$  right from the start, introducing it as a new arbitrary variable in the standard manner of proving universal quantification. Then, the predicate to be established by inductively is “parameterised” by this  $l$ , so the statement  $Q(l, m)$  doesn’t actually need a nested quantification. Despite  $(\text{IH})_2$  not containing a universal quantification, the proof only requires it at the specific  $l$  we already introduced. This makes the second proof slightly simpler, but it would not work if we ever needed the induction hypothesis at any other value of  $n$ .

(b) Suppose  $k$  is a positive integer that is not prime. Then  $2^k - 1$  is not prime.

Let  $k$  be an arbitrary positive integer. We consider two cases:

- $k = 1$ . The statement holds because  $2^1 - 1 = 1$  is not prime.
- $k \geq 2$ . Assume that  $k \geq 2$  is not prime. Hence, it is of the form  $m \cdot n$  for natural numbers  $m, n$  greater than or equal to 2. It follows from the previous item that  $2^k - 1 = 2^{m \cdot n} - 1 = (2^n - 1) \cdot \sum_{i=0}^{m-1} 2^{i \cdot n}$ ; and, since  $2^n - 1 \geq 2^2 - 1 = 3$  and  $\sum_{i=0}^{m-1} 2^{i \cdot n} \geq 1 + 4 = 5$ , we have that  $2^k - 1$  has a non-trivial decomposition. Hence it is not prime.

2. Prove that

$$\forall n \in \mathbb{N}. \forall x \in \mathbb{R}. x \geq -1 \implies (1+x)^n \geq 1+n \cdot x$$

We prove  $\forall n \in \mathbb{N}. P(n)$  for  $P(n)$  the statement

$$\forall x \in \mathbb{R}. x \geq -1 \implies (1+x)^n \geq 1+n \cdot x$$

by the Principle of Induction.

**Base case:**  $n = 0$ . The statement  $P(0)$  reduces to

$$\forall x \in \mathbb{R}. x \geq -1 \implies 1 \geq 1$$

and holds vacuously.

**Inductive step:**  $n = k+1$ . Let  $k$  be an arbitrary natural number, and assume  $P(k)$ ; i.e. assume the Inductive Hypothesis

$$\forall x \in \mathbb{R}. x \geq -1 \implies (1+x)^k \geq 1+k \cdot x \quad \textcircled{\text{IH}}$$

We need show that  $P(k+1)$  also holds; i.e. that

$$\forall x \in \mathbb{R}. x \geq -1 \implies (1+x)^{k+1} \geq 1+(k+1) \cdot x$$

To this end, we let  $y$  be an arbitrary real number, assume further that

$$y \geq -1 \quad \textcircled{1}$$

and proceed to show that

$$(1+y)^{k+1} \geq 1+(k+1) \cdot y \quad \textcircled{2}$$

From  $\textcircled{\text{IH}}$ , by instantiation and Modus Ponens using  $\textcircled{1}$ , one concludes that

$$(1+y)^k \geq 1+k \cdot y$$

and from this, since by  $\textcircled{1}$  we have  $1+y \geq 0$ , it follows that

$$(1+y)^{k+1} = (1+y)^k \cdot (1+y) \geq (1+k \cdot y) \cdot (1+y) = 1+(k+1) \cdot y + k \cdot y^2$$

Thus, from the fact that  $k \cdot y^2 \geq 0$ , ② holds.

3. Recall that the Fibonacci numbers  $F_n$  for  $n \in \mathbb{N}$  are defined recursively by  $F_0 = 0$ ,  $F_1 = 1$ , and  $F_{n+2} = F_n + F_{n+1}$  for  $n \in \mathbb{N}$ .

a) Prove Cassini's Identity: For all  $n \in \mathbb{N}$ ,

$$F_n \cdot F_{n+2} = F_{n+1}^2 + (-1)^{n+1}$$

We prove

$$\forall n \in \mathbb{N}. F_n \cdot F_{n+2} = F_{n+1}^2 + (-1)^{n+1}$$

by the Principle of Induction.

**Base case:**  $n = 0$ . We have that

$$F_0 \cdot F_2 = F_1^2 + (-1)^1$$

because  $F_0 = 0$  and  $F_1 = 1$ .

**Inductive step:**  $n = k + 1$ . For any natural number  $k$ , assume the Induction Hypothesis

$$F_n \cdot F_{k+2} = F_{k+1}^2 + (-1)^{k+1}$$

which can be rearranged to the following form by subtracting  $(-1)^{k+1}$ :

$$F_{k+1}^2 = (-1)^k + F_n \cdot F_{k+2} \quad (\text{IH})$$

We need show that

$$F_{k+1} \cdot F_{(k+1)+2} = F_{(k+1)+1}^2 + (-1)^{(k+1)+1}$$

i.e. that

$$F_{k+1} \cdot F_{k+3} = F_{k+2}^2 + (-1)^k$$

for which one calculates as follows:

$$\begin{aligned} F_{k+1} \cdot F_{k+3} &= F_{k+1}^2 + F_{k+1} \cdot F_{k+2} && (F_{k+3} = F_{k+1} + F_{k+2}) \\ &= (-1)^k + F_n \cdot F_{k+2} + F_{k+1} \cdot F_{k+2} && (\text{by } (\text{IH})) \\ &= (-1)^k + F_{k+2}^2 && (F_{k+2} = F_k + F_{k+1}) \end{aligned}$$

b) Prove that for all natural numbers  $k$  and  $n$ ,

$$F_{n+k+1} = F_{n+1} \cdot F_{k+1} + F_n \cdot F_k$$

We prove that

$$\forall k \in \mathbb{N}. P(k)$$

for  $P(k)$  the statement

$$\forall n \in \mathbb{N}. F_{n+k+1} = F_{n+1} \cdot F_{k+1} + F_n \cdot F_k$$

by the Principle of Induction.

**Base case:** We need show that

$$\forall n \in \mathbb{N}. F_{n+1} = F_{n+1} \cdot F_1 + F_n \cdot F_0$$

which holds because  $F_1 = 1$  and  $F_0 = 0$ .

**Inductive step:** For an arbitrary natural number  $k$ , assume the Induction Hypothesis

$$\forall n \in \mathbb{N}. F_{n+k+1} = F_{n+1} \cdot F_{k+1} + F_n \cdot F_k \quad (\text{IH})$$

We need show that

$$\forall n \in \mathbb{N}. F_{n+(k+1)+1} = F_{n+1} \cdot F_{(k+1)+1} + F_n \cdot F_{k+1}$$

i.e. that

$$\forall n \in \mathbb{N}. F_{n+k+2} = F_{n+1} \cdot F_{k+2} + F_n \cdot F_{k+1} \quad (1)$$

To this end, we let  $m$  be an arbitrary natural number and proceed to show the equivalent identity:

$$F_{(m+1)+k+1} = F_{m+1} \cdot F_{k+2} + F_m \cdot F_{k+1} \quad (2)$$


Indeed, instantiating the universally-quantified Induction Hypothesis (IH) for the natural number  $m + 1$ , one has that


$$F_{(m+1)+k+1} = F_{(m+1)+1} \cdot F_{k+1} + F_{m+1} \cdot F_k$$

from which one further calculates as follows:

$$\begin{aligned} & F_{(m+1)+1} \cdot F_{k+1} + F_{m+1} \cdot F_k \\ &= F_m \cdot F_{k+1} + F_{m+1} \cdot F_{k+1} + F_{m+1} \cdot F_k && (F_{(m+1)+1} = F_m + F_{m+1}) \\ &= F_m \cdot F_{k+1} + F_{m+1} \cdot F_{k+2} && (F_{k+2} = F_k + F_{k+1}) \end{aligned}$$

to conclude (2).

 This is an example of a proposition that could also be established by nested induction: rather than show (1) directly for an arbitrary  $n \in \mathbb{N}$ , we could do another base case for  $n = 0$  and inductive case for  $n = m + 1$ . It's not always obvious when this is required, but quite often results in a lengthier, but simpler proof.

 If either of  $k$  or  $n$  is positive, this identity gives a way of expanding  $F_{n+k}$  as a sum of products of Fibonacci numbers – a useful property whenever the index is a sum.

c) Deduce that  $F_n \mid F_{l \cdot n}$  for all natural numbers  $n$  and  $l$ .

We prove that

$$\forall l \in \mathbb{N}. P(l)$$

for  $P(l)$  the statement

$$\forall n \in \mathbb{N}. F_n \mid F_{l \cdot n}$$

by the Principle of Induction.

**Base case:** We need to show that

$$\forall n \in \mathbb{N}. F_n \mid F_{0 \cdot n}$$

i.e. that

$$\forall n \in \mathbb{N}. F_n \mid 0$$

which holds because we know that every integer divides 0 from §1.2.1(b).

**Inductive step:** For an arbitrary natural number  $l$ , assume the Induction Hypothesis

$$\forall n \in \mathbb{N}. F_n \mid F_{l \cdot n} \quad (\text{IH})$$

We need to show that

$$\forall n \in \mathbb{N}. F_n \mid F_{(l+1) \cdot n}$$


i.e. that

$$\forall n \in \mathbb{N}. F_n \mid F_{l \cdot n + n}$$

To this end, let  $n \in \mathbb{N}$  be an arbitrary natural number. We first consider the case when  $n = 0$ : we have  $F_0 \mid F_{l \cdot 0 + 0}$  from the fact that  $0 \mid 0$  (see §1.2.1(a)). Otherwise, we can express  $F_{l \cdot n + n}$  as  $F_{l \cdot n + (n-1)+1}$  and expand using §4.2.3(b) as follows:

$$\begin{aligned} & F_{l \cdot n + (n-1)+1} \\ &= F_{l \cdot n + 1} \cdot F_{(n-1)+1} + F_{l \cdot n} \cdot F_{n-1} && \text{(by §4.2.3(b))} \\ &= F_{l \cdot n + 1} \cdot F_n + k \cdot F_n \cdot F_{n-1} && \text{(by (IH), } \exists k \in \mathbb{Z}. F_{l \cdot n} = k \cdot F_n) \\ &= F_n \cdot (F_{l \cdot n + 1} + k \cdot F_{n-1}) \end{aligned}$$

Thus,  $F_{(l+1) \cdot n} = k' \cdot F_n$  for  $k' = F_{l \cdot n + 1} + k \cdot F_{n-1}$ , showing that  $F_n \mid F_{(l+1) \cdot n}$ , as required.

 Words like “deduce” and “conclude” are a dead giveaway that you should be using properties you showed in a previous part of the question, so you should always try to transform the proposition or play around with your assumptions until a previous lemma could be applied – this step often takes care of the “hard part” of the proof. In this exercise the inductive step gave us  $F_{l \cdot n + n}$ ; since the index is a sum of two natural numbers with  $n$  positive, we notice that the previous identity can be applied to expand the term into two more “manageable” subterms.

- d) Prove that  $\text{gcd}(F_{n+2}, F_{n+1})$  terminates with output 1 in  $n$  steps for all positive integers  $n$ .

We prove that

$\forall n \in \mathbb{N}. \text{gcd}(F_{n+2}, F_{n+1})$  terminates with output 1 in  $n$  steps

by the Principle of Induction.

**Base case:** We need to show that

$\text{gcd}(F_3, F_2)$  terminates with output 1 in 1 step

Since  $F_3 = 2$  and  $F_2 = 1$  and  $1 \mid 2$ , the algorithm terminates with the base case of  $F_2 = 1$  after one step.

**Inductive step:** For an arbitrary natural number  $k$ , assume the Induction Hypothesis

$\text{gcd}(F_{k+2}, F_{k+1})$  terminates with output 1 in  $k$  steps (IH)

We need to prove that

$\text{gcd}(F_{k+3}, F_{k+2})$  terminates with output 1 in  $k + 1$  steps

By the definition of Fibonacci numbers,  $F_{k+3} = F_{k+2} + F_{k+1}$ . Since  $F_{k+2} \geq F_{k+1}$ , this is a valid quotient-remainder decomposition of  $F_{k+3}$  so by the Division Theorem we have that  $\text{quo}(F_{k+3}, F_{k+2}) = 1$  and  $\text{rem}(F_{k+3}, F_{k+2}) = F_{k+1}$ . As  $F_{k+1}$  is positive,  $F_{k+2} \nmid F_{k+3}$  and  $\text{gcd}(F_{k+3}, F_{k+2})$  steps to  $\text{gcd}(F_{k+2}, \text{rem}(F_{k+3}, F_{k+2})) = \text{gcd}(F_{k+2}, F_{k+1})$ . By the (IH), this terminates with output 1 in  $k$  steps; thus, starting with the additional computation step,  $\text{gcd}(F_{k+3}, F_{k+2})$  terminates with output 1 in  $k + 1$  steps.

e) Deduce also that:

(i) for all positive integers  $n < m$ ,  $\text{gcd}(F_m, F_n) = \text{gcd}(F_{m-n}, F_n)$ ,

and hence that:

(ii) for all positive integers  $m$  and  $n$ ,  $\text{gcd}(F_m, F_n) = F_{\text{gcd}(m,n)}$ .

Firstly, we prove the following statement equivalent to (i):

For all positive integers  $n$  and natural numbers  $k$ ,

$$\text{gcd}(F_{n+k+1}, F_n) = \text{gcd}(F_{k+1}, F_n)$$

We make use of the following corollary/restatement of [Theorem 61](#), which allows us to use properties of Euclid's Algorithm in reasoning about gcds:

For all positive integers  $m$  and  $n$ ,  $\text{gcd}(m, n) = \text{gcd}(m, n)$ .

In particular, we can adapt the recursive case of the definition of  $\text{gcd}$  into:

$$\forall m, n \in \mathbb{Z}^+. \text{gcd}(m, n) = \text{gcd}(\text{rem}(m, n), n) \quad \textcircled{1}$$

and the previous part §4.2.3(d) (shifted to positive integers) into:

$$\forall m \in \mathbb{Z}^+. \gcd(F_{m+1}, F_m) = 1 \quad (2)$$

Now, let  $n$  be a positive integer and  $k$  a natural number. Then,

$$\begin{aligned} \gcd(F_{n+k+1}, F_n) &= \gcd(F_{n+1} \cdot F_{k+1} + F_n \cdot F_k, F_n) && \text{(by §4.2.3(b))} \\ &= \gcd(\text{rem}(F_{n+1} \cdot F_{k+1} + F_n \cdot F_k, F_n), F_n) && \text{(by (1))} \\ &= \gcd(F_{n+1} \cdot F_{k+1}, F_n) && \text{(by §2.1.3(a))} \\ &= \gcd(F_{k+1}, F_n) && \text{(by §3.2.3 and (2))} \end{aligned}$$

Secondly, we prove the following statement from which (ii) follows:

for all positive integers  $l$ ,  $P(l)$

where  $P(l)$  is the statement:

for all positive integers  $m, n$ ,  
if  $\text{gcd0}(n, m)$  terminates in  $l$  steps then  $\gcd(F_m, F_n) = F_{\gcd(m, n)}$

for  $\text{gcd0}$  the function from §3.3.3. The proof is by the Principle of Induction.

**Base case:** Let  $m, n$  be arbitrary positive integers. Assume that  $\text{gcd0}(m, n)$  terminates in 1 step. Then  $m = n$  and  $\gcd(F_m, F_n) = F_m = F_{\gcd(m, n)}$ .

**Inductive step:** Let  $l$  be an arbitrary positive integer, and assume the Induction Hypothesis  $P(l)$ . Further, let  $m, n$  be arbitrary positive integers, and assume that  $\text{gcd0}(m, n)$  terminates in  $l + 1$  steps. Then, for  $p = \min(m, n)$  and  $q = \max(m, n)$ ,  $\text{gcd0}(m, n) = \text{gcd0}(p, q - p)$  and  $\text{gcd0}(p, q - p)$  terminates in  $l$  steps. Thus, by the Induction Hypothesis, we have that  $\gcd(F_{q-p}, F_p) = F_{\gcd(q-p, p)}$ . Finally, since by the previous item,  $\gcd(F_m, F_n) = \gcd(F_q, F_p) = \gcd(F_{q-p}, F_p)$  and  $F_{\gcd(q-p, p)} = F_{\gcd(q, p)} = F_{\gcd(m, n)}$  we are done.

One can intuitively deduce that property (ii) holds because we are performing the simplified Euclid's Algorithm (with repeated subtraction rather than remainder) on the indices of the Fibonacci number via a repeated application of property (i). This is indeed the case, but formulating this into a proof is far from obvious. Given that this is an exercise sheet on inductive proofs, we could try doing induction on  $m$  or  $n$ , only to notice that we can't make use of the inductive hypothesis in any meaningful way. Indeed, the "repetition" that we're trying to capture has nothing to do with the numerical value of  $m$  or  $n$  directly, but rather the number of times we have to apply property (i) to compute their gcd. Given  $m, n \in \mathbb{Z}^+$ , we either cannot apply (i) because  $m$  and  $n$  are equal, or we can apply it once to get  $\gcd(F_{m-n}, F_n)$ , recursively apply it  $l$  more times to get  $F_{\gcd(m-n, n)}$ , and then "unapply" one step of  $\text{gcd0}$  to get  $F_{\gcd(m, n)}$ .

Extracting a strong enough induction hypothesis from this intuition is still nontrivial and requires us to explicitly refer to the termination of  $\text{gcd0}$ . Moreover,  $m$  and  $n$

are universally quantified in the induction statement and the required property  $\gcd(F_m, F_n) = F_{\gcd(m,n)}$  is made dependent on a termination hypothesis that refers to the induction variable  $l$ , rather than relating the two with a conjunction. This means that when proving the inductive case, we can *assume* that  $\text{gcd0}(n, m)$  terminates in more than one step, and execute one step of the algorithm manually by applying property (i). It may take several attempts to construct sufficiently strong induction hypotheses, and as this exercise shows, they are not always as direct as case-analysing on a positive/nonnegative integer that is quantified over in the proposition.

f) Show that for all positive integers  $m$  and  $n$ ,  $(F_m \cdot F_n) \mid F_{m \cdot n}$  if  $\gcd(m, n) = 1$ .

Since  $m$  and  $n$  are coprime, §4.2.3(e) gives:

$$\gcd(F_m, F_n) = F_{\gcd(m,n)} = F_1 = 1$$

implying that  $F_m$  and  $F_n$  are themselves coprime. From §4.2.3(c) we know that  $F_m \mid F_{m \cdot n}$  and  $F_n \mid F_{m \cdot n}$ . This, together with coprimality of  $F_m$  and  $F_n$  and §3.2.2 implies that  $F_m \cdot F_n \mid F_{m \cdot n}$ , as required.

g) Conjecture and prove theorems concerning the following sums for any natural number  $n$ :

(i)  $\sum_{i=0}^n F_{2 \cdot i}$

After some test cases we conjecture the following identity:

$$\sum_{i=0}^n F_{2 \cdot i} = F_{2n+1} - 1$$

and prove it by the Principle of Induction.

**Base case:**  $n = 0$ . The sum consists of a single term  $F_{2 \cdot 0} = F_0 = 0$ , which equals  $F_{2 \cdot 0 + 1} - 1 = F_1 - 1 = 0$ .

**Inductive step:**  $n = k + 1$ . We assume the Induction Hypothesis

$$\sum_{i=0}^k F_{2 \cdot i} = F_{2k+1} - 1 \quad (\text{IH})$$

and prove that

$$\sum_{i=0}^{k+1} F_{2 \cdot i} = F_{2(k+1)+1} - 1$$

We can calculate as follows:

$$\begin{aligned} \sum_{i=0}^{k+1} F_{2 \cdot i} &= F_{2 \cdot (k+1)} + \sum_{i=0}^k F_{2 \cdot i} \\ &= F_{2k+2} + F_{2k+1} - 1 \\ &= F_{2k+3} - 1 = F_{2(k+1)+1} - 1 \end{aligned} \quad (\text{by } (\text{IH}))$$

(ii)  $\sum_{i=0}^n F_{2 \cdot i+1}$



We conjecture the following identity:

$$\sum_{i=0}^n F_{2 \cdot i+1} = F_{2n+2}$$

and prove it by the Principle of Induction.

**Base case:**  $n = 0$ . The sum consists of a single term  $F_{2 \cdot 0+1} = F_1 = 1$ , which equals  $F_{2 \cdot 0+2} = F_2 = 1$ .

**Inductive step:**  $n = k + 1$ . We assume the Induction Hypothesis

$$\sum_{i=0}^k F_{2 \cdot i+1} = F_{2k+2} \quad (\text{IH})$$

and prove that

$$\sum_{i=0}^{k+1} F_{2 \cdot i+1} = F_{2(k+1)+2}$$

We can calculate as follows:

$$\begin{aligned} \sum_{i=0}^{k+1} F_{2 \cdot i+1} &= F_{2 \cdot (k+1)+1} + \sum_{i=0}^k F_{2 \cdot i+1} \\ &= F_{2k+3} + F_{2k+2} \\ &= F_{2k+4} = F_{2(k+1)+2} \end{aligned} \quad (\text{by } (\text{IH}))$$

(iii)  $\sum_{i=0}^n F_i$

We conjecture the following identity:

$$\sum_{i=0}^n F_i = F_{n+2} - 1$$

We can prove this by induction as before. Instead, we derive it from the previous two results by case-analysis on  $n$ :

**Case**  $n = 2k$ . If  $k$  is 0, the sum is  $0 = F_{0+2} - 1$ . Otherwise, the sum consists of the first  $k$  even Fibonacci numbers plus the first  $(k - 1)$  odd Fibonacci numbers:

$$\sum_{i=0}^{2k} F_i = \left( \sum_{i=0}^k F_{2 \cdot i} \right) + \left( \sum_{i=0}^{k-1} F_{2 \cdot i+1} \right) = F_{2k+1} + F_{2k} - 1 = F_{2k+2} - 1$$

**Case**  $n = 2k + 1$ . The sum consists of the sum of the first  $k$  even Fibonacci numbers plus the first  $k$  odd Fibonacci numbers:

$$\sum_{i=0}^{2k+1} F_i = \left( \sum_{i=0}^k F_{2 \cdot i} \right) + \left( \sum_{i=0}^k F_{2 \cdot i+1} \right) = F_{2k+1} - 1 + F_{2k+2} = F_{(2k+1)+2} - 1$$

### 4.3. Optional exercises

1. Recall the `gcd0` function from §3.3.3. Use the Principle of Mathematical Induction from basis 2 to formally establish the following correctness property of the algorithm:

For all natural numbers  $l \geq 2$ , we have that for all positive integers  $m, n$ , if  $m + n \leq l$  then `gcd0`( $m, n$ ) terminates.

As suggested, we proceed by Mathematical Induction from basis 2.

**Base case:** We need show that for all positive integers  $m, n$ , if  $m + n \leq 2$  then `gcd0`( $m, n$ ) terminates. To this end, we let  $m$  and  $n$  be arbitrary positive integers, and assume that  $m + n \leq 2$ . Then,  $m = n = 1$  and `gcd0`( $m, n$ ) terminates.

**Inductive step:** Let  $l$  be an arbitrary natural number greater than or equal 2, and assume the Induction Hypothesis

For all positive integers  $m, n$ , if  $m + n \leq l$  then `gcd0`( $m, n$ ) terminates. Ⓜ

We need show that for all positive integers  $m, n$ , if  $m + n \leq l + 1$  then `gcd0`( $m, n$ ) terminates. To this end, we let  $a, b$  be arbitrary positive integers, assume that  $a + b \leq l + 1$ , and proceed to prove that `gcd0`( $a, b$ ) terminates.

We consider three cases.

- If  $a = b$ , then `gcd0`( $a, b$ ) terminates.
- If  $a < b$ , then `gcd0`( $a, b$ ) = `gcd0`( $a, b - a$ ). Moreover, by the Inductive Hypothesis Ⓜ, we have that

if  $a + (b - a) \leq l$  then `gcd0`( $a, b - a$ ) terminates,

and since

$$a + (b - a) = b \leq l + 1 - a \leq l$$

it follows that `gcd0`( $a, b - a$ ) terminates and therefore that so does `gcd0`( $a, b$ ).

- If  $b < a$ , then `gcd0`( $a, b$ ) = `gcd0`( $a, a - b$ ). Moreover, by the Inductive Hypothesis Ⓜ, we have that

if  $b + (a - b) \leq l$  then `gcd0`( $a, a - b$ ) terminates,

and since

$$b + (a - b) = a \leq l + 1 - b \leq l$$

it follows that `gcd0`( $a, a - b$ ) terminates and therefore that so does `gcd0`( $a, b$ ).

2. The set of *univariate polynomials* (over the rationals) on a variable  $x$  is defined as that of arithmetic expressions equal to those of the form  $\sum_{i=0}^n a_i \cdot x^i$ , for some  $n \in \mathbb{N}$  and some *coefficients*  $a_0, a_1, \dots, a_n \in \mathbb{Q}$ .

(a) Show that if  $p(x)$  and  $q(x)$  are polynomials then so are  $p(x) + q(x)$  and  $p(x) \cdot q(x)$ .

Let  $p(x) = \sum_{i=0}^m a_i \cdot x^i$  and  $q(x) = \sum_{j=0}^n b_j \cdot x^j$  be polynomials, and assume without loss of generality that  $m > n$ . For simplicity, we extend the coefficients  $a_i$  and  $b_j$  to all natural indices, with  $a_i = 0$  for  $m < i$  and  $b_j = 0$  for  $n < j$ . Then, the sum  $p(x) + q(x)$  is a polynomial (of degree  $m$ ) because it is of the form:

$$p(x) + q(x) = \sum_{i=0}^m (a_i + b_i) \cdot x^i$$

where the coefficients  $a_i + b_i$  are rational numbers since  $\mathbb{Q}$  is closed under addition.

For the product  $p(x) \cdot q(x)$ , we calculate using the distributivity of multiplication over addition:

$$\begin{aligned} p(x) \cdot q(x) &= \left( \sum_{i=0}^m a_i \cdot x^i \right) \cdot \left( \sum_{j=0}^n b_j \cdot x^j \right) \\ &= \sum_{i=0}^m \left( a_i \cdot x^i \cdot \sum_{j=0}^n b_j \cdot x^j \right) \\ &= \sum_{i=0}^m \sum_{j=0}^n a_i \cdot x^i \cdot b_j \cdot x^j \\ &= \sum_{i=0}^m \sum_{j=0}^n a_i \cdot b_j \cdot x^{i+j} \end{aligned}$$

The number of terms in the sum of a fixed degree  $d$  will be equal to the number of ways one can construct  $d$  as a sum of an  $i \leq m$  and a  $j \leq n$ ; for example there will be at most one term of degree 0 or  $m+n$ , two terms of degree  $1 = 1+0 = 0+1$  and  $m+n-1 = m+(n-1) = (m-1)+n$ , three of degree 2 and  $m+n-2$  and so on. Terms of the same degree can be combined, with their coefficients getting added together. Using our extended coefficient indexing, the coefficient of the term of degree  $k$  can be concisely expressed as:

$$c_k = \sum_{j=0}^k a_j \cdot b_{k-j}$$

As expected,  $c_0 = a_0 \cdot b_0$  (the constant terms),  $c_{m+n} = a_0 \cdot b_{m+n} + \dots + a_m \cdot b_n + \dots + a_{m+n} b_0 = 0 + \dots + a_m \cdot b_n + \dots + 0$  (most of the coefficients are “out of range” and are 0) and  $c_n = a_0 \cdot b_n + a_1 \cdot b_{n-1} + \dots + a_n \cdot b_0$  ( $n$  nonzero coefficients). Since these are all rational numbers, the product of two polynomials is indeed a polynomial (of degree  $m+n$ ) because it is of the form:

$$p(x) \cdot q(x) = \sum_{k=0}^{m+n} c_k \cdot x^k$$

- (b) Deduce as a corollary that, for all  $a, b \in \mathbb{Q}$ , the linear combination  $a \cdot p(x) + b \cdot q(x)$  of two polynomials  $p(x)$  and  $q(x)$  is a polynomial.

Every rational number  $a$  can be seen as a polynomial of degree 0, with its only coefficient being  $a$ . Thus,  $a \cdot p(x)$  is a product of polynomials and hence is a polynomial. The sum of two such expressions is still a polynomial, so we can conclude that the linear combination  $a \cdot p(x) + b \cdot q(x)$  of two polynomials for  $a, b \in \mathbb{Q}$  is a polynomial.

- (c) Show that there exists a polynomial  $p_2(x)$  such that  $p_2(n) = \sum_{i=0}^n i^2 = 0^2 + 1^2 + \cdots + n^2$  for every  $n \in \mathbb{N}$ .<sup>1</sup>

*Hint:* Note that for every  $n \in \mathbb{N}$ ,

$$(n+1)^3 = \sum_{i=0}^n (i+1)^3 - \sum_{i=0}^n i^3$$

The required polynomial is

$$p_2(n) = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$$

We show that this is a sum of squares for any  $n \in \mathbb{N}$  by induction.

**Base case:**  $n = 0$ . The polynomial reduces to 0, which is the sum of the square number  $0 = 0^2$ .

**Inductive step:**  $n = k + 1$ . Assume the Induction Hypothesis:

$$p_2(k) = \frac{1}{3}k^3 + \frac{1}{2}k^2 + \frac{1}{6}k = \sum_{i=0}^k i^2 \quad (\text{IH})$$

We need to prove that

$$p_2(k+1) = \sum_{i=0}^{k+1} i^2$$

The polynomial expands as follows:

$$\begin{aligned} p_2(k+1) &= \frac{1}{3}(k+1)^3 + \frac{1}{2}(k+1)^2 + \frac{1}{6}(k+1) \\ &= \frac{1}{3}k^3 + k^2 + k + \frac{1}{3} + \frac{1}{2}k^2 + k + \frac{1}{2} + \frac{1}{6}k + \frac{1}{6} \\ &= \left( \frac{1}{3}k^3 + \frac{1}{2}k^2 + \frac{1}{6}k \right) + k^2 + 2k + \frac{1}{3} + \frac{1}{2} + \frac{1}{6} \\ &= \sum_{i=0}^k i^2 + (k^2 + 2k + 1) \quad (\text{by (IH)}) \\ &= \sum_{i=0}^k i^2 + (k+1)^2 = \sum_{i=0}^{k+1} i^2 \end{aligned}$$

Thus  $p_2(k+1)$  is the sum of consecutive squares, as required.

<sup>1</sup>Chapter 2.5 of *Concrete Mathematics* by R.L. Graham, D.E. Knuth and O. Patashnik looks at this in great detail.

As is usual with existence proofs, the hard work is done behind the scenes and we start off the formal proof by magically producing a witness that just so happens to satisfy the required property. The required witness for the existence was calculated from the supplied hint:

$$\begin{aligned}
 (n+1)^3 &= \sum_{i=0}^n (i+1)^3 - \sum_{i=0}^n i^3 \\
 &= \sum_{i=0}^n (i^3 + 3i^2 + 3i + 1) - \sum_{i=0}^n i^3 \\
 &= \left( \sum_{i=0}^n 3i^2 + 3i + 1 \right) + \sum_{i=0}^n i^3 - \sum_{i=0}^n i^3 \\
 &= \sum_{i=0}^n 3i^2 + 3i + 1 = 3 \cdot \sum_{i=0}^n i^2 + \sum_{i=0}^n 3i + 1
 \end{aligned}$$

Rearranging, we get that

$$\begin{aligned}
 \sum_{i=0}^n i^2 &= \frac{1}{3} \left( (n+1)^3 - \sum_{i=0}^n 3i + 1 \right) \\
 &= \frac{1}{3} \left( n^3 + 3n^2 + 3n + 1 - \left( n + 1 + \frac{3}{2}(n^2 + n) \right) \right) \\
 &= \frac{1}{3} n^3 + n^2 + n + \frac{1}{3} - \frac{1}{3} n - \frac{1}{3} - \frac{1}{2} n^2 + \frac{1}{2} n \\
 &= \frac{1}{3} n^3 + \frac{1}{2} n^2 + \frac{1}{6} n
 \end{aligned}$$

Now, we suspect that this is the right answer, but the formal proof should start with the statement of the answer followed by a proof that it satisfies the required property. This is especially important in this case, when the proposed witness was calculated using the (unverified) hint; separately proving that the polynomial is a sum of squares makes our answer independent of the hint. The formal proof may well be done using a different technique (in this case, induction), but it should not present any unpleasant surprises since our proposed witness is almost certainly correct.

Of course, the statement for this question is a rather obfuscated way of saying “find a formula for the sum of the first  $n$  square numbers”. You may already have it memorised as

$$\sum_{i=0}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

Multiplying things out indeed leads to the formula for the polynomial  $p_2(n)$  we had above. Even if we recognise this shortcut (instead of deriving it from the hint), we still need to prove that the formula works – this is still best accomplished using induction.

- (d) Show that, for every  $k \in \mathbb{N}$ , there exists a polynomial  $p_k(x)$  such that, for all  $n \in \mathbb{N}$ ,  $p_k(n) = \sum_{i=0}^n i^k = 0^k + 1^k + \cdots + n^k$ .

*Hint:* Generalise the hint above, and the similar identity

$$(n+1)^2 = \sum_{i=0}^n (i+1)^2 - \sum_{i=0}^n i^2$$

For  $k \in \mathbb{N}$ ,  $P(k)$  be the statement

There exists a polynomial  $p_k(x)$  such that, for all  $n \in \mathbb{N}$ ,  $p_k(n) = \sum_{i=0}^n i^k$ .

We prove this by the Principle of Strong Induction.

**Base case:** The polynomial needs to satisfy  $p_0(n) = \sum_{i=0}^n i^0$ ; since  $i^0 = 1$ , this is simply equal to  $p_0(n) = n + 1$ , which is a polynomial.

**Inductive step:** Assume the Strong Induction Hypothesis: for all  $0 \leq l \leq k$ ,

there exists a polynomial  $p_l(x)$  such that, for all  $n \in \mathbb{N}$ ,  $p_l(n) = \sum_{i=0}^n i^l$ . (IH)<sub>S</sub>

We need to show that  $P(k+1)$  holds, that is

there exists a polynomial  $p_{k+1}(x)$  such that, for all  $n \in \mathbb{N}$ ,  $p_{k+1}(n) = \sum_{i=0}^n i^{k+1}$

The required witness of existence is

$$p_{k+1}(n) = \frac{1}{k+2} \left( (n+1)^{k+2} - \sum_{j=0}^k \binom{k+2}{j} \cdot p_j(n) \right) \quad (\text{E})$$

This is indeed a polynomial since:

- $p_j(n)$  is a polynomial for all  $0 \leq j \leq k$  by the Strong Induction Hypotheses, and  $\sum_{j=0}^k \binom{k+2}{j} p_j(n)$  is a linear combination of polynomials which is a polynomial;
- $(n+1)^{k+2}$  can be expanded using the Binomial Theorem into a sum of powers of  $n$  with binomial coefficients, so it too is a polynomial;
- the sum of two polynomials is a polynomial, and  $\frac{1}{k+2}$  is a rational coefficient.

We prove that  $p_{k+1}(n) = \sum_{i=0}^n i^{k+1}$  for all  $n \in \mathbb{N}$  by induction on  $n$ .

**Base case:** As before,  $p_{k+1}(0) = 0$ .

**Inductive step:** Assume the Induction Hypothesis

$$p_{k+1}(n) = \sum_{i=0}^n i^{k+1} \quad (\text{IH})$$

and prove that

$$p_{k+1}(n+1) = \sum_{i=0}^{n+1} i^{k+1}$$

First, we note the following two calculations:

$$\begin{aligned}(n+2)^{k+2} &= ((n+1)+1)^{k+2} = \sum_{i=0}^{k+2} \binom{k+2}{i} (n+1)^i && \text{(Binomial Theorem)} \\ &= (n+1)^{k+2} + (k+2) \cdot (n+1)^{k+1} + \sum_{i=0}^k \binom{k+2}{i} (n+1)^i && \text{(extract two summands)}\end{aligned}$$

$$\begin{aligned}&\sum_{j=0}^k \binom{k+2}{j} \cdot p_j(n+1) \\ &= \sum_{j=0}^k \binom{k+2}{j} \cdot \sum_{a=0}^{n+1} a^j = \sum_{a=0}^{n+1} \sum_{j=0}^k \binom{k+2}{j} \cdot a^j && \text{(by } \textcircled{\text{IH}}_S \text{ and distributivity)} \\ &= \sum_{j=0}^k \binom{k+2}{j} \cdot (n+1)^j + \sum_{a=0}^n \sum_{j=0}^k \binom{k+2}{j} \cdot a^j && \text{(extract last summand)}\end{aligned}$$

Combining the two, we have that

$$\begin{aligned}(n+2)^{k+2} - \sum_{j=0}^k \binom{k+2}{j} \cdot p_j(n+1) \\ = (n+1)^{k+2} + (k+1) \cdot (n+1)^{k+1} - \sum_{a=0}^n \sum_{j=0}^k \binom{k+2}{j} \cdot a^j && \textcircled{1}\end{aligned}$$

Now we are ready to expand the polynomial of the inductive step:

$$\begin{aligned}&p_{k+1}(n+1) \\ &= \frac{1}{k+2} \left( (n+2)^{k+2} - \sum_{j=0}^k \binom{k+2}{j} \cdot p_j(n+1) \right) \\ &= \frac{1}{k+2} \left( (n+1)^{k+2} + (k+2) \cdot (n+1)^{k+1} - \sum_{a=0}^n \sum_{j=0}^k \binom{k+2}{j} \cdot a^j \right) && \text{(by } \textcircled{1}) \\ &= \frac{1}{k+2} \left( (n+1)^{k+2} - \sum_{a=0}^n \sum_{j=0}^k \binom{k+2}{j} \cdot a^j \right) + (n+1)^{k+1} \\ &= \frac{1}{k+2} \left( (n+1)^{k+2} - \sum_{j=0}^k \binom{k+2}{j} \cdot \sum_{a=0}^n a^j \right) + (n+1)^{k+1} \\ &= \frac{1}{k+2} \left( (n+1)^{k+2} - \sum_{j=0}^k \binom{k+2}{j} \cdot p_j(n) \right) + (n+1)^{k+1} && \text{(by } \textcircled{\text{IH}}_S) \\ &= p_{k+1}(n) + (n+1)^{k+1} = \sum_{i=0}^n i^{k+1} + (n+1)^{k+1} = \sum_{i=0}^{n+1} i^{k+1} && \text{(by } \textcircled{\text{E}} \text{ and } \textcircled{\text{IH}})\end{aligned}$$

Thus, we have shown (by the nested Mathematical Induction) that our definition of

$p_{k+1}(n)$  by ⑤ indeed satisfies  $p_{k+1}(n) = \sum_{i=0}^n i^{k+1}$  for all  $n \in \mathbb{N}$ . Then, by the outer Strong Induction, we can conclude that there exists a polynomial  $p_k(n)$  for all  $k \in \mathbb{N}$  that satisfies  $p_k(n) = \sum_{i=0}^n i^k$  for all  $n \in \mathbb{N}$ .

♪ Once again, we found the witness ⑤ by calculating backwards from the (conjectured) generalisation of the hint

$$(n+1)^k = \sum_{i=0}^n (i+1)^k - \sum_{i=0}^n i^k$$

We *could* prove that this holds, but we can also use it without proof to derive the witness, as long as we then formally show that the witness is correct. Given that the property is only used behind the scenes as an “educated guess”, it will not invalidate the proof even if the conjecture is actually incorrect. The calculation of the witness is as follows:

$$\begin{aligned} (n+1)^{k+2} &= \sum_{m=0}^n (m+1)^{k+2} - \sum_{m=0}^n m^{k+2} \\ &= \left( \sum_{m=0}^n \sum_{j=0}^{k+2} \binom{k+2}{j} \cdot m^j \right) - \sum_{m=0}^n m^{k+2} && \text{(Binomial Theorem)} \\ &= \left( \sum_{j=0}^{k+2} \sum_{m=0}^n \binom{k+2}{j} \cdot m^j \right) - \sum_{m=0}^n m^{k+2} && \text{(commute summation)} \\ &= \sum_{j=0}^{k+1} \sum_{m=0}^n \binom{k+2}{j} \cdot m^j && \text{(subtract last summand)} \\ &= \sum_{m=0}^n \binom{k+2}{k+1} \cdot m^{k+1} + \sum_{j=0}^k \sum_{m=0}^n \binom{k+2}{j} \cdot m^j && \text{(extract last summand)} \\ &= (k+2) \sum_{m=0}^n m^{k+1} + \sum_{j=0}^k \binom{k+2}{j} \cdot \sum_{m=0}^n m^j && \text{(binom. coefficient)} \\ &= (k+2) \sum_{m=0}^n m^{k+1} + \sum_{j=0}^k \binom{k+2}{j} \cdot p_j(n) && \text{(\textcircled{H}}_S) \end{aligned}$$

We rearrange this to get  $\sum_{m=0}^n m^{k+1}$  and set that as the witness formula for  $p_{k+1}(n)$ .

♪ This proof is a rather involved example of a nested, mixed induction proof: we do strong induction over  $k \in \mathbb{N}$  and mathematical induction over  $n \in \mathbb{N}$  when proving that our proposed witness ⑤ for  $p_{k+1}(n)$  (the inductive case of the outer induction) is correct. The strong induction hypothesis  $\textcircled{H}_S$  is used throughout the proof, both in the derivation of the witness and the proof of its correctness.

♪ Note that we haven’t actually constructed a closed-form expression for  $p_k(n)$ , but a recursive algorithm for computing it from formulae for lower degrees. Importantly, we established that the recursive expression is indeed a polynomial using the clos-



ure properties proved in earlier parts. This is sufficient to prove that there exists a polynomial expression for  $\sum_{i=0}^n i^k$ , but of course one has to do quite some additional work to extract the degree and the coefficients of the polynomial from the recursive construction. The general, closed-form expression is known as Faulhaber's Formula and features the Bernoulli numbers, a rather irregular-looking sequence of rational numbers used throughout mathematics; for instance,  $B_{14} = \frac{7}{6}$ ,  $B_{15} = 0$ ,  $B_{16} = -\frac{3617}{510}$ .