MATH 323 DISCRETE MATHEMATICS, *Handout 2*

Mathematical proof, especially proof by induction

Laurence Barker, Mathematics Department, Bilkent University,
version: 26 Feburary 2009.

In this handout, we discuss two fundamental techniques of mathematical argument which will be used frequently in the course: *proof by contradiction*, which is fairly straightforward, and *proof by mathematical induction*, which is a little more subtle. We give some illustrative applications to graph theory. We shall also give an appendix concerning some subtleties behind the notion of mathematical induction.

I suggest that you learn the proofs of the theorems, because the proofs can sometimes be adapted to solve problems which the theorems do not cover. Besides, I shall sometimes ask for proofs of the theorems in the exams.

Please let me know of your corrections to these notes.

♣ ♡ ♢ ♠

Our aims are different. In any handouts, and also in class, *my* aim is to help you to learn the material. Part of *your* aim is to do well in the exams. So I might be inclined to comment on material which I — and maybe you too — find interesting. Whereas you might might prefer to have revision notes. Well, you will have to write your revision notes for yourself. That is part of the learning process.

Some of you may come to these notes just one or two days before Midterm I. Then I guess you are not interested in the course. That is okay. There is much more to life than discrete mathematics, and I guess you have your own priorities. However, in that case, I am under no obligation to help you, and you will probably not find these notes very useful.

♣ ♡ ♢ ♠

One of the purposes of the course is to develop the skill of mathematical proof. Certainly, you already of some experience of this, but mostly in areas of mathematics that have what used to be called a *calculus*: the calculus of probability, the calculus of linear algebra and, most famously, the area that is nowadays simply called "calculus"; the differential and integral calculus. In this course, though, the mathematical machinery is less powerful, and greater effort has to be expended on explaining ideas.

A proof is a communication to a reader, a very clear explanation. Obvious statements are those which are already very clear without any explanation. In a proof, every piece of the proof must be obvious in view of the previous pieces. When composing a proof — this includes justifications of solutions to particular problems — you should try to imagine what effect your words and equations will have on the reader. I suggest that you imagine your reader to be an ordinary first-year undergraduate. Even in exams, you should write as a communication not to me, but to someone at your own level of knowledge and experience. My marking is based on the criterion: *how helpful would this answer be towards explaining the argument to other students in the class?*

Sometimes, when students visit me for an autopsy of their exam papers, they claim that, although their explanation was incomprehensible, they really did see the ideas behind a question. "Look", they say, "how could I have written this strange cryptic things if the right idea had not been at the back of my mind?" But that is irrelevant. The quality of a proof depends on how clearly it conveys the ideas.

When checking a proof, ask yourself what objections your imaginary reader might raise. Can the reader ask: *What does this mean?* There is a huge difference between "Let x be an integer" and "Assume that x is an integer" and "Therefore x is an integer". If you have written only "x is an integer" then the reader may have difficulty in guessing the status of the statement.

Or, can the reader ask: *Why does this follow?* If one omits an important step then, of course, the reader will find the argument difficult or impossible to follow. On the other hand, if one rambles on about matters that are trivial or irrelevant, then that will obscure the main ideas and it will make the task more difficult for the reader. Just how much should one write in order to consider a proof complete? Well, that is often very hard to judge. With practise, one gets better at recognizing the point where the argument is already as clear as one can make it, and any further comments will just cause obscurity.

1: Proof by contradiction

Argument by contradiction is a technique that is used in ordinary everyday reasoning, not just in mathematics. Until recently, it was usually referred to by its latin name, *reductio ad absurdum* — reduction to absurdity — and actually the older name seems to give a better description of the idea. We assume the negative of what we wish to prove, and then we make deductions until we arrive at something impossible.

The blacksmith knows that, every Saturday, his wife visits either the butcher or the baker. The path the butcher is muddy and, if she visits him, then her shoes will be muddy when she returns home. Last Saturday her shoes were not muddy. Therefore she visited the baker. "What makes you think I visited the baker?" she asks him. The blacksmith replies "Suppose, for a contradiction, that you did not visit the baker. Then you would have visited the butcher, and your shoes would have been muddy. But your shoes were not muddy. This is a contradiction, as required." Or, for a more realistic dialogue, he replies "Did you fly over the muddy path to the butcher on your broomstick?" But that sarcastic reply is just another way of applying the same technique: he has reduced the counter-hypothesis to an absurdity.

The proof of the next result is a mathematical application of the same technique. Again, you may detect a trace of sarcasm in the attitude behind the argument that we shall be presenting.

First, though, some terminology: recall that the set of integers $\mathbb{Z}$ is the set of numbers $..., -2, -1, 0, 1, 2, 3, ...$. The set of rational numbers $\mathbb{Q}$ is the set of numbers having the form a/b where a and b are integers and $b \neq 0$. The set of real numbers $\mathbb{R}$ can be viewed a line stretching infinitely in both directions. All the rational numbers are real numbers, but the following theorem tells us that there exist real numbers that are not rational; such numbers are called *irrational*.

Theorem: *The real number $\sqrt{2}$ is irrational.*

Proof: Suppose, for a contradiction, that there exist integers a and b, with $b \neq 0$, such that $\sqrt{2} = a/b$. Then we may chose a and b to be positive and as small as possible. We have $a^2 = 2b^2$, which is even, so a is even, say, $a = 2\alpha$ for some positive integer α. Then $2\alpha^2 = b^2$,

which is even, so b is even, say, $b = 2\beta$ for some positive integer β. Then $\alpha^2 = 2\beta^2$, in other words, $\sqrt{2} = \alpha/\beta$. But $\alpha < a$ and $\beta < b$, which contradicts the condition that a and b are as small as possible. $\square$

To give an example from graph theory, let us first state two other results without proof.

Theorem: *Given a connected graph T with n vertices and e edges, then $n = e + 1$ if and only if T does not have a cycle.*

A graph T satisfying those two equivalent conditions is called a **tree**.

Proposition: *Let G be a graph with vertices $x_1, ..., x_n$ and with e edges. Then the sum of the degrees of the vertices is $2e = d(x_1) + ... + d(x_n)$.*

We can now give our second illustration of the technique under discussion.

Proposition: *Any tree has a vertex with degree* 1.

Proof: Let T be a tree with vertices $x_1, ..., x_n$ and e edges. By the definition of a tree, the number of vertices is $n = e + 1$. Suppose, for a contradiction, that every vertex has degree $d(x_i) \geq 2$. Then, by the previous proposition,

$$2n - 2 = 2e = d(x_1) + ... + d(x_n) \geq 2n\,.$$

This is a contradiction, as required. $\square$.

One could quibble about the language. Arguably, the inequality $2n - 2 \geq 2n$ is not so much a *contradiction* as a *fallacy*. But anyway, we succeeded in reducing the counter-hypothesis to an absurdity.

Actually, the technique is not really needed for the latest proposition. There are other ways of expressing the same proof. A crucial application of the technique to graph theory will be given in Section 3.

2: Mathematical induction

We shall prove the following proposition.

Proposition: *For all positive integers n, we have $1^2 + 2^2 + ... + n^2 = n(n + 1)(2n + 1)/6$.*

Let $S_n = 1^2 + 2^2 + ...n^2$ and $T_n = n(n + 1)(2n + 1)/6$. We are to show that $S_n = T_n$.

We first note some supporting evidence. By direct calculation,

$$S_1 = 1 = T_1\,, \qquad S_2 = 5 = T_2\,, \qquad S_3 = 14 = T_3\,.$$

As another piece of supporting evidence, we have $1^0 + ... + n^0 = 1 + ... + 1 = n$ and $1^1 + ... + n^1 = 1 + ... + n = n(n + 1)/2 = n^2/2 + n/2$, so it is reasonable to guess that $S_n = An^3 + Bn^2 + Cn$ for some constants A, B, C. Using the values of S_1, S_2, S_3 to solve for A, B, C, we find that $A = 1/3$ and $B = 1/2$ and $C = 1/6$. Observing that $T_n = An^3 + Bn^2 + Cn$, it is reasonable to surmise that $S_n = T_n$.

The kind of reasoning used here is called *retroductive reasoning*. Some further comments on this kind of reasoning appear at the end of this section. However, the reasoning is not purely deductive. We have not yet proved that $S_n = T_n$.

Some further evidence could be gained by checking the equality for some more values of n. Thus,

$$S_4 = 30 = T_4 \ , \qquad S_5 = 55 = T_5 \ , \qquad S_6 = 91 = T_6 \ .$$

But this is still retroductive reasoning, and we have not yet proved the above proposition.

First proof of the proposition: Let $S_n = 1^2 + 2^2 + ...n^2$ and $T_n = n(n+1)(2n+1)/6$. We are to show that $S_n = T_n$. First note that

$$S_1 = 1 = T_1 \ , \qquad S_2 = 5 = T_2 \ , \qquad S_3 = 14 = T_3 \ .$$

Let us consider what would happen if we were to continue with such calculations, stopping only if we find a counter-example. If the equality $S_n = T_n$ fails for some n then, eventually, we will find a counter-example and we shall stop. But, if the equality holds for all n, then we will continue forever. Now let us imagine that, for some given n, we have just finished checking the equation $S_{n-1} = T_{n-1}$. Since

$$T_n - T_{n-1} = \frac{n}{6}\big((n+1)(2n+1) - (n-1)(2n-1)\big) = n^3 = S_n - S_{n-1}$$

it follows that, when we come to check the next case, we shall find that the equality $S_n = T_n$ does hold for that given value of n. Therefore, we shall continue forever, and the equality $S_n = T_n$ must hold for all positive integers n. □

Is the argument that we have just given really a proof of the proposition? Is it a very clear explanation? Well, yes, it is a very clear explanation. It is a proof.

But it is a very clumsy proof. To prove a simple assertion about integers, we had to describe a fantasy about immortal mathematicians calculating forever. It had to be written very carefully so as to make sure that it is clear. If one tries to shorten the argument, there is a danger that the reader may misunderstand it as the circular argument: assuming that $S_{n-1} = T_{n-1}$ for all $n \geq 2$, then $S_n = T_n$ for all $n \geq 1$.

Before giving a better proof, let note some terminology and then make an obvious remark. The set of natural numbers, denoted $\mathbb{N}$, is the set of numbers $0, 1, 2, 3, ...$ (including 0). Thus, the natural numbers are the sizes of the finite sets (the empty set is finite). The set of positive integers, denoted $\mathbb{Z}^+$, is the set of numbers $1, 2, 3, 4, ...$ (not including zero).

Remark: (The Well-Ordering Principle.) *Any non-empty set of natural numbers has a smallest element.*

The Well-Ordering Principle is behind the proof of the above proposition.

Second proof of the proposition: Let $S_n = 1^2 + 2^2 + ...n^2$ and $T_n = n(n+1)(2n+1)/6$. We are to show that $S_n = T_n$. Suppose, for a contradiction, that $S_n \neq T_n$ for some n. Let n be the smallest positive integer such that $S_n \neq T_n$. Since $S_1 = 1 = T_1$, we have $n \geq 2$. Since n is as small as possible, $S_{n-1} = T_{n-1}$. We have

$$T_n - T_{n-1} = \frac{n}{6}\big((n+1)(2n+1) - (n-1)(2n-1)\big) = n^3 = S_n - S_{n-1} \, .$$

Combining this will the equality $S_{n-1} = T_{n-1}$, we deduce that $S_n = T_n$. This is a contradiction, as required. □

Note 1: it was important to tell the reader that we were arguing by contradiction. If we had omitted the sentence "Suppose, for a contradiction, that $S_n \neq T_n$ for some n", then that would have made the argument much harder to follow.

Note 2: I was tempted to write "By the Well-Ordering Principle, there exists a smallest integer n such that $S_n \neq T_n$." But there is no need for that, because the argument is already entirely clear without mentioning that principle. As a matter of fact, we argued in a similar way in Section 1, when proving that $\sqrt{2}$ is irrational, and we felt no need, there, to explicitly mention the Well-Ordering Principle.

But still, there is something awkward about the proof. The usual form of a proof by contradiction might run something like this:

> We are to show that all crows are black. Suppose, for a contradiction, that some crow is not black. Blah blah blah. Therefore all lions are red. Yada yada yada. Therefore some lion is not red. This is a contradiction, as required.

Everything in the section "Blah blah... not red" takes place in the fantasy world of the negative assertion. Nothing in that world is reliable. Back in reality, when the proof is complete, we have discovered nothing whatsoever about lions.

But the second proof of the proposition had the following form.

> We are to show that all crows are black. Suppose, for a contradiction, that some crow is not black. Blah blah blah. Therefore all lions are red. Yada yada yada. Therefore all crows are black. This is a contradiction, as required.

Here, everything in the section "Blah blah... are black" takes place in the fantasy world of the negative assertion. In particular, the line "Therefore all crows are black" resides in the realm of fantasy. So, at that point, the argument is not yet complete. We do need that final line "This is a contradiction, as required". The final line explains to the reader that, since we have destroyed the fantasy where our desired assertion is false, our desired assertion must be true.

The reasoning, here, is valid but convoluted and awkward. So let us try, once more, to give a good proof of the above proposition. We shall make use of another obvious remark.

Remark: (Principle of Mathematical Induction.) *For each natural number n, let P_n be a statement. Suppose that P_0 holds and that P_{n-1} implies P_n for all $n \geq 1$. Then P_n holds for all n.*

To prove an assertion using this principle, there are two parts, sometimes called the *base step*, where we check that P_0 holds, and the *induction step*, where we check that P_{n-1} implies P_n for all $n \geq 1$.

Unlike proof by contradiction, the technique of proof by induction does not correspond to any process in everyday reasoning. I cannot give any illustration involving blacksmiths and butchers and bakers. It is something that has to be learned in a formal way. To get familiar with it, you will need exercise and practise.

Third proof of the proposition, version 1: Let $S_n = 1^2 + 2^2 + ...n^2$ and $T_n = n(n+1)(2n+1)/6$. We are to show that $S_n = T_n$. We argue by induction.

Base Step: We have $S_n = 1 = T_n$.

Induction Step: Suppose that $n \geq 2$ and that $S_{n-1} = T_{n-1}$. We have

$$T_n - T_{n-1} = \frac{n}{6}\big((n+1)(2n+1) - (n-1)(2n-1)\big) = n^3 = S_n - S_{n-1}\,.$$

Combining this will the equality $S_{n-1} = T_{n-1}$, we deduce that $S_n = T_n$. $\square$

A proof is not a computer program! The principle was stated in terms of the set of natural numbers $\mathbb{N}$ (the initial statement was called P_0) but, in the proof, we applied it to the set of positive integers $\mathbb{Z}^+$ (the initial statement was the case $n = 1$). Subroutines of computer programs tend not to work very well when one changes a convention about numbering or parametrization. But our imaginary reader is a human being, not a computer.

If you wish to play safe in the exam, then I suggest you stick to the above form as a template. I recommend it, even if you think you are the brightest student in the class. I will not look down on you for it, because I do understand that good grades are important. However, the terms "base step" and "induction step" belong to introductory literature for babies. Those terms never appear in more advanced texts. Here is a more sophisticated version of the argument.

Third proof of the proposition, version 2: Let $S_n = 1^2+2^2+...n^2$ and $T_n = n(n+1)(2n+1)/6$. We are to show that $S_n = T_n$. We argue by induction. We have $S_1 = 1 = T_1$. Now suppose that $n \geq 2$ and that $S_{n-1} = T_{n-1}$. We have

$$T_n - T_{n-1} = \frac{n}{6}\big((n+1)(2n+1) - (n-1)(2n-1)\big) = n^3 = S_n - S_{n-1}\,.$$

Combining this with the equality $S_{n-1} = T_{n-1}$, we deduce that $S_n = T_n$. □

A mathematical proof is not a magic ritual. Admittedly, it does look a bit like a ritual. Routine techniques tend to become standardized, and then the set phrases, such as "We argue by induction", do tend to seem very much like the incantations of a priest. Nevertheless, everything that appears in a proof is there for the sake of clarity.

The most important line in both versions of the proof is, indeed, that set phrase "We argue by induction." The reader needs this. If we have neglected to inform her that we were arguing by induction, then, again, our argument might have looked like that circular argument: assuming that $S_{n-1} = T_{n-1}$ for all $n \geq 2$, then $S_n = T_n$ for all $n \geq 1$.

When I mark exam questions that need mathematical induction, my first step is to scan through the text looking for the word "induction". If I find it then, starting with full marks, I subtract for mistakes and omissions. If I do not find it then, starting from zero, I add marks for anything that might still somehow be of some help to the reader.

Here is an even more sophisticated proof of the proposition.

Third proof of the proposition, version 3: This comes from an easy inductive argument. □

At a higher level, that would be an excellent version of the proof: with only seven words, it puts a spear right through the heart of the matter. But, in this course, our imaginary readers are first-year undergraduate students. They need more detail.

Every argument by induction, as in the third proof, can be reformulated as an argument by contradiction, as in the second proof. There are times when a proof can be expressed using mathematical induction but, somehow, it is easier to use argument by contradiction. The decision is a matter of judgement and, to some extent, a matter of personal preference. The term *mathematics* arose as a contraction of the phrase *the mathematical arts*.

♣ ♡ ♢ ♠

We mention that the term *mathematical induction* has always been intended as a joke. To explain the joke, we must first mention that there is a form of reasoning, nowadays often called *retroductive reasoning*, traditionally called *inductive reasoning*, which is often employed in situations where purely deductive reasoning does not suffice. Retroductive reasonaing typically

involves a theoretical justification together with an enumeration of items of empirical evidence. For example, part of the early evidence for Newton's Inverse-Square Law was that, firstly, the Inverse-Square Law is suggested by theoretical considerations involving gravitational flux through the surface of a sphere with a mass at its centre, secondly, the Inverse-Square Law implies Kepler's Laws concerning the solar orbits of Mercury, Venus, Earth, Mars, Jupiter and Saturn. Retroductive reasoning is used throughout science, including mathematics. Indeed, mathematicians use it as a way of assessing speculations and conjectures which they cannot yet prove or refute. We saw an illustration of it at the beginning of Section 2. However, some mathematicians — and many philosophers of mathematics — have sometimes liked to pretend that mathematical reasoning is exclusively deductive. Sure enough, mathematical proof really is exclusively deductive. The name *mathematical induction* is a deliberate irony, as if to suggest that, in mathematics, even the induction is deductive! (Warning: I am worried that some students might find this information confusing. So, if you cannot understand this paragraph, please ignore it.)

3: Illustration: Planar graphs

We give another illustration of a proof by induction and a proof by contradiction.

Theorem: (Euler's formula for planar graphs) *Given a connected planar graph with n vertices, e edges and f faces, then $n - e + f = 2$.*

Proof 1: We argue by induction on e.

Base step: Any connected graph with no edges has exactly one vertex and exactly one face. Hence $n - e + f = 1 - 0 + 1 = 2$.

Induction step: Suppose now that $e \geq 1$, and that the assertion holds for all connected planar graphs with $e - 1$ edges. Let G be a connected planar graph with e edges. We shall construct a connected planar graph G' with $e' = e - 1$ edges. Letting n' and f' be, respectively, the number of vertices and the number of faces of G' then $n' - e' + f' = 2$.

First suppose that G is a tree. A proposition in Section 1 tells us that G has a vertex x with degree 1. Let G' be the graph obtained by removing x and its edge. Obviously G' is a connected planar graph. We have $n' = n-1$ and $f' = f$, so $n-e+f = (n'+1)-(e'+1)-f' = n'-e'+f' = 2$.

Now suppose that G is not a tree. Then G has a cycle. Let ϵ be an edge on a cycle, and let G' be the graph obtained by removing ϵ. Obviously G' is planar. Since ϵ is on a cycle, G' is connected. We have $n' = n$. The two faces of G with xy on their boundary are replaced by a single face of G', so $f' = f - 1$. So $n - e + f = n' - (e' + 1) + (f' + 1) = n - e + f = 2$. □

The induction step was not really necessary in the case where G is a tree. Let us give another proof, which deals with that case in a different way. This time, let us omit the baby-talk "base step" and "induction step".

Proof 2: We argue by induction on f. Let G be a connected planar graph with n vertices, e edges, f faces. If $f = 1$ then G cannot have any cycles, in other words, G must be a tree. By one of the definitions of a tree, $n = e + 1$. Hence $n - e + f = (e + 1) - e + 1 = 2$.

Suppose now that $f \geq 2$, and suppose that the assertion holds for all connected planar graphs with $f - 1$ faces. Then G must have a cycle. In other words, G cannot be a tree. The argument now proceeds as in the last paragraph of the previous proof. □

In passing, let us note a corollary which follows immediately from the theorem.

Corollary: *Let G be a planar graph. Then all the drawings of G on a plane have the same number of faces.*

Let us also mention that the theorem is of tremendous significance in pure mathematics. We have defined a planar graph to be one that can be drawn on the plane, but we could equally well have defined a planar graph to be one that can be drawn on a sphere. A planar graph on a sphere divides the sphere into faces that can be deformed into disks. In some sense, the theorem is really a theorem about the 2-dimensional sphere: we say that the Euler characteristic of the 2-sphere is 2. In a similar way, if we draw a graph on a 2-dimensional torus (the surface of a doughnut), and if all of the faces can be deformed into discs, then $n - e + f = 0$. We say that the Euler characteristic of the 2-torus is 0. The notion of the Euler characteristic was one of the seeds that led to algebraic topology. And the algebraic machinery behind that, called homological algebra, because an area of study in itself, with many applications outside of topology.

♣ ♡ ♢ ♠

We cannot use the theorem to prove that a given graph is planar. But we can sometimes use it to prove that a given graph is non-planar. The rough idea is that, if $n - e + f \neq 2$, then the graph cannot be planar. However, for a non-planar graph, the "number of faces" f is not defined. We shall need an intermediate result.

Proposition: *Let G be a planar graph with n vertices and e edges, with $e \geq 3$. Then $e \leq 3n-6$.*

Proof: Consider the pairs (ϵ, F) where ϵ is an edge on the boundary of a face F of G. Each edge is on the boundary of two faces, so the number of such pairs is $2e$. But, letting $\overline{c}$ be the average number of edges on the boundary of a face, and letting f be the number of faces, then the number of pairs (ϵ, F) is $\overline{c}f$. Combining the equality $2e = \overline{c}f$ with the equality $n - e + f = 2$, we arrive at $n - e + 2e/\overline{c} = 2$, in other words,

$$n - 2 = e(\overline{c} - 2)/\overline{c}\,.$$

The condition $e \geq 3$ implies that each face has at least three edges. So $3 \leq \overline{c}$, hence $n-2 \geq e/3$. □

Corollary: *The graph K_5 is non-planar.*

Proof: Suppose, for a contradiction, that K_5 is planar. The number of vertices is $n = 5$ and the number of edges is $e = 10$. By the latest proposition, $10 \leq 3.5 - 6 = 9$. This is a contradiction, as required. □

Let us try to use that method to show that $K_{3,3}$ is non-planar. Here, the number of vertices and the number of edges are, respectively, $n = 6$ and $e = 9$. The inequality $e \leq 3n - 6$ becomes $9 \leq 3.6 - 6 = 12$. Nothing has been proved here. Well, when trying to solve a problem, the standard theorems are sometimes not enough. It helps to know the proofs of the theorems, so that one can adapt the arguments.

Corollary: *The graph $K_{3,3}$ is non-planar.*

Proof: Suppose, for a contradiction, that $K_{3,3}$ is planar. Let n and e be the number of vertices and the number of edges, as before. Every face must have at least 4 edges. So, in the notation of the proof of the latest proposition, $\overline{c} \geq 4$. From an inequality in that proof, we obtain

$n - 2 \geq e/2$, in other words, $e \leq 2n - 4$. We have $e = 9$ and $n = 6$, hence $9 \leq 2.6 - 4 = 8$. This is a contradiction, as required. □

Let us mention that a deep theorem of Kurutowsky asserts that a graph is non-planar if an only if, in a certain sense — up to homotopy — the graph has K_5 and $K_{3,3}$ residing within it. A precise statement of the theorem can be found in the textbook.

Homework II

The diagrams for this homework, and some other comments and details, were given in class.

1: Show that any tree with at least one edge has at least two vertices with degree 1.

2: Show that, if we remove any edge from K_5, then the resulting graph is planar. Do likewise for $K_{3,3}$.

3: Show that the Peterson graph is non-planar. Letting G be a graph formed by the 16 vertices and 32 edges of a 4-dimensional cube, show that G is non-planar.

4: Consider a planar graph with n vertices, e edges, and let c be a positive integer such that every face has at least c edges. State and prove an inequality relating n and e and c, and show that your inequality is strong enough to cover all four non-planar graphs considered above: K_5 and $K_{3,3}$ and the two graphs is question 3.

Homework III

With one modification, these are Exercises 4.1.2 and 4.1.14 and 4.2.12 from the book.

1: Using mathematical induction, show that $\sum_{i=1}^{n} i2^i = 2 + (n-1)2^{n+1}$.

2: For which positive integers does the inequality $3^n < 4n!$ hold? Justify your answer using mathematical induction.

3: The **Fibonacci numbers** are the numbers F_n such that $F_0 = 0$ and $F_1 = 1$ and $F_{n+2} = F_{n+1} + F_n$. The next few Fibonacci numbers are

$$F_2 = 1\ , \quad F_3 = 2\ , \quad F_4 = 3\ , \quad F_5 = 5\ , \quad F_6 = 8\ , \quad F_7 = 13\ , \quad F_8 = 21\ .$$

Show that $F_0 + F_1 + ... + F_n = F_{n+2} - 1$.

4: Second-order homogeneous recurrence relations

I include a discussion of this topic because it gives a good illustration of how proof by induction can go wrong if used without proper care.

Later in the course, we shall be discussing first-order recurrence relations

$$ax_{n+1} + bx_n = f(n)$$

and second order recurrence relations

$$ax_{n+2} + bx_{n+1} + cx_n = f(n)$$

where f is some given function sending each natural number n to a complex number $f(n)$, and a, b, c are complex numbers with $a \neq 0$. We are find a formula for the complex numbers X_0,

x_1, x_2, ... Such recurrence relations usually have a **general solution** with some unknown parameters. To determine a *particular solution*, we need some initial conditions, say, given values for x_0 and maybe for x_1.

In the special case where $f(n) = 0$, the recurrence relation is said to be **homogenous**. This special case is important, because the method will be as follows:

Step 1: Find the general solution to the homogeneous recurrence relation $ax_{n+1} + bx_n = 0$ or $ax_{n+2} + bx_{n+1} + cx_n = 0$.

Step 2: Find any particular solution to the relation $ax_{n+1} + bx_n = f(n)$ or $ax_{n+2} + bx_{n+1} + cx_n = f(n)$.

Step 3: Combine the above to find the general solution to $ax_{n+1} + bx_n = f(n)$ or $ax_{n+2} + bx_{n+1} + cx_n = f(n)$.

Step 4: Solve for the unknown variables to find the particular solution to $ax_{n+1} + bx_n = f(n)$ or $ax_{n+2} + bx_{n+1} + cx_n = f(n)$ which satisfies the initial conditions.

In these notes, we shall be discussing only the first step. Further discussion of the other three steps will appear in lectures.

Let me mention that the above recurrence relations are the discrete analogues of the differential equations

$$p\,\frac{dy}{dx} + qy = f(x)\ , \qquad\qquad p\,\frac{d^2y}{dx^2} + q\,\frac{dy}{dx} + ry = f(x)\ .$$

In the theory and in the applications, there are many connections between the "discrete" recurrence relations and the corresponding "continuous" differential equations.

♣ ♡ ♢ ♠

The first-order recurrence relation $ax_{n+1} + bx_n = 0$ is very easy to solve. The general solution has the form

$$x_n = A(-b/a)^n$$

where A is an unknown parameter. Of course, if we are given the value of x_0, then solving for A is again very easy, indeed, $A = x_0$.

Before turning to an example of a second-order recurrence relation, let me mention some other versions of the Principle of Mathematical Induction. One very useful version is as follows:

Remark: (Strong version of the Principle of Mathematical Induction.) *For each natural number n, let P_n be a statement. Suppose that P_0 holds and that P_n holds whenever P_0, P_1, ..., P_{n-1} hold. Then P_n holds for all n.*

Thus, in the induction step, we may assume that P_0, ..., P_{n-1} hold, and we are to deduce that P_n holds. Another way of stating the principle is as follows.

Remark: (The Principle Mathematical Induction, Again.) *For each natural number n, let P_n be a statement. If P_0 holds and if each P_n implies P_{n+1}, then P_n holds for all n.*

Thus, in the induction step, we may assume that P_n holds, and we are to deduce that P_{n+1} holds. The version that we shall be using below is the next one.

Remark: (Yet Another Version of The Principle Mathematical Induction.) *For each natural number n, let P_n be a statement. Suppose that P_0 and P_1 hold and that P_{n+2} holds whenever P_{n+1} and P_n hold. Then P_n holds for all n.*

When using this version, the base step is to show that P_0 and P_1 hold. In the induction step, we assume that P_n and P_{n+1} both hold, and then we are to deduce that P_{n+2} holds.

Problem: *Suppose that $x_{n+2} - 5x_{n+1} + 6x_n = 0$. Show that there exist complex numbers A and B such that $x_n = A2^n + B3^n$.*

Answer: Using induction, we shall show that, putting $A = 3x_0 - x_1$ and $B = x_1 - 2x_0$, then $x_n = A2^n + B3^n$.

Base Step: We have $A2^0 + B3^0 = A + B = (3x_0 - x_1) + (x_1 - 2x_0) = x_0$ and $A2^1 + B3^1 = 2A + 3B = 2(3x_0 - x_1) + 3(x_1 - 2x_0) = x_1$. So the assertion holds in the cases $n = 0$ and $n = 1$.

Induction Step: Now fix n and suppose that $x_n = A2^n + B3^n$ and $x_{n+1} = A2^{n+1} + B3^{n+1}$. Then

$$\begin{aligned} &A2^{n+2} + B3^{n+2} - 5x_{n+1} + 6x_n \\ &= A2^{n+2} + B3^{n+2} - 5(A2^{n+1} + B3^{n+1}) + 6(A2^n + B3^n) \\ &= A2^n(2^2 - 5.2 + 6) + B3^n(3^2 - 5.3 + 6) = 0 = x_{n+2} - 5x_{n+1} + 6x_n \ . \end{aligned}$$

Canceling, we deduce that $x_{n+2} = A2^{n+2} + B3^{n+2}$. □

When writing out proofs, we are allowed to play at being a magicians, pulling rabbits out of hats. Of course, we do have to explain clearly to the reader how each step follows from the previous ones. But we do not have to tell the reader how we came up with our ideas.

Actually, I found the rabbits A and B by solving the simultaneous equations $x_0 = A2^0 + B3^0 = A + B$ and $x_1 = A2^1 + B3^1 = 2A + 3B$.

A more miraculous rabbit was the question itself. Is there some general result which might enable us to find such solutions second-order homogeneous recurrence relation? Usually, proof by induction is useless for actually *finding* solutions to problems. It tends to be useful only for *proving* that a proposed solution is correct. But still, we can *guess* the solution and then use induction to prove it.

The following statement is a reasonable guess at the general solution. It is wrong, and the argument we shall give has a subtle mistake.

False Theorem: *Let a, b, c be complex numbers with $a \neq 0$, and let x_0, x_1. ..., be complex numbers such that $ax_{n+2} + bx_{n+1} + cx_n = 0$. Let α and β be the roots to the quadratic equation $at^2 + bt + c = 0$. Then there exist complex numbers A and B such that $x_n = A\alpha^n + B\beta^n$ for all n.*

Mistaken Proof: Let A and B be the solutions to the equations $x_0 = A + B$ and $x_1 = A\alpha + B\beta$. Using induction, we shall show that $x_n = A\alpha^n + B\beta^n$.

Base Step: By the definitions of A and B, the assertion holds in the cases $n = 0$ and $n = 1$.

Induction Step: Now fix n and suppose that $x_n = A\alpha^n + B\beta^n$ and $x_{n+1} = A\alpha^{n+1} + B\beta^{n+1}$. Then

$$\begin{aligned} &a(A\alpha^{n+2} + B\beta^{n+2}) + bx_{n+1} + cx_n \\ &= A\alpha^{n+2} + B\beta^{n+2} + b(A\alpha^{n+1} + B\beta^{n+1}) + c(A\alpha^n + B\beta^n) \\ &= A\alpha^n(\alpha^2 + b\alpha + c) + B\beta^n(\beta^2 + b\beta + c) = 0 = ax_{n+2} + bx_{n+1} + cx_n \ . \end{aligned}$$

Since $a \neq 0$, we can cancel, deducing $x_{n+2} = A\alpha^{n+2} + B\beta^{n+2}$. □

Alas, the assertion must be false. A counter-example is the recurrence relation

$$x_{n+2} - 4x_{n+1} + 4x_n = 0$$

which has a particular solution $x_n = n2^n$. Indeed, putting $x_n = n2^n$, we have

$$x_{n+2} - 4x_{n+1} + 4x_n = (n+2)2^{n+2} - 4(n+1)2^{n+1} + 4n2^n$$

$$= 2^n(4(n+2) - 8(n+1) + 4n) = 0\,.$$

Exercise: If you are reading this before seeing it is class, find the mistake in the "proof".

♣ ♡ ♢ ♠

What should we do when we have a desirable "theorem" and a "proof", but then we run into a counter-example? (Maybe you do not care, and you just want to get on to the correct version of the theorem. But this is a mathematics course, and the question is fundamental to the way mathematics is discovered. Besides, locating the mistake in the above "proof" will help towards understanding the correct theorem.)

Well, we do not abandon the theorem. Instead, we find out where the proof went wrong, and then we correct the theorem. Let us look at how we might solve the equations

$$x_0 = A + B\,, \qquad x_1 = A\alpha + B\beta\,.$$

There are, of course, various ways of doing this. Let us use matrices. The two given equations can be written as

$$\begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \alpha & \beta \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}\,.$$

The solution is

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \alpha & \beta \end{pmatrix}^{-1} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \frac{1}{\beta - \alpha} \begin{pmatrix} \beta & -1 \\ -\alpha & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}\,.$$

In other words,

$$A = \frac{\beta x_0 - x_1}{\beta - \alpha}\,, \qquad B = \frac{\alpha x_0 - x_1}{\alpha - \beta}\,.$$

(If you do not like matrices, or if you do not know about them then, of course, you could obtain this solution in some other way.) But the solution does not make sense when $\alpha = \beta$. The proof failed because the equations $x_0 = A + B$ and $x_1 = A\alpha + B\beta$ do not always have a solution. So the base step does not always work. The counter-example came about because the quadratic equation $t^2 - 4t + 4 = 0$ has a repeated root. So we should correct the theorem so as to deal separately with the case of a repeated root.

Theorem: *Let a, b, c be complex numbers with $a \neq 0$, and let x_0, x_1. ..., be complex numbers such that $ax_{n+2} + bx_{n+1} + cx_n = 0$. Let α and β be the roots to the quadratic equation $at^2 + bt + c = 0$. Then:*

(1) *If $\alpha \neq \beta$ then exist complex numbers A and B such that $x_n = A\alpha^n + B\beta^n$ for all n.*

(2) *If $\alpha = \beta \neq 0$ then exist complex numbers A and B such that $x_n = (A + nB)\alpha^n$ for all n.*

(3) *If $\alpha = \beta = 0$ then $x_n = 0$ for all $n \geq 2$.*

Proof: First suppose that $\alpha \neq \beta$. Let

$$A = \frac{\beta x_0 - x_1}{\beta - \alpha} , \qquad B = \frac{\alpha x_0 - x_1}{\alpha - \beta} .$$

Using induction, we shall show that $x_n = A\alpha^n + B\beta^n$. Base step: We have

$$A\alpha^0 + B\beta^0 = A + B = \frac{(\beta x_0 - x_1) - (\alpha x_0 - x_1)}{\beta - \alpha} = x_0 ,$$

$$A\alpha^1 + B\beta^1 = A\alpha + B\beta = \frac{(\beta x_0 - x_1)\alpha - (\alpha x_0 - x_1)\beta}{\beta - \alpha} = x_1 .$$

So the assertion holds in the cases $n = 0$ and $n = 1$.

Induction Step: Now fix n and suppose that $x_n = A\alpha^n + B\beta^n$ and $x_{n+1} = A\alpha^{n+1} + B\beta^{n+1}$. Then

$$a(A\alpha^{n+2} + B\beta^{n+2}) + bx_{n+1} + cx_n$$
$$= A\alpha^{n+2} + B\beta^{n+2} + b(A\alpha^{n+1} + B\beta^{n+1}) + c(A\alpha^n + B\beta^n)$$
$$= A\alpha^n(\alpha^2 + b\alpha + c) + B\beta^n(\beta^2 + b\beta + c) = 0 = ax_{n+2} + bx_{n+1} + cx_n .$$

Since $a \neq 0$, we can cancel, deducing $x_{n+2} = A\alpha^{n+2} + B\beta^{n+2}$. We have established part (1).

Now suppose that $\alpha = \beta \neq 0$. Let $A = x_0$ and $B = x_1/\alpha - A$. Base step: we have $x_0 = A = (A + 0.B)\alpha^0$ and $x_1 = (A + B)\alpha = (A + 1.B)\alpha^1$. So the assertion holds for $n = 0$ and $n = 1$. Induction Step: Fix n and suppose that $x_n = (A + nB)\alpha^n$ and $x_{n+1} = (A + (n + 1)B)\alpha^{n+1}$. Then

$$a(A + (n + 2)B)\alpha^{n+2} + bx_{n+1} + cx_n$$
$$= a(A + (n + 2)B)\alpha^{n+2} + b(A + (n + 1)B)\alpha^{n+1} + c(A + nB)\alpha^n$$
$$= (A + nB)\alpha^n(a\alpha^2 + b\alpha + c) + B\alpha^{n+1}(2a\alpha + b) = 0 = ax_{n+2} + bx_{n+1} + cx_n$$

because $a\alpha^2 + b\alpha + c = 0$ and $\alpha = -b/2a$. Canceling, we obtain $x_{n+2} = (A + (n + 2)B)\alpha^n$. Part (2) is established.

Part (3) holds because, if $\alpha = \beta = 0$, then $b = c = 0$, and the recurrence relation reduces to $ax_{n+2} = 0$. □

A nice application, discussed on page 457 of the book, is the formula for the Fibonacci numbers,

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right) .$$

Appendix: the logical background

Grimaldi — the author of the course textbook — and I have a difference in our preference about how to introduce the notion of proof and, in particular, how to introduce the technique of mathematical induction. Of course, this is just a matter of personal inclination. As I shall explain below, he does have good reasons for his approach. If his presentation of the matter makes sense to you, then that is fine already. I include this appendix because there are certain features of the book which I, when I was a first-year undergraduate, would have found confusing.

On the whole, I think the textbook is excellent. It serves the purposes of the course very well. In many ways it is brilliant, and that is why I have selected it as the required book for

the course. But still, I can criticize certain parts of it; just as Grimaldi might well have many objections to my own efforts in the sections above.

The whole of Chapter 2 is devoted to logic, and he introduces "mathematical proof" at the same time. Section 2.5 has a discussion of esoteric principles such as "the rule of universal specification" and "the rule of universal generalization". At the end of that section, we have exercises such as 2.5.17: "Prove the following result in three ways (as in Theorem 2.4): if n is an odd integer, the $n + 11$ is even." The exercise is gibberish. It is *obvious* that if n is odd then $n + 11$ is even. As we noted above, a proof explains an assertion by breaking it down into a sequence of obvious pieces. If the assertion is already obvious, then we do not need a proof, and anyway, proof would be inapplicable. So it is unnecessary and also *impossible* to prove that if n is odd then $n + 11$ is even. And then, the whole of Chapter 3 is devoted to set theory, and only in Chapter 4 can we get onto the Principle of Mathematical Induction, with its prerequisite notions of a *statement* and a *set*.

In my view, it is unhelpful and actually perverse to discuss logic and set theory when introducing mathematical proof. To do so can only reinforce the impression that mathematical proof is just magic ritual. This point was expressed nicely by the great mathematician René Thom. Quoting him from memory, "If the fox knows that the hen is in the hen-house, and the hen-house is in the field, then the fox knows that the hen is in the field." His point is that, to find the hen, the fox does not begin by making a study of logic and set theory.

The situation becomes quite baffling in Section 4.1 of the book, where the Well-Ordering Principle is stated. I think Grimaldi is stating it as an obvious remark, although there is a preliminary discussion where he mentions that the principle does not hold for $\mathbb{Q}$ or $\mathbb{R}$, and I am not sure whether he is trying to justify the principle in some way; anyway, I cannot see any line of deductive reasoning here. He then presents the Principle of Mathematical Induction as a theorem, and he "proves" it using the Well-Ordering Principle. But, since the Induction Principle is just as obvious as the Well-Ordering Principle, why is the Induction Principle awarded a proof while the Well-Ordering Principle is merely stated? One could equally well argue it the other way around, just stating the Induction Principle and then "proving" the Well-Ordering Principle. Frankly, Section 4.1 of the book seems quite incoherent to me. A student, upon reading it, might very reasonably conclude that mathematical proof really is just a meaningless game, or some kind of magic ritual.

Actually, there is a rationale behind the presentation in the book but, to understand it, one needs to know some history. Cantor, working on the theory trigonometric series in the 1870s, found that he needed to invent two kinds of numbers: *cardinal numbers*, which express sizes of sets, and *ordinal numbers*, which express sizes of well-ordered lists. Well-ordering means that, for any non-empty set of items on the list, there is an earliest item in that set. His cardinal numbers dealt with all sets, finite or infinite. His ordinal numbers dealt with all well-ordered lists, finite or infinite. The natural numbers are the finite cardinal numbers, and they are also the finite ordinal numbers. But his interest was in the infinite case.

Cantor's theory of the infinite was ignored for a while, but gradually, as its usefulness became apparent, it came to be seen as interesting but controversial. In 1888, Dedkind defined a *finite set* to be a set X such that every injective function $X \to X$ is bijective. (The termonology here will be introduced later in the course). That allowed him to *define* the natural numbers to be the sizes of finite sets. From his definition, he then proved that the natural numbers satisfy the following conditions:

P1: For each natural number n, there exists a natural number Sn.

P2: There exists a natural number 0 such that $0 \neq Sn$ for for all natural numbers n,

P3: Given natural numbers n and m such that $Sn = Sm$, then $n = m$.

P4: Given statements P_n for all natural numbers n, and supposing that P_0 holds and each P_n implies P_{Sn}, then P_n holds for all n.

Addition of natural numbers can be defined by the recursive condition $n + Sm = S(n+m)$. Writing $1 = S0$, then $Sn = n + 1$. It is an easy exercise to prove that $2 + 2 = 4$, where 2, 3, 4 are defined to be the natural numbers $2 = S1$ and $3 = S2$ and $4 = S3$. Multiplication can be defined by the condition $n.Sm = n.m + n$. It can be shown how the above conditions imply all of the familiar properties of the natural numbers, such as $n + m = m + n$ and $nm = mn$. These ideas were pursued in 1889 by Peano, who took the conditions P1, P2, P3, P4 to be not just a characterization of the natural numbers but actually the *definition* of the natural numbers. Those four conditions are nowadays called the **Peano Axioms**.

Gradually, over the next few decades, axiomatic theories of sets were developed and, to support the study of sets, axiomatic theories of logic were developed too. The hope was that set theory and logic would provide a secure foundation for Cantor's theory of the infinite. There were some genuine difficulties to deal with. The naive notion of a set as "a collection of objects" is inadequate, because it leads to various paradoxes, the most famous being the following one:

Russell's Paradox: Let S be the set whose elements are the sets T such that T is not an element of T. If S is an element of S then, from the definition of S, we deduce that S is not an element of S. But, if S is not an element of S then, from the definition of S, we deduce that S is an element of S.

Of course, the familiar properties of the natural numbers were never in doubt. The foundations were needed so that mathematicians could work with the infinite, proving theorems such as: every vector space has a basis; this holds for all vector spaces, not just the finite-dimensional vector spaces which you may have studied in a linear algebra course. The motive for proving equalities such as $2 + 2 = 4$ and $a + b = b + a$ was just to check that the foundations make sense. Having shown how the foundations are adequate to recover the ordinary arithmetic of the natural numbers, one may then proceed, with some confidence, towards more interesting things that reside in the realm of the infinite.

It can be shown that there is a unique total ordering relation $\leq$ on the natural numbers such that $n \leq Sn$ for all natural numbers n. (Again, the terminology here will be introduced later in the course). Axiom P4 can be replaced by the following axiom.

P4*: Given a non-empty set X of natural numbers then, with respect to the relation $\leq$, the set X has a smallest element.

It can be shown that, assuming axioms P1, P2, P3, then the axioms P4 and P4* are equivalent to each other. That is to say, each of them implies the other. Of course, P4 is the Principle of Mathematical Induction, and P4* is the Well-Ordering Principle.

During the 20th century, the controversies settled down. Every professional pure mathematician has to know some axiomatic set theory. It is needed, for instance, to prove the above-mentioned theorem about every vector space having a basis. (Cantor invented the ordinal numbers as a way of generalizing mathematical induction to the infinite. That idea is now largely obsolete because, in the 1930s, it was replaced by a result called Zorn's Lemma. But

Zorn's Lemma is so surprising that no-one would be able to believe it without first making a study of axiomatic set theory.)

It turned out that logic — the study of logical principles — is irrelevant to most of mathematics. However, the work of the logicians did turn out to be useful in unexpected ways. As well as having some minor applications in pure mathematics, logic also has some deep applications to computer science, particularly in the theory of algorithms. (This is something far more substantial than the trivial applications of logic which, if time permits, we might discuss in class: the design of the half-adder and the full-adder. If time does not permit then, anyway, some of you will already have seen that material in courses on electronic design.)

In the light of the above history, Grimaldi's presentation makes sense. His use of the Well-Ordering Principle in his "proof" of the Principle of Mathematical Induction can be seen as a recognition of the fact that, assuming P1, P2, P3, then P4* implies P4. But, as we noted above, it works just as well the other way around: assuming P1, P2, P3, then P4 implies P4*.

One can understand what Grimaldi is doing. He is trying to give his readers just a little taste of some deep culture. So he has compromised, telling half the story. In Chapter 2, he has presented a little bit of logic, in Chapter 3 he has presented a little bit of set theory and, in Chapter 4, he has touched upon the abstract definition of the natural numbers. His reason for emphasizing logic is that it will become genuinely useful later, in the study of logic networks. And, if one starts with logic, then the next thing has to be set theory, to complete the account of foundations.

Other books on discrete mathematics adopt a similar narrative. I can sympathize with the intention. Nevertheless, I think a teacher has to outline the whole story, or else say nothing. The half-story in Grimaldi — and in other textbooks — just does not make sense. In an introductory mathematics course one must work with the natural numbers without defining them precisely, but it is ridiculous to select one of the above two principles as obvious and to provide a gibberish "proof" of the other principle.

In conclusion, the Well-Ordering Principle and the Principle of Mathematical Induction are both obvious. They are not in need of proof for their own sake, moreover, it is impossible to prove them unless one begins with some kind of foundation, say, the Peano Axioms or Dedekind's definition of a finite set. In that case, the motive would be to check that the foundation is fit for its intended purpose in more advanced mathematics. One might make such a study if, for instance, one wished to work competently in some area of mathematics where Zorn's Lemma is needed.

This is not the only case where introductory texts in mathematics give the impression of trying to make easy things seem difficult. In my own first-year undergraduate days, I paid little attention to real analysis, because I did not see the point of proving apparently obvious assertions such as the Intermediate Value Theorem and the Jordan Curve Theorem. (We shall be using the Jordan Curve Theorem in graph theory, and I shall present it as an obvious remark.) As I later learned, the motive for those proofs is not to check their practical conclusions, but rather to check that the underlying definition of continuity is fit for its intended purposes in more exotic areas of geometry, topology and functional analysis.

Actually, many natural scientists and engineers make a study of real analysis only up to the point where more-or-less obvious assertions are recovered using abstract definitions. They never get to see the genuine applications of those definitions. As a result, they end up with some unfortunate misconceptions as to how pure mathematicians see proof. This is why I have such a strong negative reaction to the similarly garbled touch on foundations in the early chapters of the course text-book.