MAKING SENSE OF PROOF BY CONTRADICTION

By **Colin Foster**

I will start with an old joke:

Student: What is proof by contradiction?

Teacher: I think you know.

Student: If I knew, I wouldn't be asking. And I am asking, so that means that I don't know.

Teacher: Ah, so you did know!

There are certain topics in mathematics where 'philosophy' (in the broadest sense) is likely to intrude. Introducing negative or complex numbers is one: is mathematics discovered or invented? Another one is proof by contradiction (or *contrapositive*, see Kinnear & Sangwin, 2018, for a discussion of the difference). G H Hardy (1967, p. 94) described proof by contradiction as "one of a mathematician's finest weapons", and yet it is frequently perceived to be difficult (see Quarfoot & Rabin, 2021). People often say that the idea of a counterfactual situation is hard, and that students struggle with the logic, but actually the logic behind proof by contradiction/contraposition is common in everyday life (as in the joke), and even small children employ it frequently:

Adult: Do you think Mama is at home?

Child: No.

Adult: How do you know?

Child: If she were, her car would be out the front, and it isn't – so she isn't.

The mystery around proof by contradiction is probably not helped by the fact that students' first exposure to the idea tends to be the classic proof that $\sqrt{2}$ is irrational, which is not the simplest example:

Theorem
$\sqrt{2}$ is irrational.

Proof
Suppose that $\sqrt{2} = \frac{p}{q}$, where p and q are co-prime, positive integers. This means that $2q^2 = p^2$, which means that p must be even, so we can write $p = 2m$, where m is an integer.

Substituting this in, we get

$$2q^2 = p^2 = (2m)^2 = 4m^2$$

So, $q^2 = 2m^2$, and this means that q must *also* be even.

But, we said at the start that p and q were co-prime, so they can't both be even. So, we have a contradiction. Therefore, $\sqrt{2}$ cannot be rational.

Despite being famous, and part of the 'canon' of mathematics, this is actually a rather subtle proof to use as first exposure to proof by contradiction, because we sneaked in at the start the bit about p and q being co-prime. This isn't part of the definition of a rational number, so students often wonder why it's included. Why were we so insistent on writing the fraction in its lowest terms? (Note 1) The truth is that we only did it because we were anticipating the ending, and we knew that we would need that to get the contradiction! Really, it makes more sense to see this as a proof by infinite descent on the positive integers, and, because there is a smallest positive integer, we arrive at a contradiction. So, for these reasons, I think that this is not the best choice for a first example of a proof by contradiction (Note2). As Kinnear & Sangwin (2018) pointed out, proving the irrationality of something like log 2 is actually much easier, even though log 2 seems like a more 'advanced' number than $\sqrt{2}$:

Theorem
$\log_{10} 2$ is irrational.

Proof
Suppose that $\log_{10} 2 = p/q$, where both p and q are positive integers.

(This time there is no need to make any assumptions about p and q being co-prime.)

This means that $2 = 10^{p/q}$. So $2^q = 10^p$.

We are now basically done, because we can see that this can't possibly be right. Both p and q are integers greater than zero, so we must have a factor of a power of 5 on the right-hand side, but not on the left. (We can write $2^q = 2^p 5^p$, if we prefer, which makes this even clearer.) The qth power of 2 is never going to equal a multiple of a power of 5. So, we have a contradiction (Note 3), and so log 2 can't be rational.

It is always good in situations like this to try to unpick a little why it fell out the way it did. Proofs should be convincing, but the best ones are also enlightening. What is it about the numbers 2 and 10 that makes this happen? The crucial thing is that one is not a rational power of the other. If, instead, we used 2 and 8, no contradiction would be reached, since $\log_8 2 = \frac{p}{q}$ gives $2^q = 8^p = 2^{3p}$, which is fine, giving $q = 3p$ and therefore $\frac{p}{q} = \frac{1}{3}$. Of course, we knew this at the start because $8^{\frac{1}{3}} = 2$, and so $\log_8 2 = \frac{1}{3}$. Once you see what is going on here, this may actually be enough to convince you, at least informally, that $\log_b a$ is going to be rational if and only if a is a rational power of b (equivalent to vice versa). Similarly, with the proof of the irrationality of $\sqrt{2}$, it is well worth working it through for $\sqrt{4}$, just to see why exactly the proof fails in the case of the square root of a square number, and for $\sqrt{6}$, to see why it doesn't require the square-rooted number to be prime.

However, there is really no reason to wait until students meet logarithms before introducing proof by contradiction. A much earlier opportunity would be something like Euclid's proof that there are infinitely many primes, which can be done in lower secondary (Note 4):

Theorem
There are infinitely many primes.

Proof
Suppose, for contradiction, that there is a *finite* number of primes. Write them all down as a list. Multiply them all together and add 1. Call this number *n*.

Is *n* prime?

(i) If it is, we have a contradiction, because we've found a number, bigger than any on our list, which is prime.

(ii) If it isn't prime, then it must be a product of smaller primes. But none of the primes on our list can be factors of *n*, because all of them leave a remainder of 1 when divided into *n*. So, *n* must have at least one prime factor that we *didn't* have on our list. So, again, we have a contradiction.

This means that there must be an infinite number of primes.

I think it is much easier to get the sense of a proof by contradiction with something like this than with the classic proof of the irrationality of $\sqrt{2}$. (See also Savic, 2017, for a discussion about choice of content for introducing proofs.)

Are there even easier proofs by contradiction? There certainly are; one of the simplest is proving that there is no largest integer. Even quite small children can understand this as well as anyone, although the idea that 'you can always add 1' is not often formulated as a proof by contradiction. We can sharpen that up by supposing, for contradiction, that there is a largest possible integer. Write it down. (Pedagogically, I find the 'Write it down' instruction quite useful for making concrete what is going on.) Then, suppose I decide to add 1 to your number, and thus obtain a *larger* integer. So, you must have been wrong to think that the number you wrote down was the largest integer, because I just made a larger one! You can think of this as an iterative process, where you keep thinking you've found the largest integer, only to be foiled when I impertinently add 1 to it. But, for proof by contradiction, you only need to go through this once to establish the result.

The idea of proof by contradiction takes a bit of getting used to, so it is worth introducing it early on, when everything else that's happening in the proof isn't too taxing. "Assuming something you know ain't true" can feel wrong to students – and it should. One way to address this concern is to begin with "Suppose" rather than "Assume", and to see the whole proof as a big "If". I'm not saying my premise is true – I'm asking what follows *if it is true*. If what follows is eventually clearly a nasty contradiction or absurdity, it means that my opening statement must have been false. But students often seem to feel that it's really a bit more complicated than this. Supposing something that isn't true (even if you don't officially 'know' yet that it's false) is really quite a problematic thing to do. Once I suppose something that is false, that moves me into a counterfactual mathematical world. How on earth should I know what the rules are for operating in that world? For example, suppose that $1 + 1 = 3$. What follows from that? Should I double both sides and get $2 + 2 = 6$? Or should I add 2 to both sides and get $2 + 2 = 5$? Which step is 'correct' in this counter factual world? What *does* $2 + 2$ equal? Of course, in a sense, that is the whole point. The fact that in this world $5 = 6$ should tell us that $1 + 1 = 3$ is false. (Or maybe it's the other way round, and $1 + 1 = 3$ is even more obviously false than $5 = 6$ is?) But the idea that there are *valid* steps to perform in the counterfactual world is a bit strange. The teacher might place a tick beside each line of working, but what does that signify? Once I assume something like $\sqrt{2} = \frac{p}{q}$, how can anyone really comment on the 'correctness' of any subsequent lines? It is not that I eventually, after several *correct* steps, 'arrive' at a contradiction, as people often say – it's contradictions all the way down! It's really a matter of judgment how brazen the contradiction needs to be before I quit and consider the proof completed – how far I decide I need to go before I anticipate that the reader will accept my use of a contradiction symbol. In more advanced work, arriving at '$\sqrt{2}$ is rational' would count as a contradiction, and we would stop there!

So, the students' objections make a lot of sense. Asking a question like 'If Pythagoras's Theorem *weren't* true, what would follow from that?' is an impossible question. Saying 'We assume that everything else works as normal – we just suppose that *one* false thing' doesn't seem to help. Assuming that $1 + 1 = 3$, it is impossible to say what 2 + 2 should equal, so we're unable to proceed to the next line. Making an assumption like that 'breaks maths', and so ought to stop us in our tracks. Principles like 'multiplying both sides of an equation by the same amount preserves the equality' don't seem to work any more – or, at least, we feel unsure whether we can rely on them. This is not a weird property of arithmetic statements – it's always the case when we try to operate based on a false premise. The first few mathematical steps look reasonable just because we're not astute enough to see that everything that we are writing is false. We deceive ourselves into thinking that each line 'follows' from the previous line in the normal way, and we are kind of doing normal mathematics, but we are (wittingly) writing nonsense; we pretend to be shocked when a contradiction eventually 'appears'. Even worse, students sometimes ask, is there possibly a danger that 'two wrongs could make a right', and we could slip back into truth from our initially false starting point? These can be difficult questions to resolve to the students' satisfaction.

I think the biggest difficulties students have with proof by contradiction are not always failures to understand the basic logic, but problems knowing what they are really doing when proceeding in 'counterfactual mathematical worlds', and a general sense that the whole thing feels dubious.

Notes

1. Students sometimes feel that all that they have proved is that $\sqrt{2}$ cannot be expressed as a *simplified* fraction, but that it perhaps could be expressed as an unsimplified one, in which either the numerator or the denominator or both contained some 'decimals'.

2. The step where we say that $q^2 = 2m^2$ means that q is even perhaps *itself* needs proving? And this suggests a rather different approach to the standard proof. One way to do this is to start talking about square numbers, rather than surds (see Foster, 2012). The teacher starts by asking, "Can you find a square number that is *twice* another square number?" Students will try doubling a few square numbers and notice that none of their doubles is square, or else they will try halving squares, but they don't get any squares that way either. They may offer $0^2 = 2 \times 0^2$, which is good thinking, but you can retort that 0 is just *one* square number rather than "another square number". Students may notice that doubling *twice* works: when you double 4 you get 8, which isn't square, but when you double 8 you get 16, which is – and this *always* seems to work. So, four times a square number *always* seems to give another square number, but with doubling it *never* seems to work, and after a while someone will offer the conjecture that it is impossible, and you can ask, "Why should it be impossible?" Of course, this is equivalent to $\sqrt{2}$ being irrational, because $\sqrt{2} = \frac{p}{q}$ is equivalent to $p^2 = 2q^2$, where p and q are positive integers.

The easiest way to see why $p^2 = 2q^2$ cannot happen is by using prime factorisation (Coles, 2005). If $p = a^x b^y c^z \ldots$ where $a, b, c, \ldots$ are distinct primes and $x, y, z, \ldots$ are positive integers, then $p^2 = a^{2x} b^{2y} c^{2z} \ldots$, and we can see that all the indices have to be *even*. The problem with $2q^2$ is that we have an extra factor of 2, so in the prime factorisation of $2q^2$ it is necessarily the case that 2 will appear to an *odd* power. This means that $2q^2$ cannot possibly equal p^2, because an odd power of 2 cannot equal an even power of 2. This is nice, because immediately we see what it is about the 2 that is problematic, and why having 4 instead would be fine. The 2 is a problem not because it's prime but because *it isn't a square number*. Without doing any more work, we can see that $p^2 = 3q^2$ will fail for exactly the same reason as $p^2 = 2q^2$, as will something like $p^2 = 6q^2$, even though 6 is not prime. So, in general, $p^2 = kq^2$ will have solutions for positive integer p and q if and only if k is square. And this is equivalent to saying that $\sqrt{k}$ is rational for integer k if and only if k is square.

3. In cases like this, 'contradiction' is not exactly the right word, as we don't directly *contradict* the initial statement; however, we arrive at something that is clearly false. Maybe we should call these 'proof by absurdity' or something?

4. It has been pointed out that Euclid's original proof of this is not actually a proof by contradiction. Indeed, whether proofs are technically 'by contradiction' is often a subtle matter. Hamkins (2020, pp. xvii–xviii) commented, "I do not find proofs by contradiction …to be a natural or robust mathematical category. Such a proof, after all, might contain essentially the same mathematical insights and ideas as a nearby poof that does not proceed by contradiction." For example, uniqueness proofs, such as the uniqueness of an inverse, are often done by contradiction. You start by supposing that there are two inverses, and you call them i and j. Perhaps you should say *"two distinct inverses"*, although people often don't. Then, you go through your algebra and end up with $i = j$. What does that tell you? Students will often say that it means that they are 'the same inverse', or that 'all inverses are equal'. Is this a proof by contradiction? It doesn't seem to be, unless you were careful to state that $i \neq j$ at the start. Then, the argument would be that you *supposed* that i and j were *distinct* inverses, and now you've found that they are *equal*, which contradicts the statement that they were different. If, on the other hand, you don't bother to say that they are *distinct* inverses, then what follows from finding that they were equal? Might there not still be another inverse out there somewhere, different from your i and j – call it k – which is different from those two?

Does the fact that i and j turned out to be the same necessarily prove that there can be no others? You perhaps need instead to start by saying something like, “Suppose that i and j are *any two* inverses’. Then, concluding that $i = j$ really *means* something, because we’ve shown that *any* two things that are inverses *must* be equal to each other, and therefore that *all* inverses are equal. It may seem pedantic to belabour this, but sometimes by being sloppy about things like this we simply confuse students more.

References

Coles, A. 2005 ‘Proof & in sight’. *Mathematics Teaching*, 190, 11.

Foster, C. 2012 *Ideas for Sixth-Form Mathematics: Pure mathematics & statistics.* Derby: Association of Teachers of Mathematics.

Hamkins, J. D. 2020 *Proof and the Art of Mathematics.* MIT Press.

Hardy, G. H. 1967 *A Mathematician's Apology.* Cambridge University Press.

Kinnear, G. & Sangwin, C. J. 2018 ‘Contradiction and contrapositive, what is the difference?’ *Scottish Mathematics Council Journal*, 48, 84–90.

Quarfoot, D., & Rabin, J. M. 2021 A hypothesis framework for students’ difficulties with proof by contradiction. *International Journal of Research in Undergraduate Mathematics Education,* Advance online.

Savic, M. 2017 ‘Does content matter in an introduction-to-proof course?’ *Journal of Humanistic Mathematics,* 7(2), 149-160.

Acknowledgement

This article first appeared in *Scottish Mathematical Council Journal* 51 (2021), pp. 74-77 and is reprinted here with the kind permission of the Council.

Keywords: Proof by contradiction; Contrapositive; Irrational; Prime.

Author: Colin Foster, Department of Mathematics Education, Schofield Building, Loughborough University, Loughborough LE11 3TU.

Email: **c@foster77.co.uk**

website: **www.foster77.co.uk**

blog: **blog.foster77.co.uk**