Arthur Lei Qiu

1 Introduction

Sometimes, a picture really *is* worth a thousand words. Complex mathematical ideas can often be greatly clarified with the use of visual aids. One common type of visual aid is a "picture proof": a drawing or figure which captures the logical reasoning underlying a proof of some mathematical truth, or which supplements a written argument proving some mathematical truth.

For instance, here is a famous picture proof of the Pythagorean theorem:

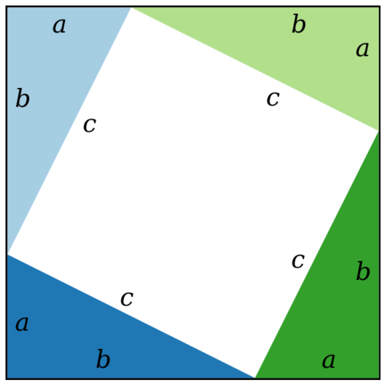


Figure adapted from Wikipedia.

The outer square (outlined in black) formed by the four coloured triangles has side length $a+b$, hence its area is $(a+b)^2$. This must equal the sum of the area of the white square and the areas of the four triangles; that is,

$$(a+b)^2 = c^2 + 4\frac{ab}{2},$$

and after a little algebraic manipulation, one obtains $a^2+b^2=c^2$. Of course, this is but one way of proving the Pythagorean theorem; many other picture proofs of the same fact have been discovered throughout history (Figure 1.1).

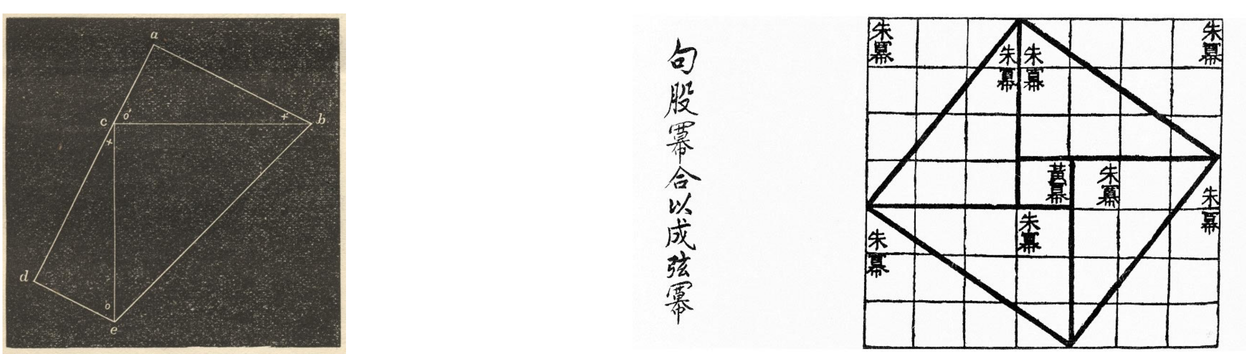


Figure 1.1: Left: A diagram from an article published in the *New England Journal of Education* (Vol. 3, No. 14, April 1, 1876) presenting a new proof of the Pythagorean theorem by James A. Garfield, better known as the 20th President of the United States. Right: An illustration of the Pythagorean theorem for the special case of a 3 : 4 : 5 triangle from the *Zhoubi Suanjing*, an ancient Chinese text dated to the 11th century BC.

As another example, recall that the n^{th} *triangular number* T_n is defined as the sum of the first n positive integers:

$$T_n = 1 + 2 + \cdots + (n-1) + n.$$

It is a well-known fact that $T_n = \frac{n(n+1)}{2}$. Here is a picture proof of this identity when $n = 4$:

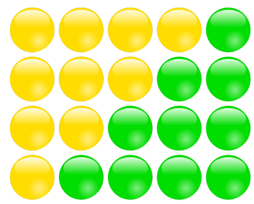


Figure adapted from Wikipedia.

There are a total of $2T_4$ circles (T_4 many yellow circles and T_4 many green circles). When arranged as in the picture, the circles form a 4×5 rectangle, which shows that $2T_4 = 4 \times 5$, hence $T_4 = \frac{4 \times 5}{2}$.

Lastly, here is a picture proof of the geometric series expansion $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 2$:

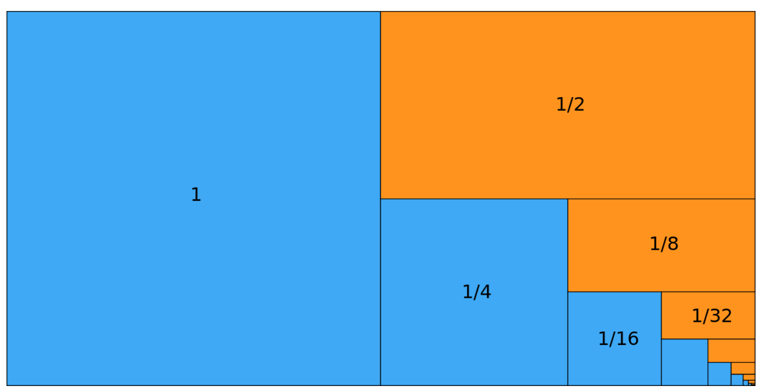


Figure adapted from Wikipedia.

Being able to draw convincing picture proofs is an incredibly useful skill for the modern mathematician. Pictures help build intuition, and can help communicate your ideas to others. At the same time, the ability to draw convincing picture proofs is a skill, and like any other skill, it requires practice. In this document, we will outline some general rules to follow when preparing picture proofs meant to supplement (or even replace) written proofs. These rules are illustrated[1] with guided examples. We will also discuss some software packages you may find useful to help make your figures.

Development of this guide was supported by the Writing-Integrated Teaching program, an initiative of the University of Toronto's Faculty of Arts and Science dedicated to helping instructors and academic units embed discipline-specific writing instruction into undergraduate courses and programs. The author is also grateful to Adam Morgan for generously allowing the adaptation of content from his *Guidelines and Tips for Making Figures* handout in the making of this document, and to Asif Zaman for permission to use content from MAT237 as examples.

[1] Pun intended.

2 Picture Proofs

2.1 Example 1: Limit points

Let's try to draw a picture proof of the following fact:

PROPOSITION 2.1. *The point $p = (2, 2)$ is a boundary point of the set $S = \{(x, y) \in \mathbb{R}^2 \mid y > x\}$.*

Any time you are asked to draw a picture proof, you should ask yourself two questions:

1. How would I **write** a formal proof of this fact?
2. How can I **translate** the key ideas of my written proof into a visual?

Let's start by addressing the first of these questions for Proposition 2.1. There are two possible approaches, depending on what one takes as the definition of "boundary point". Recall that the following statements are equivalent for a point $p \in \mathbb{R}^2$ and a set $S \subseteq \mathbb{R}^2$:

p is a boundary point of S.

$$\Updownarrow$$

For every $\varepsilon > 0$, the open ball $B_\varepsilon(p)$ of radius ε centred at p has non-empty intersection with both S and $\mathbb{R}^2 \setminus S$.

$$\Updownarrow$$

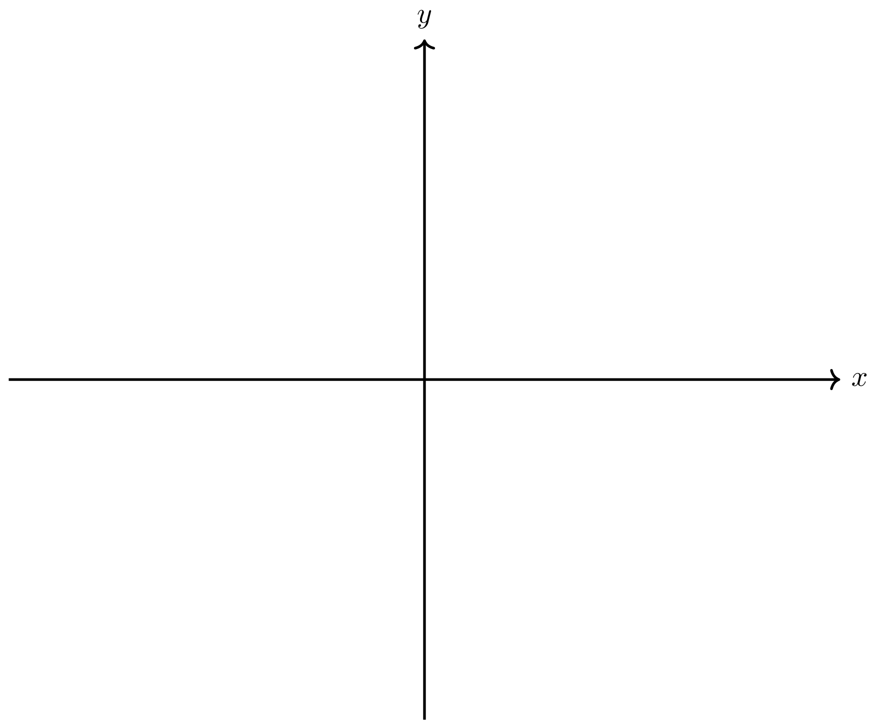
There exist sequences $(x_k)_{k\in\mathbb{N}}$ in S and $(y_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^2 \setminus S$ such that both x_k and y_k converge to p as $k \to \infty$.

To make sure we're all on the same page, let us briefly outline how each proof would go.

Proof sketch (ball definition). Fix $\varepsilon > 0$. Show that $B_\varepsilon(p) \cap S$ is non-empty by explicitly giving an element of it. Do the same for $B_\varepsilon(p) \cap (\mathbb{R}^2 \setminus S)$. Conclude using the arbitrariness of ε. ■

Proof sketch (sequence definition). Explicitly define two sequences $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^2$. Verify that $x_k \in S$ and $y_k \in \mathbb{R}^2 \setminus S$ for all $k \in \mathbb{N}$. Show that $\lim_{k\to\infty} x_k = p$ and $\lim_{k\to\infty} y_k = p$. ■

Just like how the steps of a written proof will depend on which one of these definitions we choose, so too will a picture proof. However, there will be some commonalities to start with. For instance, any picture proof will start with roughly the same "background canvas": the space on which we will draw our picture proof. Proposition 2.1 is a statement about the relationship between a point and subset of $\mathbb{R}^2$—but more than that, it is a statement about a point with *specific* coordinates and a subset defined by a *specific* inequality of coordinates. A written proof will have to refer to these coordinates at some point, so our picture proof should start by drawing the coordinate axes of a two-dimensional plane.



As simple as this may seem, there are already several points worth mentioning.

- **Conventions and consistency.** Humans can only accurately make drawings of one-, two-, or three-dimensional objects.[citation needed] In these dimensions, the components of $\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{R}^3$ are often given distinguished labels.
  - An arbitrary element of $\mathbb{R}$ is often written as x, so the letter x usually refers to its only component. Sometimes the letter t is used instead when one wants to think of $\mathbb{R}$ as representing "time" (such as in a parametrization of a curve describing an object's motion).
  - An arbitrary element of $\mathbb{R}^2$ is often written as (x, y), so the letters x and y usually refer to the first and second components, respectively.
  - An arbitrary element of $\mathbb{R}^3$ is often written as (x, y, z), so the letters x, y, and z usually refer to the first, second, and third components, respectively.

  We labelled the axes x and y for two reasons. The first is that it matches with the convention above. The second (and arguably more important!) reason is that *the problem statement used these letters.*

  If a problem is given to you with stated notation already, then you should stick with that notation when writing a proof or drawing a figure. For example, if Proposition 2.1 were stated such that we were asked to show that $(2, 2)$ is a boundary point of $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 > x_1\}$, then we should label our coordinate axes x_1 and x_2 instead of x and y. The reader will expect that the notation used in a fact statement is the same as the notation used in a proof of that statement, so choose your notation in a way that is consistent with previously introduced notation.

Rule

If a problem is given to you with stated notation already, then you should stick with that notation when labelling elements of your figure.
If a problem is given to you without stated notation, then you may come up with your own notation, which should be sensible and align with conventions.

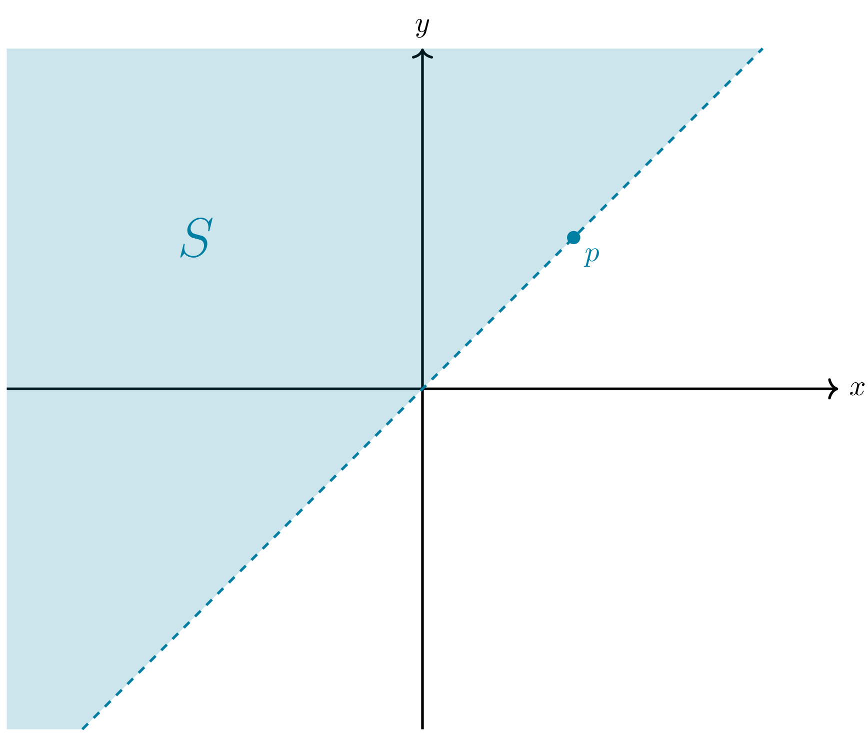
If we were illustrating a three-dimensional statement, we could label the axes (x, y, z), or (x_1, x_2, x_3), or another set of three labels. The choice will depend on whether the problem statement already introduced notation for the coordinates of $\mathbb{R}^3$, among other factors.

- As our drawing gets more complicated, we will likely have to introduce more labels to distinguish objects. Importantly, we should make sure there are **no notational clashes**. For instance, since we used (x, y) as coordinates, we should think of those letters as reserved only for that purpose; we should *not* use x and y to refer to specific numbers or points elsewhere in the picture proof, for instance. Similarly, if we use (x_1, x_2, x_3) as coordinates in a three-dimensional drawing, then we should definitely *not* label a sequence of points in $\mathbb{R}^3$ as $x_1, x_2, x_3, x_4, \ldots$, as this would get confusing. Instead, we should search for a different label.

Rule

Once an element of your drawing has been labelled with a given letter or symbol, that letter or symbol should not be used to mean something different elsewhere.

Let us continue by drawing the set S and the point p.



Note that we are abiding by a convention specifically for subsets of $\mathbb{R}^2$:

Convention

The boundary of an open subset of $\mathbb{R}^2$ is typically drawn as a dashed $(---)$ or dotted $(\cdots)$ curve. The boundary of a closed subset of $\mathbb{R}^2$ is typically drawn as a solid curve (—).

Bear in mind that there are sets which are both open and closed, as well as sets which are neither open nor closed. For such sets, you must decide how to draw their boundary in a way that will not confuse the reader.

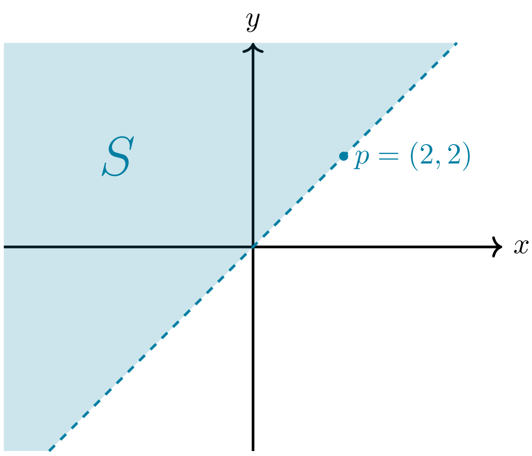
At this point, it is worth introducing perhaps the most important rule of picture proofs:

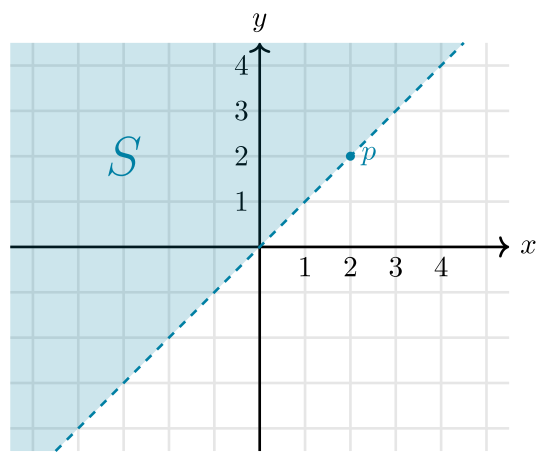
Rule

A good picture proof must accurately reflect whether a quantity is "arbitrary" or "special".

We will explore this more later, but for now, let us note how this rule is already reflected in our drawing so far:

1. The set $S = \{(x, y) \in \mathbb{R}^2 \mid y > x\}$ is given to us; it is *not* arbitrary. As such, our picture proof should match the description of this set. We would not be off to a great start if we accidentally shaded in the set $\{(x, y) \in \mathbb{R}^2 \mid y < x\}$, or drew some other random set.
2. Similarly, the point $p = (2, 2)$ is given to us; it is *not* arbitrary. Since we have not yet introduced a sense of scale to the x- and y-axes (i.e., how far one unit distance is), we have some freedom in deciding where to draw p; however, we should at the *very least* ensure that the indicated point appears in the upper right quadrant of $\mathbb{R}^2$ in our picture proof, and on the line $y = x$. Once we have drawn p, the drawing's scale is fixed, and we should ensure that all other features of the drawing are consistent with this scale. (For instance, if we had to draw the point $2p$, then it should be drawn twice as far from the origin as p.) To make the scale clear, we could label the coordinates of p, or introduce a coordinate grid. These options are shown separately below.





Both options have pros and cons to them. For instance, introducing a coordinate grid makes the sense of scale clearer to the reader; however, it may be difficult to draw such a grid using pen and paper without the figure becoming cluttered. Below, we will take a compromise between these two approaches by giving the coordinates of p and introducing a grid, while omitting the tick marks on the x- and y-axes.

At this point, the two ways in which we can prove Proposition 2.1 diverge in their approaches. Let us consider them separately.

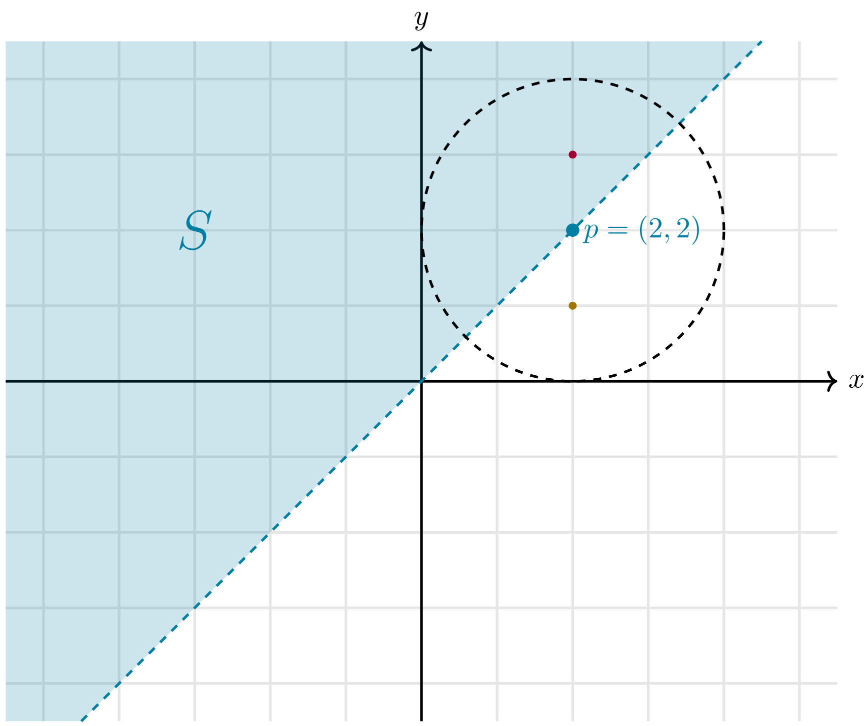
2.1.1 The ball approach

Recall that this method of proof went as follows:

Proof sketch (ball definition). Fix $\varepsilon > 0$. Show that $B_\varepsilon(p) \cap S$ is non-empty by explicitly giving an element of it. Do the same for $B_\varepsilon(p) \cap (\mathbb{R}^2 \setminus S)$. Conclude using the arbitrariness of ε. ■

The written proof involves a statement about open balls of arbitrarily small radius $\varepsilon > 0$, and yet we can only draw finitely many. One workaround for this is to draw just *one* such ball, but ensure that it is sufficiently "generic" to capture the idea for arbitrary ε. (You could also draw a sequence of open balls which get smaller and smaller, but this comes at the risk of cluttering up your drawing.)

Here is a first attempt at illustrating this proof by depicting $B_\varepsilon(p)$ and two points lying in $B_\varepsilon(p) \cap S$ and $B_\varepsilon(p) \cap (\mathbb{R}^2 \setminus S)$:

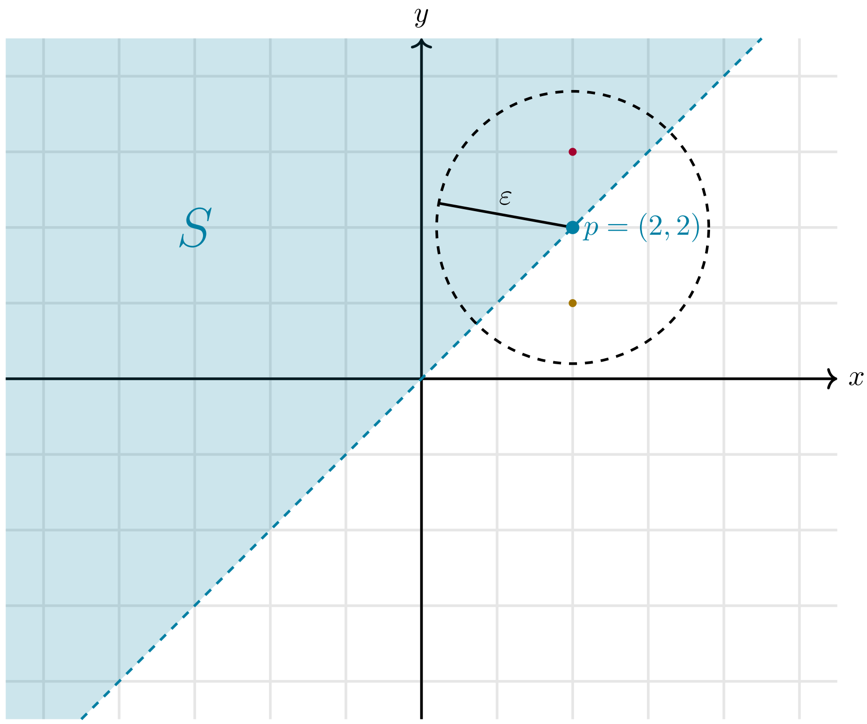


However, here we must ask whether we satisfied the rule about whether our picture proof accurately reflects which quantities are "arbitrary" and which ones are "special". For one, we have drawn the ball so that it lies exactly tangent to the x- and y-axes. This might lead a viewer to wonder: "*Is there something important about the fact the ball depicted is tangent to the coordinate axes?*

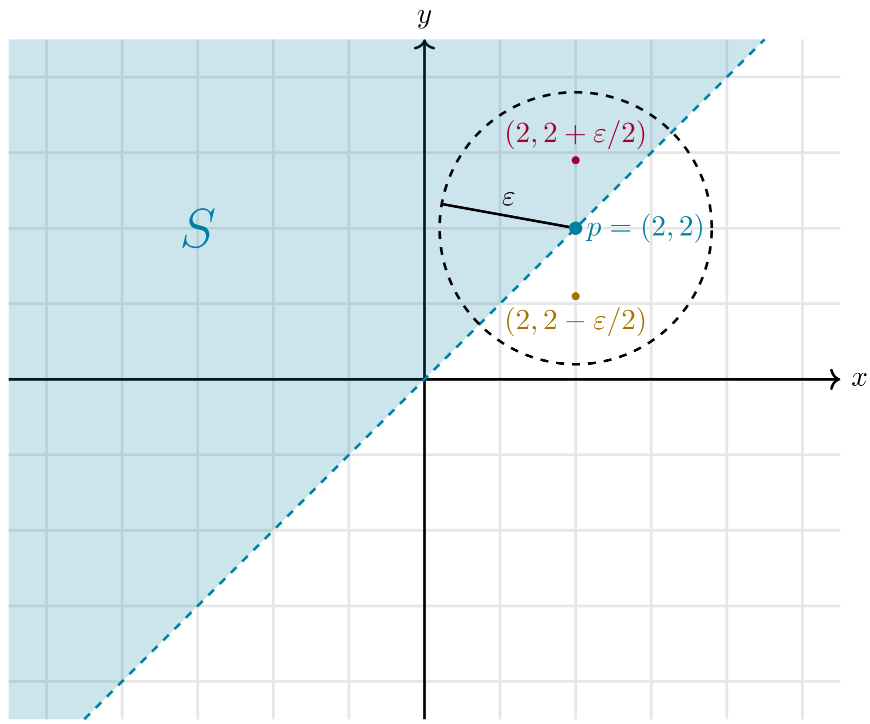
Of course, the answer is no, but this is not accurately captured in the picture. To fix this, let us make two changes:

- Shrink the ball slightly (after all, the main idea of the proof is that the argument must hold for balls of arbitrarily small radius), and
- Label a radius of the ball by ε *without* giving ε a specific value, to emphasize its genericity.

Making these changes produces the picture below.



We are still not done yet, though. A written proof will require us to specify the point in $B_\varepsilon(p) \cap S$ and the point in $B_\varepsilon(p) \cap (\mathbb{R}^2 \setminus S)$ explicitly—e.g., by giving their coordinates (which will depend on ε). On the other hand, while our picture depicts these two points, it does not explain how to come up with their coordinates! Since we have drawn these points as lying on the same vertical line as p, plausible formulas for the points above and below p might look like $(2, 2 + \varepsilon/2)$ and $(2, 2 - \varepsilon/2)$, respectively. To finish our picture, we will label our points with these formulae. (To make the formulae visually consistent with the size of the ball, we will also have to shift the points slightly.)



This is more or less a complete picture proof: we have drawn and labelled all the important objects that come into play in a way so that a viewer could feasibly reproduce the written proof based off the picture. You might reasonably ask questions like:

- The sets $B_\varepsilon(p) \cap S$ and $B_\varepsilon(p) \cap (\mathbb{R}^2 \setminus S)$ are both important to the proof; why didn't we label those?
- Should my picture proof somehow explain why the two points lie inside of $B_\varepsilon(p) \cap S$ and $B_\varepsilon(p) \cap (\mathbb{R}^2 \setminus S)$ respectively as claimed?

The truth is that there is no one correct answer for these types of questions. You will have to use your judgement to decide what the most important aspects of the written proof are that you wish to highlight in the picture, and what details to suppress. Here is one guiding principle to help you decide:

Rule

Picture proofs should be labelled with the same key quantities that would appear in a written proof.

Note that this is only a rule of thumb: sometimes you may want to illustrate an idea which would require lots of notation to communicate in writing. If you try to put too much notation in a single picture, it runs the risk of becoming cluttered and difficult to separate the key steps of the proof from the relatively minor details. In our final picture, we have abided by this rule in a minimal way: the most important quantities are the point p, the set S, the radius ε, and the points $(2, 2 \pm \varepsilon/2)$, and we have decided that the sets $B_\varepsilon(p) \cap S$ and $B_\varepsilon(p) \cap (\mathbb{R}^2 \setminus S)$ are not important enough to warrant their own labels.

2.1.2 The sequence approach

Let us recall how this proof outline goes.

Proof sketch (sequence definition). Explicitly define two sequences $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^2$. Verify that $x_k \in S$ and $y_k \in \mathbb{R}^2 \setminus S$ for all $k \in \mathbb{N}$. Show that $\lim_{k\to\infty} x_k = p$ and $\lim_{k\to\infty} y_k = p$. ■

Roughly speaking, you can think of "arbitrary" as corresponding to the quantifier "for all" ($\forall$), and "special" as corresponding to the quantifier "there exists" ($\exists$). Using the sequence approach, we have to show that there *exist* two sequences with certain properties. As such, when drawing these two sequences, they should be considered "special".

To illustrate this, let us start off with a *bad* picture proof using the sequence definition.

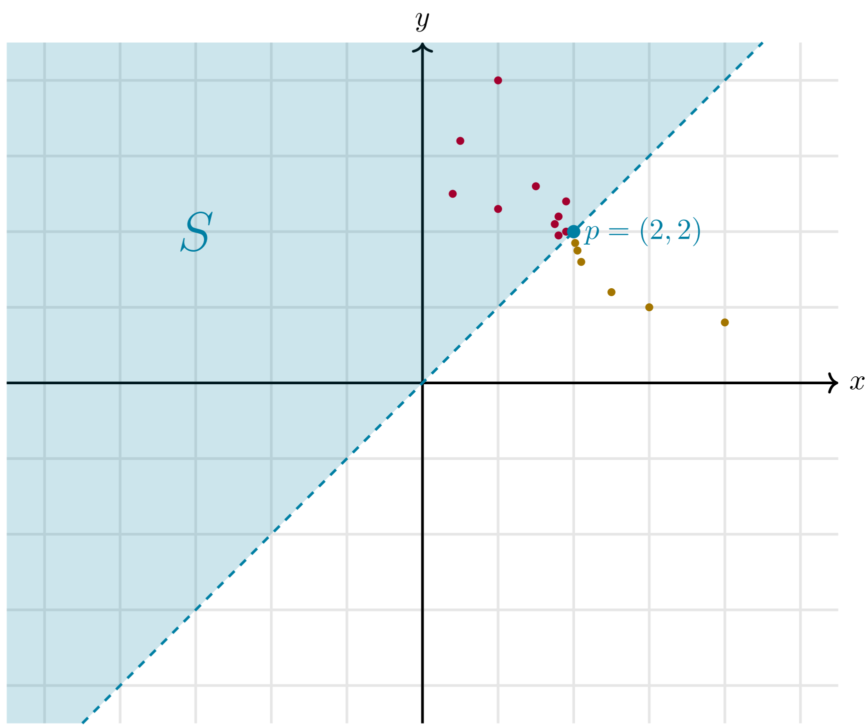
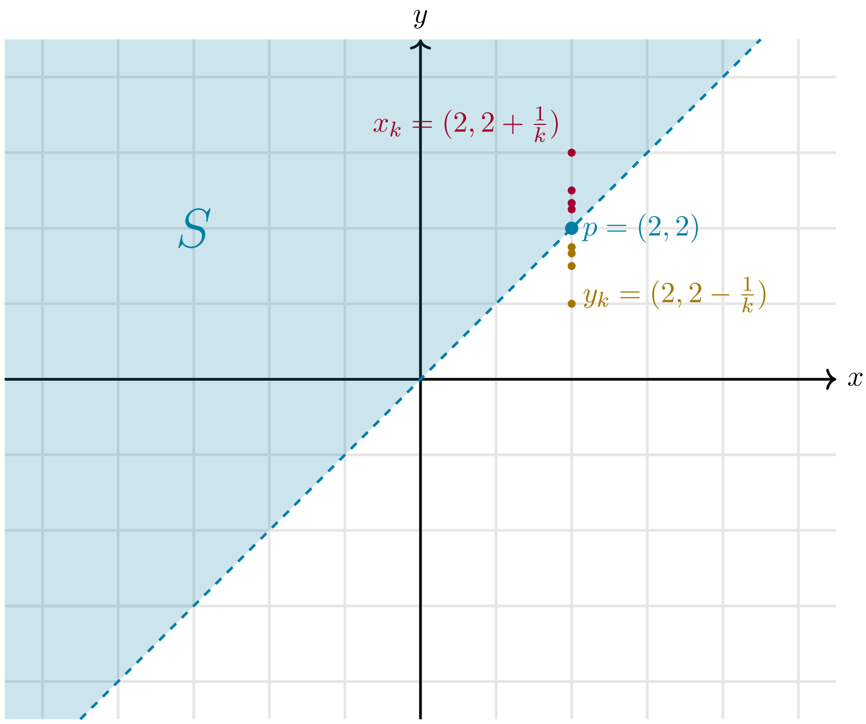


Figure 2.1: A bad picture proof of Proposition 2.1 using sequences.

On the surface, this might seem okay: we have depicted the sequence (x_k) lying in S and the sequence (y_k) lying inside $\mathbb{R}^2 \setminus S$, and both look like they converge to p. But there is a serious problem: the sequences we drew are too *generic*. In the written proof, explicitly defining (x_k) and (y_k) would typically involve giving a formula for their coordinates for each k—but, for instance, I certainly don't know a formula which describes the red sequence.

In order to fix this, we have to decide once and for all what our sequences x_k and y_k should be. There are many valid choices; one such choice is depicted in the picture below. (See if you can fill in the written proof using just this picture—note that you might not be able to do so had we not labelled x_k and y_k and given their formulae!)



Warning: Be careful about using colour in mathematical figures. While it can be very useful in the right circumstances, bear in mind that some of your audience members may not be able to perceive or distinguish certain colours; for instance, studies suggest that red–green colourblindness affects up to 1 in 12 men and 1 in 200 women. This website demonstrates how colourblindness can affect visuals, and has suggestions on colourblind-friendly palettes to use.

Also ask yourself: will your picture be printed out? If so, your reader may lose the ability to distinguish colours should your picture be rendered in black/white or grayscale.

2.2 Example 2: Path-connectedness

First, let us recall an important definition from topology.

Definition 2.2. *A subset $S \subseteq \mathbb{R}^n$ is **path-connected** if, for any $x, y \in S$, there exists an interval $[a, b] \subset \mathbb{R}$ and a continuous curve $\gamma\colon [a, b] \to \mathbb{R}^n$ such that $\gamma(a) = x$, $\gamma(b) = y$, and $\gamma(t) \in S$ for all $t \in [a, b]$.*

Less formally, a set S is path-connected if any two points in it can be joined by a continuous curve which lies entirely within S. This informal description lends itself well to picture proofs, and is what we will try to illustrate for a picture proof of the following statement:

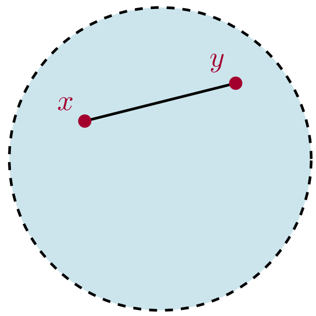
Proposition 2.3. *An open ball in $\mathbb{R}^2$ is path-connected.*[2]

Note the generality of Proposition 2.3: it does not depend on the radius of the ball or the point that the ball is centred at. As such, neither a written proof nor a picture proof should directly depend on these quantities.

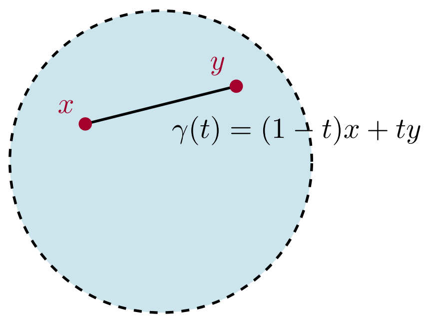
The proof idea of Proposition 2.3 is simple enough: given two points x, y inside an open ball B, connect them with the line segment joining them. (Actually proving that every point on this

[2]This statement is true in all dimensions, of course, but the picture in 1 dimension is not very interesting, and while the picture in 3D is doable, the proof's main idea is already captured by the 2D case.

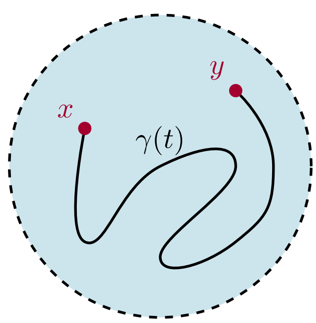
line segment lies inside B requires the triangle inequality, but we will not be focusing on this in the picture). Below is a corresponding illustration.



Given how simple this statement and its proof is, we might call it a day here. But if we wanted to be thorough, we could add a few labels. For one, we could give the curve a name (γ), and we could give a formula for the function parametrizing the line segment. One possible choice among many is $\gamma(t) = (1-t)x + ty$ for $t \in [0,1]$; this yields the following picture.

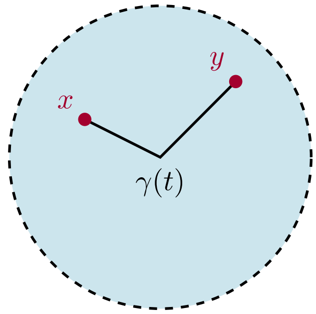


For comparison, here is a *bad* picture proof:



This picture is a bad picture proof for the same reason that Figure 2.1 is a bad picture proof: namely, in the written argument, one would have to explicitly construct the curve γ, e.g., by specifying what $\gamma(t)$ is for each time t in the parametrization interval. But from this picture, it is virtually impossible to tell what the formula for $\gamma(t)$ might look like.

Connecting the points x and y via the line segment joining them is not the only valid way of proving Proposition 2.3, however. Another perfectly valid method is to first connect x to the centre of the ball with a line segment, then connect the centre of the ball to y with another line segment. A basic illustration of this alternative approach is below.



This picture does lack two things, though:

- It might be hard for a viewer to recognize that the "turning point" of the curve is exactly the centre of the ball. As such, it might be worth labelling the centre somehow.
- Since this curve is a bit harder to describe than the straight line segment connecting x and y, it is definitely worth including a formula for $\gamma(t)$ in the picture.

EXERCISE 2.4. Finish the picture proof above by labelling the centre and finding a formula for a curve which describes $\gamma(t)$. You may need to use piecewise notation.

EXERCISE 2.5. Draw a picture proof of the following fact: *the set* $\{(x, y) \in \mathbb{R}^2 \mid x \neq 0\}$ *is not path-connected.*

EXERCISE 2.6. Draw a picture proof of the following fact: *the union of two path-connected sets with non-empty intersection is path-connected.*

Let us summarize some of the important points to remember when drawing picture proofs.

Guidelines for Drawing Picture Proofs

- First, come up with a written proof; translate it into a picture proof *afterwards*. This helps you decide what the most important aspects of the proof are that should appear in your illustration.
- Picture proofs should communicate the same key ideas as a written proof.
- Picture proofs should be labelled with the same quantities that would appear in a written proof.
- Picture proofs should accurately reflect whether a quantity is "arbitrary" or "special".
- Objects should be drawn to scale whenever a sense of scale is important for the proof.
- Colour can be helpful in small doses, but overusing colours can make pictures visually confusing and alienate some viewers.

3 Technology and Tools for Making Figures

Below are some options you might want to consider when deciding how you want to draw your pictures. The computer applications we mention are quite popular and well-documented, so with some web searching you can quickly figure out how to do pretty much anything you want with any of them (and for this reason, this document will not give you a comprehensive tutorial on how to use these).

- **Good old fashioned pen-and-paper.** But please make sure your drawing and any handwriting on it is legible, including after possibly being scanned into a computer!

- **Electronic tablets.** On most tablets, you can download notetaking applications which include features that make drawing mathematical figures a bit easier; for instance, some applications will detect when you are trying to draw a circle with a stylus and replace your handwriting with a perfect circle, and similarly with straight lines or polygons.
- **Desmos (`https://www.desmos.com/calculator`):** Good for generating pictures in $\mathbb{R}^2$ including graphs, curves, and subsets. As of 2025, Desmos has also released a beta version of Desmos 3D (https://www.desmos.com/3d).
- **Math3D (`https://www.math3d.org/`):** Good for generating pictures in $\mathbb{R}^3$ including curves, surfaces, and vector fields.
- **Geogebra:** has a graphing calculator similar to Desmos (https://www.geogebra.org/graphing) and a 3D plotter similar to Math3D (https://www.geogebra.org/3d). However, Geogebra tends to be quite laggy as the complexity of the objects drawn increases.
- **TikZ:** If you use LaTeX, you might consider learning how to use the TikZ package, which is what was used to generate the figures in this document. The syntax can be rather hard to learn, but online tutorials such as Overleaf's are available, and if you want to know how to draw something in TikZ, chances are someone else has asked how to do so on Stack Exchange.