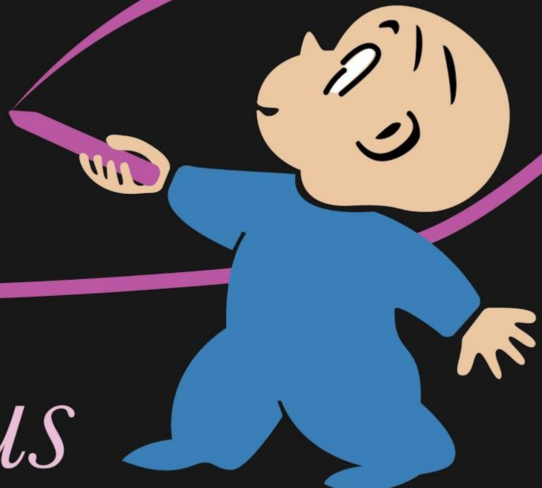
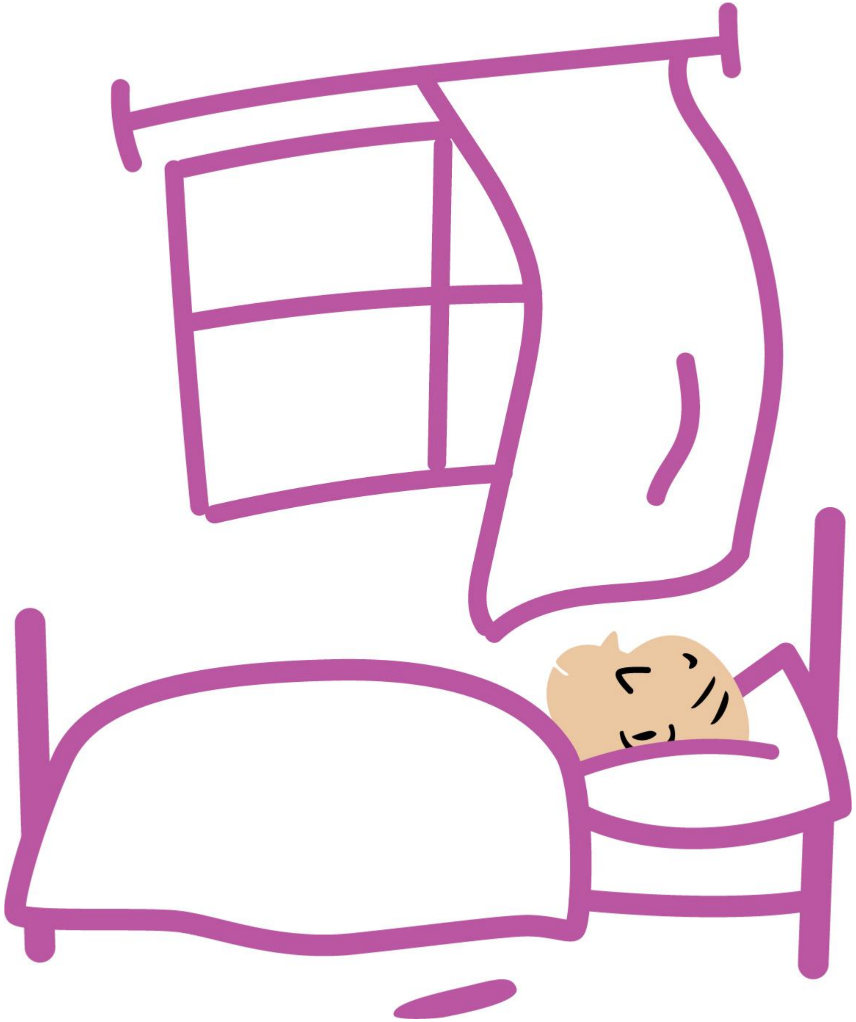
HAROLD *and the* PURPLE CRAYON...

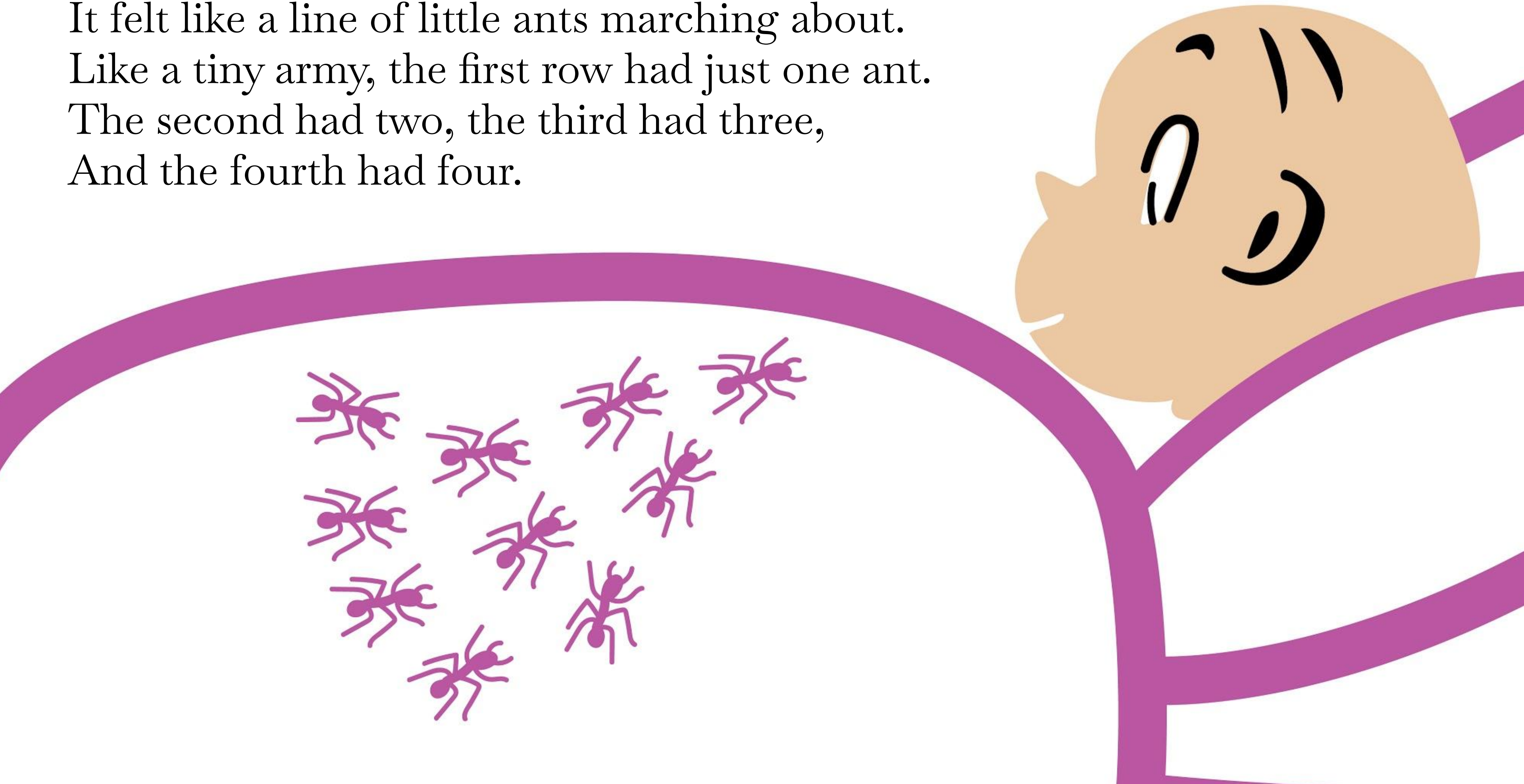
learn multivariable calculus

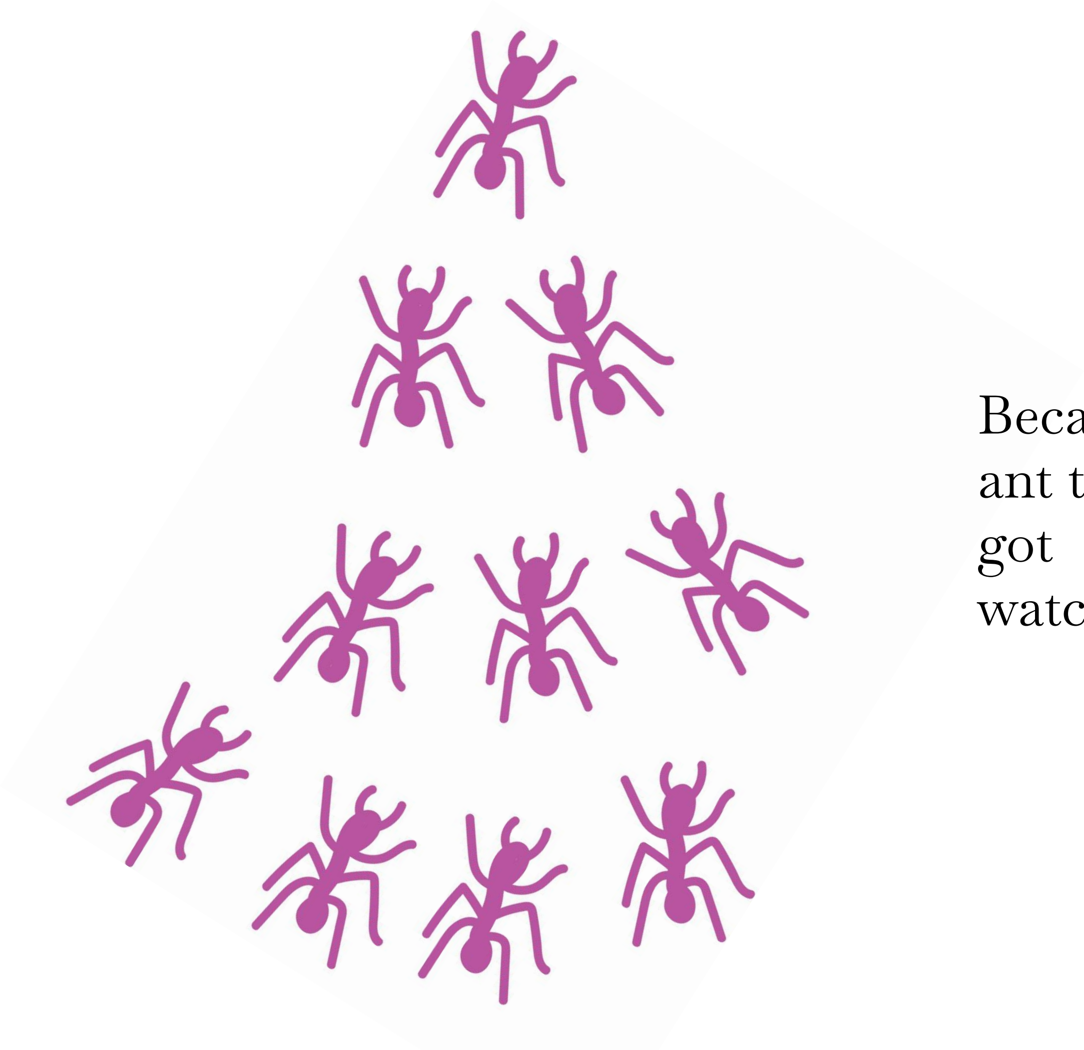




The window-curtain brushed
Harold's cheek as it danced
in the cool night breeze.

It felt like a line of little ants marching about.
Like a tiny army, the first row had just one ant.
The second had two, the third had three,
And the fourth had four.





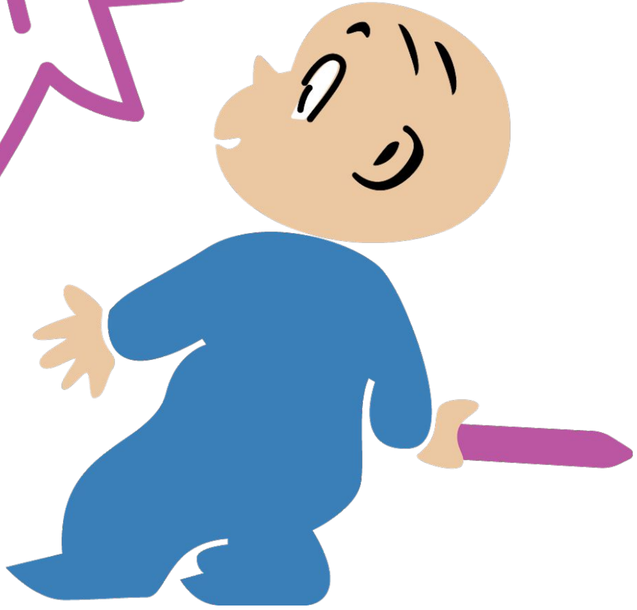
Because each row had one more ant than the one before it, the rows got wider and wider as Harold watched them march.

Harold wondered how many ants had marched past, so he wrote down what he had seen.

$$1+2+3+4+\dots=\infty\,?$$

It seemed to Harold that there were infinite ants, and they would soon fill his room!

*note 1



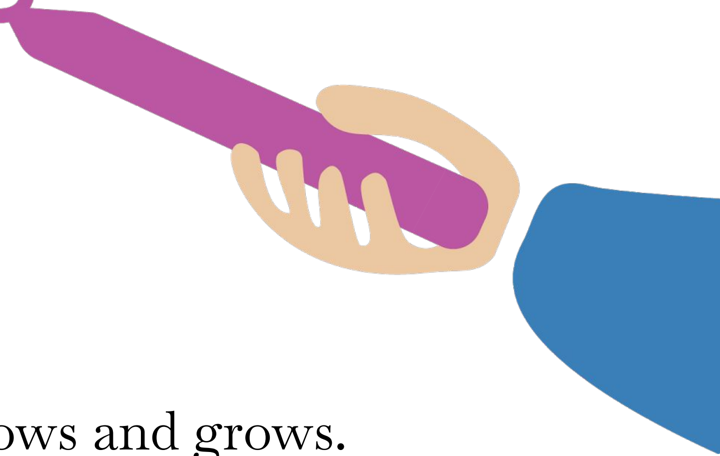
He saw that that every row was always one ant longer than the one before it. Harold thought, even with a *really high* number for *n*, the next row would always have more ants!

*notes 2–4 (incl. a–b)

Harold smartly wrote his thinking like this:

$$a_n = n$$

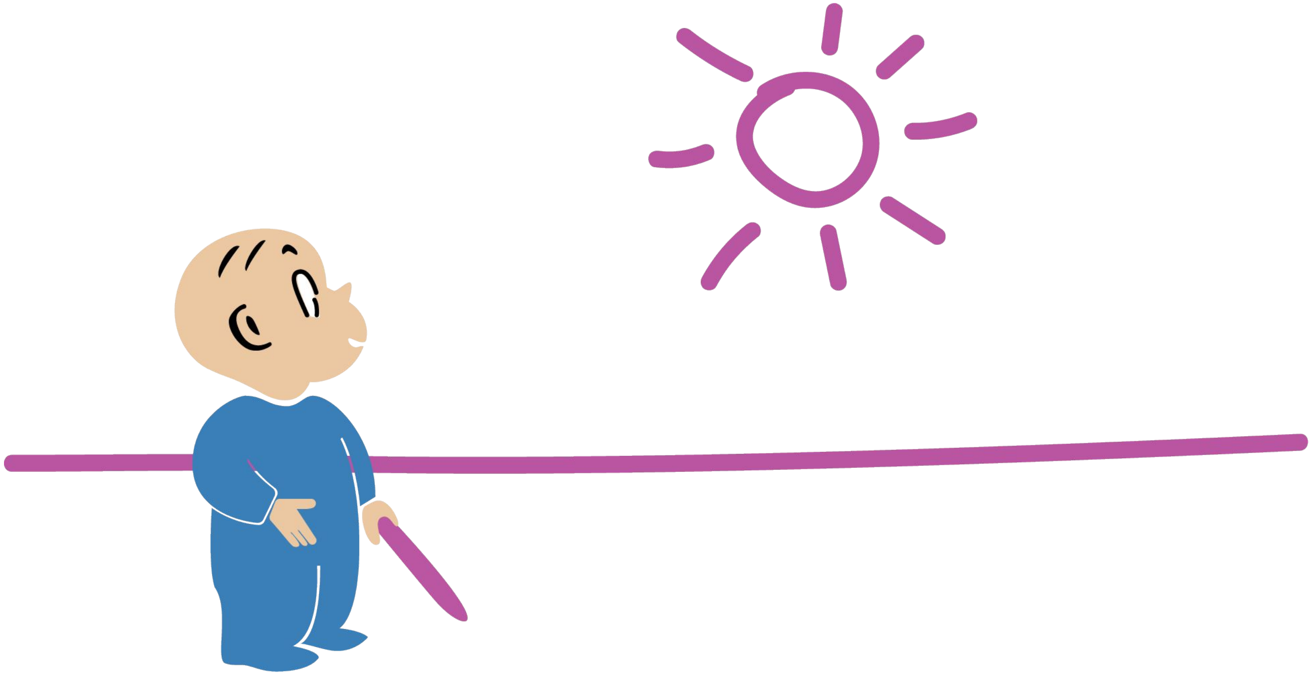
$$\lim_{n \to \infty} a_n = \infty$$



So a_n doesn't have a limit, it just grows and grows.

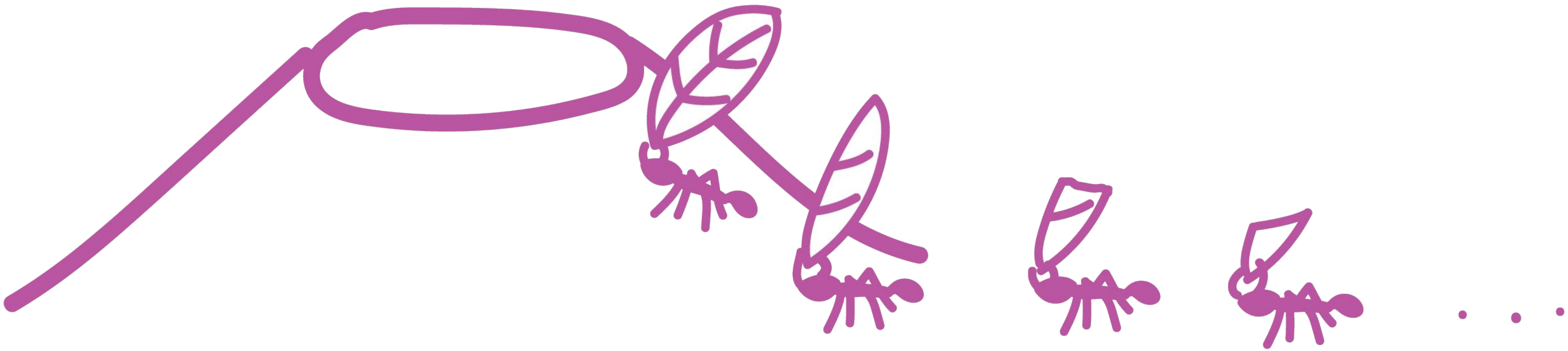
*notes 3–5, 10

When Harold awoke from a long night spent dreaming of counting ants, he stepped outside and breathed in the fresh morning air. "What if there weren't infinite ants?" Harold thought. He figured that if each row was smaller than one before it, then maybe he could count every ant before dozing off to sleep. But how much smaller?



"G! G!" Harold heard a bird chirp. Geometric series! Instead of waiting for ants to find him, this time Harold drew an anthill so they would come on their own.

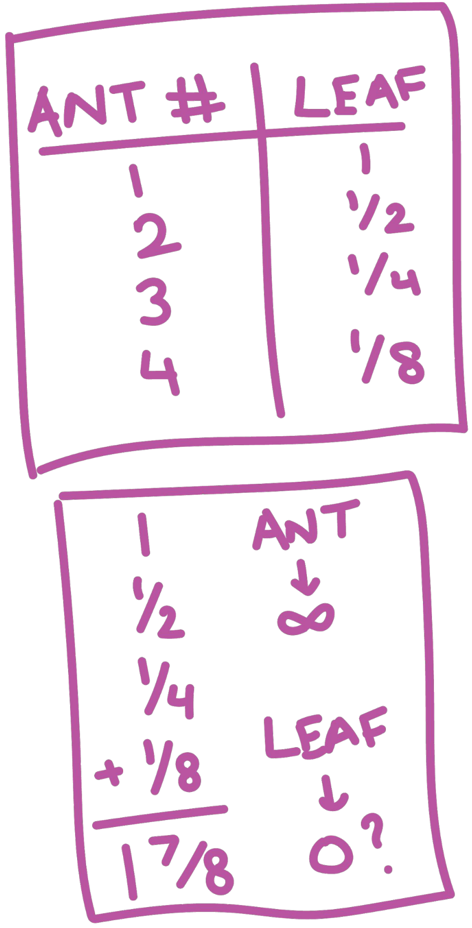
One by one, ants marched down into the anthill, carrying pieces of leaves to feed their queen. The first one had a whole leaf, the second ant half a leaf, the third a quarter, and so on.

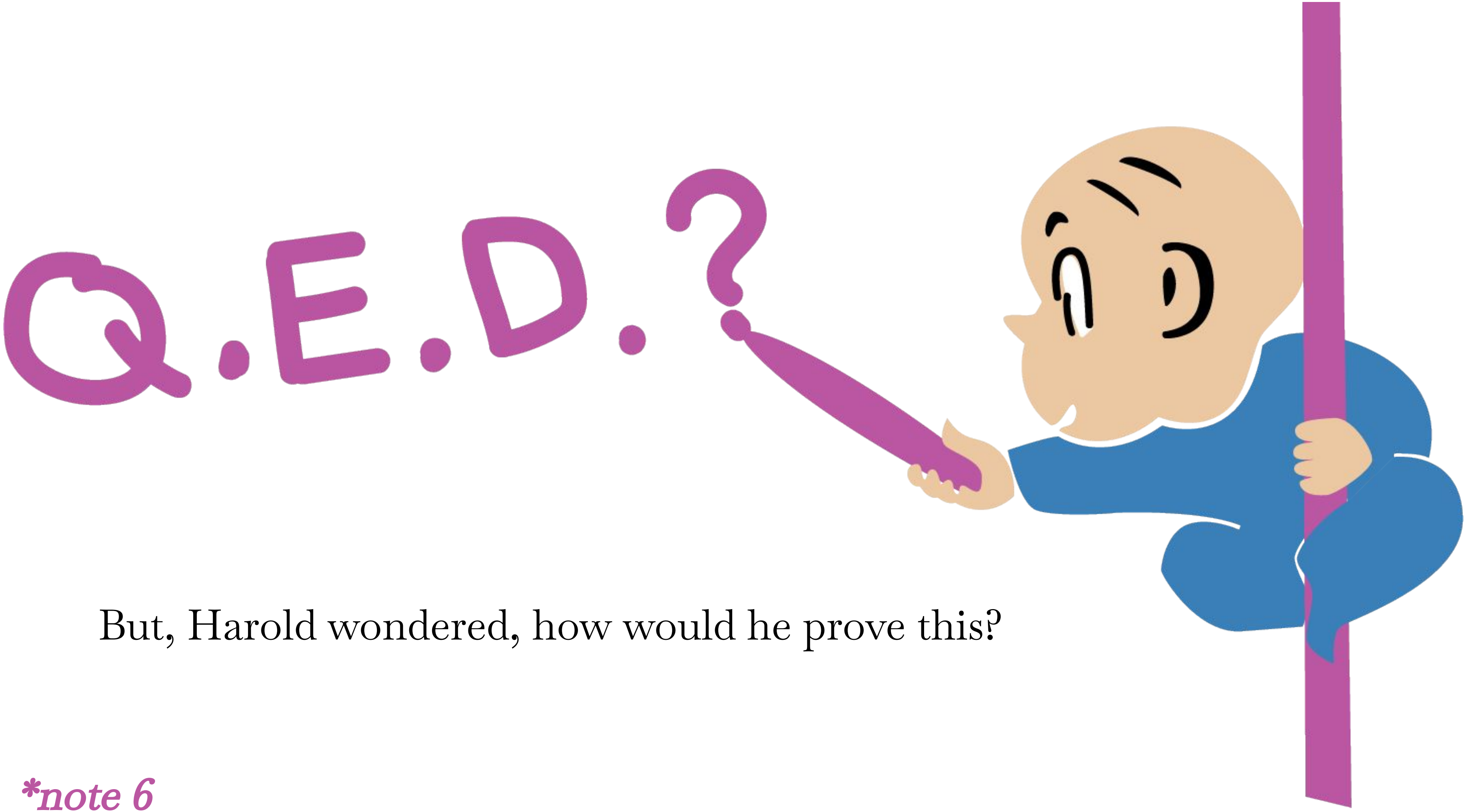


Harold wrote down what he'd seen. As he watched more and more ants disappear down the chute, their leaf-pieces shrank and shrank until they were hard to see. Like with the previous night's ants, Harold wrote what he was thinking:

$$\ell_n = \frac{1}{2^n}$$

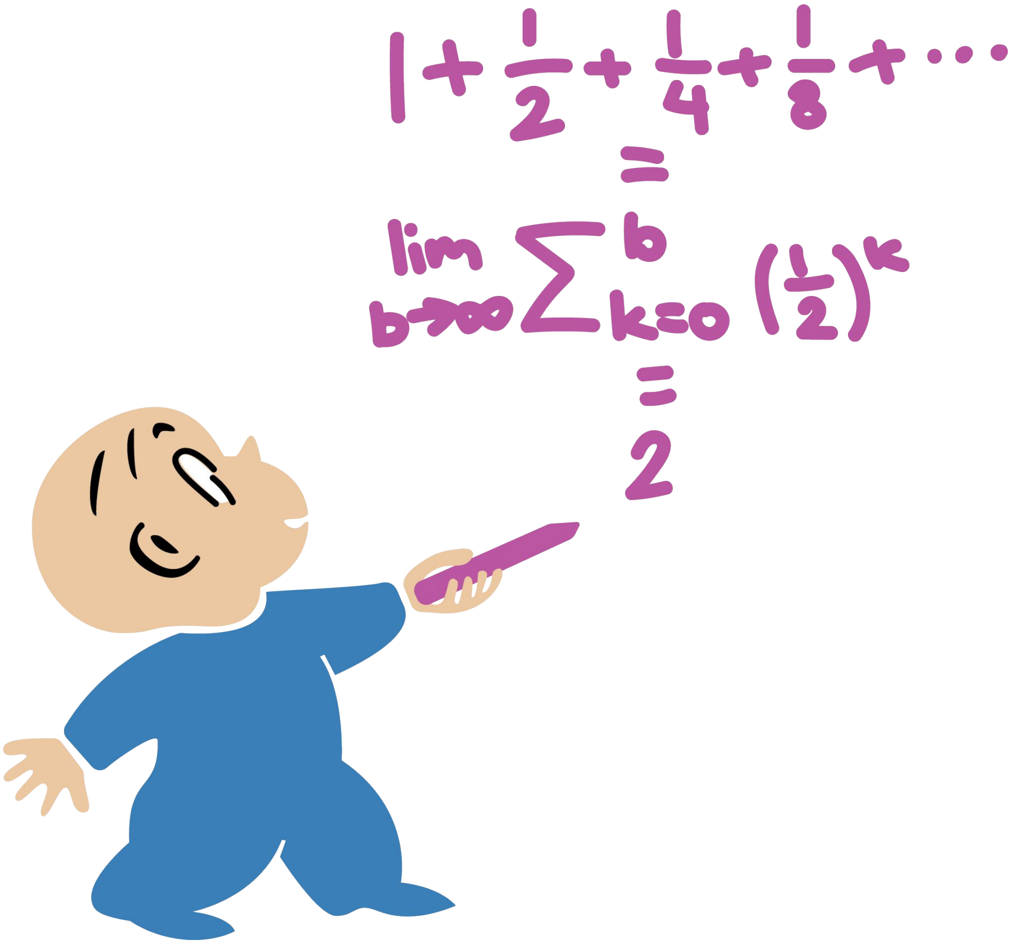
$$\lim_{n \to \infty} \ell_n = 0$$





But, Harold wondered, how would he prove this?

*note 6

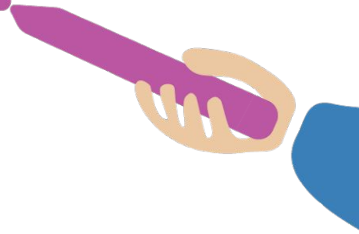


Harold could tell the pieces of leaf were getting smaller, and looked like they might eventually disappear. But he still wanted to know how many leaves the ants were bringing to the queen. He thought he could write down every piece that came by, but what if the ants never stopped marching? Harold thought of another way, using the sum of a geometric sequence.

*note 8

But Harold noticed that he was only counting natural values of *k*. What would happen, he wondered, if *k* could be any real number? To do this, Harold wrote this *improper* integral:

$$\int_0^{\infty} \left(\frac{1}{2}\right)^x dx = \lim_{b \to \infty} \int_0^{b} \left(\frac{1}{2}\right)^x dx$$



*note 7

He did this because it doesn’t make sense to write:

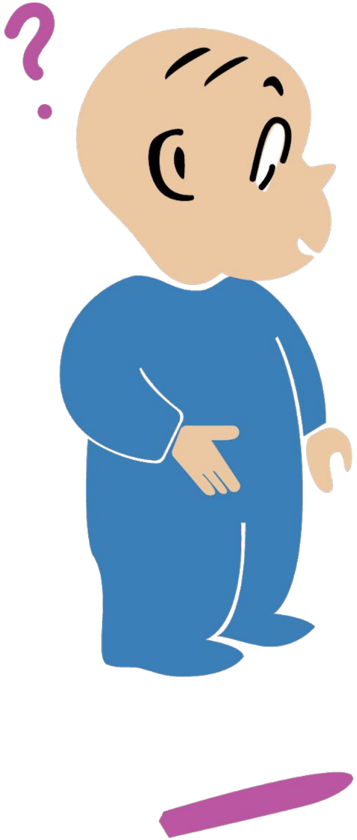
$$\left[\frac{1}{\ln(2)\cdot 2^x}\right]_0^\infty = \frac{1}{\ln 2}\left(\frac{1}{2^\infty} - \frac{1}{2^0}\right)$$

$$2^\infty = ?$$

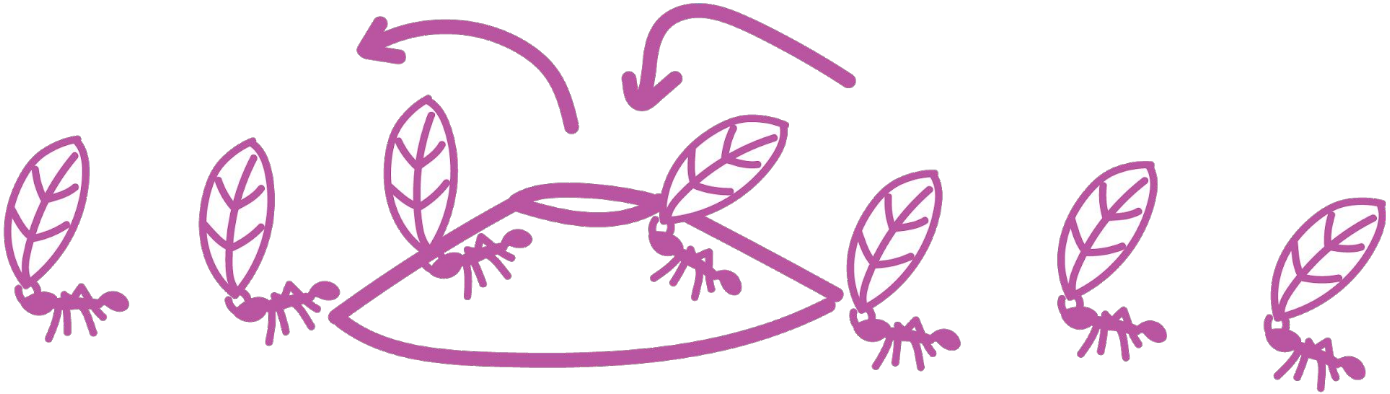
because infinity isn’t a number—it’s an *idea*. So we take limits, like Harold.

Harold thought it was interesting that both the sum and the integral of his function $f(x) = 0.5^x$ converged over $[1,\infty)$. It struck him that perhaps all functions with this property converged—an *integral test*, one might say.

*note 11



Harold left the 0.5^k anthill and stumbled upon a very peculiar one. After an ant brought one leaf in, an ant would leave the hill with another!



ANT	LEAF
1	1
2	-1
3	1
4	-1

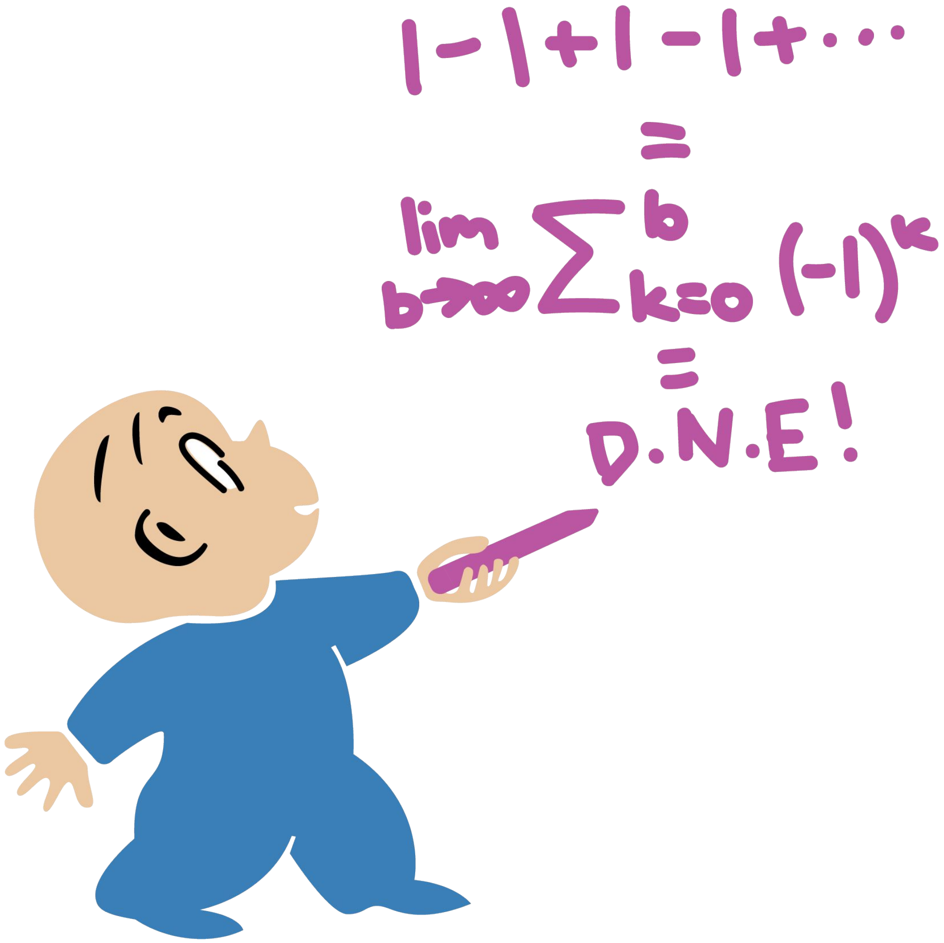
Note to self:

1: Put leaf in

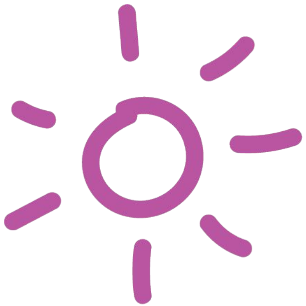
-1: Take leaf out

Like with the other anthill, Harold wrote down what he saw. It seemed to him that the anthill wasn't gaining or losing leaves overall, just in between "1" ants and "-1" ants.

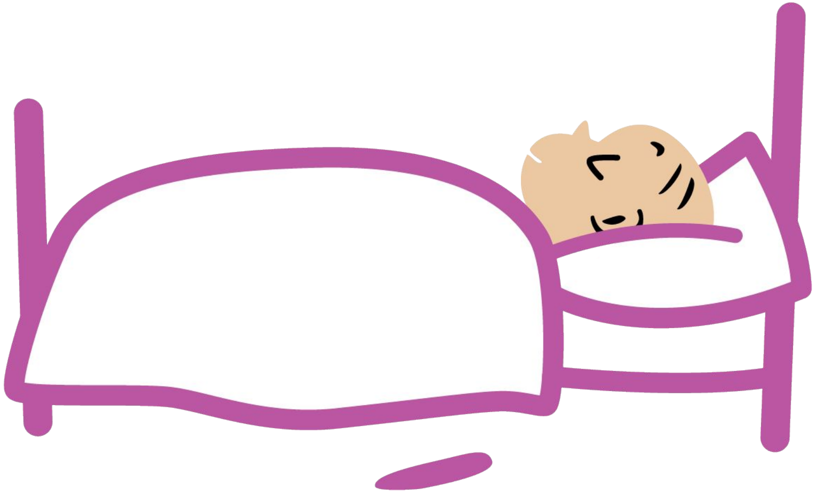
Since the sum has no limit, Harold thought it *divergent*. However, unlike the first $a_n = n$ anthill, the sum never increases or decreases beyond 0 or 1. Harold smartly said that this antill had *bounded divergence*.



*note 9



Harold suddenly felt hot from standing in the sun all morning, so he drew himself a bed and napped right where he was. He dreamt, to his chagrin, of ants bringing leaves in ever-more complex patterns. Alas, there was much to be learned from holes in the sand.



ANT k brings $\dfrac{k^3+2k^2-3}{k^5-7k+1}$ leaves

Harold wondered if the ants simply dividing a finite number of leaves into smaller and smaller pieces, or if they would eat up the whole forest. That is, he wondered if the series was convergent or divergent.

At high values of k, Harold thought that the lower-order terms would become irrelevant, and so he thought that as k approaches infinity, the differs little from k^{-2}.

$$\frac{k^3+2k^2-3}{k^5-7k+1} \rightarrow \frac{k^3}{k^5} = \frac{1}{k^2}$$

$$\frac{k^3+2k^2-3}{k^5-7k+1} \div \frac{1}{k^2} = \frac{k^5+2k^4-3k^2}{k^5-7k+1}$$

$$\text{as } k \to \infty, \quad \frac{1+\left(\frac{2}{k}\right)-\left(\frac{3}{k^3}\right)}{1-\left(\frac{7}{k^4}\right)+\left(\frac{1}{k^5}\right)} \to 1$$

Since k^{-2} converges (it's a *p*-series), and the limit of the quotient converges to 1, Harold could tell by the Limit Comparison Test that the ants weren't eating up the whole forest because their series converges.

*note 12

After resolving his fear the the ants were eating up his forest, Harold began thinking about the anthill where ants came in and out with leaves. He wondered, what if the ants put in and took out differently sized pieces of leaf?

Harold imagined the ants like this:

ANT 1 takes 1 leaf.
ANT 2 brings ½ leaf.
ANT 3 takes ⅓ leaf.
ANT 4 brings ¼ leaf.

⋮

ANT k brings $\frac{(-1)^k}{k}$ leaves

It made sense to him that the anthill was losing leaves overall, because none of the ants bringing leaves in could compensate for the leaves being taken out.
Harold also thought that the series was convergent because the pieces of leaf being put in and taken out seemed to be approaching zero.

$$\lim_{k \to \infty} \frac{(-1)^k}{k} = 0$$

*note 10

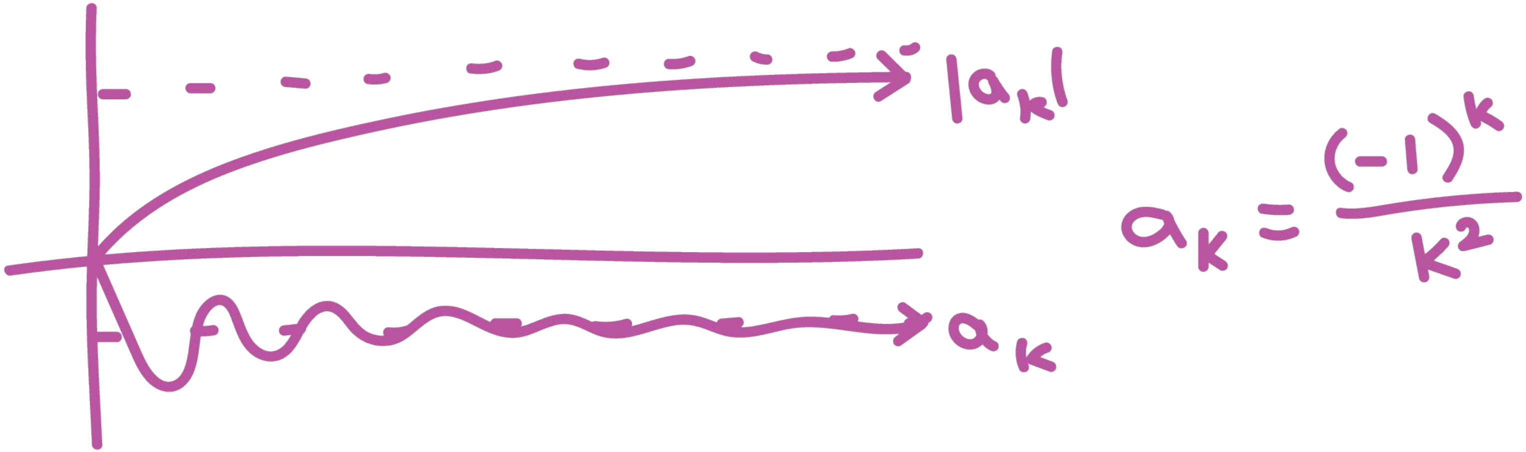
$$a_k = \frac{(-1)^k}{k}, \quad |a_k| = \frac{1}{k}$$

Harold thought, then, what would happen if the ants put in leaves according to the same rule, but didn't take any out. Immediately, since he's so astute, Harold realized that $/a_k/$ was the harmonic series, a notorious divergent series! So, it seemed to him, a_k only converges sometimes, that is, it is *conditionally convergent*.

*note 14

If that pattern was conditionally convergent, then, Harold dreamt, there must be some patterns that aren't. Some patterns must converge whether the ants are putting in or taking out leaves. He recalled the first dream-pattern, the convergent series he mistook for a divergent series.

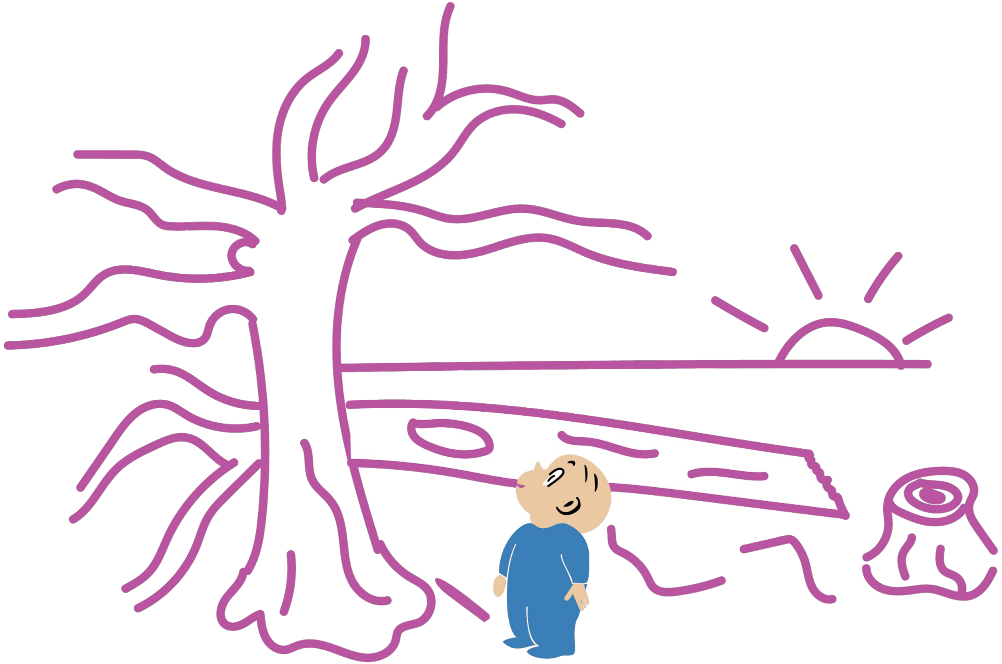
$$a_k = \frac{(-1)^k}{k^2}, \quad |a_k| = \frac{1}{k^2}$$



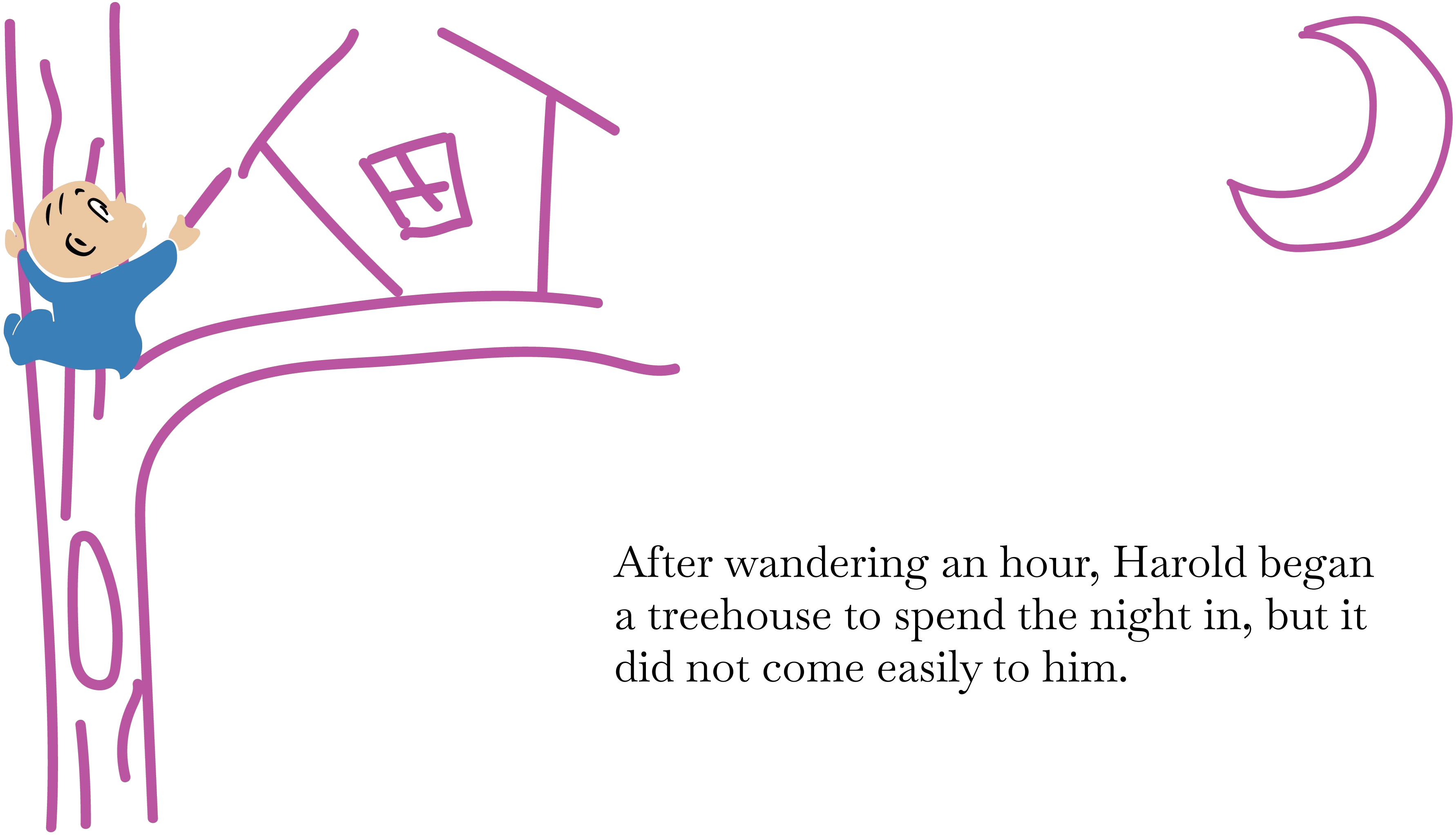
Some strange images appears in Harold's head, and from them, he knew that both a_k and $/a_k/$ were convergent. So, he concluded, series of this kind *converge absolutely*.

*note 13

Harold awoke from his late-afternoon nap with a fright. The sun was setting, and he'd nowhere to spend the night yet!

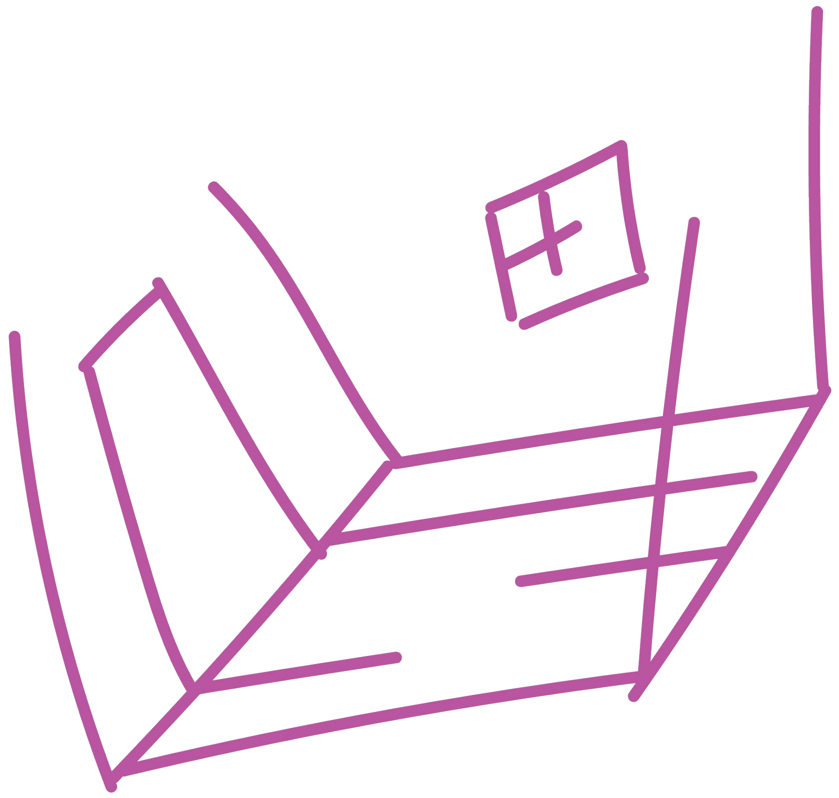


Finished with the ants, Harold now began a small adventure to find a place to sleep. He walked to the forest he'd drawn a few days before, and stepped over mossy fallen boughs and crunched on stray twigs as the setting sun's low rays dappled his blue pajamas.

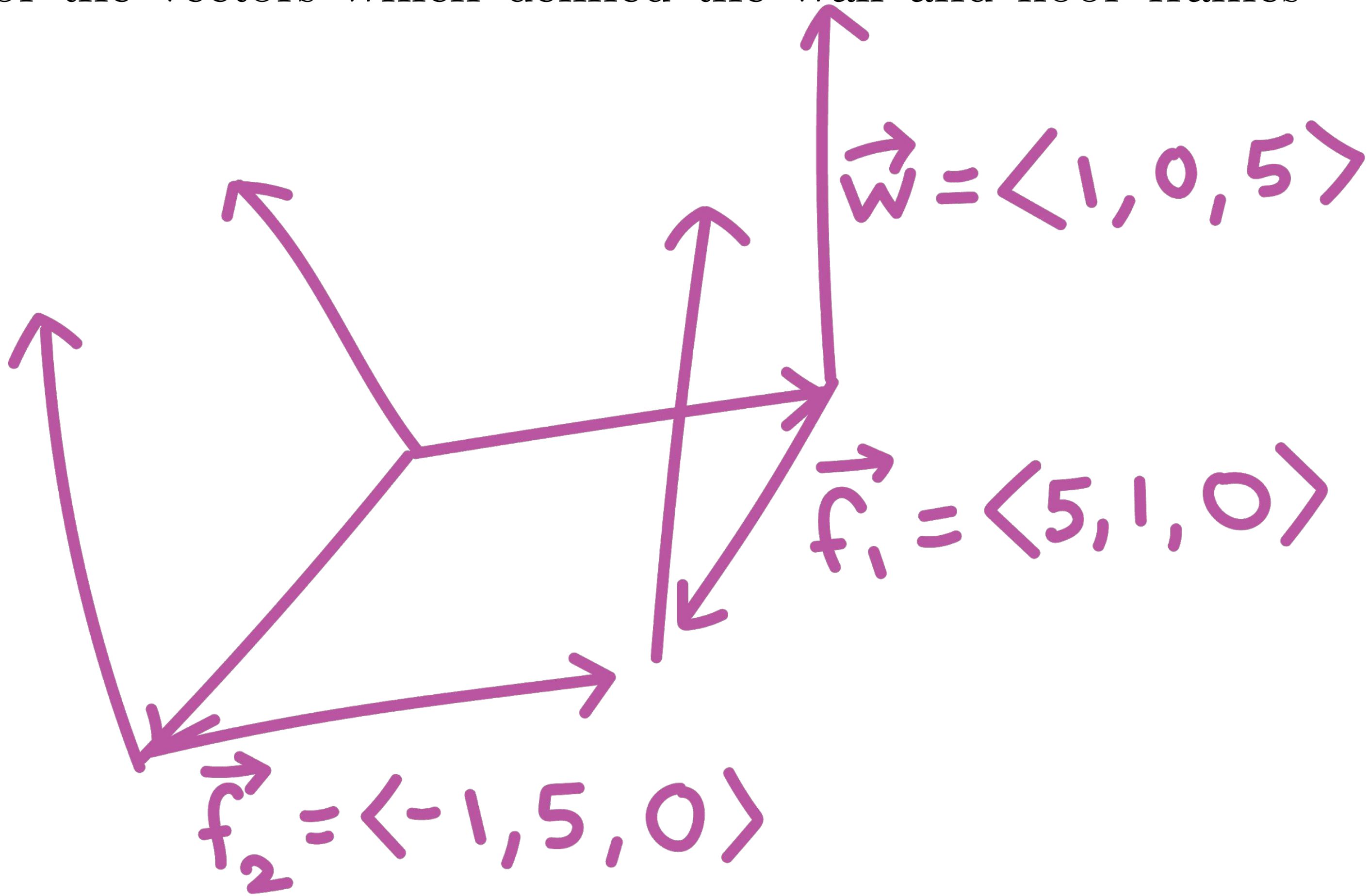


After wandering an hour, Harold began a treehouse to spend the night in, but it did not come easily to him.

The floor frame kept coming out crooked and the windows didn’t sit right.



Then Harold realized all he needed for the frame to be square was to make the dot product of the vectors which defined the wall and floor frames equal zero.



$$\vec{f_1} \cdot \vec{f_2} = -5+5+0=0 \quad \checkmark$$

$$\vec{f_1} \cdot \vec{w} = 5+0+0=5 \quad \times$$

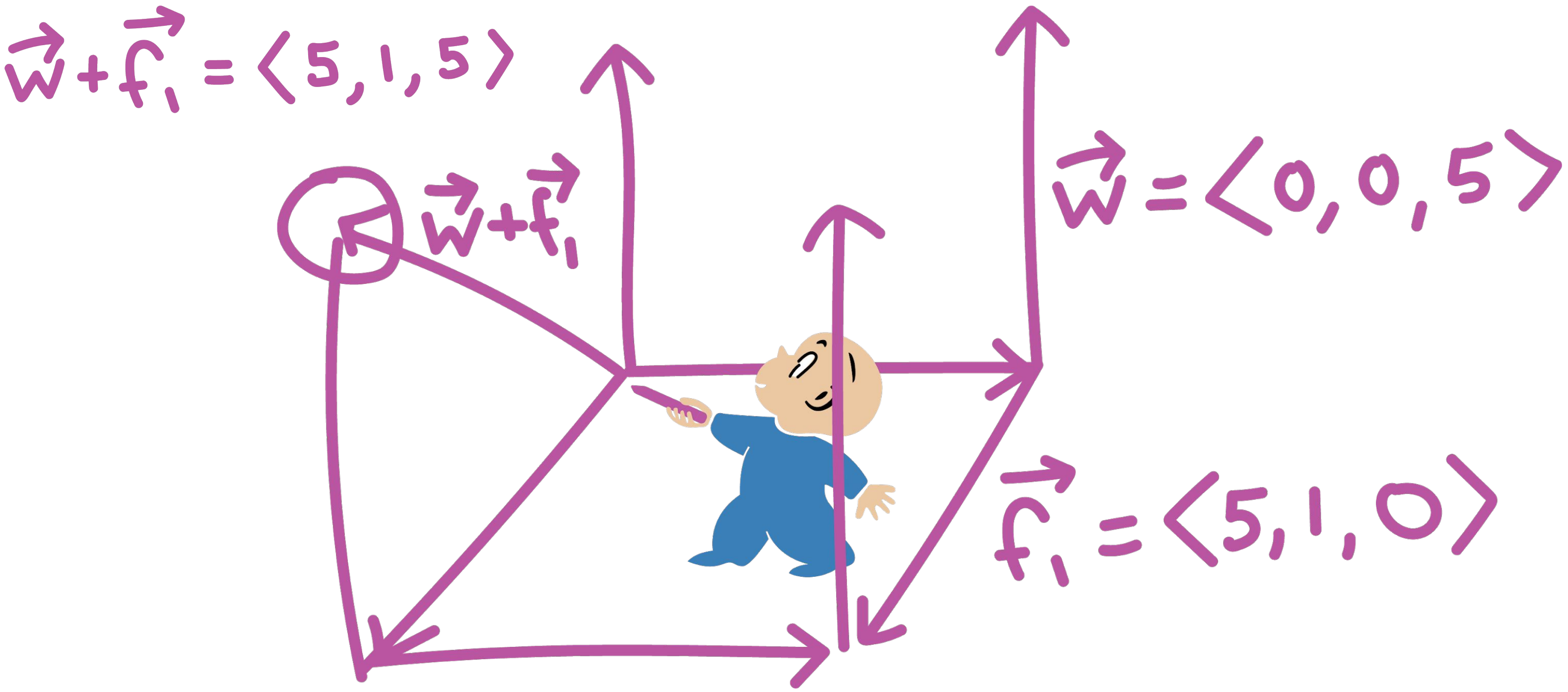
$$\vec{f_2} \cdot \vec{w} = -1+0+0=-1 \quad \times$$

Harold saw that the floor was square, but that the walls were were not square to the floor. To fix the wall frames, Harold realized he could give the vector *w* zeros for its *x* and *y* coordinates, which he thought made sense because walls should tilt neither left nor right, nor forward nor backward.

He change *w* to this:

$$\vec{w} \stackrel{\checkmark}{=} \langle 0,0,5 \rangle$$

*note 15



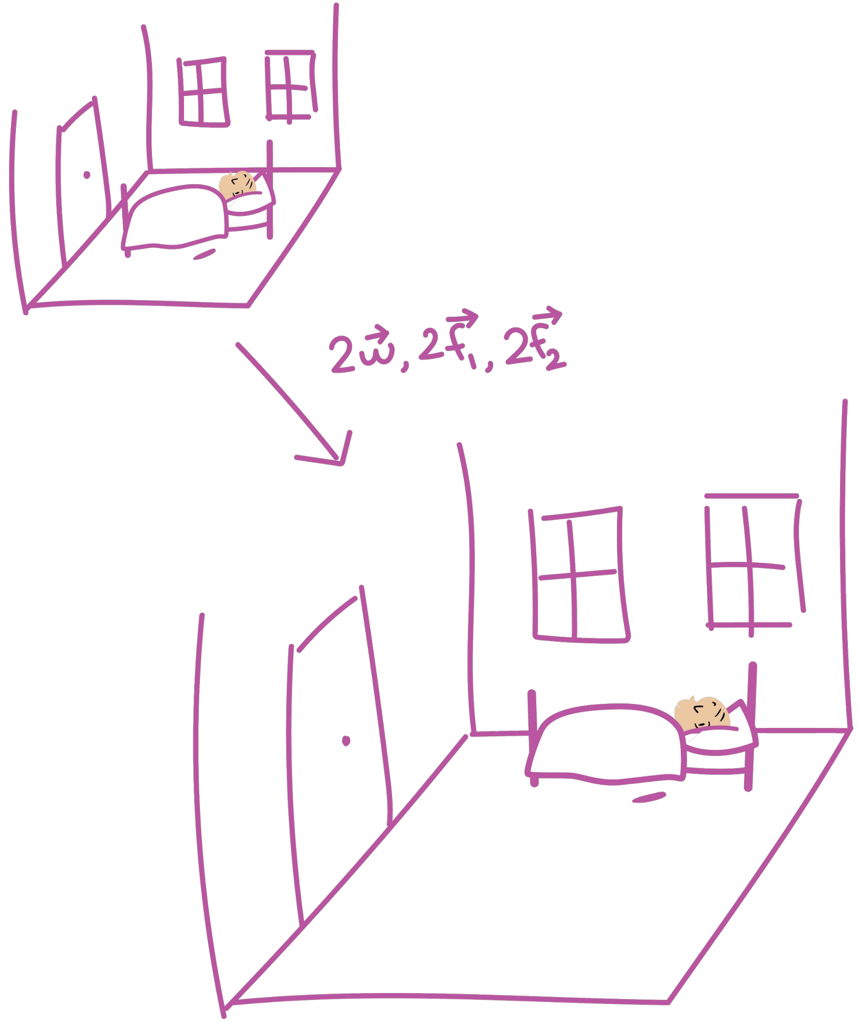
To find the opposite edge of his treehouse's roofline, Harold simply adds vector f_1 and w to get a nice diagonal vector. The head of this new vector points right where Harold needs to start the roof.

*note 16

But even with square walls and floor, Harold didn't like that there was just room enough for him and his bed.

Where would his friends sleep?

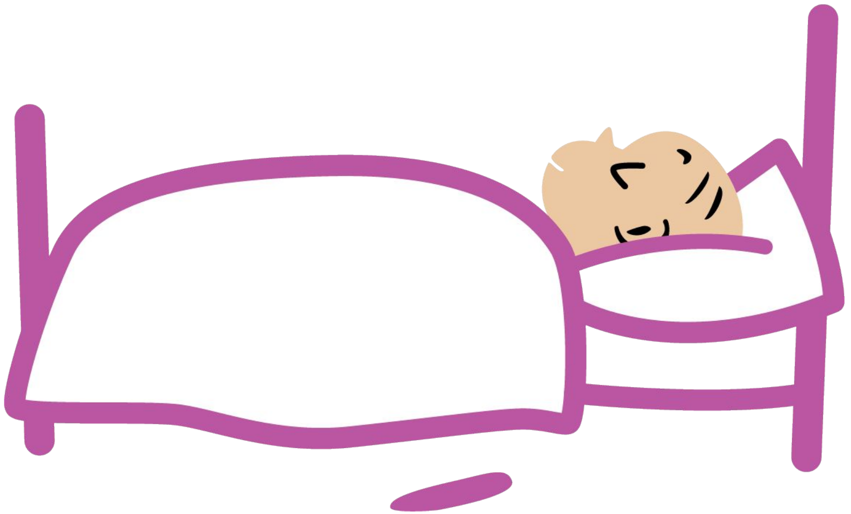
Just by taking each corner on his treehouse and multiplying them by two, Harold made enough space for friends to come, too.

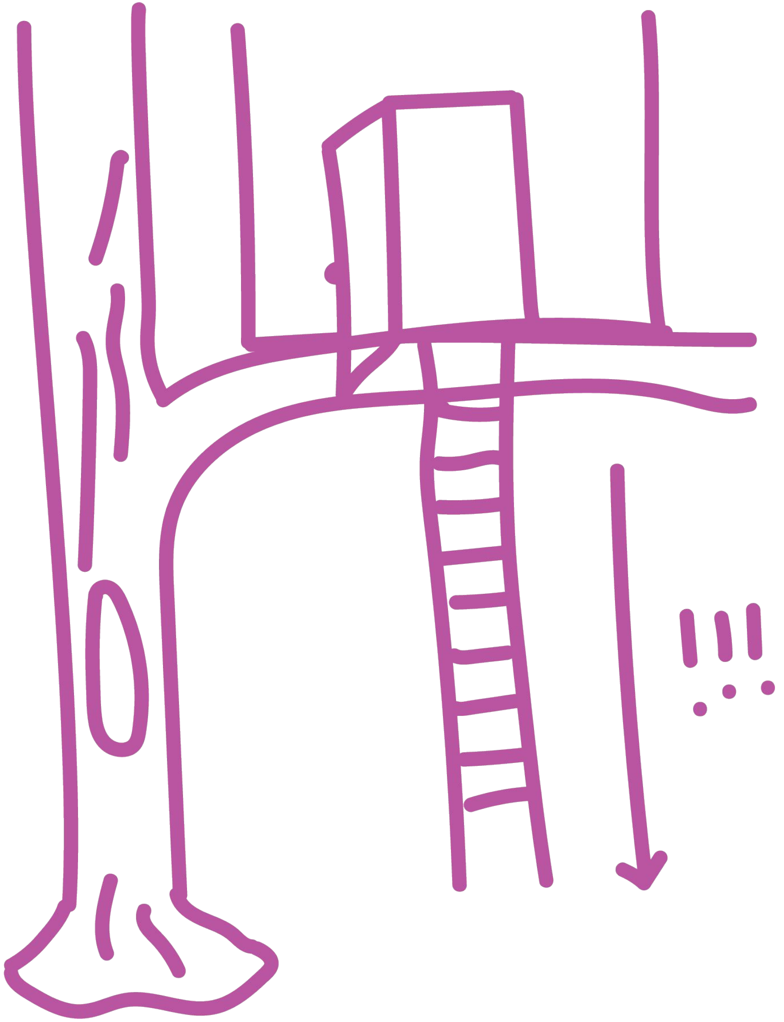


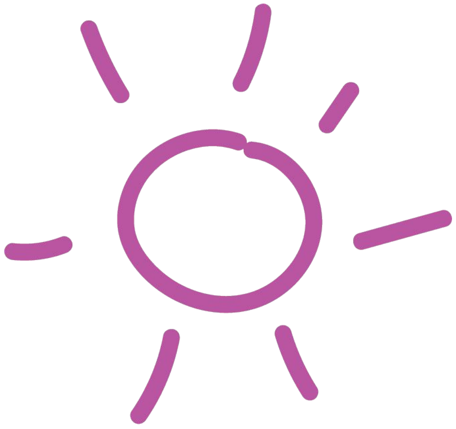
*note 17

After such tiring work,
Harold was ready for bed.

He tucked himself in under the soft white sheets
and nodded off into the world of dreams.

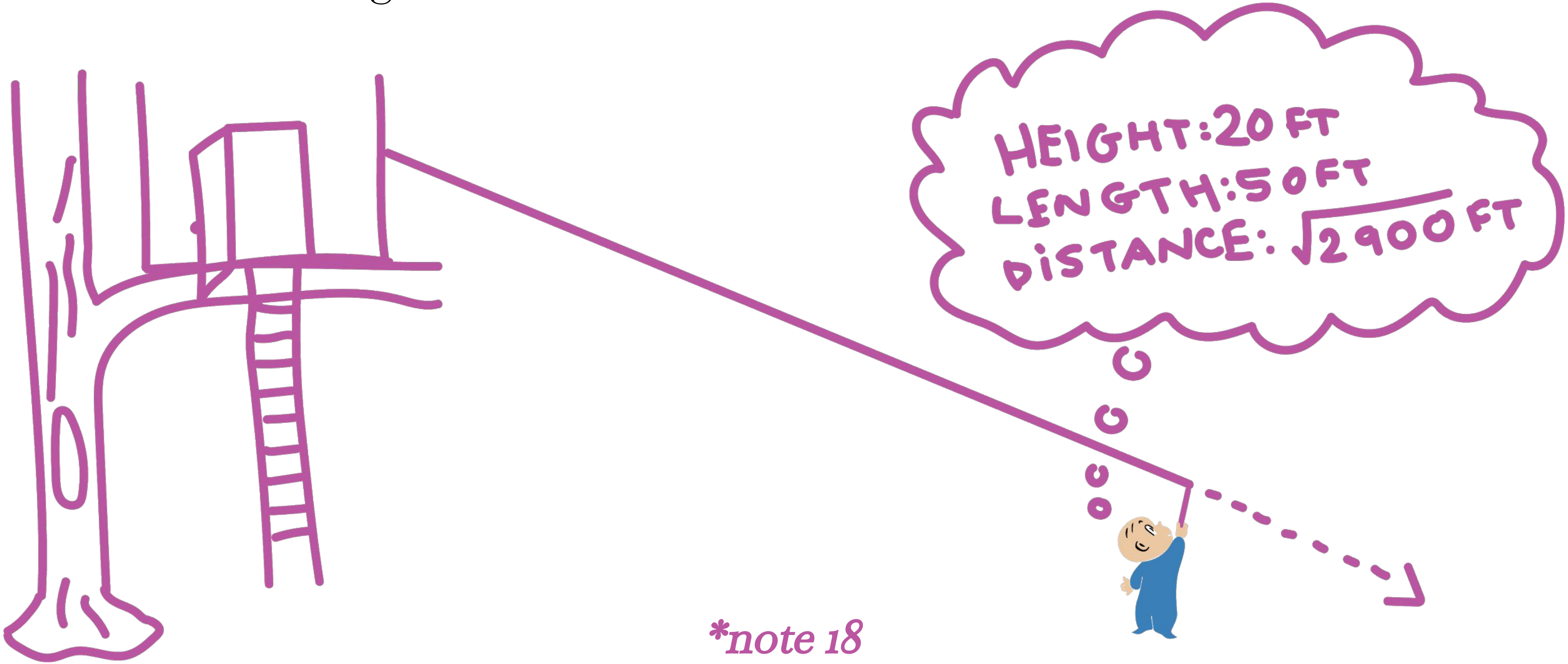






In the morning, Harold went
to the door to go down, but
the ladder scared him.
It was so far down!

Harold thought that a zip-line would be a fun and safe way to leave in the morning. But how long would it have to be? Harold thought some more.



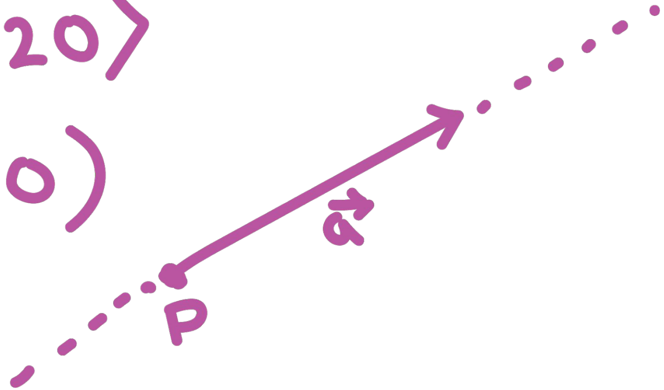
*note 18

For fun, Harold imagined where his zipline would go if it could continue forever. He assumed it would go off into space, past the treehouse, but he wondered where on Earth he would emerge if he could pass through the ground.

If he knew the equation of zip-line's line, maybe he could figure it out! Wherever the line emerged from the Earth's surface, the coordinates would fulfill the equation.

$\vec{a} = \langle 50, 0, 20 \rangle$

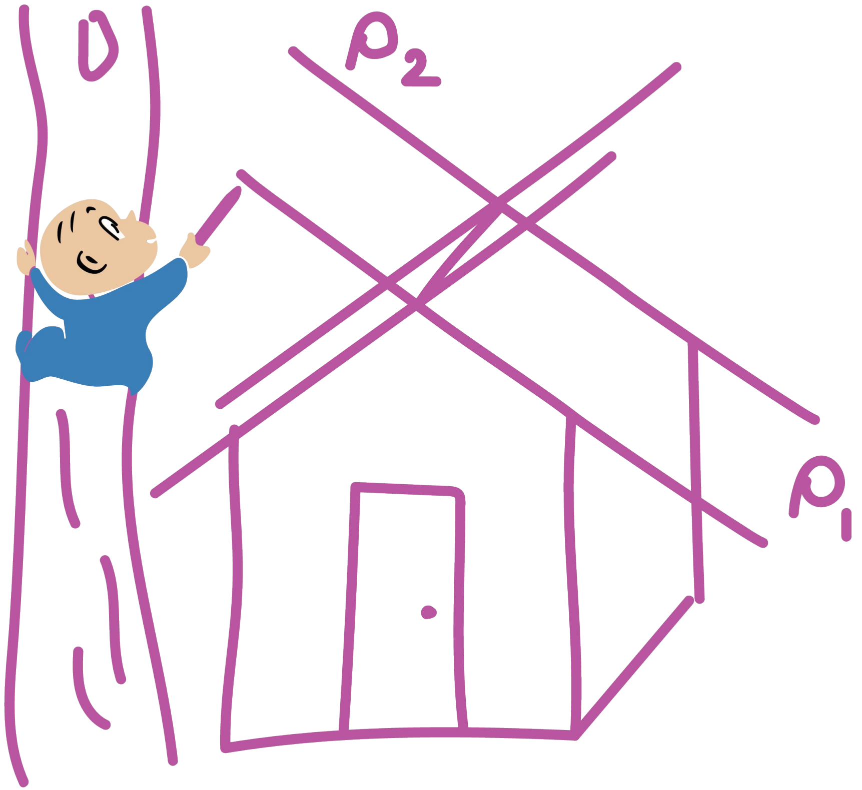
$P = (3, 7, 0)$



$x(t) = 3 + 50t$
$y(t) = 7$
$z(t) = 20t$

*note 19

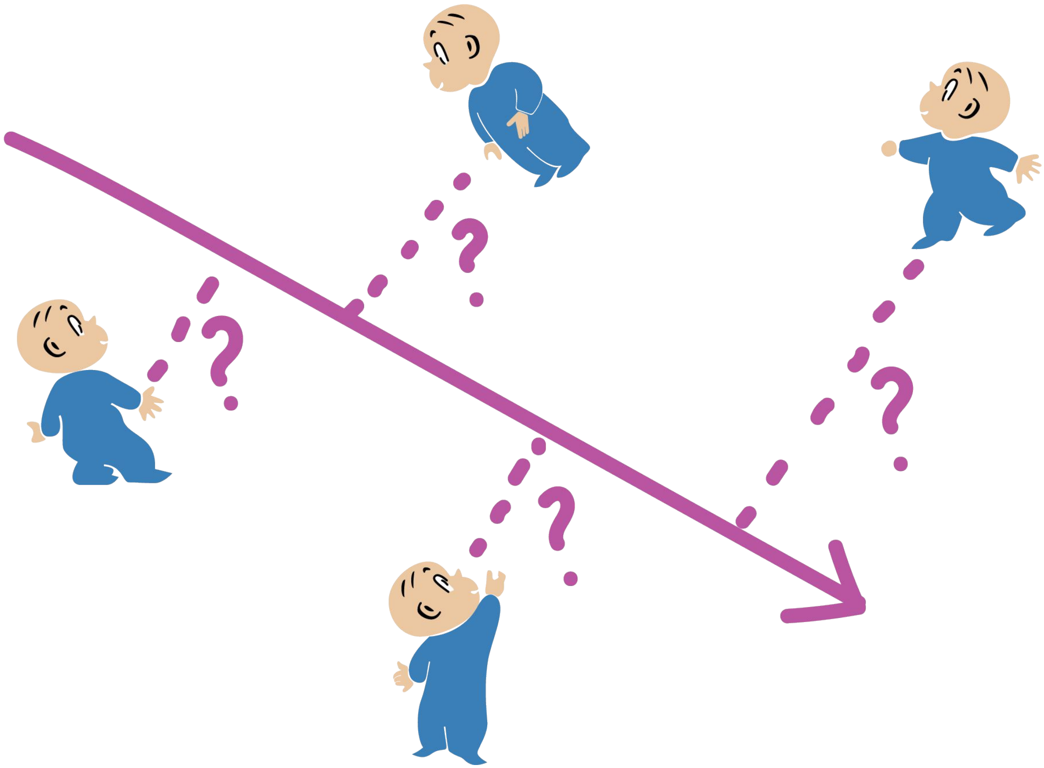
Just then, Harold realized there was a
more important problem at hand:
it had begun to rain, and he'd forgotten to
draw a roof for his treehouse!



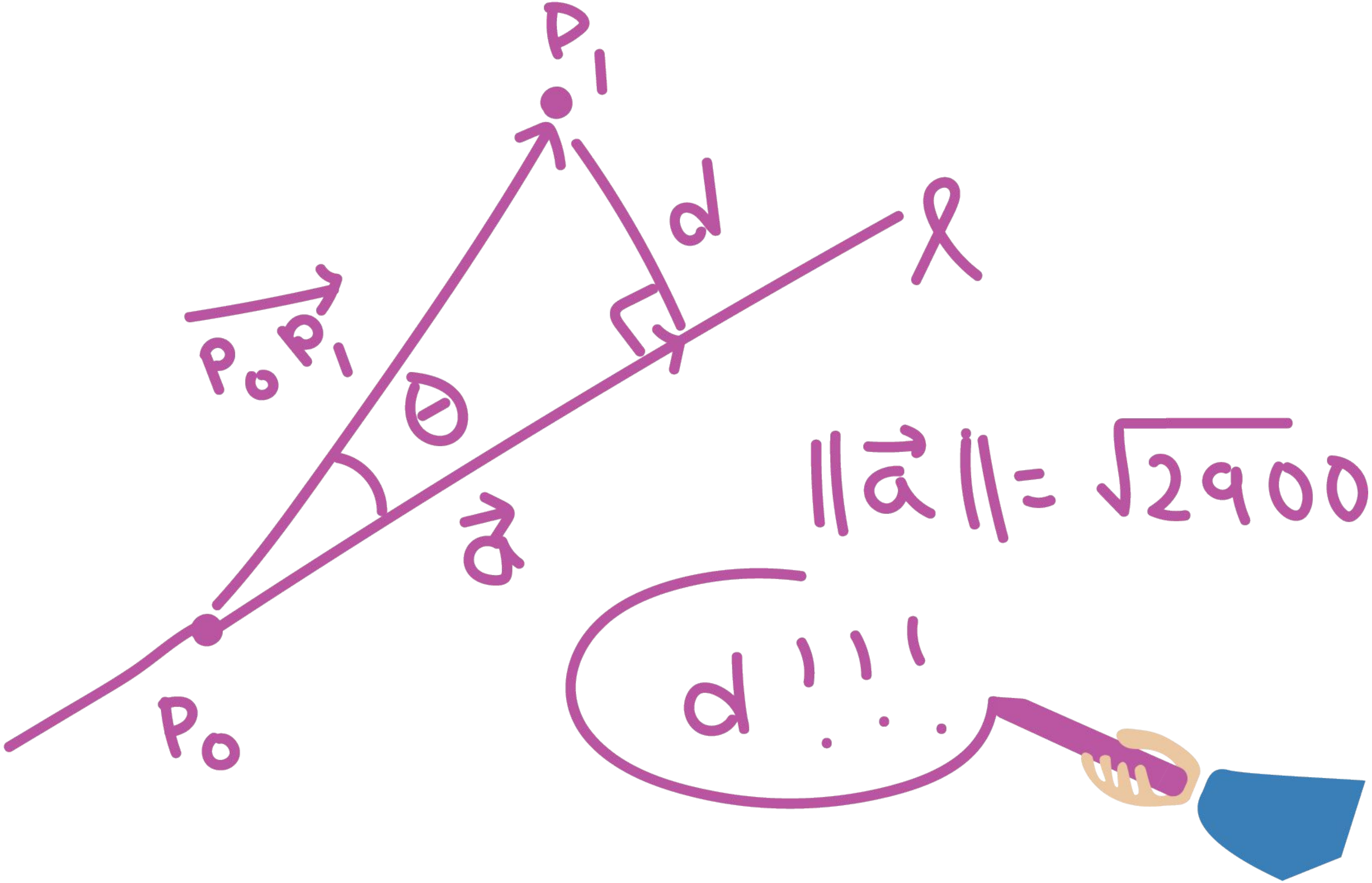
Quickly, Harold drew up two planes that intersected to form a roof.

*note 20

Having saved his treehouse, Harold could get back to the zip-line. Now that there was line in his world, he wondered how far things were from it. Which way would he walk to find the zip-line fastest?



*note 21

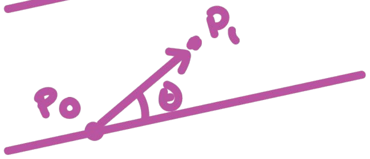


P_1
d
ℓ
$\overrightarrow{P_0P_1}$
θ
$\vec{a}$
P_0
$\|\vec{a}\| = \sqrt{2900}$
d!!!

Now, Harold could tell anyone, anywhere in the world, how far they were from his zip-line; all they need to do was:

1. Pick a point on zipline

2. Draw a vector from you to P_0



3. Find length of the vector

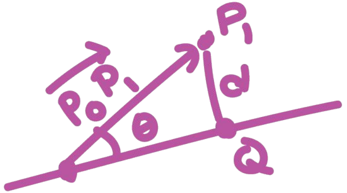
$$\|\overrightarrow{P_0P_1}\| = ?$$

4. Take dot product of $\overrightarrow{P_0P_1}$ and the zipline's direction vector $\vec{d}$

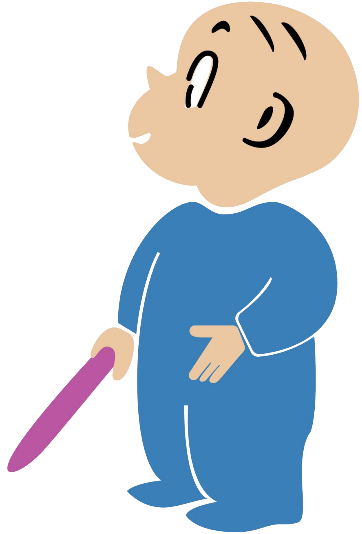
$$\overrightarrow{P_0P_1} \cdot \vec{d} = \cos\theta \|\overrightarrow{P_0P_1}\| \|\vec{d}\|$$

5. Solve for θ

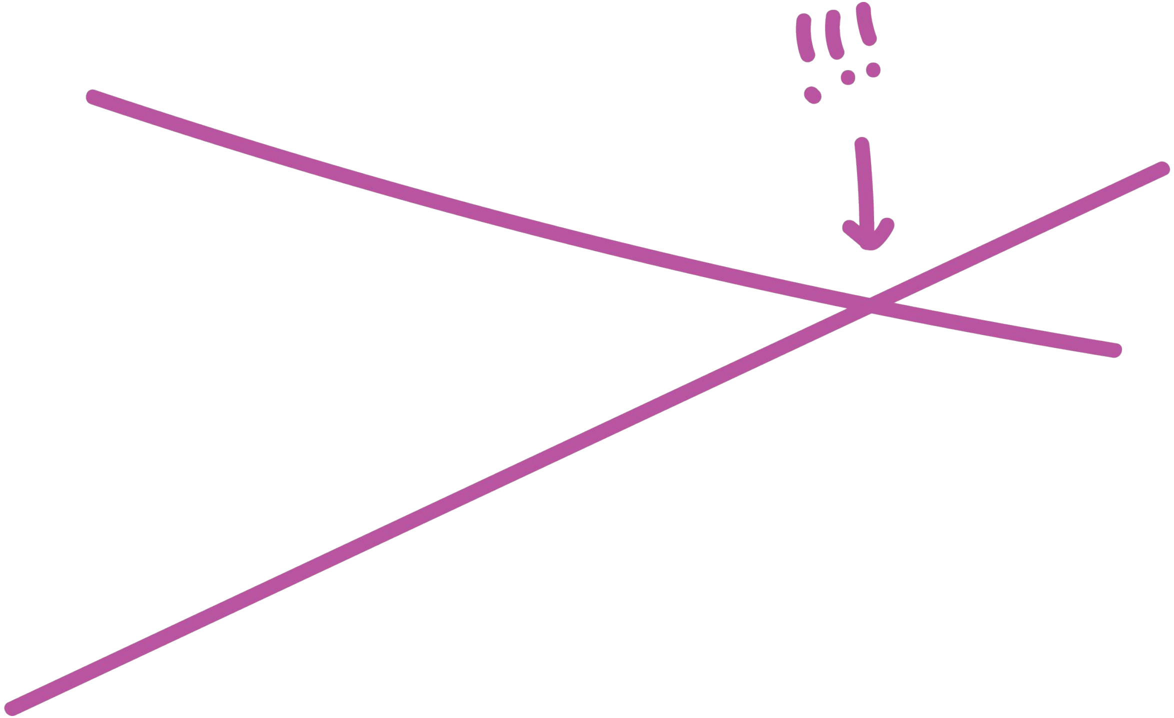
6. Use trig to find distance from P_1 to Q



$$d(P_1, Q) = \|\overrightarrow{P_0P_1}\| \sin\theta$$



Because the zip-line was such a success, Harold thought he should draw other zip-lines to get around the forest. He worried, though, that they might run into each other.



$$l_0 : f_0(t) = (3+50t, 7, 20t)$$

$$l_1 : f_1(t) = (10+3t, 1+70t, 30t)$$

Before he drew his new zipline, Harold thought he would check if its parametric equations would intersect with the old zipline.

Lo and behold, since there is no solution, the lines do not intersect! Harold could build his second zip-line without fear of collision.

$$3+50t=10+3s$$
$$50t=7+3s$$
$$t=\frac{7+3s}{50}$$
$$t=\frac{7+2t}{50}$$
$$50t=7+2t$$
$$48t=7$$
$$t=\frac{7}{48}$$

$$7=1+70s$$
$$6=70s$$
$$s=\frac{6}{70}$$

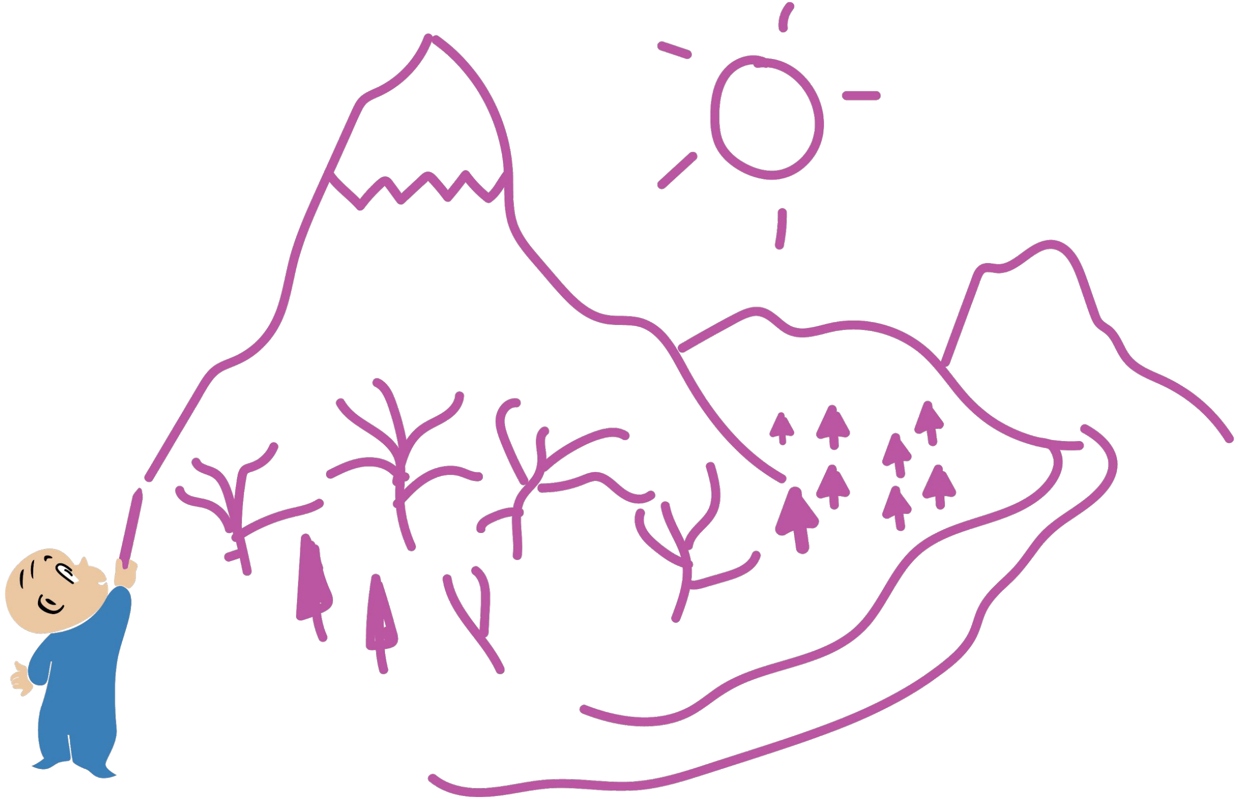
$$20t=30s$$
$$s=\frac{2t}{3}$$

$$\frac{6}{70}=\frac{2t}{3}$$
$$140t=18$$
$$t=\frac{9}{70}$$

$$\frac{7}{48}\neq\frac{9}{70}$$

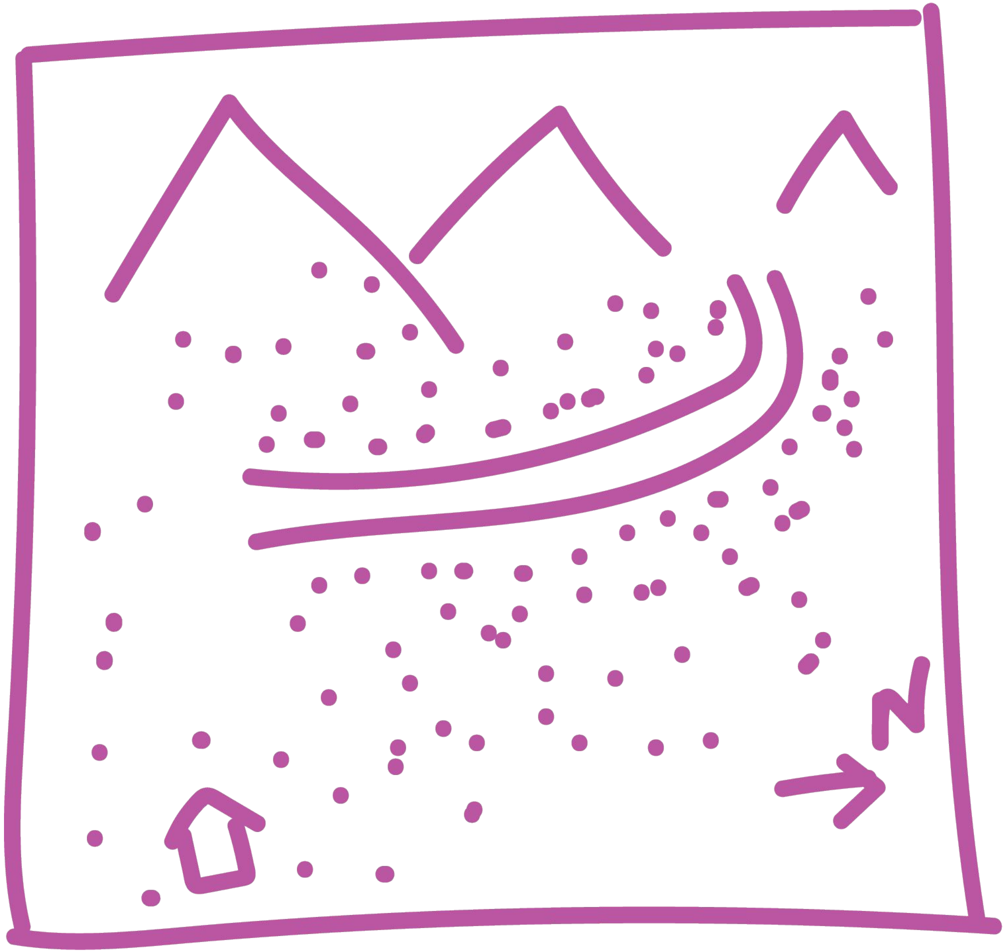
NO SOLUTION

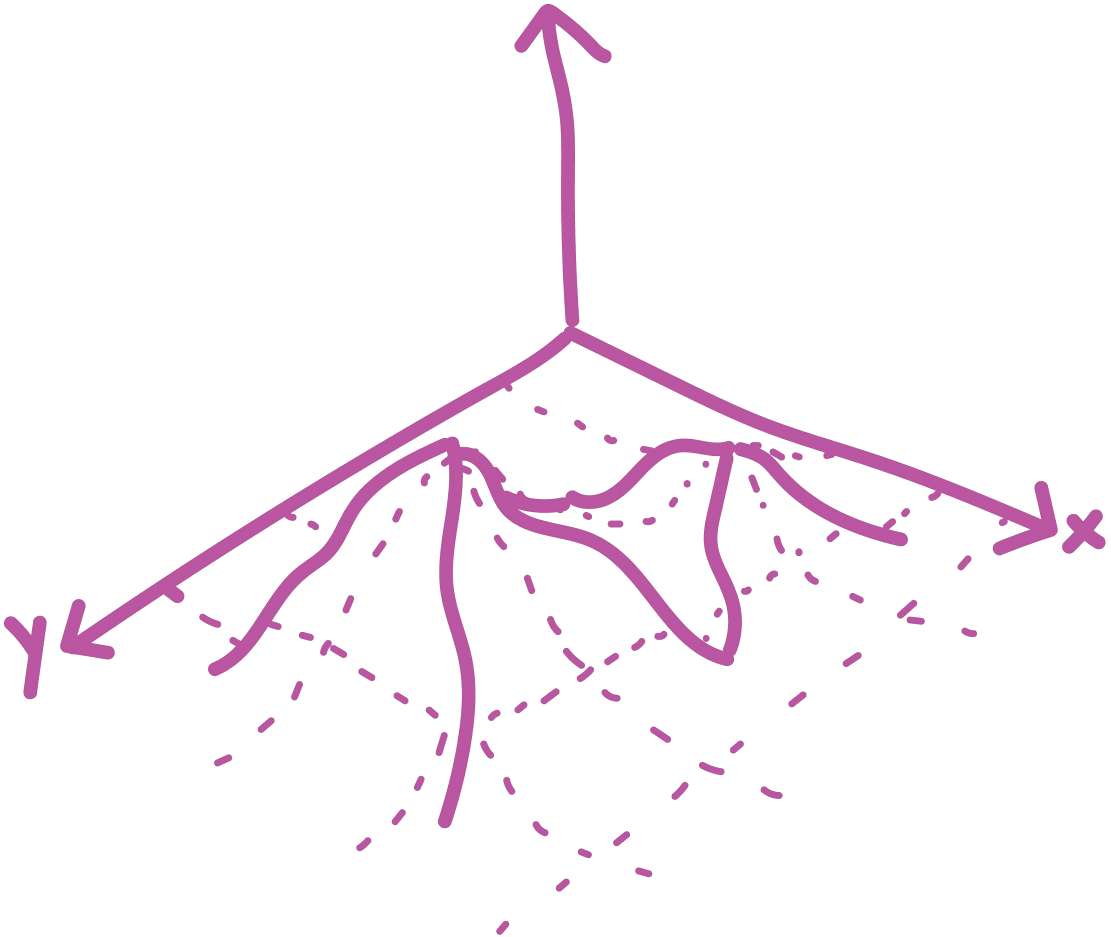
*note 22



Although the forest around his treehouse was beautiful, it was small.
So, Harold drew more and more trees, a river, and mountains in the distance.

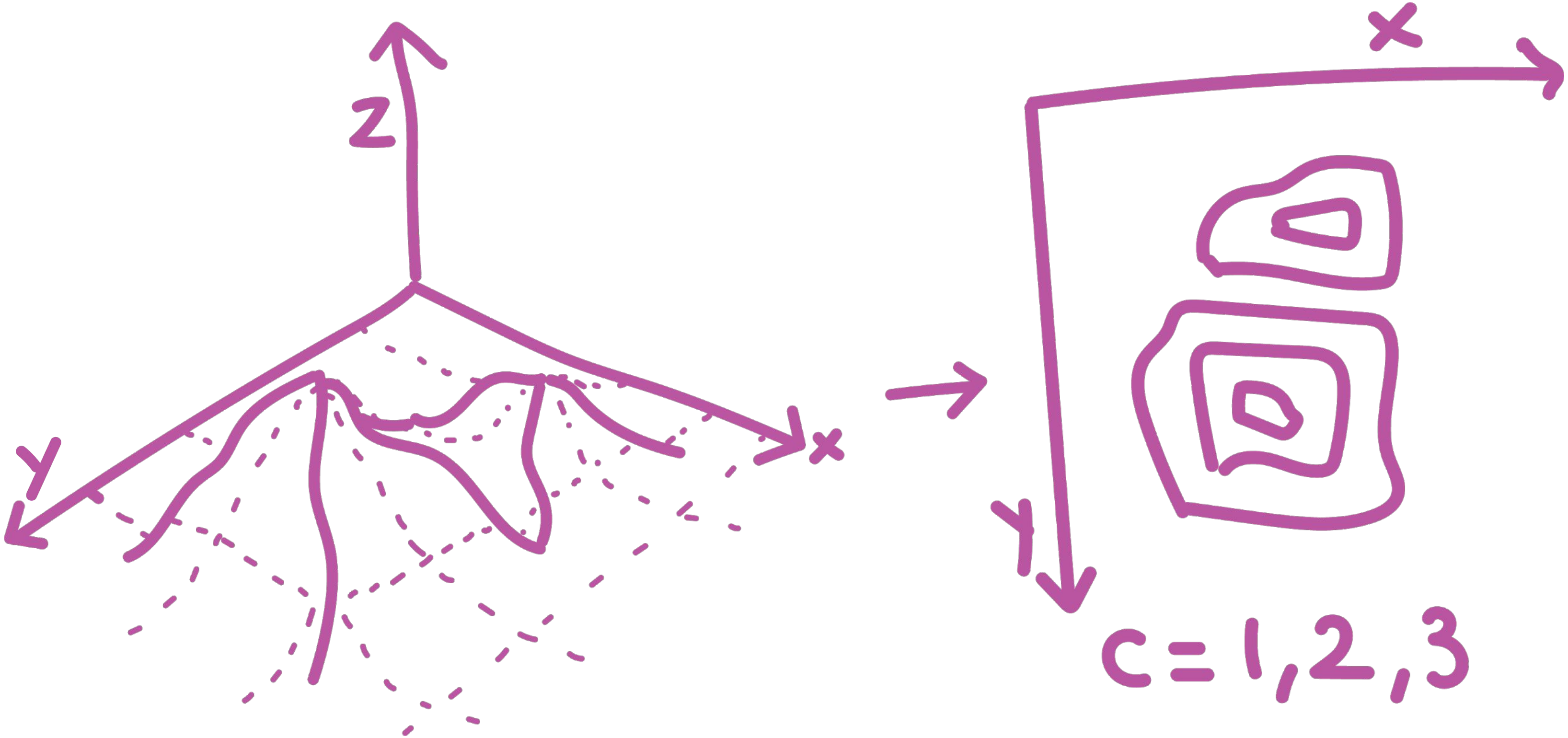
Harold knew the terrain since he'd drawn it, but he wanted his friends to be able to explore, too. So, Harold drew up a map. Somehow, Harold wanted to represent the slope of the mountains, so his friends knew where was steep and where was flat.





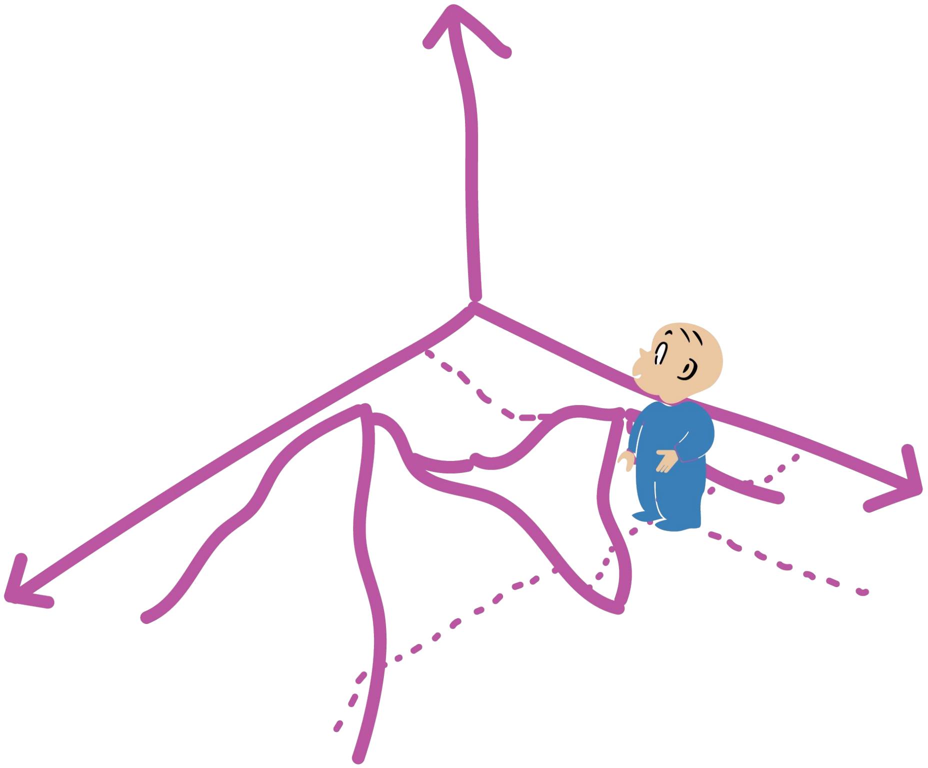
Harold imagined that there was a function $f(x,y)$ that gave the height of the terrain at some coordinate (x,y). Then, he imagined taking the derivative with respect to x or y to find the slope at that same point, in the x or y direction.

*note 23

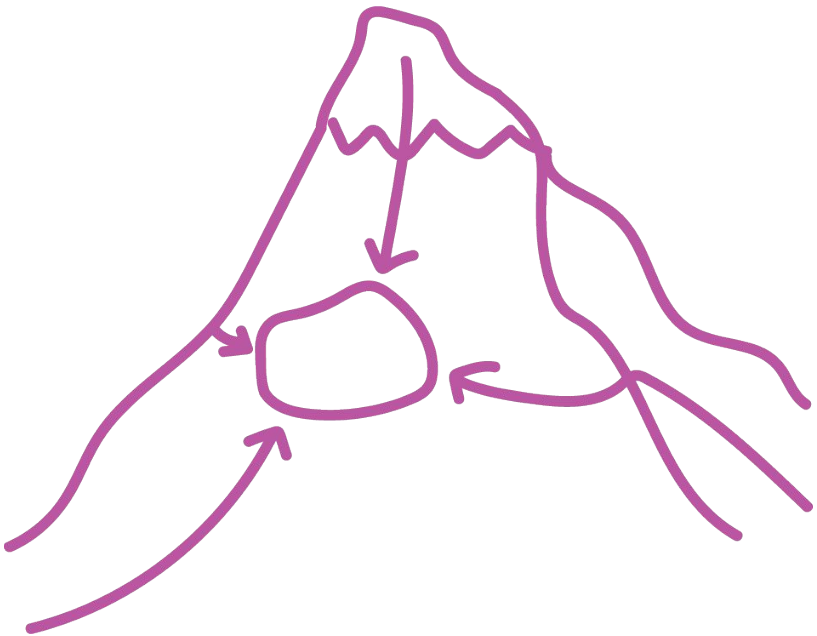


By drawing curves that show where the terrain is the same height, Harold can show three dimensions on a flat map!

Harold is at a point (1,3) on his map. He knows that $f_x(1,3)$ tells him how steep it is going east or west, and that $f_y(1,3)$ is how steep it is north or south. But by differentiating one of partials with respect to the other variable, like $f_{xy}(1,3)$, Harold can tell whether the east-west terrain gets steeper as he moves north or south. Similarly, $f_{yx}(1,3)$ tells Harold if the north-south terrain gets steeper as he moves east or west.



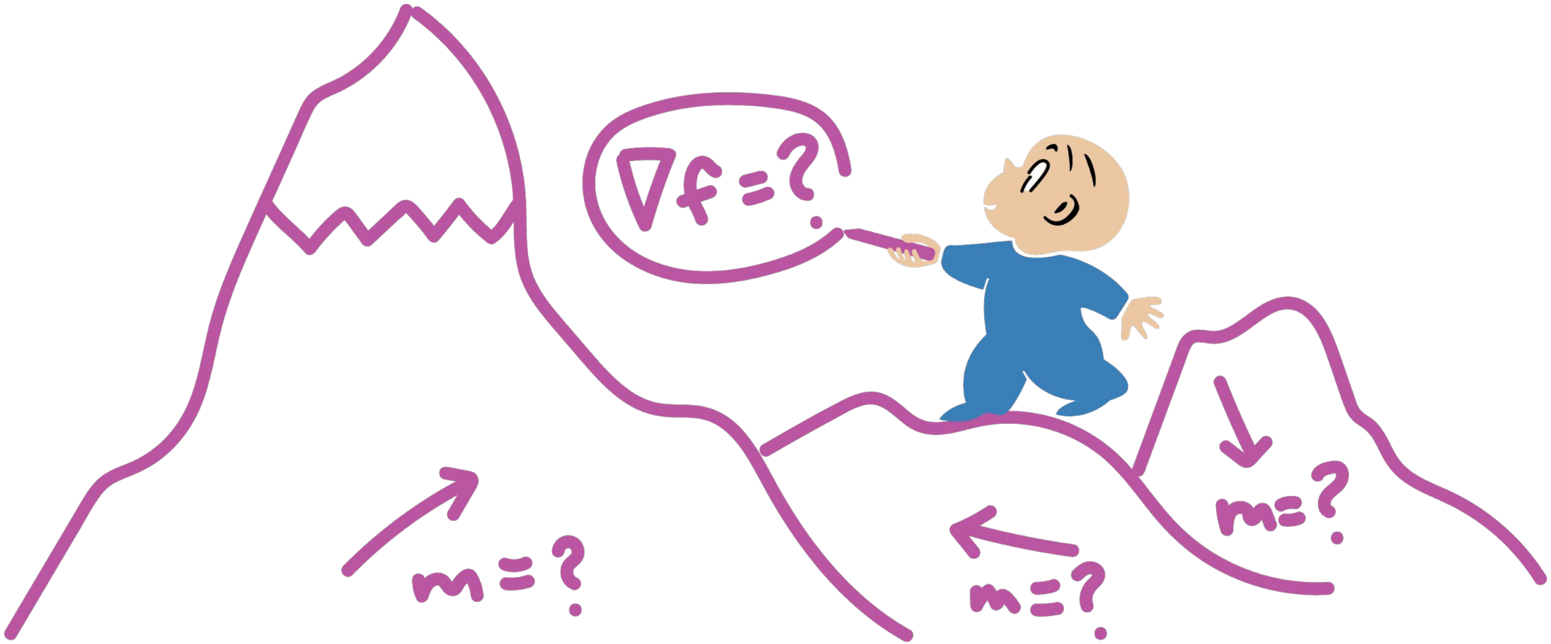
*note 25



While Harold knew that the mountains and terrain he'd drawn were continuous, he thought that other ones might not be. He imagined a mountain defined by $g(x,y)$ with a cave in it, where the the limit at a point differed depending on which side of the cave you arrive. Since the mountains are 3D, there are many ways to approach the cave. Harold could come from below, and have to stop under the cave's maw. Or he could come from the top, and stop before falling down. But the terrain is only continuous if all of these paths yield the same limit.

*note 26

But, Harold wondered, what if I want to know how steep the slope is in another direction, not just north-south or east-west? After all, the mountain has a in every direction, or else it wouldn't be continuous.



*note 27

From *note 27*, Harold also discovered the *gradient* of a function. The gradient of a function *f* is a vector whose components are the partial derivatives of *f*, and in this sense it is the full derivative of *f*. It tells Harold how quickly *f* is changing at a certain point.

when $f:\mathbb{R}^2 \to \mathbb{R}$,

$$\nabla f(\vec{a}) = \langle f_x(\vec{a}), f_y(\vec{a}) \rangle$$

But Harold wants to know how fast the slope is changing *in a certain direction*, so that he knows which routes to recommend when his friends come to the mountain. For this, he needs the *directional derivative*. What follows is one of of Harold's practice problems he did to prepare for mapping the mountain and choosing trails based on the directional derivatives.

*note 28–29

rate of change of
$f(x,y) = xy\sin(x+y)$
at $P=(-1,-1)$
in direction of $\langle 2,3\rangle$

$\vec{u} = \langle \frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}} \rangle$

$f_x(x,y) = y\sin(x+y) + xy\cos(x+y)$

$f_y(x,y) = x\sin(x+y) + xy\cos(x+y)$

$\downarrow$

$f_x(-1,-1) = \cos(-2) - \sin(-2)$

$f_y(-1,-1) = \cos(-2) - \sin(-2)$

$\downarrow$

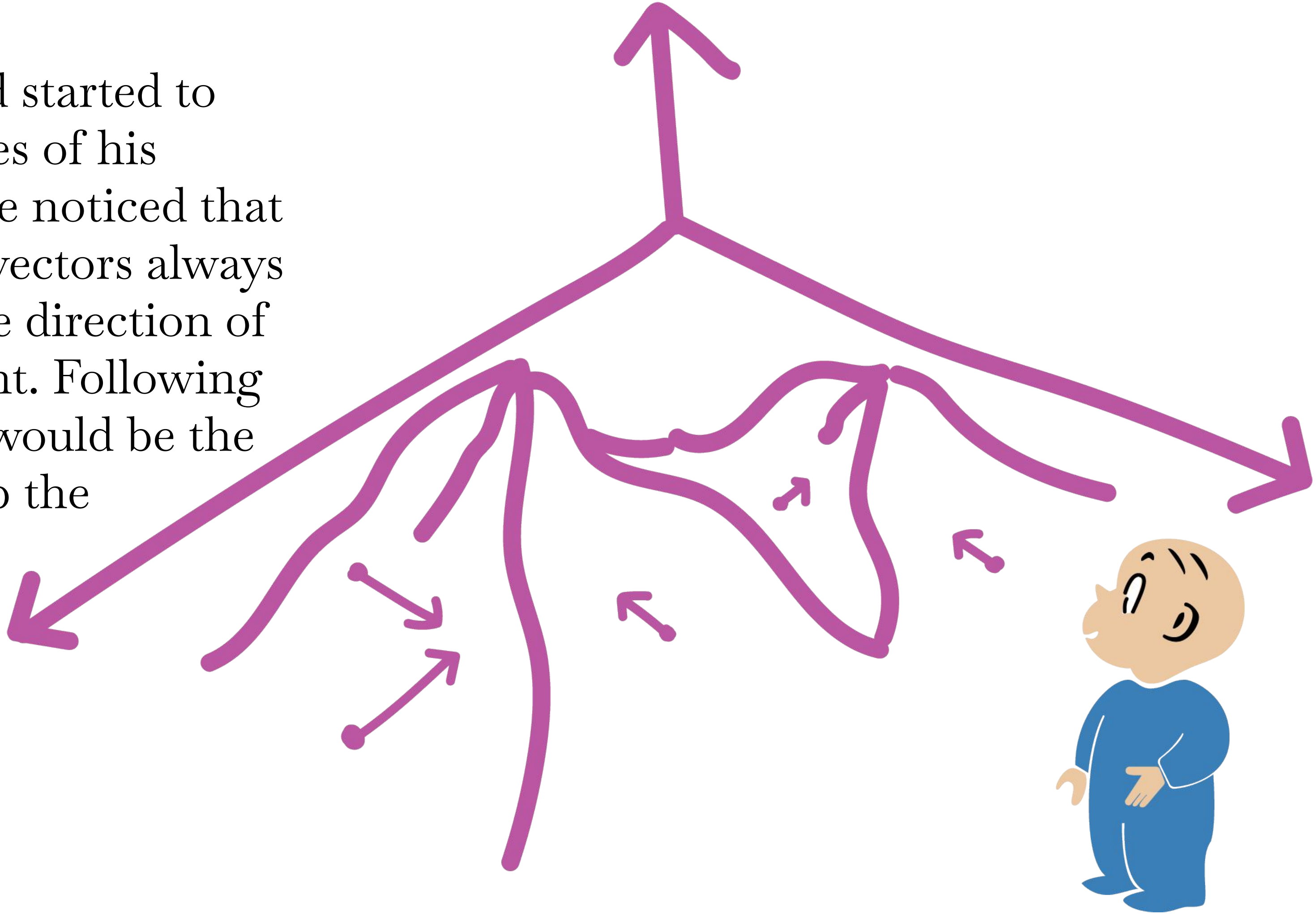
$\nabla f(P) = \langle \cos(-2) - \sin(-2), \cos(-2) - \sin(-2) \rangle$

$$f'_{\vec{u}}(-1,-1) = \nabla f(P) \cdot \vec{u}$$

$$= \langle \frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}} \rangle \cdot \langle \cos(-2) - \sin(-2), \cos(-2) - \sin(-2) \rangle$$

$$\approx 0.684$$

When Harold started to map the slopes of his mountains, he noticed that the gradient vectors always pointed in the direction of steepest ascent. Following the gradient would be the fastest way up the mountains!

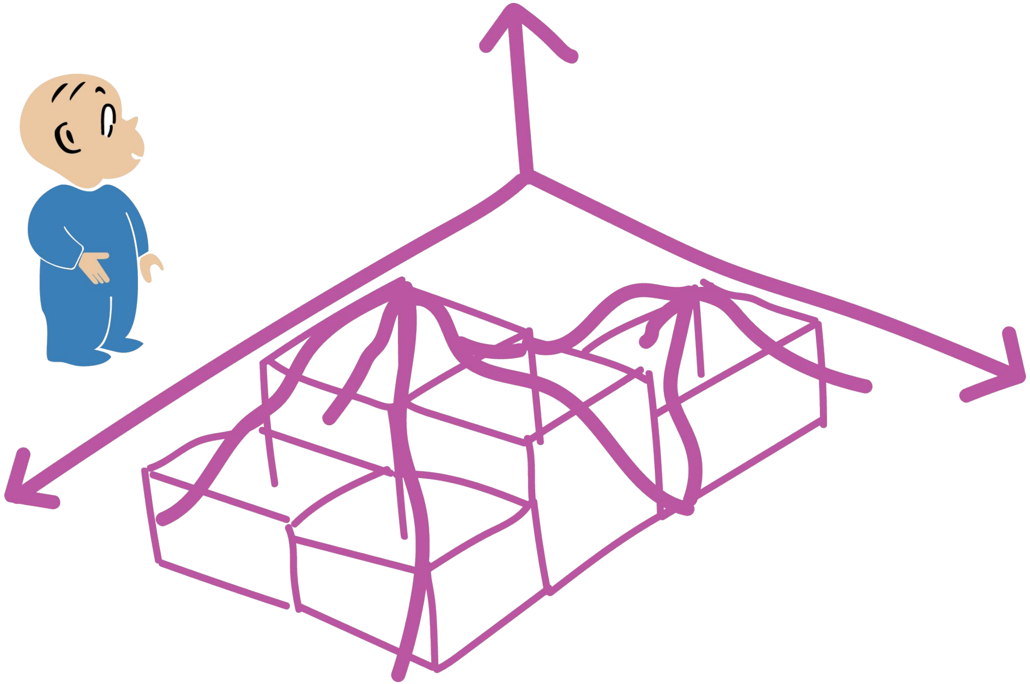


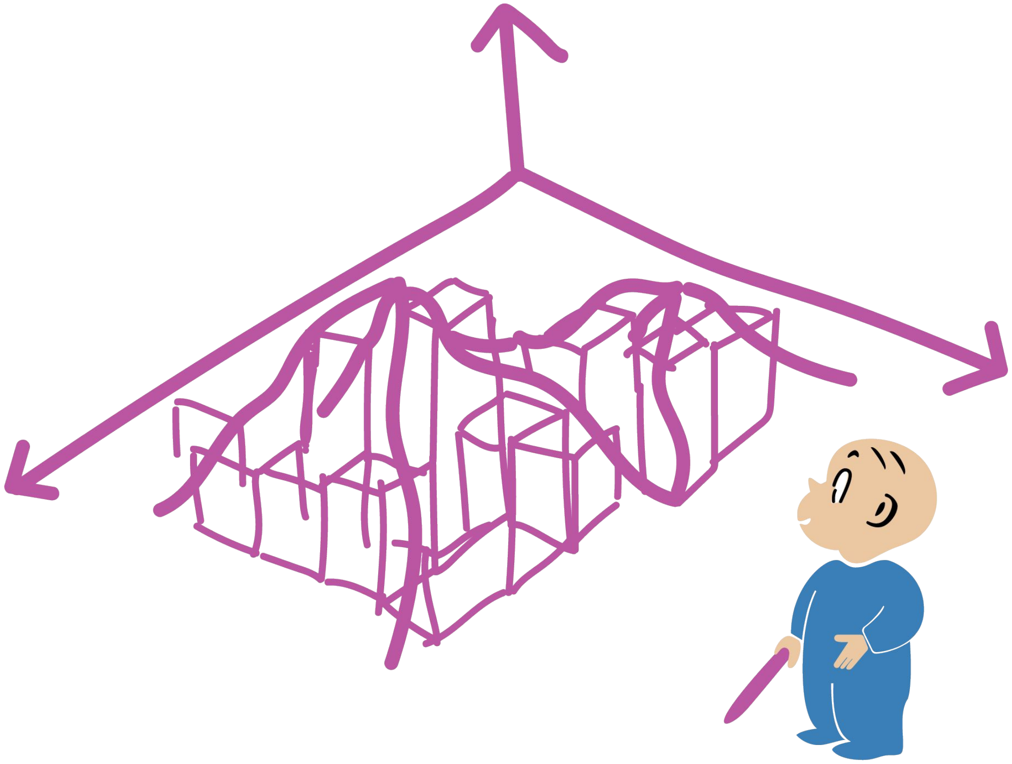
*note 30

After making his map, and even drawing some friends with whom to explore the mountains on foot, Harold was still eager to understand them more.

Harold knew the shape of the mountains like the back of his hand, since he'd drawn them, but still, he wondered, how much rock were they made of?

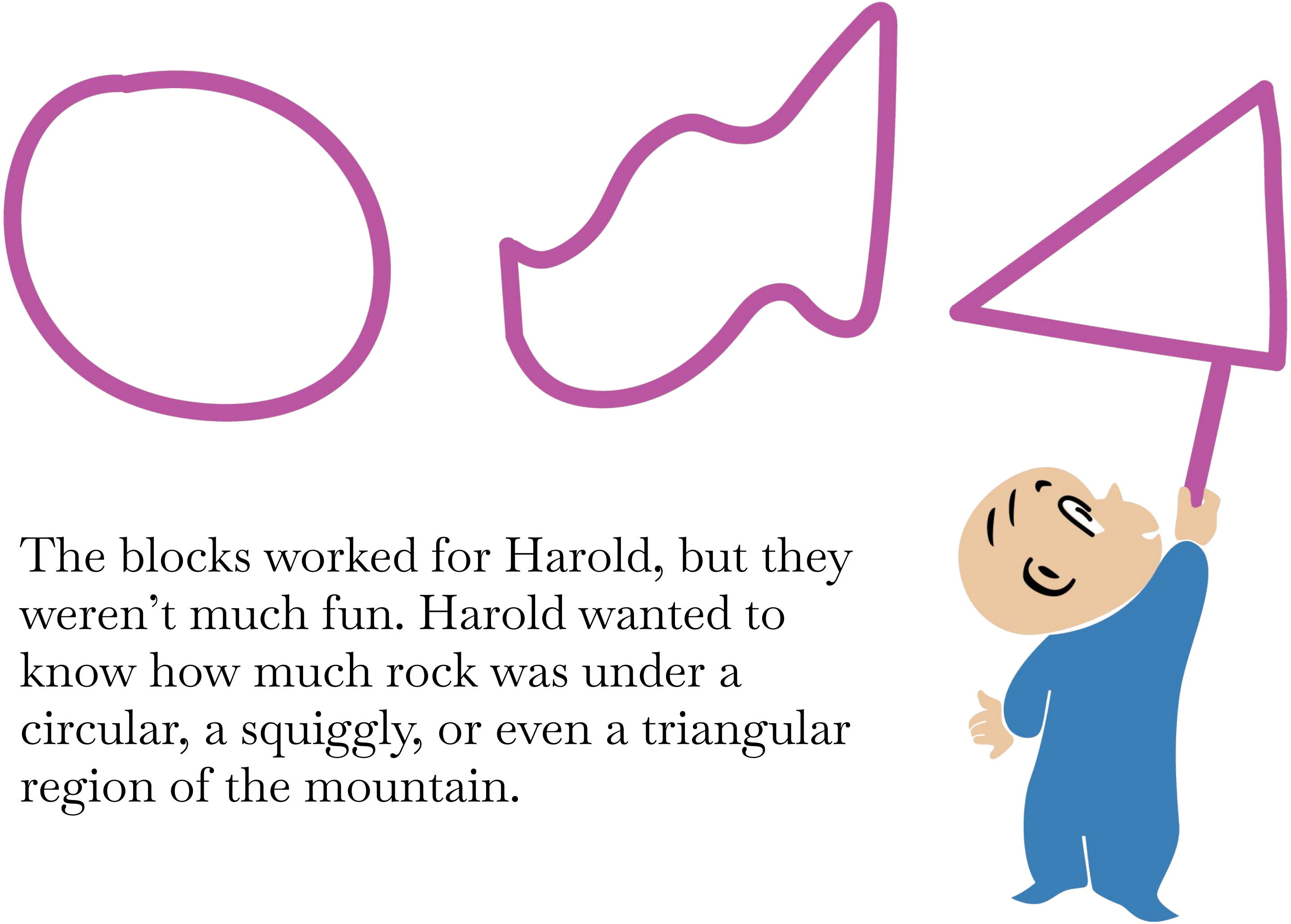
Harold returned to his 3D map to answer this question. First, he thought he would draw some blocks that were around the same size as the mountains. He could use basic arithmetic to find their volumes, just multiplying base by height.





Harold thought that if he broke up—*partitioned*—the mountains into more blocks, he would get a better estimate of the volume. Harold's estimate approached an integral in two dimensions: a *double integral*, perhaps.

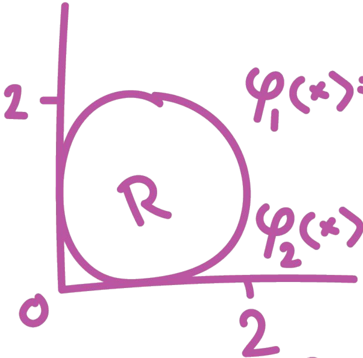
*note 31–33



The blocks worked for Harold, but they weren't much fun. Harold wanted to know how much rock was under a circular, a squiggly, or even a triangular region of the mountain.

*note 34

As practice, Harold tried finding the volume under a circle of a function $f(x,y) = xy$. It was so hard to take one of the integrals that he had to to use a computer. His work follows, and it's not neat!



$$\varphi_1(x) = 1+\sqrt{1-(x-1)^2}$$

$$\varphi_2(x) = 1-\sqrt{1-(x-1)^2}$$

$$\iint_R xy\,dy\,dx = \int_0^2 \left(\int_{1-\sqrt{1-(x-1)^2}}^{1+\sqrt{1-(x-1)^2}} xy\,dy \right) dx$$

$$= \int_0^2 \left[\frac{xy^2}{2} \right]_{y=1-\sqrt{1-(x-1)^2}}^{y=1+\sqrt{1-(x-1)^2}} dx$$

$$G(x) = \int_{1-\sqrt{1-(x-1)^2}}^{1+\sqrt{1-(x-1)^2}} xy\,dy$$

$$G(x) = \frac{x\left(1+2\sqrt{1-(x-1)^2}+(1-(x-1)^2)\right)-x\left(1-2\sqrt{1-(x-1)^2}+(1-(x-1)^2)\right)}{2}$$

$$= \frac{x\left(4\sqrt{1-(x-1)^2}\right)}{2}$$

$$= 2x\sqrt{1-(x-1)^2}$$

$$\int_0^2 G(x)\,dx = \cdots \left[\frac{(x+1)(2x-3)\sqrt{-x(x-2)}}{3} + \arcsin(x-1) \right]_0^2 = \pi$$

notes

1. **Function definition of sequence:** If we look closely at Harold notes, we can see that he has begun to intuit the function definition of a sequence. The sequence of ants he has observed can be thought of as a function which takes in a positive integer, in his case the row's index, and outputs a real number, the number of ants in that row. Formally, a sequence is a real-valued function defined on the set of positive integers (Definition 11.2.1).
2. **Define Natural Numbers:** For ants to march in rows like those which Harold observes, they must be whole, since partial ants are, well, dead. In other words, we can only count ants by *natural numbers*. These numbers are "natural" because they are plainly visible: just look at your fingers, assuming no firework accidents have befallen them.
3. **Definition of limit of a sequence:** In math terms, we say that $\lim_{n\to\infty} a_n = L$ if for each $\epsilon > 0$, there exists a positive integer ***K*** such that if ***n* ≤ *K***, then $|a_n - L| < \epsilon$ (Definition 11.3.1). But this is almost hieroglyphic. In other words, we say that for any small amount (ϵ) from the limit (***L***), there is another small amount (δ) close to the input (***a***) of the function which defines the sequence (***f(x)***) such that when ***x*** is within δ of ***a***, ***f(x)*** is within ϵ of the limit ***L***. Even more simply said, a sequence has a limit if, after a certain point, every term is very close to a number, then that number is the limit of that sequence.

The next page has notes that should be between notes 2 and 3, and I will denote them a and b.

a. **Least Upper Bound:** If we examine the ants Harold has been looking at as the set {1, 2, 3, 4, 5,...}, we can see that there is no *upper bound*. That is, there is no number $\boldsymbol{M}$ that is greater than or equal to every element of the set; we could try picking 9,733, but as Harold knows, the next row will just have 9,734 ants. So we say that the set of ants is unbounded above. But another set, say $A = (-\infty, 0)$, is different. Picking an $\boldsymbol{M}$ that is greater than or equal to every element is easy—just let $\boldsymbol{M}$ = 1. But one could just as easily let $\boldsymbol{M}$ = 893 or 7922; clearly there are infinite possible $\boldsymbol{M}$s. Now let's imagine the set of all upper bounds, which in this case would be $[0, \infty)$. The smallest element of this set, which we term the *least upper bound*, is 0. And this makes sense because 0 is greater than every element of $\boldsymbol{A}$, but if we decrease it by any amount, this will cease to be the case; and so 0 really is an *upper bound* of $\boldsymbol{A}$, and it is also the *least* of these.

b. **Greatest Lower Bound:** Sets, like Harold's set of ants, are not bounded above, but are bounded below. Just as we imagined $\boldsymbol{M}$ to be the number greater than or equal to all elements of a set, we can imagine $\boldsymbol{J}$ to be the number which is *less than or equal to* every element of a set. For Harold's set, $\boldsymbol{J}$ could be 0, or it could be -10, but in either case $\boldsymbol{J}$ is less than every element; we can write the set of these lower bounds as $(-\infty, 1]$. Thus, 1 is the *greatest lower bound* of Harold's set. That is, if it were increased by any amount, it would no longer be a lower bound, and if it were decreased by any amount, it would no longer be the greatest lower bound.

4. Ok, so that's what a sequence which has a limit is defined as. But how would we define a sequence that doesn't have a limit? Enter the negation. When we define something, we state that it is something. We can negate this statement by saying that it is *not* something. Suppose we say "Socrates is a man." Clearly, the negation is "Socrates is not a man." But what about more complicated statements with ifs and thens like the definition of the limit of sequence? Take the statement "If he is rich, then he is happy." The negation of this would be someone who is rich but not happy. So, we say that "He is rich *and* he is *not* happy" is the negation of "If he is rich, then he is happy." Generally, the negations of statements of the form "If A, then B" are "A and not B."
5. Now that we know how to negate statements, we can define when a sequence doesn't have a limit. This will help us understand why Harold's ants are unending. We say that a sequence doesn't have a limit if, after any point early or late in the sequence, the terms are not close to a single number. Furthermore, we term these types of sequences as "divergent."
6. A proof is a claim, which Harold has already, that is supported by statements which are true, and which are connected by logical links. A proof is not always beautiful like the world Harold draws for himself, but a proof always makes sense. For example, suppose we wanted to prove that for any two real numbers r and s where $r > s$, there is always another real number t which lies between them; this means that, no matter how close r and s are, there is always space between them. Our proof might go like this: "Let $a = r-s$. For any real b such that $0 < b < a$, we know that $s + b < r$ since $a + s = r$ and b is smaller than a. Because $b > 0$, we also know that $s < s+b$. Together, these two facts tell us that $s < s+b < r$. Now let $t = s+b$.
Clearly, $s < t < r$, and thus we have proven that there exists a t between r and s."

7. Given the integral $\int_a^b f(x)dx$, and if as b tends to infinity, the integral approaches a finite limit L: $\lim_{b\to\infty} \int_a^b f(x)dx$. We write this alternately as $\int_a^\infty f(x)dx$, and say that the improper integral $\int_a^\infty f(x)dx$ converges to L.

8. When $|x| < 1$, $\sum_{k=0}^\infty x^k = \frac{1}{1-x}$. First, Harold sets $S_n = \sum_{k=0}^n x^k$, and then recognizes that he can rewrite $\sum_{k=0}^n x^k$ as $lim_{n\to\infty} S_n$. But this is just shorthand for $1 + x + x^3 + \cdots + x^n$. Harold then finds $xS_n = x + x^2 + x^3 + \cdots + x^{n+1}$. When Harold subtracts S_n from xS_n, he gets this: $(x + x^2 + x^3 + \cdots + x^{n+1}) - (1 + x + x^2 + \cdots + x^n)$. By cancelling out terms, Harold finds that . $xS_n - S_n = x^{n+1} - 1$. Solving for S_n, Harold gets $S_n = \frac{1-x^{n+1}}{1-x}$. At last, taking the limit of S_n as $n \to \infty$, the numerator shrinks to zero (since $|x| < 1$), and so $\lim_{n\to\infty} S_n = \frac{1}{1-x}$. This is, Harold defined above, the same as $\sum_{k=0}^\infty x^k = \frac{1}{1-x}$.

9. Imagine the series $\sum_{k=0}^\infty (-1)^k$, which we write as $s_n = (1) + (-1) + (1) + \cdots + (-1)^n$. Clearly, even when we let n be a very large number, the series does not approach a single value; rather, when n is even, $s_n = 1$ and when n is odd, $s_n = 0$.

10. The kth term of a convergent series tends to zero, or more specifically, if $\sum_{k=0}^\infty a_k$ converges, then as $k \to \infty$, $a_k \to 0$. For example, let $a_k = \frac{1}{2^k}$, which we know converges; also, we can see that for a large value of k, a_k is very small. The converse of this theorem is also very useful. If $a_k \not\to 0$ then $\sum_{k=0}^\infty a_k$ diverges. For example, suppose $a_k = \frac{k}{k+1}$. As k approaches infinity, we can see that a_k approaches 1, since the difference between numerator and denominator becomes proportionally irrelevant. Since $a_k \not\to 0$, we know that $\sum_{k=0}^\infty \frac{k}{k+1}$ diverges (Theorem 12.2.5–6).

11. **Integral Test:** If f is continuous, positive, and decreasing on $[1,\infty)$, then $\sum_{k=1}^{\infty} f(k)$ converges if and only if $\int_1^{\infty} f(x)dx$ converges. See "Additional Notes" for proof.
12. **Limit Comparison Test:** Let $\sum a_k$ and $\sum b_k$ be series with positive terms. If $\lim_{k\to\infty} \frac{a_k}{b_k} = L$ and L is positive, then $\sum a_k$ converges if and only if $\sum b_k$ converges.
13. **Absolute Convergence:** If $\sum |a_k|$ converges, then $\sum a_k$ converges. We say $\sum a_k$ for which this is true are absolutely convergent.
14. **Conditional Convergence:** Conditional convergence is the negation of absolute convergence, and so according to note 4, all we have to do is write "A and not B" where there was "If A, then B." Thus, if $\sum a_k$ converges and $\sum |a_k|$ diverges, $\sum a_k$ is conditionally convergent.
15. **Perpendicular Vectors:** Two vectors $\vec{a}$ and $\vec{b}$ are perpendicular if and only if $\vec{a} \cdot \vec{b} = 0$, since $\vec{a} \cdot \vec{b} = \|\vec{a}\|\|\vec{b}\| \cos\theta$. When $\theta = \frac{\pi}{2}$, $\vec{a} \cdot \vec{b} = 0$ (Definition 13.3.7–8).
16. **Adding Vectors:** Let $\vec{a} = \langle a_1, a_2, a_3 \rangle$ and $\vec{b} = \langle b_1, b_2, b_3 \rangle$. We define $\vec{a} + \vec{b}$ as $\langle a_1 + b_1, a_2 + b_2, a_3 + b_3 \rangle$ (Definition 13.2.1).
17. **Scalar Multiplication:** Let $\vec{a} = \langle a_1, a_2, a_3 \rangle$ and $\alpha \in \mathbb{R}$. We define $\alpha\vec{a}$ as $\langle \alpha a_1, \alpha a_2, \alpha a_3 \rangle$ (Definition 13.2.2).
18. **Distance Formula:** The magnitude of a vector $\|\vec{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2}$ (Definition 13.2.3).
19. **Definition of a line:** Given $\vec{a} = \langle a_1, a_2, a_3 \rangle$ and a point $P = (x_0, y_0, z_0)$, we can define a line as the set of points which fulfill these parametric equations: $x(t) = x_0 + a_1 t$, $y(t) = y_0 + a_2 t$, and $z(t) = z_0 + a_3 t$ (Definition 13.5.3).

20. **Definition of a plane:** Given $\vec{a} = \langle a_1, a_2, a_3 \rangle$, a point $P = (x_0, y_0, z_0)$ and an arbitrary point $Q = (x, y, z)$, we can define a plane $\boldsymbol{\rho}$ as the collection of points which fulfill $a_1(x - x_0) + a_2(y - y_0) + a_3(z - z_0) = 0$.

21. When we say "distance" we really mean "shortest distance." First, imagine a line ℓ and a point off to one side of it $\boldsymbol{P_1}$. We'll call the distance from ℓ to $\boldsymbol{P_1}$ "$\boldsymbol{d}$". Now imagine the point $\boldsymbol{Q}$ on ℓ which forms the vector $\overrightarrow{P_1Q}$ such that $\overrightarrow{P_1Q}$ is perpendicular to ℓ. Next, we choose an arbitrary point $\boldsymbol{P_o}$ somewhere on ℓ, and draw the vectors $\overrightarrow{P_0P_1}$ and $\overrightarrow{P_0Q}$, which is parallel to ℓ from $\boldsymbol{P_o}$ to $\boldsymbol{Q}$. All we need to do to find $\boldsymbol{d}$ is to find the distance from $\boldsymbol{Q}$ to $\boldsymbol{P_1}$, which we can do with trigonometry. Letting $\boldsymbol{\theta}$ be the angle between $\overrightarrow{P_0P_1}$ and $\overrightarrow{P_0Q}$, we can see that $\sin\theta = d(P_1, Q) \div \|\overrightarrow{P_0P_1}\|$. Since we stated that $d(P_1, Q) = d(P_1, \ell)$, we can now say that $d(P_1, \ell) = \|\overrightarrow{P_0P_1}\| \sin\theta$. To solve this, all we have to do is calculate the magnitude of $\overrightarrow{P_0P_1}$, which we know from note 18, and to find $\sin\theta$, which we can do by rearranging the equation from note 15.

22. In note 19 we defined a line as the set of points which fulfill three parametric equations, and thus all points on some line ℓ have the form $(x(t), y(t), z(t))$. We can therefore define a line ℓ_0 as the function $f_0(t) = (x(t), y(t), z(t))$, and another line ℓ_1 as $f_1(t) = (u(t), v(t), w(t))$. Given that we have defined lines as collections of points which fulfill certain conditions, conceptually, all it means for two lines to intersect is for them to share a point. In other words, ℓ_0 and ℓ_1 intersect if and only if there exist real $\boldsymbol{t_o}$ and $\boldsymbol{t_1}$ such that $f_0(t_0) = f_1(t_1)$. If we expand this with the coordinates of the points which constitute two two lines, we say $(x(t_0), y(t_0), z(t_0)) = (u(t_1), v(t_1), w(t_1))$; all coordinates have to match at a single point for there to be an intersection.

23. **Partial Derivatives:** Let f be a function of two variables x and y. The partial derivatives of f with respect to x and with respect to y are the functions f_x and f_y defined by setting $f_x(x,y) = \lim_{h\to 0} \frac{f(x+h,y)-f(x,y)}{h}$ and $f_y(x,y) = \lim_{h\to 0} \frac{f(x,y+h)-f(x,y)}{h}$, provided that these limits exist. Geometrically, this is equivalent to the slope in either the x or y direction at a given cross-section ($x = x_0$ or $y = y_0$) of the surface.

24. **Level Curves:** Given a function $f : \mathbb{R}^2 \to \mathbb{R}$and some real z, the level curve at height c is the collection of points $\{(x,y) : f(x,y) = c\}$. While this can seem like some multidimensional mumbo-jumbo, it's not too hard to understand when we think of a level curve like an altitude line on a topographical map. Just taken individually, a level curve shows where the function is a certain height (it is true that the idea of "height" becomes less clear when the function is defined from $\mathbb{R}^3 \to \mathbb{R}$), but when we look at many level curves all spaced evenly, e.g. $c = \{1, 1.2, 1.4, 1.6, \ldots\}$ they can visualize the shape of the function quite well. When two level curves are close to each other, it means that the function increased over a short distance, which means the function is steeper in that region. Likewise, distant level curves indicate flatness.

25. **Equality of Mixed Partials:** $f_{xy}(x,y) = f_{yx}(x,y)$. Let's demonstrate this with an example where $f(x,y) = x^2 \cos y$.. First we have to get the first-order partial derivatives $f_x(x,y)$and $f_y(x,y)$. Treating y as a constant, we get $f_x(x,y) = 2x\cos y$. Treating x as a constant, we get $f_y(x,y) = -x^2 \sin y$. To find $f_{xy}(x,y)$, we must now differentiate $f_x(x,y)$ with respect to y: treating x as a constant, $(f_x(x,y))_y = f_{xy}(x,y) = -2x\sin y$. Similarly, differentiating $f_y(x,y)$ with respect to x we get $(f_y(x,y))_x = f_{yx}(x,y) = -2x\sin y$.
As we can see, since $-2x\sin y = -2x\sin y$, $f_{xy}(x,y) = f_{yx}(x,y)$.

26. **Multivariable Continuity:** A function ***f*** of several variables is continuous at a point $c = (c_1, c_2, c_3, \ldots)$ if and only if the limit exists along every path, and the limit of all paths agree. For example, $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \frac{x^2 y}{x^4+y^2}$. If $\lim_{(x,y)\to(0,0)} f(x,y)$ exists, we should be able to obtain the same limit even as we approach (0,0) with different paths, just like how for a single-variable function, the limit must be the same as it is approached from both the positive and negative side for it to exist. So, we will approach (0,0) along two paths, ***y = x*** and ***y = x²***. When ***y = x***, we have $f(x,y) = f(x,x)$ which is $\frac{x^3}{x^4+x^2}$, which simplifies to $\frac{x}{x^2+1}$. Next, $\lim_{x\to 0} \frac{x}{x^2+1} = 0$ When ***y = x²***, $f(x,y) = f(x,x^2)$, which is $x^4 \div 2x^4$. Thus, $\lim_{x\to 0} \frac{x^4}{2x^4} = \frac{1}{2}$. Clearly, as we approach (0,0) with different paths, we find different results, and so therefore ***f*** does not have a limit at (0,0).

27. **Gradient Definition of Differentiability:** A function ***f*** is differentiable at ***a*** if and only if there exists a vector $\vec{y}$ such that $f(\vec{a}+\vec{h}) - f(\vec{a}) = \vec{y}\cdot\vec{h} + o(\vec{h})$. If such a vector $\vec{y}$ exists, we say that the gradient $\nabla f(\vec{a}) = \vec{y}$.

28. **Limit Definition of Directional Derivative:** Given a function ***f*** and a unit vector $\vec{u}$, the directional derivative of ***f*** in the direction of $\vec{u}$ at $\vec{a}$ is $f'_{\vec{u}}(\vec{a}) = \lim_{h\to 0} \frac{f(\vec{a}+h\vec{u}) - f(\vec{a})}{h}$, provided the limit exists.

29. **Gradient Definition of Directional Derivative:** If ***f*** is differentiable at $\vec{a}$, then the directional derivative of ***f*** in the direction of any unit vector exists, and $f'_{\vec{u}}(\vec{a}) = \nabla f(\vec{a})\cdot\vec{u}$. Proof that this is equivalent to the limit definition is provided in "Additional Notes"

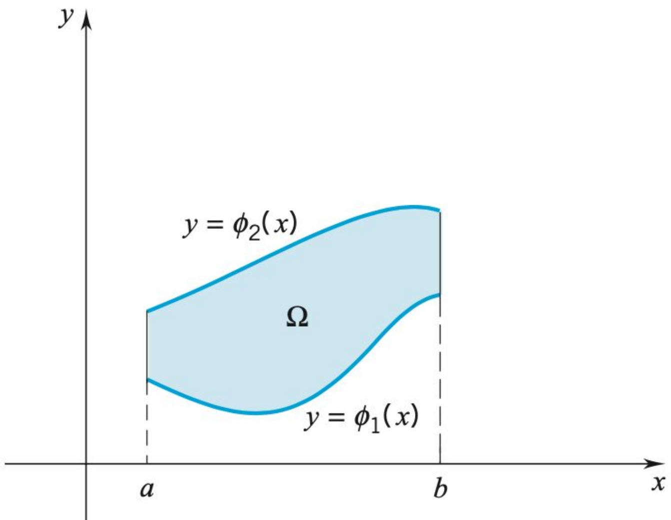
30. **Gradient Direction is Steepest Slope:** A differentiable function ***f*** at a point $\vec{a}$ increases most rapidly in the direction of $\nabla f(\vec{a})$This implies that ***f*** decreases most rapidly in the direction opposite $\nabla f(\vec{a})$, since what is steep uphill is also steep downhill, trivially. And so the gradient at a particular point indicates which direction is the "path of most resistance," so to speak.

31. **Upper Sum:** Given the partition $\boldsymbol{P}$ of a rectangle $\boldsymbol{R}$, the upper sum $U_f(P) = \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \cdot (\text{area of } R_{ij})$ where $\boldsymbol{M_{ij}}$ is the maximum of $\boldsymbol{f}$ within one of the sub-rectangles defined by the partition. We can think of $\boldsymbol{M_{ij}}$ as the "high point" of the portion of the surface under which we are finding the volume. But we should clarify what is meant by "partition." The partition $\boldsymbol{P_1}$ of an interval $[\boldsymbol{a},\boldsymbol{b}]$ is the set $\{x_0, x_1, x_2, \ldots, x_i\}$ such that $a = x_0 \leq x_1 \leq \cdots \leq x_{n-1} \leq x_n = b$. Now let's partition the interval $[\boldsymbol{c},\boldsymbol{d}]$ with another set $P_2 = \{y_0, y_1, y_2, \ldots, y_j\}$. Each partition we just created divides an interval into smaller intervals, but how would we go about dividing a rectangle into smaller rectangles? Why, by taking the product of two partitions of intervals, which is how we define the partition $\boldsymbol{P}$ of our region of integration $\boldsymbol{R}$ as $P = P_1 \times P_2$.

32. **Lower Sum:** With the same partition $\boldsymbol{P}$ of the same rectangular region $\boldsymbol{R}$ as in note 31, the lower sum $L_f(P) = \sum_{i=1}^{m} \sum_{j=1}^{n} m_{ij} \cdot (\text{area of } R_{ij})$ where $\boldsymbol{m_{ij}}$ is the minimum of $\boldsymbol{f}$ within one of the sub-rectangles defined by the partition. We can think of this as the "low point" of the portion of the surface under which we are finding the volume.

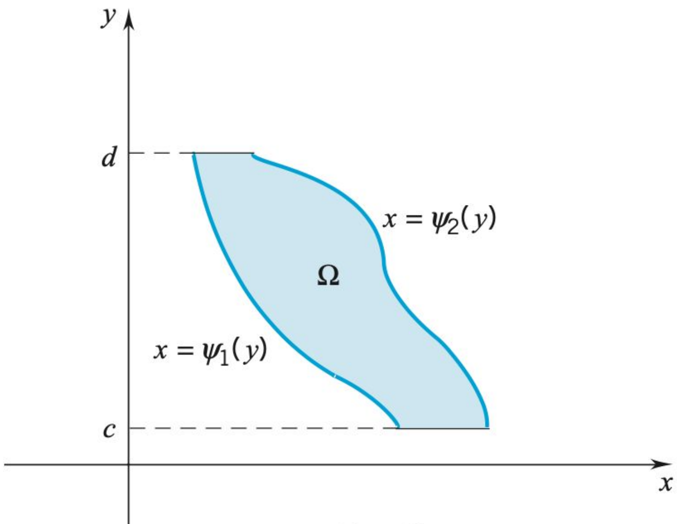
33. **Double Integrals:** The upper and lower sums give us a good idea about how large we expect the volume under a particular region of a surface to be, but they are just the extremes, not what the volume actually is. But, because they are by definition the highest and lowest the values the volume could be, we know that the actual volume must be between them. Like how with a single variable function we might use left and right Riemann sums to approximate the integral, we are doing the same thing here. So, the double integral, which we can think of as the volume under a surface, is the unique number which, no matter how many pieces we partition $\boldsymbol{R}$ into, always lies between the upper and lower sums. Mathematically, we say $L_f(P) \leq \iint_R f(x,y)\, dydx \leq U_f(P)$ for all partitions $\boldsymbol{P}$. Furthermore, we say that if $\boldsymbol{f}$ is continuous on $R = [a,b] \times [c,d]$, then $\iint_R f(x,y)\, dydx = \int_a^b \int_c^d f(x,y)\, dydx$

34. **Non-Rectangular Regions:** Suppose that region of integration $\boldsymbol{R}$ (Ω in the figure below) is defined by $[\boldsymbol{a},\boldsymbol{b}]$ and two curves $y = \phi_1(x)$ and $y = \phi_2(x)$. Thus, $\iint_R f(x,y)\,dydx = \int_a^b \int_{\phi_1(x)}^{\phi_2(x)} f(x,y)\,dydx$. In the figure below, this is a "Type I Region." We can also define $\boldsymbol{R}$ with respect to $\boldsymbol{y}$ instead of $\boldsymbol{x}$: $\iint_R f(x,y)\,dxdy = \int_c^d \int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\,dydx$, which we call a "Type II Region."



Type I

ϕ_1, ϕ_2 continuous



Type II

ψ_1, ψ_2 continuous

addtional notes

Proof of the Integral Test

Lemma 1: Let a_n be an increasing sequence which is bounded above. Then a_n converges to the least upper bound of $\{a_1, a_2, a_3, \ldots\}$.
Lemma 2: If f is continuous, positive, and decreasing on $[1,\infty)$, then $\int_1^\infty f(x)dx$ converges if and only if the sequence $a_n = \int_1^n f(x)dx$ converges.
Note: M is a natural number that is greater than or equal to all terms s_n for all n.

"⇒" Suppose $\sum_{k=1}^\infty f(k)$ converges. Then, for all n since $f(k)$ is positive for all k, $\int_1^n f(x)dx \le \sum_{k=1}^{n-1} f(k) \le \sum_{k=1}^\infty f(k)$. Since f is positive, $\int_1^n f(x)dx$ is an increasing sequence. Therefore, by Lemma 1, $\int_1^n f(x)dx$ converges. By Lemma 2, this implies that $\int_1^\infty f(x)dx$ converges, since f is continuous, positive, and decreasing.

"⇐" Suppose $\int_1^\infty f(x)dx$ converges. Then, for all n, $\sum_{k=2}^n f(k) \le \int_1^n f(x)dx \le \int_1^\infty f(x)dx = M$. Since $f(x)$ is positive, $\sum_{k=2}^n f(k)$ is increasing and bounded above. So by Lemma 1, $\sum_{k=2}^n f(k)$ converges. So, $\sum_{k=1}^\infty f(k)$ converges. Therefore $\sum_{k=1}^\infty f(k)$ converges if and only if $\int_1^\infty f(x)dx$ converges.

We have thereby proven both directions, and thus the integral test is sound.

Proof that the limit and gradient definitions of the directional derivative are interchangeable

Proof. Because we stated that function f is differentiable at vector $\vec{a}$, we know that the gradient $\nabla f(\vec{a})$ exists and

$$f(\vec{a}+h\vec{u}) - f(\vec{a}) = \nabla f(\vec{a}) \cdot h\vec{u} + o(h\vec{u})$$

Dividing by h, we get

$$\frac{f(\vec{a}+h\vec{u})-f(\vec{a})}{h} = \nabla f(\vec{a}) \cdot \vec{u} + \frac{o(h\vec{u})}{h}$$

Since $\frac{o(h\vec{u})}{h} \to 0$, therefore

$$\frac{f(\vec{a}+h\vec{u})-f(\vec{a})}{h} \to \nabla f(\vec{a}) \cdot \vec{u}$$