How to draw a cube on a flat sheet of paper

Jack Sarick

Contents

1 Introduction

We live in a world of three physical dimensions[1], yet we view it almost entirely in two dimensions. Our screens have length and width but no depth. Despite having 50% less dimensions, screens still convey complex three dimensional images, like videogames or movies. This is a fact most of us take for granted in our day to day lives, but it is not a simple miracle to perform. In the case of videogames, and any other three dimensional model viewed on a two dimensional screen, this is done via projection. As an avid connoisseur of videogames, this is something I've thought about for a while, and find it a fundamentally interesting concept.

1.1 What is projection?

Projection isn't a particularly hard idea by itself. It is to take a set of data (typically numbers) and map it into a subset. Mapping is just taking one set of data and uniformly transforming it into another set. For example, the set $(1, 2, 3, 4)$ can be mapped to $(1, 4, 9, 16)$ using the function $f(x) = x*x$. Another case would be $(1, 2, 3, 4) \rightarrow (1)$ using $f(x) = 1$. Projection is a specific case of mapping. In practical terms, that means for a cube that has been projected onto a plane, every point on the cube has a corresponding point on the plane.

[1]For all intents and purposes of this paper, time does not exist, and will be utterly ignored

There are, typically, two kinds of projection when dealing with viewing objects in higher dimensions (e.g. viewing a cube on a flat screen): orthographic and isometric. Orthographic is when the viewer is a single point, like a camera taking a picture of a painting[2]. This is the more common in today's day and age, as it feels more natural because it is how we as humans see. The other kind is isometric, which is when the viewer is not a single point. A rough example would be a photocopier, as it is able to take a picture of plane. Another way of thinking of the difference between the two is that for orthographic projection, the projection converges on a single point at the viewer, and for isometric the projection runs perpendicular to the viewer.

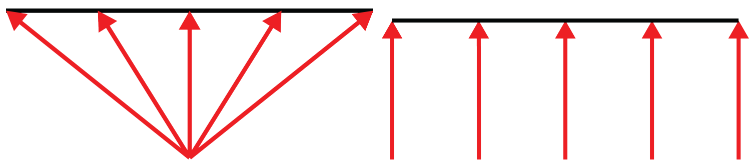


Figure 1: Rays signify perspective lines, black line represents the cross-sectional portion of the world being viewed. Orthographic is on the left, isometric pictured right.

This diagram doesn't fully represent the projections. For example, the orthographic diagram pictures all the rays originating from a single point, but the astute reader will note that when they take a picture, the picture most certainly is not a single point. In the case of real world application, the viewport (what the final product actually looks like) would actually be a small cross section of those rays just before that converging point[3]. This will discussed in more detail later in the paper.

2 Projection of 2 dimensions to 1

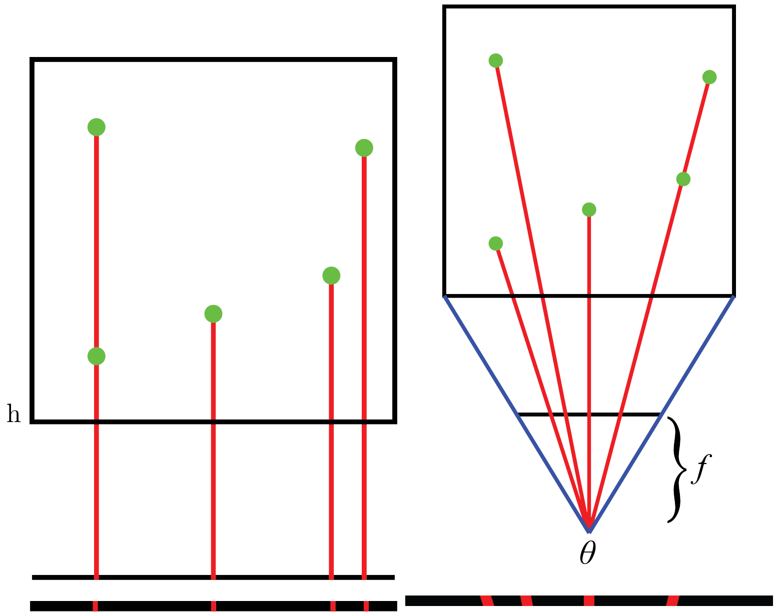
To start out simple, let's reduce the amount of dimensions to something more manageable. It will allow for simpler math, and greater insight in the future.

2.1 Visualization

Imagine a flat whiteboard with a bunch of flat magnets of different shapes and sizes on it. If you view the white board from the side, on same plane as it, all that can be seen is the sides of the magnets as little line segments.

[2]Not really because the sensor of a camera is still a plain, but there is still a focal point so for most intents and purposes this example stands up

[3]Or in the case of human eyes and some telescopes, after the point. This causes the image to be upside down, and requires it to be flipped. In the case of telescopes, this is done with a smaller mirror, and with eyes it is flipped by our brain (To be more precise, the optical nerve basically plugs into the brain upside down, so that the image arrives an upside down version of the already upside image. This is actually a very interesting process and if I took bio I'd probably write an IA on this. Alas, this is a math paper.)



(a) The diagram on top illustrates what isometric projection looks like. The bottom shows what the projection would approximately look like.

(b) The diagram on top illustrates what orthographic projection looks like, and the purpose of angle of view, and focal length in an orthographic projection. The bottom shows what the projection would approximately look like.

Figure 2: Comparison and visualization of orthographic and isometric projections

2.2 Mechanics

There are two main components in play here. The projection set, where all the points ”exist” and the projected set, where the transformed points are. In the case of 2D $\rightarrow$ 1D, the projection set is a plane, and the projected set is a line[4]. To reference a previous diagram, 1, the projection set is the black line, and the projected set is where the red rays originate from.

2.3 Equations

2.3.1 Isometric

Isometric projections are relatively straightforward, in a very literal sense. For any given projection set, the function to convert a 2D point to a 1D one would be[5]

$$p(x, y) = x$$

[4] In almost all practical purposes it is a line segment, because infinitely long lines take a while to calculate.

[5] I am using p as the function because f will be a variable later, and p makes much more sense as the name of a *projection* function.

Quite simple, but it makes sense. Simply dropping a dimension was the whole purpose to begin with, so why complicate the process? In the real world, where the size of the view port is limited, the equation gets a slight improvement.

$$p(x,y) = \begin{cases} x, & \text{if } x_1 < x < x_2 \\ undefined, & \text{otherwise} \end{cases}$$

Where x_1 and x_2 are the bounds, this simply says that if the point falls outside the view port, it cannot be seen. Not a particularly complicated idea, but a necessary one.

2.3.2 Orthographic

Orthographic is substantially more complex. The viewer is a single point, but the projected set is still a line segment, so there are a couple more variables needed: angle of view (θ) and focal length (f).

As seen in 2b, angle of view and focal length are relatively straight forward. To put it in non math terms, angle of view is how wide the view is, and focal length is how far away the observer is from the area that is being projected on. Calculations become a little harder when these new variables are introduced. It is easier to start from a completed equation and work backwards to understand it.

$$p(x,y) = \begin{cases} f\frac{x}{y+f}, & \text{if } |f\frac{x}{y+f}| < f\tan\frac{\theta}{2} \\ undefined, & \text{otherwise} \end{cases}$$

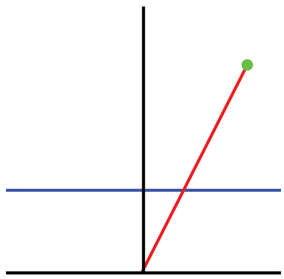


Figure 3: With the observer at the origin, the blue line is $y = f$ and the green point is the point being projected.

The equation can be seen as two smaller equations: $f\frac{x}{y+f}$ and $f\tan\theta$. Each does a separate task, and combine to form a single equation. The first one is the one that actually calculates the projected point. Only the x value of the point is needed. As shown in , this x value is the intersection of a line formed between the point and the observer, and the focal point.

There are two ways to approach this solution, either the observer or the focal point is at the origin. The former means that the focal line is at $y = f$, but the y value of all points must be shifted up f. In the latter, all points remain where they are, the focal line is at $y = 0$, but the origin is at $(0, -f)$. Following the first method[6], the point exists at the intersection of $f = y$ and a line between $(0,0)$ and $(x, y+f)$. For all lines that pass through the origin, $y = mx$. Slope is equal to $\frac{y_2-y_1}{x_2-x_1}$, but the first point is always $(0,0)$, so it can simply be reduced to $\frac{y+f}{x}$. Given slope $\frac{y+f}{x}$ and a $y = f$, putting the two together gets $f = x_p\frac{y+f}{x}$ where x_p is the projected point. This can be rearranged into $x_p = f\frac{x}{y+f}$. Taking the product of f, and the quotient of x and

[6] I've gone with the first method because it makes more sense to me, but they get the same result.

$y + f$ results in a single value that represents how far away the point appears to be from the observer.

To determine if the point is in view is straightforward once the value is known. The width of the field of view is equal to the base of the triangle formed by the observer and the angle of view, and a height equal to the focal length. Bisecting along its height makes the math easier, yielding a right angle triangle with one side as f, and an angle of $f\frac{\theta}{2}$ adjacent to that side. To find the base of this new triangle it would be $\tan\frac{\theta}{2} = \frac{f}{base}$, which simplifies to $f \tan\frac{\theta}{2} = base$. The base of this triangle is equal to half the width of the projected line. As long as the magnitude of the projected point (x_p) is less than half the width of v line, then it is on the line[7]. In equation form, this is $|f\frac{x}{y+f}| < f\tan\frac{\theta}{2}$

Once both halves of the equation have been created, putting them together is as simple as writing them together and using the correct notation.

$$p(x, y) = \begin{cases} f\frac{x}{y+f}, & \text{if } |f\frac{x}{y+f}| < f\tan\frac{\theta}{2} \\ undefined, & \text{otherwise} \end{cases}$$

3 Projection of 3 dimensions to 2

Stepping up a dimension is not quite as easy as it sounds. There is a good reason that highschools almost exclusively teach math in two dimensions or less and the vast majority of people never get beyond that. It is only one more dimension, but it is arguably three times harder. Instead of dealing with the single relationship of x and y, there are now all three relationships between x, y and z[8]. This price hangs on great power, however. If a single dimension can be added, then another one can be as well, and another and another and another ad infinitum. Once the code is cracked, it remains cracked for good. With that said, let's add another dimension

3.1 Visualization

This is, oddly enough, easier to visualize because we live in three dimensions and see in two on a regular basis. Though it is easier to visualize, it also becomes substantially harder to put into a paper because papers are two dimensional. That means that from here on, I can no longer represent the projection set. To do so, I would have to project the values onto the two dimensional plane of the paper. As such, I simply cannot display the projection set[9].

[7] This is because half the line is negative, and half of all numbers are negative. If you halve the amount of numbers x can be, you must also halve the amount of numbers that can be on the projected

[8] This is an opinion, but still a valid concern. Anyone who says that three dimensional math is easy doesn't spend enough time actually living in three dimensions

[9] As an aside, this is why I believe that handing in an actual paper is ridiculous. While writing this on my computer, I spun up a couple models that were interactive. They were still on a 2D screen, but you could move them, and adjust them, and play with them. These kinds of things are only possible with modern technology, and have no comparable counterparts in the physical world.

3.2 Equations

3.2.1 Isometric

Isometric projection remains almost entirely the same.

$$p\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ y \end{bmatrix}$$

A simple equation, but quite an illuminating one.[10] It shows that there will in fact be a pair of coordinates in the end, as opposed to a single number seen previously.

3.2.2 Orthographic

Orthographic can be a bit daunting. The previous equation took two inputs and returned one. Given three the desired result would be two, though it's not as simple as just dropping one like isometric. x, y, and z all change one another, so it's not particularly simple. Breaking the projected surface into two axis helps to clarify the situation. Using the previous orthographic 2D → 1D function as p_{2d} it can be seen that $x = p_{2d}(x, z)$ and $y = p_{2d}(y, z)$. Putting them together yields

$$p\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} p_{2d}(x, z) \\ p_{2d}(y, z) \end{bmatrix}$$

Before trying this out, this can be double checked against the isometric function set. Taking the original isometric 2D → 1D function and doing the same same process provides the exact same results. It is also simple enough that it can be solved by hand without effort.

$$p\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} p_{2d}(x, z) \\ p_{2d}(y, z) \end{bmatrix}$$

$$p\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} p_{2d}(x, z) \rightarrow x \\ p_{2d}(y, z) \rightarrow y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

Isometric function, not orthographic

This is a handy little pseudo-proof that all projections can be broken down into a bunch of two dimensional projections no matter how complex it might initially seem[11]. With that in mind, three dimensions no longer seems particularly intimidating.

Re-factoring leaves behind a clean and concise function. With it, a collection of three dimensional points can be quickly projected onto a two dimensional

[10] Note that matrices are now being used. Technically they should have been used for 2D → 1D, but the result was a single value they would have overcomplicated things unnecessarily.

[11] Well sort of. This breaks down lower than the number two, and with any non-integers. In a general statement within this paper, this is true enough to not get in the way

$$p\left(\begin{bmatrix}x\\y\\z\end{bmatrix}\right)=\begin{bmatrix}p_{2d}(x,z)\\p_{2d}(y,z)\end{bmatrix}$$

$$p\left(\begin{bmatrix}x\\y\\z\end{bmatrix}\right)=\begin{bmatrix}f\frac{x}{z+f}\\f\frac{y}{z+f}\end{bmatrix}$$

$$p\left(\begin{bmatrix}x\\y\\z\end{bmatrix}\right)=\frac{f}{z+f}\begin{bmatrix}x\\y\end{bmatrix}$$

Back to orthographic functions. If/otherwise statements left out for the time being

plane. To prove it works, 4 is a render of a cube with side length ten centred at the origin, and a focal length of 2.

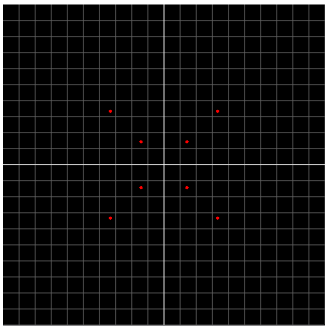


Figure 4: I wrote the code behind the renderer myself. It is available online at `goo.gl/MMOEPW`

Something a little strange happens when this formula gets applied to the real world: angle of view comes naturally. Since the size of the screen is limited, and focal length changes, angle of view becomes a function of those two variables. From the start, and two of these variables could have been chosen, the only reason that focal length and angle of view were the ones used was that I happen to find them the easiest ones to use. Given any two, the third is no issue to solve for. In the example, focal length is 2, and the viewport is 20 by 20. This means that both the vertical and horizontal angles of view are about 160 degrees. In the sample in 4 angle of view is never calculated, though it does exist. It can get away with this, because if it draws a point outside the view port, the point simply doesn't get seen.

3.2.3 Observer

Despite 4 supposedly picturing a cube, it does not look particularly cube like. Lines could be added, but it would still look a lot like two squares of different sizes. This is primarily because we rarely look at cubes head on, and cubes aren't normally floating points. One major pitfall of the function up to this point is that the observer cannot move. In the real world, this would be a problem. Nothing would be much fun if you couldn't look around. To add variables for an observer would require a lot of time and effort. It is possible, though not easy. Luckily, with a little thought, there is an alternative that requires much less thought.

Take the cube in the sample render, for example. Say that observer moves

10 units to the right, 12 backwards[12], and rotates 30 degrees clockwise. Instead of complicating the equation to account for a moving observer, move everything else opposite to the given actions. In this case it would be 10 units left, 12 back[13], and 30 degrees counter-clockwise. The rotation is a bit tricky because it needs to be rotated around the observer. Conceptually, rotating two three dimensional points around each other is exactly like projection, just a little different. The same way that projection can be broken up into smaller two dimensional problems, so can rotations. To rotate x_1, y_1 around x_2, y_2 by angle θ

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos(\theta) * (x_1 - x_2) - \sin(\theta) * (y_1 - y_2) + x_2 \\ \sin(\theta) * (x_1 - x_2) + \cos(\theta) * (y_1 - y_2) + y_2 \end{bmatrix}$$

Not exactly concise, but it works. That is a general equation, however. As long as the point is rotating around the origin, a tool called a rotation matrix can be used. It allows for compact and speedy calculations to be done.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Much cleaner, thanks to matrix multiplication. This is all that is needed, as every rotation in three dimensional space can be broken down into at most three consecutive two dimensional rotations. With that in mind, 5 is the new render with the earlier transformations applied

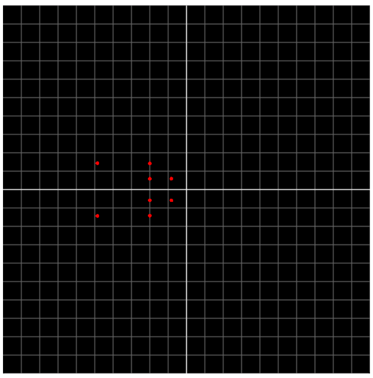


Figure 5: Point size does not scale with distance

4 Generalizing

The equations outlined are relatively specific equations for specific types of projection, but they are revealing about the broader field. For example, isometric and orthographic were presented as two separate equations, though this is not

[12]Going positive to negative, the Y axis is top to bottom of the screen, X is right to left, and Z is coming out of the screen to behind it

[13]It seems counter-intuitive that the opposite of left is right, yet the opposite of backwards is still back. Stand facing a friend, and each take a step to your respective lefts. You'll be two steps apart. To achieve the same effect along the other axis, you must each step backward.

the truth. Isometric projection is simply orthographic projection with an infinite focal length. This can be seen in 7 (Appendix) visually, and is equally apparent in the equations.

Orthographic Equation

$$p\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \frac{f}{z+f}\begin{bmatrix} x \\ y \end{bmatrix}$$

$$p\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \frac{\infty}{z+\infty}\begin{bmatrix} x \\ y \end{bmatrix}$$

$$p\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = 1\begin{bmatrix} x \\ y \end{bmatrix}$$

$$p\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ y \end{bmatrix}$$

Isometric

Figure 6: Demonstrates that isometric projection is simply orthographic with a focal length of infinity

This seems quite strange considering that in the initial definitions, the diagrams looked fundamentally different. Consider what focal length actually is. It is how far away the observer is from the scene, in some sense, thus it is also how far away the vanishing point is. The vanishing point is when a pair of parallel lines perpendicular to the projected plane eventually converge[14]. For a focal length of infinity, the lines never cross. This can actually been seen in the one of the earlier diagrams, 2. Note that in the isometric example, the two leftmost points line up, as they should, yet in the orthographic example they are in separate, and the two rightmost points, which certainly aren't in line, appear as a single point. Another way to think about it is in relation to the angle of view. As discussed earlier, it is inversely proportional to the focal length. It makes a right angle triangle with focal length and view width, such that focal length is the height, view width is width, and angle of view is the angle opposite the base. For any real view width and a focal length of infinity, the angle would be 0, or in more common parlance, a line. A line would make it an isometric projection, and that brings about the same conclusion.

This shows how those are the same in two and three dimensions, but the equation and method work for any number[15]. Starting from any projection function that reduces the dimension by one as p_x, a general formula for n number of dimensions can be made as p_n (dimensions listed as d_1 through d_n). The

[14]I tried to draw a diagram of this phenomenon, but I always ended up with an optical illusion that only confused the point. Instead, try looking down a set of railroad tracks. They certainly seem to converge way off in the distance.

[15]It doesn't work particularly well for fractions, though some creative thinking is required.

minimum number of dimensions will be the output of the original function.

$$p_n\left(\begin{bmatrix} d_1 \\ d_2 \\ ... \\ d_n \end{bmatrix}\right) = \begin{cases} p_x\left(\begin{bmatrix} d_1 \\ d_2 \\ ... \\ d_x \end{bmatrix}\right), & \text{if } n = x \\ p_{n-1}\left(\begin{bmatrix} d_1 \\ d_2 \\ ... \\ d_{n-1} \end{bmatrix}\right), & \text{otherwise} \end{cases}$$

A large function no doubt, certainly intimidating at first. The recursion can be a bit confusing at first, but it is no more than a representation of all the logic up to this point. All it says is that for a projection of n to x, reduce by a single dimension each time, unless it is the number of dimensions for the original function. In this way, any method of projection (orthographic, isometric, or otherwise) can be used.

5 Conclusion

Projection is a subtly fascinating topic. Depending on how far down the rabbit hole you fall, all areas of math are covered. Trigonometry and algebra are used to build a strong foundation, calculus is needed to discuss surfaces, matrices and vectors are used for any dimension larger than one, topology is useful for complex shapes, and so much more. It is field of great depth and even greater use. Every single time we interact with screens, projection is at play behind the scenes. Sometimes it is obvious, sometimes less so. More often than even screens, our eyes function the these very principles. The world as we see it is literally seen through projection. To just barely scratch the tip of the subject was an awe inspiring process of creation and discovery. Using just basic principles of logic and math, one can easily play with dimensions and the sort. It is one of those rare fields where all of math seems to mesh perfectly, and fits so well into the real world.

5.1 Extension

There was so much more I could have done. For starters, I only covered vertexes. Lines and faces are an entirely different level, ones that take entire university courses to cover. Building off of that, actually rendering models with colour and light would be a fascinating area to work with. In a similar vein, more complex drawing functions could be discussed, ones that dealt with the way light actually works, and how it might deform near large sources of gravity, though that is heading a bit too far into the realm of physics. Off in the other direction, more pure math could have been introduced. Dimensions beyond the typical three, and fractional dimensions have some interesting and complex features that would definitely be worth exploring. There are also edge cases, like focal lengths of zero, that raise unusual questions. Deciding what I would and would not include in this paper was a huge challenge, as a result of the field be so vast.

5.2 Errors

This topic is quite advanced. The math included was, for the most part, not beyond a highschool level. As such, some concepts could not be sufficiently explained. For example, the way in which dimensions are discussed is not perfect, though it works in this context. There is nothing wrong in the paper[16], but not everything is entirely correct either. This is a result of a need to simplify, in part to fit both the assignment requirements and the material taught in the course. Some peculiar features can come of a result of manipulating the matrices had they been used throughout, yet this would have left the the guidelines for writing this. Abstract concepts such as infinity also needed to be confined so that the amount of tangential conversations would be kept to a minimum. I also tried to use minimal external sources. This means that I was able to do a lot more independent and creative thinking, and all the work is mine, but it also means that I did not have any crutches to work with on the way. I stumbled a lot on the way. For example, arriving at the 2D orthographic equation took a substantial amount of time, and took countless failures to get a correct answer.

6 Appendix

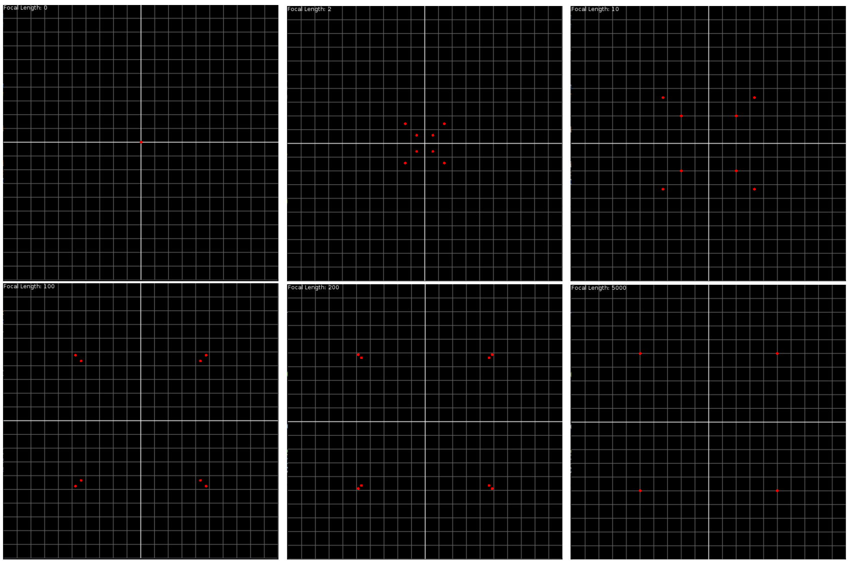


Figure 7: Comparing projections of varying focal lengths

References

[1] I. Carlbom and J. Paciorek. Planar geometric projections and viewing transformations. 10(4).

[2] Roger Thompson. *A Comprehensive Dictionary of Mathematics.*

[16] I hope

In addition to the sources listed, I also used Wolfram[17] for computations, and LOVE [18] and Lua[19] for the code behind the renderer.

[17]Both the website https://www.wolframalpha.com and the notebook tool Mathematica https://www.wolfram.com/mathematica/

[18]https://bitbucket.org/rude/love

[19]https://www.lua.org/