

Plato's Fractal Line:
A Solution to the Overdetermination Problem

0. Introduction

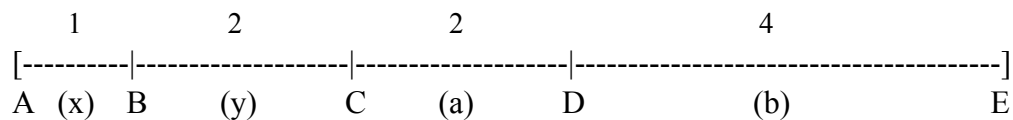
“It is like a line divided into two unequal sections.” Thus does Socrates begin one of the most famous epistemic similes, the aptly-named Divided Line. At the conclusion of Book VI of the *Republic*, Plato's Socrates leads his interlocutor Glaucon to imagine a complex mathematical system, to treat that system metaphorically, and to take that metaphor as explanatory of a certain epistemology. Once Glaucon accomplishes this mental feat of strength, Socrates moves on to *the* most famous epistemic simile—the Cave. If occasionally over-shadowed by the Cave, Plato's Divided Line remains a central representation of his epistemology. Along with Socrates and Glaucon, the reader quickly constructs an image, sets up the metaphor, and reads off the epistemic meaning: physical objects and opinion at the bottom, Forms and knowledge at the top. Yet the perspicuity with which Socrates ends the passage belies the complexity of this image and the images within it.

In an attempt to disentangle this complexity, this paper makes three claims, which correspond to the three movements Glaucon and the reader make: [1] that the Divided Line is a proto-fractal, [2] that a fractal line resolves the so-called overdetermination problem, and [3] that this understanding of the line in conjunction with an analysis of the shadow metaphor used in the passage illuminates some key aspects of Plato's epistemology. The structure of this talk follows this three-fold division. I begin with a short survey of the Divided Line passage and how it gives rise to the overdetermination problem, the seeming contradiction that arises from the fact that the middle two subsections appear to be both equal and unequal.. Having explained the

overdetermination problem, I turn to consider what fractals are and how the Divided Line conforms to that definition. In this second part of the paper, I suggest how understanding the line as a fractal may resolve the apparent tensions at the center of the line. Finally, I turn to the shadow metaphor used to describe the relationships of the subsections within the line. Here, I argue that Plato uses the shadow as a simple image for the complex relation between the ultimate objects of knowledge (i.e. Forms) and the objects of everyday experience. In the end, I hope to have persuaded at least some of you that Plato believes one can and must reach beyond a shadow of doubt to grasp some object of knowledge.

1. The Divided Line as proto-fractal

To begin this admittedly arduous journey, let us quickly recall the essentials of Socrates' description of the Divided Line. Socrates asks Glaucon to imagine a line "divided into two unequal sections" (509d6) and then to "divide each section ... in the same ratio as the line" (ὡςπερ τοίνυν γραμμὴν δίχα τετμημένην λάβων ἄνισα τμήματα, πάλιν τέμνε ἕκατερον τὸ τμήμα ἀνὰ τὸν αὐτὸν λόγον, 509d6-8). The line thus has the ratio AC is to CE as AB is to BC and as CD is to DE ($AC:CE :: AB:BC :: CD:DE$ or numerically $3:6 :: 1:2 :: 2:4$).



Moving from the purely mathematical to the allegorical, Socrates represents line section AC as the Visible and section CE as the Intelligible. Within the Visible realm, Socrates describes subsection AB as consisting of "shadows" and "reflections" (509e1), while subsection BC

consists of “the originals of these images” (510a3). Thus, *in toto*, the Visible realm consists of shadows and their respective physical objects. Within the Intelligible realm, Socrates explains that subsection CD contains “visible figures” (510d2), that is, it uses “as images the things that were imitated before” (510b3-4). In contrast, subsection DE represents the “forms themselves” (510b9), because one proceeds in investigation “without the images used in the previous subsection” (510b8). After this ontological exposition, Socrates explains that the four subsections correspond to four epistemic states or “conditions in the soul” (511d):

“Understanding for the highest, thought for the second, belief for the third, and imaging for the last. Arrange them in a ratio, and consider that each shares in clarity to the degree that the subsection it is set over shares in truth” (511d-e).

The Divided Line thus encodes an ontology and an epistemology into its mathematical structure.

I summarize those structures as such:

Ontology:

- AB - “shadows” and “reflections” (σκιάς and φαντάσματα, 509e1)
- BC - “the originals of these images” (ὅ τούτο ἔοικεν, 510a3)
- CD - “visible figures” (τοῖς ὁρωμένοις εἶδεσι, 510d2); “the things that were imitated before” (τοῖς τότε μιμηθεῖσιν, 510b3-4).
- DE - “forms themselves” (αὐτοῖς εἶδεσι, 510b9)

Epistemology:

- AB - “imaging” (εἰκασίαν, 511e1)
- BC - “belief” (πίστιν, 511e2)
- CD - “thought” (διάνοιαν, 511e)
- DE - “understanding” (νόησιν, 511e)

As many commentators have noted, however, when one maps either the ontology or the epistemology onto the mathematical structure, a contradiction arises in that the middle two subsections appear both equal and unequal. This contradiction is commonly known as the overdetermination problem. Commentators have long noted the necessary consequence of

Socrates' description that the second and third subsections must be of equal length.¹ Richard

Foley provides a paradigmatic arithmetic proof:

$$\begin{array}{ll} x/y = a/b = (x+y)/(a+b) & \text{[solve for x]} \\ x = ay/b & \text{[substitute for x]} \\ a/b = ((ay/b)+y)/(a+b) & \text{[multiply by (a+b)]} \\ a(a+b)/b = (ay/b)+y & \text{[cross-multiply the right side]} \\ a(a+b)/b = (ay+by)/b & \text{[multiply out b and extract the common y]} \\ a(a+b) = y(a+b) & \text{[divide out (a+b)]} \\ a = y & \end{array}$$

Although Socrates never explicitly marks this necessary mathematic result, it forms the first horn of our dilemma: subsections BC and CD are mathematically equal. As Socrates goes on to explain, however, each subsection represents ascending degrees of ontological truth and epistemic clarity (509d; 511d). As a result, the second horn of our dilemma appears: the middle subsections must also be unequal as they represent differing degrees of truth and clarity to the mind. Because truth, in Plato's schema, is an ontological descriptor and clarity an epistemic descriptor, the overdetermination problem applies separately to the ontology and the epistemology symbolized by the line. In short, [1] subsections BC and CD are mathematically equal, as the proof above shows; however, [2] subsections BC and CD are metaphorically unequal, as they represent differing degrees of truth and clarity to the mind.

As Richard Foley has recently reminded scholars,² any attempt to understand the line must come to terms with both versions of the overdetermination problem. Although much recent work, such as that of Nicholas Smith,³ has done much to resolve this contradiction for the ontological level of the line, Foley has demonstrated that the overdetermination problem remains

¹ Foley 2008, 2 provides an arithmetic proof, while Klein 1965, 119 offers a geometric version.

² Foley, Richard. 2008. "Plato's Undividable Line: Contradiction and Method in *Republic* VI." *JHP* 46.1: 1-23.

³ Smith, Nicholas. 1981. "The Objects of 'Dianoia' in Plato's Divided Line." *Apeiron* 15.2: 129-37.

a difficulty for the epistemological level of the line as well. Foley himself believes that this issue “is an intentional feature of the divided line analogy,” intended to compel readers toward deeper philosophical engagement with the text. I agree with Foley on this last point, although I hold that this apparent contradiction in fact rests upon an equivocation, which arises from a misunderstanding of the line as a fractal.

This leads me to my first claim, that the divided line is a proto-fractal. Before I argue for this, however, let us briefly consider what a fractal is. In simplest terms, a fractal is a self-similar mathematical pattern. As Kenneth Falconer puts it: a fractal “contains copies of itself at many different scales.”⁴ The study of fractals was pioneered by Benoît Mandelbrot some fifty years ago.⁵ Although fractal geometry is a modern pursuit, Mandelbrot would be the first to remind us that fractals are as old as nature herself;⁶ Mandelbrot offers the example of a mountain, where each peak, hill, and rock resemble the mountain as a whole. Thus, while the category of “fractal” did not exist for Plato, there is no reason to believe that fractal-like systems could not have been investigated in the ancient world. I contend that the Divided Line presents precisely such an investigation. In order to support this claim, let us consider a modern linear fractal, such as the so-called “middle third Cantor set”:



[*ad infinitum*]

⁴ Falconer 2003, xviii

⁵ See Mandelbrot 1977 and Mandelbrot and Blumen 1989, 3-16.

⁶ Mandelbrot, Benoit B. *The Fractal Geometry of Nature*. Macmillan, 1983.

This image presents four iterations of a simple function whereby a line segment is divided equally into 3 sub-segments and then the middle sub-segment is removed, thus producing two new line segments. The function is then performed on these two line segments at F^1 to produce four line segments at F^2 , and so forth and so on, *ad infinitum*. This example demonstrates five essential conditions of a fractal as defined by Falconer:

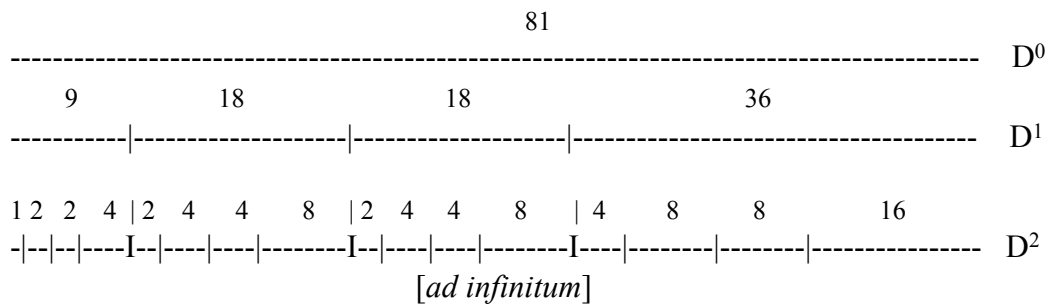
- [1] a fractal is self-similar; that is, F^1 relates to F^0 just as F^2 relates to F^1 ,
- [2] it has a fine structure, so that it retains its detail at any scale,
- [3] it has a simple definition,
- [4] it is obtained by a recursive procedure,
- and [5] it is awkward to describe its local geometry.⁷

Let us turn now to the Divided Line and consider whether it meets these five criteria.

Socrates' initial description of the mathematical structure of the line, though short, suggest that the Divided Line highly resembles a fractal. The line clearly conforms to the first condition, as self-similarity is the defining characteristic of the line: the division of the whole is the same as the divisions of the parts (*ἀνὰ τὸν αὐτὸν λόγον*, 509d8). This opening description also proves to meet condition three, as Socrates can describe the function in so few words and the function itself is clearly no more complex than that of the middle third Cantor set. Next, by asking Glaucon to perform two mental steps, to imagine a line divided unequally and to divided each section by that same ratio, Socrates' description also meets condition four. As for condition five, the overdetermination problem itself suggests the line meets this criterion, though I shall return to this point more fully in a bit. For now, suffice it to say that it is difficult to describe the local geometry of the Divided Line. This leaves only the second condition. Does Plato's Divided Line have a fine structure, that is, does it have an infinite depth of detail? The short answer is

⁷ Falconer 2003, xviii. He adds, [6] "the geometry ... is not easily described in classical terms" and [7] "its size is not quantified by the usual measures such as length," although I find these to be sufficiently covered by [5].

“no, but.” As found in the *Republic*, the line does not present a fine structure because its details are lost after the second iteration of the function. However, the mathematical apparatus does offer the possibility of a line that does meet this criterion. Clearly one need only continue Socrates’ described recursive procedure *ad infinitum*, as in the example of the excluded middle, to produce a full fractal:



Note that creating a divided line function with fine structure also makes the local geometry as difficult to describe as the Cantor set, insofar as every point contains an infinite number of points within/beneath it. In short, we can say that Plato’s Divided Line represents a proto-fractal. Socrates establishes the mathematical apparatus for creating a fractal, but only uses this apparatus to construct an image with two degrees of complexity. Nonetheless, using the concept of a fractal as a heuristic device can, I argue, point us toward an answer to the overdetermination problem.

The conceptual beauty of a fractal lies in its ability to represent geometrically that subtle movement of the mind sensed when viewing images such as these:

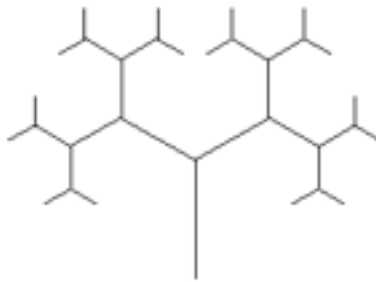


Rubin’s Vase



Duck/Rabbit

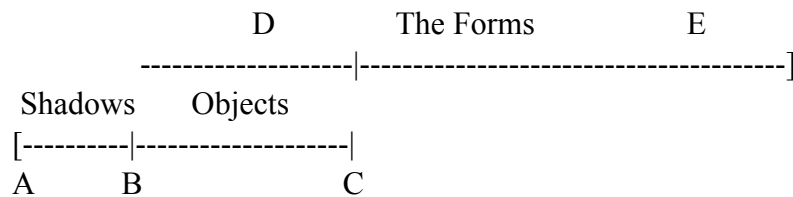
Optical illusions of this sort create multi-stable perception wherein our minds are able to form two different, though stable, perceptions from a single image. Put simply, the image on the right is *both* a duck *and* a rabbit. Similarly, in a fractal, any section represents both a source and a derivation, or put otherwise, a whole and a part. Consider a simple tree fractal:



The first two “branches” are also simultaneously “trunks” for two of their own respective branches. In other words, fractals create unified conceptual systems composed of some binary relationship, whether that be wholes and parts, sources and derivations, or trunks and branches. The defining characteristic of fractals lies in that while there is only one absolute whole, or source, or trunk, each part that derives from it perfectly resembles that whole, if one zooms into that level of complexity. This is precisely what it means for a fractal to be self-similar—as you continue to zoom in, it appears that you are still looking at the whole. For example, the right-hand side of F^1 in the middle third Cantor set (item three) appears identical to F^0 , merely smaller. Thus, in an *absolute* sense, F^1 is a part or a derivation, but in another, *contextual* sense, F^1 is also a source/whole. It is this consequence of the self-similarity of fractals that makes it difficult to describe their local geometry, which simply means that it is difficult to fix and define any one point in a fractal, given the multivalency of that point.

If Plato was indeed attempting to create a quasi-fractal, that is, a fractal with all of the essential aspects except the fine detail of infinite iterations, then an interpreter ought to see the

central sections as both “branches” and “trunks.” To put it more precisely, on the ontological level, the physical objects are simultaneously the *imaged* objects of shadows and the *imaging* objects of Forms, and on the epistemic level, they are the concrete objects of belief (πίστις) and the abstract objects of thought (διάνοια). Below I attempt to capture this fractalized conception of the Divided Line graphically:



With this line there are two horizontal movements and one vertical movement: the horizontal movements represent material changes, while the one vertical movement represents an aspectual change. Thus, on both the line’s ontological and epistemic levels of meaning the middle two subsections are in an *absolute sense* equal, but in a *contextual sense* they are unequal, because of the mental shift in how that section is treated.

2. The Overdetermination Problem(s)

Armed with our understanding of the mathematical structure of the line, let us turn to confront both the ontological and epistemic overdetermination problems head-on. First, the ontological problem. Socrates clearly associates three of the four subsections with a class of objects: AB (imagination) includes shadows and reflections; BC (belief), physical objects; and DE (understanding), the Forms. CD (thought), however, the first subsection of the intelligible realm, has been a point of much scholarly contention. Many interpreters feel that in order to

satisfy the parallelism of Plato's structure, the objects of thought must be intermediaries between the Forms and the visibles, frequently dubbed simply "lower noetics".⁸ The intermediaries, under this view, image the Forms just as shadows image physical objects and physical objects image the intermediaries. As Nicholas Smith and others have recently argued, however, this interpretation lacks full textual support.⁹ Socrates explicitly states that the *διάνοια* subsection uses "as images the things that were imitated before" (510b). Thus, both subsections contain the same *stuff* but the mind treats that material in different *aspects*.¹⁰ This recapitulates the distinction between the horizontal and vertical movements in my re-conceived version of the line. Although there are four parts to the line, Plato offers a tripartite ontology comprising of shadows, objects, and Forms. In my opinion, reading the line as a fractal most clearly and effectively makes sense of this arrangement. Foley himself concedes that such a three-class division resolves the ontological overdetermination problem; the more sticky issue is the epistemic problem.

At the heart of the epistemic side of the overdetermination problem lies Socrates' concluding remarks in Book VI. Here he sums up the meaning of the divided line:

"there are four such conditions in the soul, corresponding to the four subsections of our line: Understanding for the highest, thought for the second, belief for the third, and imaging for the last. Arrange them in a ratio, and consider that each shares in clarity to the degree that the subsection it is set over shares in truth"

καί μοι ἐπὶ τοῖς τέτταρσι τμήμασι τέτταρα ταῦτα παθήματα ἐν τῇ ψυχῇ γιγνόμενα λαβέ,
νόησιν μὲν ἐπὶ τῷ ἀνωτάτῳ, διάνοιαν δὲ ἐπὶ τῷ δευτέρῳ, τῷ τρίτῳ δὲ πίστιν ἀπόδος καὶ τῷ

⁸ See Foley 2008, ___ for an extensive bibliography on all those who endorse a Platonic theory of intermediates.

9

¹⁰ Smith succinctly summarizes his position thusly, "while the class of objects at the level of *πίστις* has the same extension as that of *διάνοια*, the two classes have different intensions" (132).

τελευταίῳ εἰκασίαν, καὶ τάξον αὐτὰ ἀνὰ λόγον, ὥσπερ ἐφ' οἷς ἐστὶν ἀληθείας μετέχει, οὕτω ταῦτα σαφηνείας ἡγησάμενος μετέχειν (511d-e).

As noted earlier, truth is Socrates' ontological criterion and clarity is his epistemic one.¹¹ As Foley rightly notes, the soundest interpretation of the phrase “in a ratio” identifies that ratio as the initial mathematical ratio. But herein lies the problem—although the two middle subsections must be equal in length, Socrates hierarchically ranks the epistemic states they represent. Thus, to Foley, the epistemic states of mind are simultaneously equal and unequal.

A divided line fractal resolves this issue. If the linchpin for the epistemological overdetermination problem is the shared ratio, then our mathematical solution to the ontological problem ought to apply equally as well to the epistemic problem. Once again, we must distinguish between an absolute and a contextual sense when considering the line's equality. In the absolute mathematic sense, the middle two subsections are necessarily equal, whether they represent ontological classes or epistemic states of mind. Under my interpretation, this absolute equality is meant to picture the fractal structure of the line as a whole. The middle subsections must be equal to signal the hinging role physical objects play in Plato's epistemology.

Once this absolute mathematical structure is formed, however, and the reader enters into the line, like a prisoner in the Cave, the subsections represent unequal degrees of clarity and truth. Context is key. This is most fully seen in the Cave, where at each level people believe they are dealing with the ultimate objects of knowledge, whether they are cave shadows, cut-outs, or outside shadows. At each step, the philosopher must realize the larger context within which these

¹¹ Socrates had stated right before that “as regards truth and untruth, the division is in this proportion: As the opinable is to the knowable, so the likeness is to the thing it is like” (510a). He clearly is treating truth as an ontological descriptor with epistemic overtones. This foreshadows the deep connection the ontological and epistemic levels to the line share.

objects fall. Although both subsections are essentially equal in that they are composed of the same class of objects, *διάνοια* is clearly superior insofar as our would-be philosopher here begins to turn away from the visible, concrete world to the realm of abstraction, the realm of the Forms. As noted earlier, while the material is the same (an absolute equality), one can see that material in two ways (an aspectual inequality): [a] as objects of knowledge or [b] as means to knowledge. As Plato believes the first is false, it is clearly superior to understand physical objects in the latter sense. Hopefully, reading the line as a proto-fractal has illuminated how the Divided Line can present these two levels of meaning. This structure merely suggests *how* the same class of objects may be treated as objects of knowledge or as means to knowledge. In order to gain some insight into *why* physical objects play such a central role in Plato's epistemology, let us now finally consider the primary metaphor Socrates uses to describe the relationships between the various subsections, the shadow.

3. Shadows of an Epistemology

Although shadow imagery stands out more clearly in Socrates' allegory of the Cave, the shadow also lurks in the Divided Line. When Socrates describes the class of images that compose the lowest subsection, he includes "shadows, then reflections in water and in all close-packed, smooth, and shiny materials, and everything of that sort" (*μὲν τὰς σκιὰς, ἔπειτα τὰ ἐν τοῖς ὕδασι φαντάσματα καὶ ἐν τοῖς ὅσα πυκνὰ τε καὶ λεῖα καὶ φανὰ συνέστηκεν, καὶ πᾶν τὸ τοιοῦτον*, 509e-510a). The shadow then returns as Socrates attempts to describe the confusing third subsection using an analogy with geometry: "these figures that they make and draw, of which shadows and reflections in water are images, they now in turn use as images, in seeking to

see those others themselves that one cannot see except by means of thought” (αὐτὰ μὲν ταῦτα ἃ πλάττουσιν τε καὶ γράφουσιν, ὧν καὶ σκιαί καὶ ἐν ὕδασιν εἰκόνες εἰσίν, τούτοις μὲν ὡς εἰκόσιν αὖ χρώμενοι, ζητοῦντες δὲ αὐτὰ ἐκεῖνα ἰδεῖν ἃ οὐκ ἂν ἄλλως ἴδοι τις ἢ τῇ διανοίᾳ, 510e-511a).

The shadow becomes the primary motif to describe the relationship between the subsections of both the Intelligible and Visible realms. Subsections AB and CD are both related to their counterparts as a shadow is related to its object. Thus, an analysis of the shadow motif is one key element to understanding the epistemology of the Divided Line. Yet in order to see how the shadow illuminates Plato’s theory of knowledge, we must first consider how exactly shadows relate to their objects.

I suggest that shadows be defined as *authentic, yet obscure* representations of their objects, but let’s unpack this a bit. Because a shadow depends for its existence on its object and because a shadow must and can only represent that object, there is necessarily an authentic relation between object and shadow. Put simply, if there is a shadow, then there is an object whence it derives. Mirrors similarly produce authentic images, yet shadows uniquely produce images that are also obscure. Unlike a reflection, a shadow is opaque; it hides much more than it reveals. The shadow is thus both insubstantial in the ontological sense, in that it has no mass or substance, but also obscure in an epistemological sense—it offers limited epistemic access to its object.

The shadow motif within the Divided Line reveals the hurdles one must overcome in one’s progression through the Line. As Socrates describes them, the physical shadows of subsection AB represent their respective physical objects, and the intellectual shadows of subsection CD likewise represent the Forms. Both sets of shadows share an intimate connection

with their respective objects. The physical shadows present an authentic image of their objects, though they are a two-dimensional representation of three-dimensional objects. Likewise, the intellectual shadows grant true access to the Forms, though they are dependent on the visible realm. Yet these *authentic* shadows are also lacking in their representations. Both sets of shadows fail to capture their object's essence—the physical shadows lack ontological matter, and the intellectual shadows lack philosophic purity. In this way the shadows also *obscure* their objects, failing to reveal the essence or fullness of their object. Therefore, as authentic yet opaque representations, shadows allow Plato to picture an epistemology where knowledge is both tenuous and attainable.

Shadows illustrate the immensely difficult epistemic task of knowing the Forms insofar as they are *opaque* representations of their respective objects, yet they also exemplify the possibility of epistemic success, as they are also *authentic* images. The linchpin of this epistemology is the class of physical objects in subsections BC and CD. On the one hand, they are the objects from which the shadows of AB derive; on the other hand, they are themselves shadows of the objects from which they derive, the Forms. As noted earlier, these visible objects represent both the *imaged* objects of shadows and the *imaging* objects of Forms. As a would-be philosopher moves through the epistemic stages of the line, this duality proves essential. Having recognized that physical shadows are merely images and seeing their imaged objects, the philosopher understands that if there is a shadow, then there is an object and that shadow gives some access to that object. As the philosopher begins to turn away from the visible world, perhaps through geometry as in Socrates example, he begins to see that physical objects are themselves intellectual shadows; they are obscure yet authentic representations of some final

objects of knowledge. He knows now that the Forms must exist (because there is no shadow without its object) and that he has some access to them. If physical objects were mere mirror-like images of the Forms, the horizontal movement from thought to understanding would be fairly easy, and both subsections of the intelligible realm would have equal clarity (and thus be of equal length). Yet if physical objects shared nothing with the Forms, the horizontal movement from thought to understanding would be impossible. Because visible objects relate to the Forms as a shadow relates to its object, the attainment of true philosophical understanding is both difficult and possible.

This picture of Plato's epistemology may leave you with a final question: why did Plato make this all so complicated? I believe a clue is found at the conclusion of Socrates' epistemic digression at 534a:

“But as for the ratio between the things these are set over and the division of either the opinable or the intelligible section into two, let us pass them by, Glaucon, lest they involve us in arguments many times longer than the ones we've already gone through.”

τὴν δ' ἐφ' οἷς ταῦτα ἀναλογίαν καὶ διαίρεσιν διχῆ ἑκατέρου, δοξαστοῦ τε καὶ νοητοῦ, ἐῷμεν, ὧ Γλαῦκων, ἵνα μὴ ἡμᾶς πολλαπλασίων λόγων ἐμπλήσῃ ἢ ὅσων οἱ παρεληλυθότες (534a).

Here, I agree with Richard Foley, who supposes that, “Plato is surreptitiously hinting that the serious reader should analyze what these further difficulties might be.”¹² In an attempt to turn his readers' heads away from the shadows dancing on the cave wall, Plato asks us to analyze the math and the imagery of his simile in order to tease out its full meaning. Although Plato didn't have the vocabulary available to him to describe his fractal fully, I believe he offers a clear enough picture to follow his thinking. In this fractal filled with shadows Plato encodes a deft and nuanced epistemology where one can reach beyond a shadow of doubt and grasp knowledge.

¹² Foley 2008, 18.

Select Bibliography

Falconer, Kenneth. 2004. *Fractal Geometry: Mathematical Foundations and Applications*. John Wiley & Sons.

Foley, Richard. 2008. "Plato's Undividable Line: Contradiction and Method in Republic VI." *Journal of the History of Philosophy* 46 (1): 1–23. doi:10.1353/hph.2008.1842.

Mandelbrot, Benoit B. 1983. *The Fractal Geometry of Nature*. Macmillan.

Morrison, J. S. 1977. "Two Unresolved Difficulties in the Line and Cave." *Phronesis* 22 (3): 212–31. doi:10.2307/4182019.

Smith, Nicholas. 1981. "The Objects of 'Dianoia' in Plato's Divided Line." *Apeiron: A Journal for Ancient Philosophy and Science* 15 (2): 129–37. doi:10.2307/40913474.

———. 1996. "Plato's Divided Line." *Ancient Philosophy* 16: 25–46.