I propose here a new interpretation of the “Final Argument” for immortality in Plato’s *Phaedo*. Although the “Final Argument” is sometimes taken to encompass, even, Socrates’ entire reply to Cebes (95a–107b), I am most interested in the very last and brief consideration of this section—which I shall call the “Ultimate Final Argument” (106c9–107a1)—where Socrates argues that the soul, because it is immortal, is imperishable. This is the argument in response to which Cebes, at last, relinquishes any degree of doubt and declares Socrates’ view to be “absolutely necessary” [πολλὰν ἀνάγκη]. Thus, the work of the dialogue is completed here;¹ and a sound interpretation of the dialogue must explain why this is so. I shall maintain that this brief passage gives what might be called a “cosmological argument” for the everlastingness of the soul, a kind of argument which has analogues in other dialogues. My goal in this paper is to explain what Plato thought was needed for this argument, and how he meant to supply it, and, in the course of my doing so, the argument and its actual weaknesses will become clear as well.

¹ Note that there is really no basis in the text for holding, as is commonly thought, that Socrates or his interlocutors persevere in doubt, or that the conclusion is provisional. It is true that Simmias, who is portrayed as weak-minded throughout the dialogue, and presumably recognizing his own admittedly human weakness, expresses his own need to look at the arguments again τὴν ἄνθρωπινὴν ἀσθένειαν ἀτιμᾶσαι, ἀναγκάζομαι ἀπιστίαν ἐτι ἔχειν παρ’ ἐμοῦτῷ περὶ τῶν εἰρημένων, (107b1–3); yet in reply Socrates encourages him to do so, because then he will become fully satisfied, as Socrates and Cebes already are οὐδὲν ἐπιτίθετε περαιτέρω, (b9). Note also that the investigation Socrates seems to be encouraging at the end is merely one that proceeds back “toward first principles,” and not “away from” them, and thus it would not be capable of altering the dialogue’s conclusion, or its force (compare 101d–e). On Cebes’ being represented as the strong and Simmias as the weak reasoner of the dialogue, see David Sedley, “The dramatis personae of Plato’s *Phaedo*,” in T. Smiley, ed., *Philosophical Dialogues: Plato, Hume, and Wittgenstein* (British Academy: Oxford, 1995), 1–26.

Correspondence to: Institute for the Psychological Sciences, 2001 Jefferson Davis Highway, Suite 511, Arlington, VA 22101.
The Objections of Simmias and Cebes. The first step in construing the Final Argument is to understand precisely what it is meant to accomplish, since only then can we discern its method and correctly judge its success. In order to discern what it is meant to accomplish, we need to pay attention to the objections of Simmias and Cebes, and especially that of Cebes, and try to determine, from details of the text, the precise weight and significance that Plato assigns to these objections.

So that we can have both objections before us, here is Socrates’ elegant summary of them:

Simmias, I believe, is doubtful and afraid that the soul, though more divine and lovelier than the body, may still perish before it, being a kind of attunement \( \text{ἐν ὀρμονίας ἐξέ} \) \( \text{ούσα} \). Whereas Cebes, I thought, agreed with me in this much, that soul is longer-lived than body; but he held that no one could be sure \( \text{ἂν ἀνθῆλον πάντι} \) whether the soul, after wearing out many bodies time and time again, might not then perish itself, leaving its last body behind, and whether death might not be just that, the perishing of the soul—since body, at any rate, is perishing incessantly and never stops \( \text{ἄει ἄπολλυμενον οὐδὲν παυεται}. \)

Now I want to draw attention to three details of the text which signal that Plato does not himself consider these arguments to be weighty.

The first is that Socrates makes the agreement of the objections with each other a condition of those objections having weight; yet then it becomes clear, in the statement of those objections, that they are radically in conflict with each other.

After Simmias states his objection, Socrates says, using a courtroom metaphor, that “before answering I think we should first hear from Cebes here what further charge he has to bring against the argument \( \text{τι σὺ ὃδε ἔγκαιλε τῷ λόγῳ} \)” and that “when we’ve heard from them both, either we should agree with them, if it seems they’re at all concordant \( \text{SPATH} \) \( \text{σὺ προσά} \); or if not, we should at that point take up advocacy of our argument \( \text{ἦδη} \)."
The implicit suggestion is that witnesses who contradict each other discredit each other; their lack of agreement makes their testimony worthless.

Plato has Simmias and Cebes state their objections in a way that makes it clear that they are in disagreement with each other on the most fundamental point. Indeed, almost Cebes’ first words, when he states his objection, is that he disagrees with Simmias: “I don’t agree with Simmias’ objection [οὐ συγχωρῶ τῇ Σιμμίου ἀντιλήψει], that soul isn’t stronger and longer-lived than body: because I think it far superior in all those ways.” And afterwards the nature of the disagreement becomes sharp and clear: Simmias premises his objection on the view that the soul perishes with or before the body (compare 86b2); whereas Cebes’ starting point is that the soul perishes after the body (87d1, e2, that is, except in the last case).

Similarly, Simmias, in holding that the soul is an “attunement” or “harmony,” regards the organization of the soul as depending upon that of the body, whereas Cebes, in picturing the soul as a weaver of its own clothing, makes the organization of the body depend upon that of the soul. Needless to say, these are strikingly different views. Now, although we might think that an argument is weaker to the extent that it can be attacked from a variety of views, it seems clear that Plato’s supposition, rather, is that, if there were indeed a weakness in Socrates’ argument, which two persons had together succeeded in noticing, then they should at least be able to agree on what that is. And yet, there is no possibility that Simmias’ and Cebes’ objections could come into alignment.

The second detail to be noticed is that both Simmias and Cebes admit that their objections are based on (as a Wittgensteinian might say) “pictures” (εἰκότα, “liabilities,” even “likenesses” or “impressions”) rather than arguments. “I too, like Simmias, need a kind of image [εἰκόνος . . . τινος],” Cebes says, when he introduces his objection. After Simmias gets refuted, he confesses that the notion that the soul is an attunement was just a misleading impression he had

5 Plato, *Phaedo*, 86c5.
picked up (μοι γέγονεν ... μετὰ εἰκότος τινὸς). He was led astray by a deceptive picture, as he recognizes can happen in geometry, when one does not take care to attend only to reasons and demonstrations. At one point the dialogue’s dramatic narrator, Echecrates, even chimes in and confesses that “that notion, that our soul is a kind of attunement, has a strange hold on me, now as it always has done [καὶ νῦν καὶ ἂν].” He adds, “I very much need some other argument that will convince me once again, as if from the start.”

What Plato is suggesting through these details is that the attraction of the “attunement” view is not a rational influence. Tellingly, it exerts its power always, even after it was shown wrong before, and it needs to be repeatedly put aside; its deceptiveness is standing by, always ready to mislead us. This is why Phaedo, in the narrative frame, referring back to the episode, speaks of Socrates not simply as arguing in reply to the objections, but also as “giving curative treatment” to everyone present [ὁς ἐὰν ἡμᾶς ἱόσατο]. It is this susceptibility to being deceived and entranced by images that Socrates is referring to also, surely, when in the misology passage he advises that we should not aim to “shift the blame from oneself to arguments.” “Let’s not admit into our soul,” he says, “the thought that there’s probably nothing healthy in arguments [τῶν λόγων ... οὐδὲν ὑγιὲς]; but let’s far rather admit that we’re not yet healthy ourselves [ἡμεῖς ὁπως ὑγιῶς ἔχομεν].” So it is Socrates’ view, apparently, that certainly Simmias’ objection and probably also to some extent Cebes’, have a nonrational origin; these objections get their force insofar as we are tempted to rely illicitly on “pictures” rather than reasons—a “human weakness” that, we might suspect, Plato is supposing is a consequence of the incarnate condition of the soul.

The third detail we should notice is that both objections are put forward in an obviously inconsistent manner. Simmias proposes that the soul is an attunement entirely dependent on the body, but in his

---

10 Plato, *Phaedo*, 89a5.
13 And, this surely is the “human weakness” ἡ ἀνθρωπίνη ἀσθένεια that Simmias, with better and more sober self-knowledge, refers to at 107b1.
next breath he says that he agrees with the Recollection Argument, that the soul existed prior to the body that it animates.\footnote{Plato, \textit{Phaedo}, 92b–e.} Clearly, Plato is portraying him as deeply confused. Cebes shows a similar inconsistency: he thinks that a good comparison for the relation of soul to body is that of a weaver who dons the very coat that he weaves, and yet Cebes is prepared to say that, from dying and being reborn over and over again, the soul might wear out and perish. This is as if someone were to suggest that a man’s health might be affected from his putting on and taking off his coat repeatedly. What Cebes ought to think, consistent with his basic outlook, is that something with the kind of control over the body that a soul has would presumably never be affected by anything that happened to the body.

These details indicate, then, that Plato does not regard the objections of Simmias and Cebes as actually powerful. Presumably that is the reason why Plato portrays Socrates as remaining serene throughout the episode (89a), and any emotional disturbance, and subsequent turnaround, is ascribed by Plato only to Socrates’ audience and to Echecrates:

All of us who heard them were disagreeably affected by their words, as we afterwards told one another: we’d been completely convinced by the earlier argument, yet now they seemed to disturb us again [πάλιν ἐδόκουν ἀνασταράξαι], and make us doubtful not only about the arguments already put forward but also about points yet to be raised, for fear that we were incompetent judges of anything, or even that these things might be inherently doubtful [τὰ πράγματα αὐτὰ ἀπιστα].\footnote{Plato, \textit{Phaedo}, 88c1–7.}

Socrates himself does not react in that way, not because he never believed his own earlier arguments, but because he does not think the objections of Simmias and Cebes succeed in overturning those arguments, given the qualified force that those arguments were meant to have.

Socrates does, however, allow that the objections provide an occasion to examine more thoroughly his earlier arguments (compare 84c5–7), and, it seems, he also wishes to deal with the objections as a means of probing some more serious worries that might, in some contexts, underlie those or similar objections. For instance, after
Simmias rejects the attunement view as a misleading “likeness,” on account of its patent inconsistency with the Recollection Argument; nonetheless, Plato has Socrates refute that view with a series of arguments. Why? One reason is perhaps that Socrates’ refutation depends upon, and allows him to introduce, the view that “no soul is more or less a soul than any other”—that is, that being a soul does not admit of degrees—and we shall see that Plato will rely upon this important idea later. Another reason, I believe, is that Plato is now wishing to take seriously, not simply the entrancing “picture” of the soul as an attunement, but also, and more importantly, a view from pre-Socratic natural philosophy which would analyze the soul as it would analyze any composite whole, that is, as a suitable proportion of the compounded elements of the body. That Socrates has this kind of theory in view is suggested by his language, since in his refutations he does not rely, as Simmias did, on any informal comparison of the soul with a musical instrument or on commonsense images about how we “tighten and loosen” strings. Rather, he asks such things as: “Do you think it befits an attunement, or any other compound [συνθεσία], to be in any state other than that of the elements of which it is composed?”—describing the relation supposed to hold between an “attunement” and the parts so attuned, using a fair degree of philosophical precision and generality.

As for Cebes’ objection, Plato wishes to link this as well to a view in pre-Socratic philosophy, I think: namely, the view that “all things are together” and that, as a consequence, nothing has enduring stability, and everything is in flux. This view shows up first in the context of the misology passage, where Socrates refers to those “contradiction-mongers” [άντιλογικοί] who, because they have cultivated the skill of devising contradictory arguments to any thesis whatsoever, have come to think that “there is nothing sound or secure whatever, either in

---

16 We need not suppose that Alcmaeon was the intended target. The idea of harmonious or proportional composition, and new qualities emergent from these, was used widely in explanations in pre-Socratic natural philosophy. For an account of the Phaedo as a whole which explains why the objections of Simmias and Cebes should be regarded as introducing doubts deriving from pre-Socratic philosophy in particular, see Michael Pakaluk, “Degrees of Separation in the Phaedo,” Phronesis 48 (2003): 89–115.

17 Plato, Phaedo, 86c–d.

18 Plato, Phaedo, 93a1–2.
things or in arguments \([οὔτε τῶν πραγμάτων οὐδενός οὐδὲν ᾑγίες οὔδε βέβαιον οὔτε τῶν λόγων]\); but that all realities are carried up and down, just like things fluctuating in the Euripus.”\(^{19}\) Although this warning is given to everyone in Socrates’ company, we are perhaps justified in taking it to be aimed principally at Cebes, who is presented throughout the dialogue as an especially acute reasoner. Indeed later, when Socrates is responding to Cebes in particular, the view shows up again. Socrates there explains the “method of hypothesis” to Cebes and how it requires that one keep clear about whether one is moving away from or toward starting points, and he distinguishes this from the practice of those “contradiction-mongers,” who “jumble things” and “discuss the starting point and its consequences at the same time.” He then adds, apparently in allusion to an important strand in pre-Socratic natural philosophy, “For them, perhaps, that isn’t a matter of the least thought or concern; their wisdom enables them to mix everything up together \([ικανοί γάρ ὑπὸ σοφίας ὀμοῦ πάντα κυκώντες]\), yet still be pleased with themselves.”\(^{20}\)

Why should we think that Plato wishes to diagnose Cebes’ objection as ultimately getting its force from the appeal of a view such as this? The reason is the nature of Cebes’ doubt. As is said in one of several explanations of his objection, he agrees with Socrates that the soul is “strong” \([ἰσχυρόν]\), but he denies that it is impassible in the face of death; he agrees that it is “godlike” \([θεοειδές]\), but he holds back from saying it is immortal; he accepts that it is more “long-lived” \([πολυχρόνιον]\), but he says that, for all we know, it might eventually wear out and collapse.\(^{21}\) Now—it is important to note—he has no definite reason for supposing this. As was mentioned, given his own comparison of the soul to a weaver and the body to a coat, his own view can provide no definite reason for thinking the soul might wear out eventually. People wear out coats, not coats people. His doubt is of a general and even vague nature. The body is changing always: for

\(^{19}\) Plato, Phaedo, 90c.

\(^{20}\) Plato, Phaedo, 101e4–6. These two passages need to be read in connection with other “flux” passages in the Platonic corpus, which are well canvassed in Terence Irwin, “Plato’s Heracleiteanism,” Philosophical Quarterly 27 (1977): 1–13.

\(^{21}\) Plato, Phaedo, 95c5–d1.
all we know the soul is the same.\textsuperscript{22} Who knows that the soul does not suffer (\(\pi νε\iota ν\)) something each time it is separated from the body?—We can’t rule that out.\textsuperscript{21} After all, none of us can see the soul.\textsuperscript{24} Cebes emphasizes that a “philosophical man” will be strict about this, where others perhaps would be more relaxed,\textsuperscript{25} which suggests that his doubt has an especially “philosophical” source.

In sum, Cebes shows himself unwilling to draw conclusions about what is always and is necessarily so, from generalizations which he accepts and, as regards which, he would concede, we have no reason to attribute any particular exceptions, and Plato diagnoses this as showing the influence of a view that supposes that everything is always changing. Cebes will believe that something is secretly undergoing change, unless he is shown otherwise; his antecedent and governing presumption is that flux is inherent in things, not stability. That is why Plato generally represents Cebes’ objection, not as a demand for a better demonstrative argument (\(\alpha π\delta\iota\varepsilon\iota\xi\iota\sigma\)), but as a demand that something be pointed out or shown (\(\varepsilon\pi\delta\iota\varepsilon\iota\xi\iota\sigma\)) to him.\textsuperscript{26} It is Simmias who talks about demonstrative arguments, and, as we have seen, he attributes his own doubts to his favoring likelihoods over such arguments.\textsuperscript{27} Cebes is portrayed, rather, as accepting the earlier demonstrations, but not granting that they have necessary force.\textsuperscript{28}

Here we need to attend to a curious but important difference in how Plato, in comparison with us, would approach a difficulty such as this. We would suppose that someone who had found room, as he thought, to reject the necessity of an argument’s conclusion, even if he had accepted the premises, would be finding fault in the validity of the argument, or perhaps claiming that one of its premises was merely

\textsuperscript{22} Plato, \textit{Phaedo}, 91d7.
\textsuperscript{21} Plato, \textit{Phaedo}, 88a8–9.
\textsuperscript{24} Plato, \textit{Phaedo}, 88b2–3.
\textsuperscript{25} Plato, \textit{Phaedo}, 95c1.
\textsuperscript{26} See Plato, \textit{Phaedo}, 95b9, 99d2, 100b3. The change or difference in use is not exceptionless, since \(\alpha π\delta\iota\varepsilon\iota\xi\iota\sigma\), which typically in the dialogue takes the precise sense of “a showing that something is so, through a demonstrative argument,” may also mean simply, “a showing that something is so,” in which case it can be used in much the same way as \(\varepsilon\pi\delta\iota\varepsilon\iota\xi\iota\sigma\), as I believe we see at 105e8.
\textsuperscript{27} Plato, \textit{Phaedo}, 92d.
\textsuperscript{28} Plato, \textit{Phaedo}, 87a.
contingent. Hence, we suppose that Socrates, in order to respond to Cebes, will need to devise a better argument, that is, one that is valid rather than fallacious. However, I do not believe that Plato understands Cebes’ difficulty in this way. Cebes’ difficulty, from Plato’s point of view, is that he does not grant that the world has a suitable character so as to allow any necessary conclusions to be drawn about it at all: so a response to Cebes will involve showing him that it does. That is to say, Plato thinks that he needs to allay, somehow, the suspicion that “everything is in flux” (and therefore that the soul is as well) and that he needs to point out or display how something like the soul, which so evidently is a subject of various changes (after all, it is reincarnated repeatedly, as Cebes allows and emphasizes), may nevertheless serve as an object of knowledge, that is, as the object of conviction which has the characteristics of necessity and universality.

Plato approaches Cebes’ objection in this way because he thinks of cognition as fundamentally a matter, not of skill in reasoning or facility in using a method or rules, but as perception, and thus there will be a correspondence or mirroring, he thinks, between the states we can attain in cognizing the world and what the world is like. If the world is constantly in flux, then we can never attain stable cognition, and our beliefs and thoughts too must constantly be in flux. On the other hand, if we are to reach that stable and assured condition which is “knowledge,” then we must recognize, or have pointed out to us, those things which are stable and which would serve as suitable objects of this condition. Cebes acknowledges the existence of the Forms and their stability; he himself has stable knowledge about them. This can then serve as the basis for his “cure.” Socrates’ curative

29 Note the mirroring indicated within the phrases: “οὐτε τῶν πραγμάτων οὐδενὸς οὐδὲν υγίεσ οὐδὲ βέβαιον οὔτε τῶν λόγων” nothing sound or secure whatever, either in things or in arguments, (90c3–4) and “τῶν δὲ ὀντῶν τῆς ἀληθείας τε καὶ ἐπιστήμης στερηθεὶ” deprived both of the truth of and knowledge of realities, (d6–7). Compare Cratylus, 411a1–c6; Theaetetus, 179e–180b; and especially Timaeus, 29b4–6: “the accounts we give of things have the same character as the subjects they set forth [τῶν οὐτῶν καὶ συγγενεῖς οὕτως.] So accounts of what is stable and fixed and transparent to understanding are themselves stable and unshifting.” Trans. Donald Zeyl, in ed. John M. Cooper, Plato: Complete Works (Indianapolis: Hackett, 1997).
remedy for him, then, requires that he somehow be led, starting from this point, to recognize some suitable stability in things and especially in souls.

At this point, we simply need to say something more about Plato’s “Heracliteanism,” and then we are finished with the task of understanding the nature of Cebes’ objection. Plato of course in various places shows sympathy with Heraclitus, and Aristotle even claims, famously, that it was because Plato accepted Heraclitus’ account of the sensible world that he postulated Forms as separately existing objects of knowledge. Yet at the same time in many places Plato seems to be an enemy of the view that “everything is in flux.” How do we reconcile these two positions? For present purposes, I shall accept the resolution suggested by Irwin, who, in brief, distinguishes the doctrine of “copresence of opposites” (what Irwin calls “a-change”) from the doctrine of “flux” (what he calls “s-change”). The doctrine of copresence of opposites is that sensible things have copresent opposites in connection with their relation to different things, for instance, Helen is both beautiful and ugly, that is to say, beautiful in comparison to mortals, but ugly in comparison to a god. On the other hand, the doctrine of flux is a way of accounting for the copresence of opposites. The doctrine of flux is an inference to the best explanation: according to that doctrine, the reason why opposites are copresent in things, is that things are changing constantly—since it is presumed on intuitive grounds that anything that is changing to that extent has opposite characteristics at once.

---

31 Irwin, “Plato’s Heracleiteanism.” Irwin gives the following definitions, although the notions are intuitively clear enough: \( x \) \( a \)-changes iff \( x \) is \( F \) in one aspect, not-\( F \) in another, and \( x \) is in the same condition when it is \( F \) and when it is not-\( F \) (for example, \( x \) is big in comparison with \( y \), small in comparison with \( z \)); and, \( x \) \( s \)-changes iff at time \( t_1 \) \( x \) is \( F \) and at time \( t_2 \) \( x \) is not-\( F \), and \( x \) itself is not in the same condition at \( t_2 \) as it was at \( t_1 \) (for example, at \( t_1 \) it is hot, and at \( t_2 \) it has become not-hot, be becoming colder than it was).
32 That this is indeed how Plato understands the matter is revealed at Theaetetus, 183a5–7: “if all things are in motion, any answer that can be given to any question is equally right; you may say it is so and it is not so—or ‘becomes,’ if you prefer to avoid any term that would bring these people to a standstill.” Here Plato takes the possibility of opposite predications (\( \text{ότι} \ \tau' \ \text{έχειν} \ \text{φανεῖ} \ καὶ \ \text{μὴ} \ \text{ότω} \) to be a consequence of the constant motion of the thing talked about (\( \text{πάντα} \ \text{kineῖται}, \ \gammaίγνεσθαι\)).
Irwin seems to be correct when he maintains that Plato (with suitable qualifications, as we shall see) accepts the copresence of opposites, but rejects the doctrine of flux. If that is so, then one of the things that we should expect him to be attempting, in the reply to Cebes, is to account somehow for the copresence of opposites, while blocking the inference from this to the doctrine of flux.

II

Socrates’ Autobiographical Remarks. Socrates begins his reply to Cebes by recounting his own past investigations, saying that “if any of the things I say seem helpful to you, you can use them for conviction [πρὸς τὴν πείθω] on the points you raise.”33 We should consider this, I think, as an invitation to us the readers, also, to do likewise, and that is what I shall do. There are three lessons in particular, I think, which we should draw from Socrates’ autobiographical remarks as being relevant to Cebes’ problem as we have described it.

In his autobiographical remarks, Socrates tells a story, really, of two different disenchantments. The first is his disenchantment over investigations in “natural history” [ἡ περὶ φύσεως ιστορία].34 This is the project in pre-Socratic philosophy of nature which involved explaining things by giving a narrative account of how they arose over time through processes of aggregation and separation. It is unnecessary to go into great detail here about the nature of Socrates’ disenchantment. His point seems to be a kind of Humean objection, that there is no intelligible or necessary connection between, on the one hand, a particular aggregation or separation in place, and, on the other, the formal qualities that something aggregated or separated next displays. Further, if there is no such connection, then we cannot say that the earlier process is the cause of the later, and therefore we cannot say that, in our identifying the antecedent aggregation or separation, we know the cause of the coming into existence of those formal qualities.

33 Plato, Phaedo, 96a3–4.
34 Plato, Phaedo, 96a8.
It is this last point that especially needs to be emphasized, since Socrates repeatedly complains (phrased always in the first person, of course, and with irony) that explanations of that sort do not yield knowledge: they make people who accept them believe that they know something, but in fact they do not know it (compare ἡπιστάμην, 96c4; εἴδέναι, c6, e7; ἔπισταμαι, 97b4). This suggests—the first lesson—that what is at stake in the discussion that follows, and in the Final Argument, is whether one can establish some way in which we can truly be said to know something about the natural world.

Socrates' second disenchantment is of course with the natural philosophy of Anaxagoras. He tells of how he had heard second-hand about Anaxagoras and of how, from that report, in his own mind he had drawn inferences about the way in which Anaxagoras would develop his system:

One day, however, I heard someone reading from a book he said was by Anaxagoras, according to which it is, in fact, Mind that orders and is responsible for everything [πάντων αἴτιος]. Now this was the kind of explanation that pleased me [ταύτῃ δὴ τῇ αἰτίᾳ ἰσόθην]; it seemed to me, somehow, to be a good thing that Mind should be responsible for everything. And I thought that, if that's the case, then Mind in ordering all things must order them and place each individual thing in the best way possible; so if anyone wanted to find out the reason why each thing comes to be or perishes or exists, this is what he must find out about it: how is it best for that thing to exist, or to act or be acted upon in any way?

Socrates then hurries to get Anaxagoras' book, reads it, and discovers that Anaxagoras makes no use of Mind in explaining anything, but instead gives narrative accounts involving composition, aggregation, and separation, just like the rest of the natural philosophers whom Socrates had already rejected. That Anaxagoras proceeds in this way baffles Socrates, who says that this is as if someone who could explain why Socrates refused to escape from prison by mentioning "those matters that were truly responsible for this [τὰς ὡς ἀληθῶς αἰτίος]"—that is, "that Athenians judged it better to condemne me, and therefore I in my turn have judged it better to sit here, and thought it more just to stay behind and submit to such a penalty as they may ordain"—instead, when asked to give an account, chose to speak only

---

35 Plato, Phaedo, 97b–c.
36 Plato, Phaedo, 98e1–4.
of the position and arrangement of Socrates’ bones and sinews, as Socrates is sitting.

We are meant to draw two additional lessons from this story, in my view. One is found in Socrates’ conviction that if Anaxagoras had provided explanations of the sort that Socrates was hoping for, then Socrates would have had knowledge about the world. That Plato holds this is suggested, not simply by the language of “true causes,” which we have already noted (and compare “the real cause” [τὸ αἴτιον τῶ ὄντι], 99b3), and in Socrates’ remarking that an explanation of that sort would “expound both what was responsible for something and the necessity of it [ἐπεξεργάσθαι τὴν αἴτιαν καὶ τὴν ἀνάγκην],” but also in Socrates’ explicit remark that he “made all haste to get hold of the books and read them as quickly as I could, so that I might know [ἰδεῖν] as quickly as possible what was best and what was worse.”

The other lesson we are meant to draw from this story of Socrates’ disenchantment with Anaxagoras is in Socrates’ assertion that the only thing that can be responsible for connectedness among things and their being bound together—their necessary connection—is Mind, and that therefore any explanation that succeeded in accounting for such things would need to appeal somehow to Mind. He affirms this when he summarizes his complaint about pre-Socratic natural philosophy. Such philosophers, he says, mistake the sine qua non for the cause; they confuse the instruments of rational agency, with the rational agency that makes use of those instruments:

the power [δύναμιν] by which they’re now situated in the best way that they could be placed, this they neither look for nor credit with any supernatural strength [δομινόιναν ισχύν]; but they think they’ll one day discover an Atlas stronger and more immortal than that, who does more to hold everything together. That it’s the good or binding, that genuinely does bind and hold things together [ὅς ἀληθῶς τὸ ἄγαθον καὶ δέον συνδεῖν καὶ συνέχειν], they don’t believe at all.39

Note that Socrates is saying this in his own person, and what his assertion amounts to is the principle that only an account that appeals somehow to Mind or intelligence is capable of providing knowledge of

37 Plato, Phaedo, 97e1–2.
38 Plato, Phaedo, 98b4–6.
39 Plato, Phaedo, 99c1–6.
causal relationships and connections in nature. He never abandons or rejects this principle, in this dialogue or elsewhere in the Platonic corpus (Laws 10 and the Timaeus being the most noticeable examples of it). His complaint is rather with Anaxagoras’ application of it, or rather with his failure to carry through with it.

We should therefore interpret the “second voyage” that Socrates next goes on to explain as his, in some manner, drawing upon the “true cause,” which is Mind, but without, he thinks, the hazards associated with the unsuccessful “first voyage” that Anaxagoras had had him embark on. His “second voyage” is a different, and safer, way of getting to the same end. We might speculate that its safety consists in its somehow appealing to Mind, but without, however, the need to speculate about the remote purposes or intentions of Mind.

Plato himself suggests how this sort of thing might be done, when he gives us, as an example of a good explanation, how we might account for Socrates’ sitting in prison: Why is Socrates sitting now in prison? The “hazardous” explanation, the “first voyage,” would appeal (Plato tells us) to those considerations of justice, and of what would work out for the best, that led Socrates to refuse Crito’s offer to escape and decide to remain in prison. In comparison a “safe” explanation might be something like: Socrates is sitting in prison because he chooses to sit in prison; he is sitting in prison because sitting in prison is what he is doing. It would be an appeal to something like the formal content of his decision, not his purposes in making a decision with that content. To give the formal content of an ordering decision of a Mind is a “safe” way of accounting for something by appealing to a Mind, since, on the presumption that a Mind is responsible for ordering, then a Mind is responsible for the form of some element of, or constituent in, that ordering.

It is sometimes said that there are two possible meanings of “second voyage,” and that Socrates’ discussion is indifferent as regards them. The two meanings would be: (1) to row when the winds are unfavorable; or (2) to take a roundabout route to the same location, in order to avoid some danger. But Socrates’ insistence that his alternative is a “safe” path, and his complaints about Anaxagoras’ failure to execute his program, in my view show definitively that he intends his “second voyage” in sense (2).

Sean Kelsey objects that, although the example of Socrates in prison draws a distinction between the true cause and a sine qua non; nonetheless, it is not particularly helpful, as the examples serves to illustrate only “the
To adopt such an interpretation as this is to reject, of course, the influential approach of Gregory Vlastos, who held that the phrase “second voyage” must mean an entirely new kind of investigation, not a continuation of the search for a “teleological cause.” Yet Vlastos’s arguments are not decisive, since he fails to draw an important and relevant distinction as regards teleological explanation.

After Vlastos first renders the relevant lines from the dialogue:

Well, I for my part should be delighted to learn from anyone about this sort of *aitia* (that of “the good and the fitting,” 99C5). But since I have been denied this *aitia* (ἐπειδὴ δὲ ταύτης ἐστερηθῆν) and have failed to either find it myself or learn it from another, would you like me to show you the second journey which I have been pursuing in the search for *aitia*? (99C–D2)

He then argues as follows:

Note that the reference of ταύτης in C3 (the subject of ἐστερηθῆν and also, with the implied change of case, of the infinitives ἐρεῖν, μαχεῖν) is to ταὐτῆς σίτιας in C7; what Socrates has failed to discover by his own labors or from those of others and is prepared to do without is the teleological *aitia* itself. This leaves no room for understanding him to mean (as has been done over and over again in the literature) that his “second-best journey” is (a) an alternative method of searching for teleological *aitiai* rather than (b) an alternative method of searching for *aitiai*.

However, Vlastos seems to overlook an important distinction in what he refers to as a “teleological” explanation. From what Socrates

---

obvious absurdity of not citing reason as a cause in cases in which it is presupposed to be operative” (25). However, since Socrates continues to believe that Mind is “responsible for everything,” then he is presupposing it to be operative in all of the workings of nature, and all cases of explanation in nature are parallel. Kelsey next objects that “It is difficult to see how to extract from this a more general account of the distinction itself, and in particular, an account that will allow us to see why it is not equally absurd in such cases to cite Forms as causes. For example, if Socrates has decided by an exercise of his reason to remain in prison, is it not just as absurd to cite the Form of Sitting as the cause of his sitting there as it is to cite bones and sinews?” (25). Yet presumably, to cite Socrates’ idea of sitting and to say that he has chosen to move his body in accordance with this idea would not be absurd. See Sean Kelsey, “Causation in the *Phaedo,*” *Pacific Philosophical Quarterly* 85 (2004): 21–43.


says about Anaxagoras, presumably Socrates would regard such an explanation as including three parts: (1) the claim that the cosmic Mind is “responsible for everything”; (2) the implicit premise that every mind always decides upon what it regards as best; and then (3) an account as to why some particular state of affairs would be for the best. Socrates’ remarks about a “second voyage” are surely consistent with the supposition that he continues to hold (1) and (2) while abandoning attempts to figure out (3). It even seems that his remarks are best taken in that way: the verb ἐστερήθην suggests Socrates’ being deprived of something to which he had regarded himself as having a reasonable claim or expectation; however, Socrates, as he tells us, had already accepted (1) and (2) on his own, and he was looking to learn (3) from Anaxagoras; thus ἐστερήθην naturally refers to the turnabout, whereby Socrates first anticipated learning explanations from Anaxagoras about why actual states of affairs are better than alternatives,44 but then found nothing of the sort when he read Anaxagoras’ book.45 Indeed, Socrates is quite precise about what he failed to learn, that is, “what was best and what was worse,”46 which would leave conditions (1) and (2) in place, and in that sense his “safe explanation” would also be “teleological.”

Vlastos concedes that there would be reason to understand Socrates’ “second voyage” as a different way of looking for a “teleological” explanation,

if we could assume that the earlier references to teleological aitiai as the “true” or “real” aitiai (τὰς ὀς ἀληθῶς αἰτίας, 98E1; τὸ αἰτίον τῷ ὄντι, 99B3) mean that these are for Plato not only the preferred (most fundamental, most illuminating) explanations of natural phenomena (which, of course they are throughout the Platonic corpus), but the only admissible aitiai of anything whatever

44 Plato, Phaedo, 98b1–6.
45 Plato, Phaedo, 98b7–c2.
46 Note, too, that “neither discover for myself nor learn from another” is a stock phrase in Plato. If, as one might suspect, it is added here exegetically, then that confirms that Socrates’ complaint is that he did not learn from Anaxagoras the very thing he had expected him to teach, that is, the knowledge of what would make an arrangement of things “better and worse. Vlastos’ point that the antecedent of ταὐτης is the aitia that Socrates was seeking is also not decisive, as it would be natural for Plato to say “deprived of X” when he meant “deprived of the means by which I expected I would get X.”
but, Vlastos observes, there is no evidence that Plato holds to this stronger idea.\(^4\) However, that stronger idea is not necessary; the weaker idea, which Vlastos admits is accurate and holds for Plato, will suffice, since Socrates presents his “second voyage” in the context of explaining to Cebes what is the *aitia* for “coming to be and passing away” in the natural world (περὶ γενέσεως καὶ φθορᾶς τὴν αἰτίαν, 95e9–96a1).\(^5\) Socrates is meaning to account for “natural phenomena,” including the soul’s career in the natural world, and, as Vlastos acknowledges, he does so on the principle that Mind is the only true cause in that domain. Thus, even on Vlastos’ own terms, there is nothing hindering us from holding what seems in any case to be the most natural meaning of the text, that is, that “second voyage” involves the pursuit of a “teleological” explanation by some other means.

### III

*The Final Argument.* We arrive, finally, at the Final Argument, now with this sense of what it is meant to accomplish and how: Cebes fosters a general doubt, of a philosophical character, that the soul, as something with a career in the natural world, is not the sort of thing that can remain entirely stable, without any change at all in some respect, over time and come what may. Plato supposes that the remedy for his doubt will involve “displaying” to Cebes how the soul is

---

\(^4\) Vlastos, “Reasons and Causes,” 298, n. 15. Emphasis is original.

\(^5\) Plato. *Phaedo.* 95e9–96a1. Vlastos also argues that “Forms” cannot be “teleological *aitiai*” on the grounds that Socrates says he is deprived of teleological explanations, at the same time that he says that he has been *engaged*, already, with explanations involving the Forms (περιογματευμα, 100b4): “if Socrates had thought of the Forms as teleological he would not have said that he is still “deprived” of teleological *aitiai.*” Vlastos, “Reasons and Causes,” 303–4. However, this objection seems also to be based upon Vlastos’ failure to distinguish the elements of the first-best explanation for which Socrates was hoping. Socrates might have continued to hold to the principles that Mind is responsible for everything and that it always chooses what it thinks best, while regarding himself deprived of knowledge of why actual states of affairs are best. And, there is no reason that we cannot understand him as saying that he had already been occupied with the Forms for other reasons, yet, after his disappointment with Anaxagoras, he came to see how Forms might be used as an expedient and as a second-best route to an explanation that somehow brings in Mind.
the right sort of object of knowledge. Furthermore, Plato is assuming that, insofar as things have some connection with one another, then, for that connection to be real and for an explanation of it to count as knowledge, one will need to appeal somehow to the operation of Mind. Finally, in accordance with Socrates’ talk of a “second voyage,” we take Plato to be looking for some indirect way of appealing to the operation of Mind, that is, some possibility of a teleological explanation which involves pointing merely to something that admits of being understood as the “formal content” of the ordering activity of a mind—which would then stand proxy for a first-best and full-blown teleological explanation.

Socrates begins the “Final Argument,” of course, by proposing a pattern of explanation which he regards as “safe”: a beautiful thing is beautiful on account of its sharing in the beautiful itself. This is probably the part of the Final Argument that has received the most attention by commentators: is Plato suggesting that a Form is a cause? If so, is the Form supposed to be an efficient cause or only a formal cause? But I will not dwell on it. The claim is a rather straightforward statement of the dependence of particulars on Forms for the formal existence that they have, which is a familiar idea in Plato. We need not speculate about what sort of a cause a Form is in this regard. It will be enough to say, following David Sedley, that a Form is regarded by Socrates as “responsible” for the participated reality in a particular, and leave it at that.⁴⁹

It is rather Socrates’ next claim which I think is most important.⁵⁰ In view of the importance of this passage for our purposes, it will be good to have it before us in full:

⁴⁹ See David Sedley, “Platonic Causes,” Phronesis, 43 (1998): 114–32. Plato thinks that because the particular in this way depends upon the Form, any predicates applied to particulars in virtue of their participation in a Form are derivative uses of the predicate (Plato calls this “ἐπώνυμον ἔχει” or “eponymy,” 92d9, 102b2, c10, 103b7–8); the predicate is properly and in the first instance applied to the Form, even if we infer the existence of the Form starting from our perception of particulars.

⁵⁰ Kelsey insists, most helpfully, that interpreters should be clear about where precisely in the Final Argument the autobiographical remarks of Socrates should be brought to bear. Kelsey, “Causation in the Phaedo,” 31. In my view the text which I now discuss is that place, for the reasons that I shall give.
He next asked, “If you say that that is so, then whenever you say that Simmias is larger than Socrates but smaller than Phaedo, you mean then, don’t you, that both things are in Simmias, largeness and smallness?”

“I do.”

“But now, do you agree that Simmias’ overtopping of Socrates isn’t expressed in those words according to the truth of the matter? Because it isn’t, surely, by nature that Simmias overtops him \[\text{πεφυκέναι \ Σιμμίαν ύπερεχεῖν τοῦτῳ}, \] by virtue, that is, of his being Simmias \[\text{τῷ \ Σιμμίαν \ εἶναι}, \] but by virtue of the largeness that he happens to have \[\text{τῷ \ μεγέθει \ ὁ \ τυγχάνει \ ἔχων}. \] Nor again does he overtop Socrates because Socrates is Socrates \[\text{ὄτι \ Σωκράτης \ ὁ \ Σωκράτης \ ἐστίν}], but because of the smallness that Socrates has in relation to his largeness?”

“True.”

“Nor again is he overtopped by Phaedo in virtue of Phaedo’s being Phaedo, but because of the largeness that Phaedo has in relation to Simmias’ smallness?”

“That is so.’

“So that’s how Simmias takes the name \[\text{ὄ \ Σιμμίας \ ἐπωνυμίαν \ ἔχει}] of being both small and large; it’s because he’s between the two of them, submitting his smallness to the largeness of the one for it to overtop, and presenting to the other his largeness which overtops the latter’s smallness.’

At this he smiled, and added: “That sounds as if I’m going to talk like a book. But anyway, things are surely as I say.”

He agreed.

“I say this for the following reason, wanting you to think as I do. It seems to me that not only is largeness itself never willing to be large and small at the same time, but also that the largeness in us never admits the small \[\text{οὐδὲποτε \ προσδέχεσθαι \ τὸ \ σμικρὸν}], or is it willing to be overtopped \[\text{οὐδ’ \ ἐθέλειν \ ύπερέχεσθαι}]. Rather, one of two things must happen: either it must retreat and get out of the way, when its opposite, the small, advances towards it; or else, upon that opposite’s advance, it must perish. But what it is not willing to do is to abide and admit smallness, and thus be other than what it is \[\text{οὐκ \ ἐθέλειν \ εἶναι \ ἔτερον \ ἢ \ ὑπὲρ \ ἴν}. \] Thus I, having admitted and abided smallness, am still what I am, this same individual \[\text{ἔτι \ ό \ ὁπερ \ εἰμι}], only small; whereas the large in us, while being large, can’t endure \[\text{οὐ \ τετόλισκεν} \] to be small. And similarly, the small that’s in us is not willing ever to come to be, or to be \[\text{οὐκ \ ἐθέλει \ ποτὲ \ μέγα \ γίγνεσθαι \ οὐδὲ \ εἶναι}], large. Nor will
any of the opposites, while being what it was, at the same time
come to be, and be, its own opposite. If that befalls it, either it
goes away or it perishes.\(^{51}\)

Now, it is common to observe that here Socrates wishes to extend
to particulars also a truth that he regards as holding of Forms, namely,
that, just as a Form is what it is, and entirely excludes its opposite, so a
particular insofar as it shares in a Form—which Socrates calls the
“form in us”—also excludes its opposite. A particular may share in
opposite Forms, but each “form in us,” which the particular therefore
has and in which it shares, is distinct from the other and excludes the
other.

Without doubt, Plato is proposing this idea in the passage. But, as
Socrates tells us, we are meant to draw additional things from the
passage, which pertain especially to the possibility of giving an
account of “coming to be and passing away”—after all, that was what
Socrates had promised he would talk about to Cebes in the first place
(95e8–96a1).\(^{52}\) For instance, and obviously, we are also meant to
conclude that it is because opposites exclude each other that an
increase in the one “in us” is correlated always with a decrease in the
other. This was an important idea for the Cyclical Argument, for
example, that when something becomes cooler, it simultaneously
becomes less hot, and so on. Opposites vary inversely. One might say,
then, that it is necessary that an opposite in us exclude its opposite, in
order for things to vary in degree as they do. This in itself would be an
important result because it traces variation in things back to one kind
of stability, namely, to the opposite’s keeping the character that it has
throughout the variation.

Another thing to notice in the passage is that Socrates is
concerned to distinguish between the subject which has the opposites
and the opposites themselves. Note that this is, so to speak, a
gratuitous detail that Socrates adds. It is not as being Simmias, he
insists, that Simmias is taller than Socrates, but rather in virtue of the
tallness that Simmias has. To be sure, that sort of analysis puts the
“blame” for Simmias’ being taller where Socrates thinks it belongs, that
is, on the “form in Simmias” of tallness, and therefore on the Form of

--


\(^{52}\) Plato, \textit{Phaedo}, 95e8–96a1.
Tallness, but what this also means, importantly, is that the subject is, so to speak, untouched by the conflict of opposites which the subject has. If Phaedo and Socrates were to cease to exist, Simmias would remain just as he is, even though he is no longer taller than Socrates, or shorter than Phaedo.\textsuperscript{53}

This, then, leads to another interesting detail of the passage, which is Socrates’ insistence that Simmias is taller than Socrates by the tallness in him and shorter than Phaedo by the smallness in him. In insisting on this, Socrates is denying that the relation Simmias has to other particulars and what they are is prior to the character that he has in his own right and in virtue of his relation to the Form. It is not that when a smaller person, such as Socrates, enters the room, Simmias, because he is then taller than someone, therefore acquires tallness as a consequence. It is rather that Simmias has a portion of tallness and Socrates has a portion of tallness, and it is these portions that get credit for why, when Socrates enters the room, Simmias is taller than Socrates rather than shorter.

The example of Simmias as standing between Phaedo and Socrates is meant to be a star example of the copresence of opposites, and what Socrates’ analysis is meant to show, is that we can acknowledge that opposites are copresent—indeed, we must acknowledge that, to account at all for change—without thereby concluding that the subject of change is also changing. Neither is there a conflict of opposites “all the way down” because the basis for the copresence of opposites is a clear and direct relation of dependence of a particular upon a Form, which is marked by an absolute exclusion of its opposite. That is, the Simmias example gives an analysis intended to block any inference from the doctrine of copresence of opposites to the doctrine of universal flux.

It is to the apparently gratuitous details of the passage that we must look to discern what in particular Plato is intending to convey.

\textsuperscript{53} Note that, according to Socrates, Simmias himself is called “large” or “small” only by eponymy. If we pick him out as (say) “that small (man),” we use an accidental attribute to pick out someone whose “nature” \(\pi\epsilon\varphi\kappa\epsilon\nu\alpha\iota\) is other than that (compare with a Kripkean nonrigid designator). This suggests that, in contrast, it is possible to pick out objects by terms which indicate what they are in their nature (a rigid designator)—as Socrates will claim that we do with fire in calling it “hot” or soul in calling it “alive.”
because we should presume there is some reason why he states things, or draws an emphasis, in that way in particular. Now, there is one last detail of the passage of that sort that we should attend to, and this is the language that Socrates uses to describe, or explain, how it is that an opposite in a subject excludes its opposite. Here his language becomes strikingly anthropomorphic. The opposite does not consent to be present, he says, when its opposite approaches. It would not be so daring as to remain and receive the opposite, and so on. Now we might be tempted to dismiss this as colorful language, but, again, it surely is deliberate and “gratuitous,” in the sense that Plato was not at all bound to put the point in this way.

Furthermore, we need to examine more carefully one’s sense that this language is merely colorful. For consider: what is at issue is whether an opposite really repels its opposite, in the way that, say, we feel two magnets resisting each other when we try to place their positive poles together? And, if an opposite does really exclude its opposite, what can account for this? Plato does not want to say merely that, as it happens, we do not find opposites mixing together or that this has never been seen before, and hence we have no reason to think that it will happen. He thinks it is necessary that an opposite never mix with its opposite. Observe, too, that this is a necessity that involves the connection among or relation of things: to establish something as tall is ipso facto to establish it as related in a certain way (that is, by excluding) anything else in the universe insofar as that is short. However, Plato had told us earlier that he thinks that only Mind can explain necessity or the connections among things—and that, I believe, is why he uses mind-like language here. If we say that tallness has just never, as a matter of fact, been coterminous with shortness, we explain nothing; however, if we say that it abhors shortness, or aims to exclude it, or cannot tolerate it, then, by Plato’s lights, we do explain it. If you will, Plato is giving a theory of what it is to account for something through appealing to “forces,” and his view is that, for such an explanation to be satisfactory, we have to understand a “force” as something mind-like.

That he uses anthropomorphic language here deliberately, as a “second best” explanation, to stand proxy for the most desirable, first-best sort of explanation, which would appeal directly to the ulterior purposes of a mind, is shown, I think, in the parallels between this
language here and the example he had picked earlier of the best sort of explanation. His example earlier was Socrates in prison: the “true causes” of Socrates’ being in prison are the decisions of rational agents, the Athenians and Socrates, with a view to what they regard as good. However, that example—is it coincidental?—involves exactly the same alternatives as Plato describes as holding for a “form in us.” Socrates can decide either “to stay behind and submit to the penalty” [παραμένωτα ὑπέχειν τὴν δίκην], (98e4), or “escape and run away” [φεύγειν τε καὶ ἀποδιδόσκειν], (99a3), or “submit to whatever penalty the city might impose” [ὑπέχειν τῇ πόλει δίκην ἢντιν ὄν τάττη], (99a4), depending upon whether he regards what is approaching as “opposed” to him or not. Because he regards his accepting the punishment as just, he remains in prison—and then his soul gets the benefit of the justice which it receives, whereas his body has to perish from the approach of something opposite to it.  

Note that Plato was not bound to use that example of Socrates in prison to illustrate his point. We may suspect he did so because of the analogies already pointed out and the more obvious analogy that holds as well, between Socrates in prison and Socrates’ soul in his body.

Socrates develops only one further point before he presents what I have called the “Ultimate Final Argument.” This is where he in effect develops the concept of an element and claims that the soul is an element. Again, this passage gets much attention from commentators, but for all of its elaborate comparison of a soul to life, fire to heat, and triplicity to odd the passage makes a relatively simple point and does not, for our purposes, require great attention. Briefly, Socrates says

---

54 On this point, recall the lines which bring the Final Argument to a conclusion and which constitute the true conclusion of the Phaedo as a whole, regarded as a dialogue about the nature of death; these are parallel to 102d–e: “Then when death attacks [ἐπίοντος] a person, the mortal part, it seems, dies; whereas the immortal part gets out of the way of death [ὑπεκκρίθησαν τῷ θανάτῳ], departs, and goes away intact and undestroyed.”

Note also the most basic reason that Socrates gives for a quality’s excluding its opposite, namely, self-love: “it is not willing . . . to be other than what it is” (102e3). Furthermore, after making his points about largeness and smallness “in us,” and what they are willing or unwilling to abide, Socrates shifts directly into a comparable remark about himself (“Thus I, having admitted and abided smallness”), thus inviting the comparison between his own volition, and the “volitions” inherent in the qualities of largeness and smallness.
that some things such as fire, although not themselves opposites, nonetheless always bring an opposite along with them: for example, where there is fire (not itself an opposite) there is heat. The soul, Socrates says, is the same sort of thing: although it is not itself opposite to anything, nonetheless a soul always “brings life with it.” Now just as an opposite always excludes its opposite, so things such as this, which always bring an opposite along, exclude that which is opposed to the opposite which they bring along. Thus, not simply heat but also fire excludes cold; not simply life but also a soul excludes death. A soul always brings life along with it, and life is an opposite the presence of which excludes, by a necessary force, death. It is in that precise sense, Socrates says, that a soul may be said to be immortal. *Immortal* in this precise and narrow sense does not mean “everlasting,” but rather: the sort of thing which, because of what it is and not what it has, by a necessary force excludes death, so long as it exists.

Now a common complaint against this passage is that it does not accomplish much because it establishes, at best, that a necessary property of the soul is that it be living, not that any soul which exists necessarily always exists: just as the fact that fire is necessarily hot does not guarantee that each portion of fire will burn forever. In reply it should be said that Plato wants to establish no more than that “a necessary property of an existing soul is that it is living and as such excludes being dead.” That he believes his claim is thus restricted is shown in his recognizing that he needs to argue, additionally, that a soul, given that it is immortal, is also imperishable.56

A similar complaint is the objection that the Final Argument at this point depends on an equivocation as regards the word, “immortal”: distinguish “immortal” in the technical sense of “being necessarily alive and, as such, excluding death,” from “immortal” in the popular sense of “everlasting”. According to this common objection, the Final Argument shows, at best, that the soul is “immortal” in this technical sense, whereas Plato wishes to arrive at that conclusion in the popular sense. However, once again, Plato is not guilty of such a mistake because, for his purposes, all that he wants to establish at first is that the soul is immortal in the precise sense, and then he recognizes that a

further argument is needed to establish that the soul is additionally imperishable and therefore immortal in the popular sense.

Another complaint one might wish to bring at this point is the one advanced by Vlastos, who says that assertions of the sort that “Fire is always hot” and “Soul is always living” are merely “empirical”—that is, they are not necessary—and thus the argument cannot arrive at a necessary conclusion, as Plato wants. Who knows that we won’t discover eventually some variety of fire that isn’t hot?—Vlastos insinuates—or (granted Plato’s framework) find a soul that isn’t alive?

This objection is perhaps more to the point, but Plato would not lack resources for a plausible reply. On his own terms he might naturally respond, in a Kripkean spirit, that a claim such as “Fire is necessarily hot” is necessary a posteriori, that is to say, what we know, when we know something about fire, is something necessary, even if we might be wrong in our recognition of that necessity. This would be a natural response for Plato to make, not least because he would recognize that something similar happens in our recognition of the soundness of a mathematical construction: when we succeed in recognizing it, this would be the recognition of something necessary, even if (because of “human weakness,” as Plato would say) we are sometimes wrong that a proof is sound. Indeed, in the misology passage Socrates urges his interlocutors to draw exactly the distinction between possible deficiencies in our manner of apprehending an argument and the status of that argument itself.\footnote{And in any case, on Plato’s grounds we can furthermore be confident that something must play the role of an element, if “forms in us” are distinct from the subjects in which they exist, and if there cannot be a regress of subjects all the way down. “There is, so far as I can discover, no explicit or implicit Phaedo explanation scheme for the coming-to-be of an individual,” Robert G. Turnbull objects, “Indeed, if coming-to-be is a function of the striving of individuals and is the coming-to-be of characters in individuals, it is clear that no account could be given in Phaedo terms. It is obvious that an attempt to account for the coming-to-be of individuals along these lines would involve a regress. Suppose, for instance, that $x_i$ is striving to attain $x$, making $x_i$ a kind of ‘haecceitas’ form, it is equally obvious that the individual something else is assumed. And so on.” Robert G. Turnbull, “Aristotle’s Debt to the ‘Natural Philosophy’ of the Phaedo,” Philosophical Quarterly 8 (1955): 131–43. However, the objection need not be framed in terms of haecceities: it would be enough that each time we think we have isolated a simple individual, it turns out that that individual is to be}
Furthermore, Plato might urge that there are interesting disanalogies between “fire is always hot” and “the soul is always alive,” which make the step to a claim about necessity less open to being challenged in the latter case. After all, we have some idea of what it would be for us to be wrong about fire’s being always hot: we have separate criteria for picking something out as “fire” and as “hot,” and the challenge, that perhaps we will discover fire that is not hot, has content. However, the soul in contrast is not perceptible so we seem to lack any idea of what it would be to identify a soul that was not living. It is a supposition so far without content that something identified as a soul would not be alive. Or, one might also say: we may regard the soul, simply, as whatever it is that animates a body. If we had picked out something else, which turns out not to play that role, because it is, itself, sometimes animated and sometimes not, then what has happened is not that some soul has turned out not to be alive, but that we had made an error in identification, and what we took to be the soul proved not to be so.\footnote{We saw that Plato regards a predicate which is applied to a particular as a derivative use of that predicate (it is an “eponym,” he says, see note 53 above); the primary use of any predicate is its application to a Form. However, he seems to suggest, in his discussion of fire and soul as “elements,” that uses of “hot” and “alive” as applied to things other than fire and souls are derivative uses when compared with the application of those predicates to fire and to soul themselves. That is to say, the existence of elements such as fire and soul allow for a kind of mirroring, within the sensible world, of the eponymy relationship that standardly holds between Forms and sensibles. That is why fever gets the name that it has—literally, “firey-ness”—because the heat of the sickness, the fever, πυρετός, is derived from the heat of the element fire (πῦρ) \textit{105c1–4}. We may speculate that this distinction is the source of the Aristotelian notion of paronymy, \textit{Categories} \textit{1a12–15}. His conviction that there are things that function as “elements,” then, is of a piece with his general conviction (shown in the Sun Analogy, Divided Line, and Cave Allegory) that relationships between intelligible things and sensibles are mirrored by relationships between some favored sensibles and all other sensibles.}

Presumably the chief evidence that something like fire is not an element (as we have called it), is that something identified as fire

analyzed as the accidental presence of some participated form in another individual—Plato avoids this difficulty, I think, because he is presuming that there must be elements, and that elemental stuff neither comes into existence nor perishes. His argument is, in effect, that soul stuff has the same status that is usually granted to fire-stuff, earth-stuff, water-stuff, and air-stuff (compare \textit{Phaedrus} \textit{245c–d}).
varies as regards its heat: since then the variation—fire staying the same, but heat varying in degree—would suggest that heat was only something that fire had, that is, it was in fire as in a subject, and that the fire was therefore one thing and the heat something else. Indeed, Plato argues along these lines, in the *Timaeus*, that fire is not a true element, as is commonly thought. In that dialogue he “departs” the traditional “earth, air, fire, and water” from the status of elements because of their evident variation in “purity,” and he adopts the view that fire is simply a qualification of the primordial stuff of the universe, the “Receptacle.” However, as we saw, it was a point emphasized in Socrates’ reply to Simmias that the soul does not vary in degree: one soul is not any more soul, or living, than any other. The suggestion of that passage is that those variations of degree that it does make sense to attribute to a soul amount to variations of goodness and badness, not variations in being a soul or being alive. That is to say, the only intelligible application of qualifications of degree to the soul succeed in saying something, not about the soul, but about its virtue or vice.⁵⁹

So, then, the common objections to this last step that Socrates takes, where he insists, in effect, that the soul is an “element,” are off the point; whereas Plato has ready some plausible replies in response to one objection that, it seems, is to the point.

**IV**

*The Ultimate Final Argument.* This, then, brings us to the passage with which we began and the “Ultimate Final Argument.” The state of the argument so far is this. Socrates has “displayed” or “shown” the soul to be the sort of thing such that, when we say that it is “living,” we are not picking out something that it has, but something that it is; and this what it is, he is prepared to explain in terms of its exerting a kind of “force” which excludes, necessarily, its being living

⁵⁹ Note that, understood in this way, Plato’s point is analogous to his argument for immortality in *Republic* 10. There he argues, in effect, that words that succeed in identifying an evil that is proper to the soul pick out something that succeeds only in making the soul worse, not in making it go out of existence (whereas words that pick out an evil that is commonly taken to make the soul go out of existence succeed in picking out, rather, only an evil that is proper to the body, not the soul).
at the same time being dead. In that precise sense it is immortal. In giving this analysis, he is presuming, he has thereby identified a general class, call it “soul stuff” or “immortal stuff,” which is, as it were, a natural kind and an element: immortal stuff would be the class of everything that is immortal in this precise sense.

As we have seen, he has also given an analysis of relational properties that, he thinks, blocks any inference to a doctrine of universal flux and therefore to a view that would imply that there are no natural kinds or things that possess stable, necessary properties at all. He therefore considers that any argument he now gives will not be open to doubt from Cebes. The grounds of Cebes’ doubt have been removed because soul has been “shown” to be something about which it makes sense to have conviction bearing the character of necessity and universality.

And so, with these considerations in place, he then advances the “Ultimate Final Argument,” which, as mentioned, is a cosmological argument for the everlastingness of a human soul. The first two premises depend upon soul’s being a member of a uniform kind with a necessary property:

1. The human soul is a member of the class of immortals (in the precise sense).

2. What holds of one member of this class, in virtue of its being immortal (in the precise sense), holds of all.

The argument then continues using, in effect, the method of hypothesis. Given that the soul is as described, then which of the following alternatives should next be ascribed to it qua having such a nature? Is its nature such that it carries along with it imperishability (and thus souls are imperishable), or is its nature such that it does not carry along with it imperishability? If the consequences of one of these alternatives are unsupportable, then the other alternative should be embraced:

3. A human soul, in virtue of its immortality (in the precise sense), is either such that it is imperishable or not such that it is imperishable, that is, it is perishable.
4. Suppose that a human soul, in virtue of its immortality (in the precise sense), were not such that it is not imperishable.  

5. Then everything in the class of immortals (in the precise sense) would be perishable.

At this point Plato relies on the commonsense belief that, in the world around us, things are constantly coming into existence and going out of existence. Certainly all perceptibles seem to change and perish. Thus:

6. But everything besides the class of immortals is perishable.

7. But then everything would be perishable.

Plato (I think) presumes that no true potentiality goes forever unrealized, and that a merely accidental relationship could not be sustained without exception always:

8. Thus everything would eventually perish—which is absurd.

This thought is implicit in the passage (σχολὴ γὰρ ἂν τι ἄλλο φθορὰν μὴ δέχοιτο, 106d2–3), but why would it be absurd to suppose that everything perish? Perhaps because Plato is presuming that any account or theory of the world must count as knowledge, but what is known is necessary and what is necessary holds true always—so a precondition of a theory’s counting as knowledge is that it has to hold that something always remains in existence. Or, perhaps, he is

---

60 Notice that the argument requires that we entertain this hypothesis, even if, Socrates thinks, we cannot find its consequences at all tenable. This shows that the Final Argument is not meant to depend on our being able to see, intuitively, an analytic or quasianalytic connection between the concept of immortal (in the precise sense) and imperishable, as some commentators have thought. There simply is no text in the dialogue where Socrates urges us to look for or recognize such a connection.

61 I do not think that implicit also is the thought: “then everything would have already perished in the past, and thus nothing would exist now,” because a prospective argument suffices, and elsewhere in the Phaedo, where Plato has concerns about the continued course of cycles of nature (as in the Cyclical Argument), his concern is prospective only.

62 If a chain only of perishable objects, then the chain, on this view, would somehow have to be itself necessary.
thinking of the universe as akin to an organic system which first
develops to maturity but then continues on with relative constancy
and absence of disturbance after that (which is perhaps the most
common image of the cosmos in pre-Socratic natural philosophy).
However, I rather consider that he is thinking of the construction of a
theory as a piece of workmanship, a kind of model which represents
the original,\textsuperscript{63} so that any account which makes the universe so “frail”
that it collapses and perishes would give us more reason to conclude
that the workmanship of the theorist is defective than that the universe
itself is thus so shoddily put together.

Plato next appeals to individual \textit{immortals}, which perhaps have
special force for him and his interlocutors because they suspect that
God and Forms necessarily exist. We need not consider that there is
any significance in the mention of the Form of Life rather than any
other Form, besides the fact that no special argument is needed to
introduce the Form of Life, as the explanation “the soul is living on
account of its sharing in the Form of Life” is already in the background
and presupposed:\textsuperscript{64}

\begin{enumerate}
  \item Moreover, God and the Form of Life are in the class of
  immortals.
  \item Thus, God and the Form of Life would be perishable.
  \item Thus, God and the Form of Life would eventually perish—
  which is absurd.
  \item Thus, a human soul, in virtue of its immortality (in the
  strict sense) is not perishable.
\end{enumerate}

Now with four brief observations our examination of this argument
may be concluded.

\textsuperscript{63} “That’s why one man makes the earth stay in position by means of
heaven, putting a whirl around it; while another presses down the air as a
base, as if with a flat kneading-trough” (99b): Plato vividly has Socrates
speaking here as if a theoretician, in proposing a model, were a maker of the
thing modeled.

\textsuperscript{64} Nor can it be ruled out that here God = Mind, and the significance of
the appeal to God and to the Form of Life is that these are presumed as being
the first-best and second-best entities to which one can appeal to give any
account of the soul at all. The argument would be emphasizing, then, that the
price of denying the imperishability of the soul would be the rejection of the
possibility of giving an account of anything at all.
First, as was mentioned, this is a “cosmological” argument for the eternity of the soul, which involves identifying the human soul within the class of immortal stuff and then arguing that immortal stuff, considered as a class, cannot perish—because that is the best contender for something that has imperishable status, and the universe needs to be underwritten by something imperishable, if it is to be stable and enduring. As such, the aim of the argument is to place on the same footing, on the one hand, our confidence that the universe itself will continue in existence and, on the other, our confidence that an individual soul will continue in existence. I think that for Plato this represents ground-level certainty, especially if one considers that, for him, the “universe” includes mathematical objects and the forms.

Second, as a “cosmological” argument for the eternity of the soul, the Final Argument ranks along with other such arguments in the Platonic corpus, especially the arguments in the Phaedrus (245c–e) and Laws 10 (895a–896d). However, it is interesting that, in those dialogues, the necessary property of the soul that Plato focuses on is its being self-moving, and his argument is that, if the universe is to continue in motion, then there must be a class of eternal self-moving things, within which the human soul is to be placed as well. Now, it might be that by “living,” in the Phaedo, Plato has been understanding all along something like “self-moving” and that he is not really thinking of “living” and “death” as opposite qualities, akin to “hot” and “cold.” However, if he understands “living” as “self-moving,” then a soul is immortal in the sense that it excludes of its nature being moved by

---

65 Or, put another way, from Plato’s point of view the following two fears are equally rational: upon Socrates taking the hemlock, Socrates’ soul will go out of existence; and upon Socrates taking the hemlock, the universe as a whole will go out of existence.

66 In language which finds echoes in Aristotle’s Categories 1a1–15, in the Laws the logos corresponding to the ousia of the soul is said to be “motion capable of moving itself,” and “soul” and “self-mover” are said to indicate the same ousia (896a1–5): “What’s the definition of the thing we call the soul? Surely we can do nothing but use our formula of a moment ago: ‘motion capable of moving itself.’– Do you mean the entity which we all call ‘soul’ is precisely that which is defined by the expression, ‘self-generating motion’?” Το εαυτό κινείν φῆς λόγον ἔχειν τὴν αὐτὴν οὐσίαν, ἰδίερ τούμορα ὅ δὲ πάντες ψυχὴν προσαγορεύομεν. Trevor J. Saunders, trans., in Cooper, Plato: Complete Works. Emphasis in original.
another. Furthermore, something would be “dead” insofar as it admits of being moved by another. Thus, a corpse would be “dead,” because its motion is explicable solely in terms of motion imparted to it by other bodies; a body would be “living” only indirectly, in the sense that to the extent that it is animated by a soul, to that extent its motions are not explicable solely in terms of the action of other bodies upon it.

Third, the interpretation of the argument above reveals the true weakness in the Final Argument. Suppose that *immortals* (in the precise sense) signifies a genus, and that there are two species of

And thus the *Phaedo* argument would implicitly identify the soul’s being necessarily free with the soul’s being necessarily living. That is, it would become an additional argument, in favor of the soul’s being necessarily living, that we cannot easily conceive of the soul’s not being free. In this regard, see considerations of the sort advanced by Rogers Albritton, in an essay which he begins by quoting, with approval or at least sympathy, Descartes: “Descartes held that the will is perfectly free, ‘so free in its nature that it cannot be constrained.’ ‘Let everyone go down deep into himself,’ he is reported to have said to Frans Burman, ‘and find out whether or not he has a perfect and absolute will, and whether he can conceive of anything which surpasses him in freedom of the will’” (239). Near the end of his essay, Albritton speculates that the source of the perceived conflict between freedom of will and the operation of laws of nature is our conceiving of ourselves, inappropriately, as “objects”: “I do sometimes think I see that logical inconsistency isn’t the real trouble between determinism and free will: that they aren’t in strict truth inconsistent. It then seems to me that their conflict comes rather of our being lured by the hypothesis of determinism into an alarming view of ourselves as a species of *objects*, which as such can’t be thought to *do* anything (in a sense in which we had supposed that we did things), however busy and smiling they are. Automata are just *automata*, after all, whether physical through and through or on the contrary psychophysical” (250). Rogers Albritton, “Freedom of Will and Freedom of Action,” *Proceedings and Addresses of the American Philosophical Association*, 59 (1985): 239–51. For Plato something would be an “object” or “automaton,” in Albritton’s sense, just insofar as it is a moved mover.

It should not be ruled out that a second argument, from cases, runs throughout the dialogue, as a kind of pedal tone, thus: there are only two ways in which something can go out of existence, by diminution (of a quality) or by decomposition (into its elements), but the soul does not go out of existence by diminution because “soul” does not admit of degrees and “living” is something it is, rather than something it has. Yet, neither does the soul go out of existence by decomposition because it is nonperceptible (as the Affinity Argument emphasizes). It therefore endures in the manner of an element, and any suggestion that something can happen to it that results in its nonexistence is incoherent.
immortals, one of which is imperishable because of something peculiar to that species and one which is not, and that human souls are in the latter class. Then, a human soul’s being immortal (in the precise sense) would serve to place it in the genus of “immortals”; its being in that class would not itself place it in the species of imperishable immortals; yet it would not follow that everything was perishable, since some immortals would be imperishable. The false step in the Final Argument, then, is in the inference to 5, above, and the reason it is false is that what is true of a species in a genus need not be true of every other species.

We can clarify this objection by recourse to Plato’s own analogies of threeness:odd::fire:heat::soul:living. If fire is necessarily hot, it does not follow that fire will share everything that is true of anything that is necessarily hot; from the fact that threeness is odd, it does not follow that threeness shares everything that is true of any class of odd numbers. Likewise, if a soul is necessarily immortal, it does not follow that it shares in everything that is true of some class of immortals.

Now, admittedly, we will want to show forbearance towards Plato’s misstep in these matters precisely to the extent that we recognize the philosophical achievement which is Aristotle’s careful delineation of the relationship of genus to species and also his concern to get clear about what precisely makes a genus-species definition a single definition and an essence a single essence. Moreover, it should be considered that in the Phaedo, as we have noted, Plato’s project of accounting for generation and destruction has not even gone so far as to justify the application of a genus-species hierarchical tree to anything in nature. In the dialogue he develops, rather, simply an account of elements, and he is thinking of elements as pure, simple, and not hierarchically ordered under anything else. So the objection we have considered will not have naturally arisen in this context. In a world (so far) not understood as containing true genus-species

---

69 I thank Bob Richardson for drawing my attention to this helpful way of expressing the difficulty.

70 He might therefore insist, as against the objection considered here, that it would make soul-stuff an anomaly, as being the only element which has genus-species differentiation. It would be the simpler and preferable hypothesis, then, that soul-stuff is entirely uniform, but that result would fall short of the necessity he is trying to achieve.
hierarchies, which serve as objects of knowledge, the supposition that there could be something like that would seem untoward and odd. Furthermore, our point above about self-motion perhaps should be brought to bear here as well: that is, if Plato is implicitly supposing that “living” simply amounts to the same thing in the end as “self-moving,” then he might naturally take souls (immortals) to constitute a single, uniform class, just as (it would seem) something is either fully characterized in its nature as a “self-mover” or not.  

However, a fourth and final consideration is this: we might wonder whether, when we probe Plato’s beliefs about causation and explanation, we do not discover another deep line of thought implicit in the Final Argument. Consider his conviction that only an appeal to Mind (intelligence, reason, understanding, design) is explanatory—a conviction which shows up repeatedly, for instance, in the Philebus, where he says it is “not pious” [οὐδὲ ὁσιόν], 28e2, to leave the universe to be governed by “unreason and irregularity” and that one should hold, rather, that the universe is “ordered and governed by a marvelous intelligence and wisdom” [νοῦν καὶ φρονήσιν τινα θαυμαστήν], 28d8. “Mind forever rules over everything” [αἰτὶ τοῦ παντὸς νοὺς ἀρχεῖ], 30d8, Socrates says there; it is “responsible for everything” [τοῦ πάντων αἰτίον], 30e1. In Laws 10, again, it is soul that is postulated as being “responsible for all things” [τῶν πάντων αἰτία], 896d8.

Now suppose Plato thinks something such as this: we are minds, and a mind can only understand what is akin to itself. Thus, to explain an effect would require that we regard it as the effect of a mind, or as somehow itself something mind-like, but then, to suppose that the universe as a whole is intelligible, would be to suppose that it is either the effect of a mind or itself mind-like (or both). However, if

---

71 Indeed, Socrates in the Philebus declares that “mind is in the class of things that is responsible of everything,” as if there were just one such class (νοῦς ἐστὶ γένους τῆς τοῦ πάντων αἰτία, 30d10–e1).

72 Hence, Socrates’ word-play in referring to Anaxagoras as “a teacher of the causes of existing things according to my own mind” διδάσκαλον τῆς αἰτίας περὶ τῶν ὄντων κατὰ νοῦν ἐμαυτῷ, 97d6–7. In addition, when Socrates uses himself sitting in prison as an example to illustrate the form that explanations of the universe should take (“he seemed to me to have proceeded in precisely the same way as” μοι ἔδεξεν ὁμοίωσιν πεπονθέναι ὡσπέρ, 98c2), this invites us to formulate the analogy: Socrates’ mind : his body :: the cosmic Mind : the universe as a whole.
the universe as a whole were like that, then, because of the kinship of the human mind with it, the human mind would share also in other characteristics that would reasonably be attributed to the cause of the universe as a whole, such as everlastingness.\footnote{From its start, of course, Greek natural philosophy treated the chief ἀρχή as everlasting.}

Now, the beginning of Greek natural science involved precisely the rather bold presumption that the world is a “cosmos,” that is, an ordered whole, which admits in its entirety of a single, satisfactory explanation, and that the human mind is capable of in some manner grasping that explanation. If the line of thought that I have just sketched really does underlie the \textit{Phaedo} and also Plato’s critique of the way in which the project of pre-Socratic natural philosophy was in fact executed, then another way of understanding the \textit{Phaedo} is that it is Plato’s working out of an implication for human life and practice, starting from a presupposition that he thinks we must all adopt, if there is to be coherent natural science at all. The \textit{Phaedo} would be a rudimentary treatise in “dualistic physics,” which, at one and the same time, aims to place natural science on a sound foundation and leave ample scope for the distinctive activity, autonomous existence, and divinization of the knowing human subject.

\textit{Institute for the Psychological Sciences}