

## 3.4 Plato's *Phaedo*: Theory of the Forms



*“Plato’s Dialogues seek to form, not inform, their readers.”*

### Socrates and the Theory of Forms (95e-107b)

In the last section of the “Logos” part of the dialogue, Socrates (i) offers an extended version of his intellectual biography which purports to show how he came by his dialectical approach to philosophy, (ii) introduces both a simple and more sophisticated theory of the Forms, and (iii) applies this framework in his final argument for the immortality of the soul. This section completes the discussion of Plato’s metaphysics in the *Phaedo*, by indicating how the Forms might operate *within* the physical world, as the immanent *natures* of things, in addition to their transcendent being as objects of knowledge.

### Socrates’ Philosophical Autobiography (95e-99d)

In response to Cebes’ argument, Socrates says that he must undertake a *“thorough-going examination of the causes of generation and destruction, and of being”* (cf. 95e, 96a). For Cebes’ fears go even deeper than Simmias’ fears and offer a more grave threat to philosophy, because while Cebes concedes that body and soul are different, i.e. he agrees human life is different in kind than mere physical existence, he insists this does not imply the soul will live forever. (Here again Cebes seems to represent the authentic voice of natural man, who does not blink at the finality of death, but looks it square on, if with trepidation.)

Socrates proceeds to tell the story of his own intellectual biography, reporting that as a young man he was keen on the questions of why things *“come to be, perish, and exist”*—i.e. questions about the **“becoming and being”** of things, the kinds of scientific and metaphysical questions that began with the Presocratics. He begins by listing a series of questions which concerned him then, such as (i) the relation of the physio-chemical and biological-organic aspects of living beings, (ii) the relation of ‘body’/brain and soul/thought, and (iii) the relation of sense-perception and knowledge (the empiricist theory of knowledge). (Compare these questions to those we raised concerning Atomist science and the world of human common sense at the end of Chapter 1.)

Socrates became unsatisfied with how the Milesian-type natural scientists answered these questions, however, because they reduced all explanations to strictly material and causal factors. As a result the “physicists” did not consider the very different relation idea of ‘causation’ or explanation involving *concepts* and what and why *mathematical things* are what they are, e.g. numbers and their relations (96a-97b). (This discussion shows that Socrates was not interested only in ethical questions, but that his ethical inquiries were related to his puzzlement about “the Whole,” perhaps in particular the curious difference between the ‘*number*’ *model of unity*, and that found in kinds: in number the two is a new thing, which arises out of and is different from the two ones which compose it, whereas in relation to kinds, the unity between them is exactly the same in each, cf. esp. 97ab.)

In the midst of this confusion, Socrates came upon Anaxagoras’ theory that a Divine Mind directs and is the cause of all things, acting to bring about a good end (*telos*). Socrates found this theory appealing, since it suggested that when something comes to be or is destroyed or exists, one should look for why that was *best* for that thing to be the way it was, not merely its physical causes (97b-98b). (These would be reasons or teleological explanations.)

But neither the scientific-causal or the theological-teleological way of understanding reality, at least as Anaxagoras explained it, proved ultimately satisfactory to him, though Socrates does insist, in regard to *human* conduct, that we must distinguish the “*true causes*”—the *reasons* persons have for acting—from the “*necessary conditions*” for their action—the physical basis and movements of the actions (98b-99d). Thus, Socrates appears to affirm the “two worlds” or “two aspects” theory mentioned in connection with the Affinity Argument. Whereas material things and their interactions might possibly be explained in terms of physical causation—he implies it may also not apply to living things—in the case of human beings in any case, he claims, reasons in the form of beliefs concerning what is good do function as causes of our actions, and this is related to the claim that persons are moral agents, who cannot simply be identified with their bodies. Any attempt to articulate “the whole” which reduces human actions to strictly physical events ignores the human aspect of the whole, and thereby distorts the truth.

Having discovered (i) the limitations of *scientific materialism*, which told him only about non-living nature, and being puzzled by (ii) the nature of *mathematical knowledge*, and then also (iii) discovered the limitations of *dogmatic theology*, which promised more than it could deliver, Socrates decided he would (iv) “*take refuge in logos*” and *investigate the truth of things through dialogue*. This meant, first, that he would approach his understanding of the whole by understanding, *from the perspective of the human agents in dialogue*, “what each of the beings is.” Socrates would not ignore the manifest differences among the way the beings

appear, as his Presocratic predecessors had, with the result that they ended up with fantastically partial conceptions of the whole; he would begin with “each of the beings” and attempt to understand the whole in light of them.

Furthermore, Socrates says that would take as his *hypothesis* for understanding each particular thing whatever he regarded as the most convincing human opinion or theory about it and inquire, using his method of dialectic, into what was consistent and inconsistent with this theory—though he by no means concedes that this way of approaching the truth of things is deficient as compared to the direct (empirical) approach of natural scientists or to the dogmatic approach of the theologians, each of whom, albeit in different ways, seems to be blinded by the objects of their speculations (99d-102a). (Theologians are blinded by their presumption of a Creator, natural scientists by their presumption that everything can be understood by natural science.) It is at this point that Socrates introduces the Forms.

### What is the Theory of Forms? (99d-102a)

Socrates presents the Theory of Forms in two stages: (i) a “simple” version and (ii) a “sophisticated” version. The “simple-minded” version of the Theory of Forms has four components:

1. There are Forms, F-nesses. E.g. there are the Forms of Beauty, of Goodness, of Greatness, etc.
2. If we say, “x is F,” then F is the name of a Form, F-ness. E.g. When we say, “Athens is beautiful,” beautiful is the name of the Form of Beauty.
3. A particular thing is F by virtue of “participation” in F-ness. E.g. Athens is beautiful by participating in Beauty.
4. The Form, F-ness, is itself perfectly F but things “fall short.” E.g. the Form is perfectly Beautiful, but Athens falls short.

This theory begins with (i) *words*, especially nouns, which bespeak (ii) universal *concepts* that categorize (iii) particular *things* but refer to (iv) the *Forms*. The idea is that we use the words of particular languages, e.g. “beauty,” to refer to many things, e.g. beautiful paintings, beautiful people, beautiful laws. This one word, Socrates suggests, corresponds not only to those many things and to the concept we have of that word in our minds when we speak and understand each other, but to something else as well—the Universal Form or Idea of Beauty. Thus if we grasp the true definition of beauty and we say Athens is beautiful, it would be so because Athens “*shared in*” the absolute or true meaning of beauty. (Note: the Greek word for beauty, *to kalon*, also means noble. To say “Athens is beautiful” does not just mean it is physically beautiful, e.g. the Acropolis, but it also says it is morally beautiful or noble.)

Metaphysical **nominalists** and **relativists** like Protagoras deny there are absolute ideas of justice or goodness or beauty—Athens is morally beautiful to

the Athenians, perhaps, but not to the Spartans or to Plato. They would insist that *all* words for moral concepts are culturally relative and have an “ambiguous meaning” or “*open texture*.” There may be some family resemblance between the different meanings that human beings subjectively assign to those words—the different conceptions people have of them—but there is no one ideal meaning they all ought to refer to, no one meaning they all ought to connote.

Milesian-type natural scientists also deny there are ideal standards of physical things. They were empiricists in the sense that Socrates describes in his autobiography: our concepts of physical things are generalizations we form from sense-experience, regularizes we observe in the natural world. (Atomists of course admit there are fixed meanings for natural things, e.g. the fixed meaning of water is the atomic structure that underlies what we today know as H<sub>2</sub>O; but they discard the evaluative and related words Socrates is most interested in, such as “mind” and “goodness” and “beauty.” For them emergent entities are reducible to their simpler components. See p.34.)

Socratic inquiry presupposes, as we saw earlier, that there is a *real* definition for moral ideas such as courage, temperance, justice, wisdom, beauty and goodness. Platonic realists assert there are not only normative definitions of these properties but that there exist\* *universal ideals* which these words and concepts point toward, Ideas or Forms by which we should govern our lives and find and create meaning in the world.

But it is not hard to see how the Theory of Forms might be appealing. The Platonist believes the concepts of geometry and number are not found in the physical world, but empirical things can and must be conceived in terms of them and judged to “*fall short*” of them, like the equal sticks at 74a-e. Likewise, ideals of moral beauty and justice and goodness may not be found in perfect purity in the world, but individuals and communities ought to strive to attain them, even if we always fall short of our ideals. In Plato’s view, these efforts are only truly understandable in relation to transcendent Ideals that illuminate the world of knowledge and moral life. The Theory of Forms is the metaphysical *hypothesis* which explains (i) the metaphysical ground of moral and rational order, and (ii) the knowledge we possess of these orders.

In this sense, the Theory of the Forms is the “Ariadne’s thread” Plato holds out in the *Phaedo* which may lead us out of the cave ruled by the minotaur of cynicism, nihilism and the fear of death, into the “true earth” of philosophical hope and love. The Forms are the aspects of Truth, the Whole or God—even if language obscures our understanding of them. The experience human beings have of seeking to know and act in the light of “the Universal” is how we touch immortality. Plato was committed to a Theory of Forms throughout his philosophical life, despite questions he had about it.

### Plato's Theory of Forms

The Theory of Forms says there is (i) the name or word, circle; (ii) the definition (*logos*); (iii) particular images or diagrams; (iv) knowledge of what a circle is; and (v) the abstract object or Form, both knowable and existent. The name is ambiguous, but its true meaning points to the Idea. (*7th Letter*, 342b.) Plato offers three chief arguments to defend his Theory of Forms:

- **Argument from Universals:** If people use the same word (e.g. "justice"), something must make it the same reality they are talking about, or they would be speaking private languages. But people constantly use the same word for different objects or events. Therefore there is a universal they refer to, the "one" over the many instances, the Form. This argument is rejected by relativists, but relativists must admit that there are universals, perhaps even moral universals (e.g. they must agree there is an idea of justice which is not the same as the idea of law, if they agree there can be just and unjust laws).
- **Argument from Imperfection:** We conceive of perfect, abstract things such as the Equal, Circle, Square, etc. These concepts cannot derive from sense-perception, because we realize the particular things we perceive through the senses "fall short" of corresponding perfectly to them. But they are not fictions, like centaurs or Zeus. Therefore the Forms exist, i.e. abstract entities that have the characteristics we have in mind, and they are what we are thinking of, if thinking rationally. (*Phaedo* 74-76)
- **Argument from Knowledge.** If scientific knowledge exists, that knowledge is expressed in universal, necessary truths, not merely particular, contingent truths based on generalizations from things we perceive. But it is clear that such knowledge exists—viz. the knowledge of mathematics, logic, ethics—and therefore their truths must refer to realities, i.e. the Forms. If there were no Forms, there could be no universal and necessary knowledge; there could only be general patterns of experience, which could always be mistaken or change.

Both Plato himself and his critics offered arguments against the Forms.

- **Reification:** The Theory confuses concepts with objects. The Forms are thoughts or meanings in the mind; they aren't beings, real things that exist. To think of them as things is to reify them, rather than recognize they are concepts.
- **Population:** Do all kinds of Forms exist? Are there Forms of (1) natural things like Snow, Mud? (2) artifacts like Table, Spear? (3) comparatives like Great and Small? (4) mathematical and logical concepts like Circle, Same? (5) moral concepts like Justice, Goodness? Where does a concept like Man belong? Is man an empirical concept, or does "man" imply an ideal, what man ought to be or become? Platonists have disagreed about this issue—some claiming that only Forms of types (4) and (5) exist, others that they must all exist.
- **Self-reference or "Third Man":** If there are Forms for each kind of thing for which there are names, e.g. men, and therefore e.g. a Form of Man (the 2nd Man), isn't there a Form of the new group, composed of men + Man, i.e. another Form, a 3rd Man? (This problem may be related to one below.)
- **Ambiguity:** Whatever may be true about the Forms, it seems impossible to express it unambiguously. A term like "square" refers both to the physical world (to physical squares) and to the idea of the Square (which is ideal and non-physical), similarly, e.g. justice or terms like identity or difference. Doesn't this make it impossible to express truths about the Forms in ordinary language?

What do you think of Plato's Theory of Forms? Which arguments make the most sense to you, and why?

### The Theory of Forms and the Final Proof (102b-107b)

Socrates then introduces a “*sophisticated*” variation of his Theory of Forms (102b-105b). This theory has three chief features:

1. *The character of being F in a particular thing never admits its opposite, but either (a) flees and retreats; or (b) is destroyed.* E.g. The tallness-in-Simmias, which he has compared to Socrates, does not become short, when he becomes short in comparison to Cebes. Tallness is an “accidental” property of Simmias.
2. *There are certain things which always are F and never admit the opposite of F.* e.g. Snow is always cold and never admits the Hot; fire is always hot and never cold; triads are always odd and never admit the Even. Coldness, then, is an “essential” property of snow.
3. *Therefore a thing may be F not only because it participates in the Form, F-ness, but also because it has that which always participates in F-ness.* E.g. A body may be hot not only because it “share” in Hotness, but because it has fire (which is always hot) in it; a number may be odd not only because it participates in Oddness, but because it has oneness (which is always odd) in it.

The distinction between the *accidental* properties things happen to have and the *essential* properties things have as part of their nature sets up Socrates’ last argument, in which he links the Form of Life to Soul, which is always alive.

Socrates has added something new to his picture of things, which in the “simple” theory had three kinds of entities: (1) material things, (2) Forms, and (3) characters (the form in the thing). Considering “snow” and “fire,” perhaps also “fever,” and soon, too, “soul” (*psyche*)—Socrates seems to be saying that there are also (4) natures which make the things which bear them have certain properties essentially, i.e. they appear to be the Forms ‘at work’ in things. Natures introduce an intrinsic stability to the changing sensory world, in addition to that brought about by the Forms. (We do not experience particular things *directly*, but always *as* objects of mind, constituted by percepts and by concepts which arise from our rational nature and human experience.)

Socrates uses the new theory to complete his argument (105b-107b):

1. It is soul which, when present in a body, always brings life.
2. So soul never admits the opposite of life, i.e. death.
3. Therefore soul is deathless.
4. But the deathless is also necessarily indestructible.
5. Therefore soul is indestructible, immortal.

This argument leaves two questions unanswered: (i) Why is the human soul, so closely associated in most of the preceding proofs with moral and intellectual life, now reduced to its relation to Life? (ii) If we regard the soul this way, why must we conclude that soul, which by its nature always brings life, gives way on the approach of death and destruction, withdrawing rather than perishing?

We naturally think, for example, that snow perishes rather than withdraws when heat approaches, even if nothing snowy can be anything else but cold. Might it not also be the case that nothing that has *psyche* can be anything but alive, and yet that *psyche*, to bring life, must be present *in some body*? We might agree that soul, *wherever* it exists, is deathless and that “*the God and the Form of Life*” are never destroyed, but deny *individual* souls are indestructible, if we suppose that for them to exist, they must be *in* or *of* a particular body, and that when death approaches they do not withdraw but cease to be.

This interpretation of the final argument would be consistent with the biological interpretation of Cebes’ Weaver Argument and with a view of soul as mortal, not immortal. The soul would be the unifying life-force or governing design in a living thing that re-wove its own body many times before it finally was spent and perished, together with its body. But this conception of soul, organic functioning, is very different from the one linking it to the capacity for knowledge and virtue. This conception of soul-nature seems elemental, merely being what it *is*, rather than intentional striving for what it is *meant to be*, its perfection and good. Might there be a stronger argument for the immortality of the soul, based on that aspect of human being?

And so the *logos* ends with a proof that Socrates acknowledges may have weaknesses (107b). In any case, it is still necessary to re-examine the hypothesis on which it rests—the Theory of Forms. Do the Forms exist, particularly the Forms of Beauty, Justice and Goodness, and if so, how do we know? How do material things and human beings “participate” in the Forms? Is it even possible to talk about Forms without seeming to make them into things? If there are such Forms, why don’t we all agree on them, rather than seem to view them—at least the moral Forms—so differently, as Protagoras says? These questions raise doubts about the Forms.

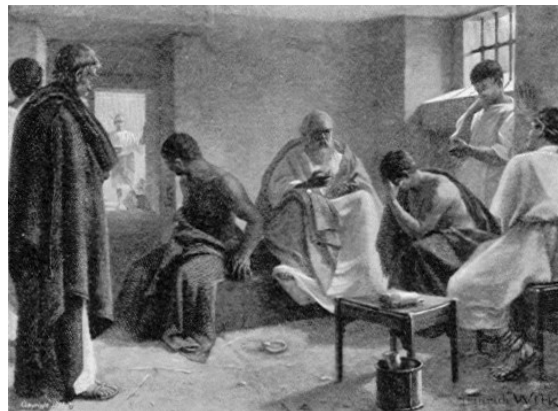
On the other hand, if something like the Forms—some kind of “transcendent” and “ideal reality”—does not exist, how can there be mathematical science or universal morality, each of which seems to point beyond the contingent particulars of the empirical, natural world? Should we *assume* the Forms exist, both to engage in philosophical inquiry, and to try to guide our lives by their light?

Socrates urges his companions to continue to philosophize and examine these assumptions with one another, with the promise that, if they do, they will eventually agree with the argument for the immortality of the soul, “*so far as it is possible for man to do so.*” But this is not the end of Socrates’ philosophical “*swan song.*” He explains that there is a final myth—a popular kind of philosophical music—he still has to tell, to complete his overall defense of “good cheer” in the face of death. One part of his final apology to his friends is complete; he must now add the other. (See 60c-61b, 84e-85b, 114d.)

### ***Mythos* and Death (107b-118a)**

Having finished his *logos*, Socrates turns to a *mythos* of “the true earth.” We live not in the pure, true world, but in the chasm (111c, the Greek word means “cave”) of the under-earth. (The Greek Hades, *A-eides*, etymologically means the world without forms or light.) The earth itself is pure and lies in the pure sky where the stars are situated, the *aether*. If anyone got to this upper limit, he would know that it is the true heaven, with the true light, and the true earth. At present however our view of things is obscured, their real being barely known. We are either blinded by the shadows we see, confusing them with reality (like the scientific materialists), or we fearfully reject them, fleeing to an other-worldly “heaven” (like the dogmatic theologians).

Socrates goes on to describe the inner earth in poetic terms, describing surging rivers that make it seem like a beating heart. He completes the myth with a Pythagorean tale of judgment in the afterlife—repeating the promise he made at the very beginning, that those who follow philosophy may be vouchsafed a happy home. (Compare the ending of the *Gorgias* 527ab.) What is the significance of this *mythos* to the teaching of the *Phaedo*? Why does Socrates include it in the conversation he has with his young friends, including Phaedo, on the last day of his life, in the final moments before his death? Should it also be considered an important part of his philosophical legacy? For someone who all his life relied on philosophy and logic, not poetry and myth, Socrates seems curiously willing, at the end, to include both *logos* and *mythos* in the philosophical life.



Socrates in prison, just before death

After cautioning that “no sensible person” would believe his words literally, Socrates adds that he thinks it is fitting to risk faith (*pistis*, “trust”) in something like this, and to repeat it to oneself as an incantation and be of good cheer in facing death. It is as if Socrates realizes the liberating power of *logos* is not enough for Phaedo and many Socratic companions: they need a *mythos* to

sustain them in moments of doubt, as the passions evoked by death arise in the dark nights of the soul.

In fact, the legacy of the *Phaedo* is not only its contribution to philosophy, but also its contribution to a new religious image of human reality, of human souls as immortal and akin to reason and God, transcending the lower realm of bodily sensation, egoism and power. The task of separating out what in the *Phaedo* is rational (*logos*) and what is mystical (*mythos*) is not easy. In any case, the rational part may not be sufficient to sustain hope in the face of death, if it is not accompanied by the actions of a life inspired by philosophical virtue and shared together in philosophical friendship.

But now the time of Socrates' own death is at hand, as this *logos* comes to an end, and the last act must be played. In yet another touch of high art, Plato has Socrates' old friend Crito ask him, "*How shall we bury you?*" and lets Socrates make light of it, saying "*Any way you can, if you can catch me and I do not escape you.*" For if Socrates *is* his soul, as he has been arguing these past many hours, what is there of him to bury but his cloak? Might he perhaps remain with us after all?

The final scene is dominated by the feelings friends would naturally have as their teacher was nearing death. When he is with them, serene and powerful, the *logos* is alive, but when he is away, it is threatened and darkness looms. His companions are cast in doubt again, as he leaves to spend his last moments with his wife Xanthippe and family. When Socrates returns, the prison guard approaches, explaining that, reluctant as he is, it is time for Socrates to take the poison. Crito urges Socrates to postpone his death, cling to life a few moments more. But Socrates orders the cup be prepared, and, after a suitable prayer, calmly drains it. At this point his companions weep, and Socrates must remind them, in the voice of ancient virtue, to control themselves. The hemlock creeps up his body, and he dies—though just before the final moment, he asks Crito to sacrifice a cock to Asclepius, the patron saint of health and medicine.

This final exchange has always puzzled Plato's readers. Is Socrates expressing thanks for a long and healthy life, indeed one as a psychotherapeutic healer himself? Is he saying that now, in leaving his mortal coil, he is attaining the true health of immortal being? Or is it an allusion to the ancient myth of Asclepius' having discovered the art of raising the dead? We do not know.

The final word of the dialogue is left to the narrator, who calls Socrates as "*the best, the wisest, and the most just*" of men (118a). There is no doubt that this *individual* man has died. Whether his *logos* will live on depends on us.