FACULDADE DE CIÊNCIAS SOCIAIS E HUMANAS
UNIVERSIDADE NOVA DE LISBOA

Masters, Slaves and Philosophers

The Origin and Fate of a Platonic Idea

Bernardo Ferro

A thesis submitted for the degree of
Doctor of Philosophy

November 2017

The following dissertation, submitted with the purpose of obtaining a Doctor's degree in Philosophy, was written under the scientific supervision of Professor Mário Jorge de Carvalho.

Masters, Slaves and Philosophers

The Origin and Fate of a Platonic Idea

Keywords

Freedom, power, philosophy, Plato, Hegel, Nietzsche, knowledge, truth, belief, doxa, master, slave, philosopher, sovereignty, mastery, slavery, imprisonment, deliverance.

Abstract

Knowledge is power, so the saying goes, and power consists in the freedom to choose and act according to one's interests. But although in everyday life this basic freedom is usually taken for granted, its true extent is far from self-evident. The debate concerning the nature and the amount of knowledge required for the actualization of freedom has a long philosophical history, whose first major chapter was written by Plato. In his dialogues, Socrates makes three momentous claims about human freedom. First, he describes human life as a peculiar form of imprisonment, and human beings as a community of prisoners or slaves. Moreover, he argues that this vital imprisonment is not due, primarily, to concrete or external constraints, but to the inner structure of human reason. Finally, to correct this situation, he calls for a global re-education of human consciousness, with direct implications in all aspects of human life.

The present study confronts Plato's project with the work of two modern authors, Hegel and Nietzsche, who have explicitly linked the notions of ignorance and truth to those of bondage and freedom — or slavery and mastery — and whose philosophical projects are also centred on the double task of educating and liberating human consciousness. Like Plato, both of them recognize a fundamental gap between human beings' usual outlook on reality and a true outlook on reality. Natural reason provides a kind of knowledge that does not get at the truth of things, but only at an appearance or semblance thereof. Furthermore, for all three authors, philosophy's main task is to correct this fault, either by raising one's usual standpoint to the standpoint of truth or by

questioning the very notion of truth such a standpoint is supposed to embody. Finally, and more importantly, Plato, Hegel and Nietzsche all agree that a philosophical reform of human cognition is a necessary condition for the attainment of a free and independent life.

Beyond these similarities, shared by other authors and doctrines, what sets Hegel and Nietzsche apart from all of Plato's heirs is the resolutive ambition of their philosophical projects. Whereas Plato raised the issue, giving rise to a long critical tradition, Hegel's and Nietzsche's explicit purpose is to move beyond the critical stage and to set humankind free. But their methods point in opposite directions, aiming at radically different forms of deliverance. Hegel builds on Plato's standpoint and sets out to fulfil its main promise. His aim is to solve reason's contradictions by eliminating human ignorance and inaugurating thereby a new, post-critical brand of philosophy. Nietzsche, on the other hand, rejects Plato's starting point and sets out to overturn his philosophical legacy. His aim is to solve reason's contradictions by challenging the very contrast between reason and unreason, and establishing thereby a new, post-rational or post-cognitive brand of philosophy.

Contents

Acknowledgements vi
Abbreviations vii
Quotations ix

1. Introduction 1
    A brief anatomy of freedom 3
    Freedom and education 17
    Origin and fate of a Platonic idea 29
    Historical affinities 53

Part I: Plato's Scientific Project 61

2. Plato's idea of truth 63
    Truth, opinion and doxa 64
    Self-confidence and self-interest 78
    Body and soul 89

3. Masters, slaves and philosophers 103
    Rhetors and philosophers 104
    Socrates and Callicles 112
    Masters and slaves 119

4. Chains and shadows 133
    Inside the cave 136
    Towards the light 147
    The self as shadow 156

5. Philosophy and deliverance 161
    The need for deliverance 163
    The possibility of deliverance 169
    Plato and Hegel 177

Part II: Hegel's Phenomenological Project 183

6. Hegel's *Phenomenology of Spirit* 185
    Truth and appearance 187
    Science and experience 199
    Further hermeneutical issues 209

7. Selfhood and otherness 219
The road to self-consciousness 220
The dialectic of self-consciousness 237
Selfhood as recognition 244

8. Freedom and servitude 251
Mastery and slavery 253
Slavery and philosophy 260
The unhappy consciousness 270

9. Towards freedom 285
Reason and Spirit 286
The trials of Spirit 303
Tragedy and comedy 317
A new beginning 324
Hegel and Nietzsche 333

PART III: NIETZSCHE'S METACOGNITIVE PROJECT 339

10. Nietzsche's idea of truth 341
Truth and untruth 344
Truth and metaphor 353
Science and wisdom 364

11. Philosophy in chains 379
The Socratic turn 380
Athens and Jerusalem 386
Masters and slaves 394
Asceticism and nihilism 403

12. Philosophy unchained 409
The will to power 410
Return to Callicles 418

13. Conclusions 427

Selected bibliography 431
Index nominum 445

Acknowledgements

First and foremost, I am deeply grateful to my supervisor, Professor Mário Jorge de Carvalho, for his unrelenting support, wisdom and encouragement.

I am also indebted to the Portuguese Foundation for Science and Technology (FCT) for the financial support granted during the past three years and to the German Academic Exchange Service (DAAD) for financing my one-year stay in Berlin, at the Humboldt University.

A special mention is due to Professor Tobias Rosefeldt, Chair of the Classical German Philosophy section of the Humboldt University's Philosophy Department, for letting me attend his weekly seminar, as well as to Professor Bernhard Thöle and to Franz Knappik for their interest in my work.

I would also like to thank Professor Jacinto Rivera de Rosales, of Madrid's Universidad Nacional de Educación a Distáncia, for his help and support.

Finally, I am grateful to my family and friends for their support and encouragement, and namely to Lucas Manarte, Martim Enes Dias and Pedro Teixeira for their help in the final stages of this work.

Abbreviations

Works by Plato

Al	*First Alcibiades*	*Phd*	*Phaedo*
Ap	*The Apology*	*Phdr*	*Phaedrus*
Cl	*Cleitophon*	*Phi*	*Philebus*
Cra	*Cratylus*	*Pr*	*Protagoras*
Cri	*Crito*	*Re*	*The Republic*
Eu	*Euthyphro*	*So*	*The Sophist*
Go	*Gorgias*	*St*	*The Statesman*
La	*Laws*	*Sy*	*Symposium*
Ly	*Lysis*	*Th*	*Theaetetus*
Me	*Meno*	*Thg*	*Theages*
Pa	*Parmenides*	*Ti*	*Timaeus*

Works by Hegel

ETW *Early Theological Writings*, trans. T. M. Knox (Philadelphia PA, University of Pennsylvania Press, 1971)

FK *Faith and Knowledge*, trans. Walter Cerf and H. S. Harris (Albany NY, State University of New York Press, 1977)

HW *Werke in zwanzig Bänden*, eds. E. Moldenhauer and K. M. Michel (Frankfurt am Main, Suhrkamp, 1986)

PM *Philosophy of Mind*, trans. W. Wallace and A. V. Miller (Oxford, Oxford University Press, 2007)

PR *Outlines of the Philosophy of Right*, trans. T. M. Knox (Oxford, Oxford University Press, 2008)

PS *Phenomenology of Spirit*, trans. A. V. Miller (Oxford, Oxford University Press, 1970)

Works by Nietzsche

A	*The Antichrist, Ecce Homo, Twilight of the Idols*, trans. Judith Norman (Cambridge, Cambridge University Press, 2005)
BGE	*Beyond Good and Evil*, trans. Judith Norman (Cambridge, Cambridge University Press, 2002)
BT	*The Birth of Tragedy and Other Writings*, trans. Ronald Speirs (Cambridge, Cambridge University Press, 1999)
GM	*On the Genealogy of Morality*, trans. Carol Diethe (Cambridge, Cambridge University Press, 2006)
GS	*The Gay Science: With a Prelude in German Rhymes and an Appendix of Songs*, trans. Josefine Nauckhoff (Cambridge, Cambridge University Press, 2001)
HAH	*Human, all-too human*: *A book for free spirits*, trans. R. J. Hollingdale (Cambridge, Cambridge University Press, 1996)
KSA	*Kritische Studienausgabe*, eds. G. Colli/M. Montinari (Berlin, Walter de Gruyter, dtv, 1988, 2nd revised ed.)
TSZ	*Thus Spoke Zarathustra*, trans. Adrian del Caro (Cambridge, Cambridge University Press, 2006)

A Few Notes on the Quotations

With regard to Plato's works, all translations, unless indicated otherwise, were taken from Harvard University Press' edition (Cambridge, MA: Loeb Classical Library, HUP). Quotations and citations are identified by their Stephanus numbers and followed, whenever relevant, by the line numbers.

With regard to Hegel's works, citations and original quotations are referenced by one of the abbreviations listed above. Translated quotations are referenced by the abbreviations of the German edition and the English edition, separated by a semicolon. Where no English edition is indicated, the translation is mine.

With regard to Nietzsche's works, citations and original quotations are referenced by one of the abbreviations listed above. Translated quotations are referenced by the abbreviations of the German edition and the English edition, separated by a semicolon. Where no English edition is indicated, the translation is mine.

With regard to the works by other authors, where no English edition is indicated the translation is mine.

Unless otherwise indicated, all quotations retain their original emphases.

... Οὐδ' ἐδέρχθης
ὀλιγοδρανίαν ἄκικυν,
ἰσόνειρον ᾗ τὸ φωτῶν
ἀλαόν γένος ἐμπεποδισμένον;

Did you not see the helpless infirmity,
no better than a dream, in which
the blind generation of men is shackled?

AESCHYLUS (?), *Prometheus Bound* 546-549

1.

INTRODUCTION

Out of all the philosophical issues comprised in the history of Western thought, the issue of human freedom — of its truth, nature and scope — is undoubtedly one of the most important. However, as is often the case with philosophical issues, it hardly ever comes to mind in the course of most people's lives. Unless one's freedom is somehow called into question and must be reasserted or redefined; or unless some practical need leads one to focus on, or to wonder about someone else's freedom; or unless one is a professional philosopher, or barrister, or magistrate, and makes a living out of dealing with such matters — the questions of whether, how and to what extent human beings are free are hardly ever asked. And this is so because the aforementioned situations are not the rule, but the exception. Most people's lives are spent dealing with more mundane issues and more immediate concerns. Although the importance of freedom is generally acknowledged, its definition is usually regarded as a theoretical problem, and even when one appreciates the full extent of its practical implications, the fact remains that a thorough answer to the questions of whether, how and to what extent human beings are free is simply not needed when, say, moving around, speaking or pouring a cup of coffee.

But even though freedom as such is not often defined or discussed, it is nonetheless an essential component of most people's lives. For although one does not need to wonder about the philosophical meaning of freedom in order to move around, or to speak, or to pour a cup of coffee, one does all of these things in the assumption that one is free to do them — or not to do them, or to do different things. Precisely because it is so familiar, the ability to think and act freely is usually merely supposed, and not explicitly

acknowledged. Just like breathing or having a body, it ranks among the most basic requirements of human experience and provides the ground whereupon most thoughts and actions are made to rest. In other words, freedom is part of the *medium* or the *atmosphere* in which human life usually unfolds, which means not only that most human beliefs are silently supported by the notion of freedom, but also that this notion is rarely, if ever, truly problematized. Whenever freedom is questioned, criticized, or even found to be lacking, one tends to perceive one's lack of freedom as a contingent loss, which calls for concrete action and for specific forms of liberation. It is at best an ethical or political problem, leading to disagreement and strife, to war and revolution. Paradoxically, however, the freedom to arrive at this diagnosis — the freedom that allows one to identify and expose one's occasional lack of freedom — is still taken for granted. Only the surface of the issue has been stirred, while its inner core remains untouched.[1]

A more serious investigation into this matter is bound to produce different and more daring results. If freedom is indeed a basic ingredient of human life, the question of whether we are as free as we take ourselves to be must have a direct bearing not only on

[1] The life-long need to take one's freedom for granted was specifically highlighted by Kant, who saw it as an expression of what he called, following A. G. Baumgarten and G. F. Meier, 'natural metaphysics' (*metaphysica naturalis*, or *Metaphysik als Naturanlage*). See notably *Kritik der reinen Vernunft* (hereinafter *KrV*) A X, B 21-22, A 843; *Kant's gesammelte Schriften*, ed. cit. (hereinafter *AA*), IV, 332, 353, 362ff., 365; IX, 17; XVI, 71; XXIV, 312, 314, 608f.; XVIII, 94; XXVIII, 302, 367f., 378; XXIX, 751ff., 757, 765, 767, 782, 937f.; XXIX, 947f., 956f. In the *Critique of Pure Reason* and elsewhere, Kant pointed out that every human being is naturally concerned with his or her practical situation. Human life is continually grounded in a specific diagnosis of who one is, what one knows, what is at stake in one's choices, etc., and this diagnosis entails a series of practical questions whose answers are irreducible to the actual contents of human experience. Therefore, 'a certain sort of metaphysics has actually been present in all human beings as soon as reason has extended itself to speculation in them, and it will also always remain there.' (*Critique of Pure Reason*, ed. cit., 147). More specifically, Kant argues that the idea of freedom is an essential part of this diagnosis. The very possibility of practical life presupposes a *metaphysics of human freedom*, fuelled by the constant need to determine whether and to what extent one is in control of one's life. Kant's overall argument can be divided into two main steps: firstly, if human beings were not already in possession of an answer to the questions of whether and to what extent they are free, however unconscious or instinctive it may be, they would not be able to interpret or inhabit the phenomenal world — and thus to move around, to speak or to pour cups of coffee; secondly, apart from having to answer these questions, they tend to answer them *in a certain way* — namely in a *positive* way. Human beings' natural attitude towards the issue of personal freedom is not interrogative, but dogmatic: they usually assume they are indeed in control of their lives and possess the freedom required to fulfil most of their daily tasks. But although this conviction is already at work in every moment of human life, it is still a mere conviction — and indeed a thoroughly *metaphysical* conviction, which is neither confirmed nor denied by human experience. There is a basic disproportion between empirical knowledge and the knowledge that is actually required to make sense of one's practical situation, which leads consciousness to fill in the gap with its own metaphysical assumptions.

this or that particular judgement or decision, but on our entire understanding of reality. When it comes to determining whether personal freedom is true or false, real or illusory, what is at stake is not merely the alternative between two simple assumptions, but the possibility of a complete transformation of our usual existential horizon. Since our general conviction regarding the freedom of our thoughts and actions provides the basis for the usual running of our lives, the abandonment of this conviction, or the endorsement of a different conviction, is sure to have profound practical consequences.

Thus construed, the debate concerning human freedom acquires a deeper ontological significance, which calls for a re-examination of what it means to be an independent, self-conscious human being. It may be, of course, that this enquiry is destined to lead nowhere. If personal freedom, like breathing, is an underlying condition of our thoughts and actions, to question its actuality may prove in the end as futile as picturing one's own death, or as absurd as Baron Münchhausen's efforts to lift himself up by his own hair. Even so, there is still a long road separating our automatic belief in freedom from the actual truthfulness of that belief. And this road is liable to reveal, if nothing else, the extraordinary complexity of the notion of freedom. In particular, it is liable to show how different definitions of freedom correspond to different outlooks on reality, and how the latter imply radically different interpretations of what is at stake in human life.

1.1. *A brief anatomy of freedom*

What is the origin of the confidence we feel with regard to personal freedom? How truthful is the assumption that we are free and how wide is its scope? The answers to these questions are not straightforward and their complexity is tied to the ambiguity of our usual definition of freedom.

To begin with, it is easy to see that the notion of freedom is not simple or univocal. Like Aristotle's being, freedom is a πολλαχῶς λεγόμενον, used in different contexts and with a variety of different meanings. Accordingly, our usual conviction regarding our own freedom is neither simple nor univocal. Like most immediate assumptions, it can be interpreted in different ways and endorsed with different degrees of enthusiasm. More

importantly, however, its apparent simplicity conceals a great deal of ignorance. Although we usually fail to notice it, the concept of freedom possesses a highly complex structure, shrouded in confusion and ambiguity. Instead of one simple assumption, it really amounts to a series of overlapping assumptions whose exact number, meaning and articulation are anything but clear. To borrow another Aristotelian concept, our usual belief in personal freedom is συγκεχυμένον: it amounts to a 'confused mass' whose different elements must be sorted out and distinguished from one another.[2] This anatomical procedure, which Aristotle names διαίρεσις, will be our guiding reference throughout the following pages.

To introduce our main subject, let us start by considering a few verses from one of Schiller's late poems, partly dedicated to the issue of human freedom:

Der Mensch ist frey geschaffen, ist frey,
Und würd' er in Ketten gebohren,
Laßt euch nicht irren des Pöbels Geschrey,
Nicht den Misbrauch rasender Thoren,
Vor dem Sclaven, wenn er die Kette bricht,
Vor dem freyen Menschen erzittert nicht.[3]

John Sigerson's skilful translation attempts to recreate, as much as possible, the style and rhyming of the original verses:

Man was created free — *is* free,
E'en though he were born in shackles.
Do not be deceived by the rabble's bray
Or idiots' abusive cackles.
Before the slave, when his chains he doth break,
Before the man who's free, O do not quake![4]

Man is free — even though he were born in shackles. In these well-known verses, Schiller seems to suggest that freedom is an essential feature of humanity, whereas

[2] Cf. *Physics* 184a21-23.
[3] *Sämtliche Werke* I, ed. cit., 214f.
[4] *Friedrich Schiller. Poet of Freedom* I, ed. cit., 335

bondage or servitude are not. He distinguishes between what might be called a primary or essential kind of freedom, which follows from the very definition of humanity, and a secondary or contingent kind of freedom, which may or may not be actualized in the course of human life. Even though all men are naturally free, some live freely while others do not. And even when personal freedom seems devoid of actuality, as in the case of life-long slavery and oppression, bondage and servitude are still somehow accidental predicates, forced upon human beings by contingent sets of circumstances.

This idea is reiterated by Schiller's vigorous defence of human freedom against the 'bray' and the 'abusive cackles' of the 'rabble', that is, against the vain utterings of those who seek to undermine personal freedom. Moreover, it is further developed both at the beginning and at the end of the poem, where Schiller argues that if men forget the true meaning of the word freedom, or lose faith in it, they are 'deprived of all their value' (*Dem Menschen ist aller Werth geraubt*)[5]. Now this last statement can be understood to have two different, albeit closely connected meanings, depending on how one construes the notion of 'value' or 'worth' in question. On the one hand, Schiller's point seems to be that those who deny men their right to freedom are cruel or unfair, and that men should be more highly valued. This is the kind of humanist claim that is generally championed in modern Western societies and that most reasonable men are ready to endorse. On the other hand, in more general terms, Schiller seems to argue that those who deny men their right to freedom are fundamentally mistaken, as they fail to acknowledge men's true value and act as though they were worth less than they actually are. In other words, they fail to recognize that human beings are *intrinsically* free, even if they are born or live in shackles. In this second reading, what is at stake is not so much human beings' actual life as their essence or nature. To question a man's freedom amounts to nothing less than to question what it means to be human.

This second claim is not as straightforward as the first one. If we consider our daily

[5] Ibidem. The 'words of faith' highlighted in the poem are three, namely Freedom, Virtue and God (or a divine Will), and Schiller argues that 'Mankind is of all [its] value bereft / If in these three words no faith is left'. Although my present concern is the notion of freedom, and particularly one's usual understanding of what it means to be free, the following chapters will show that the issue of human freedom is much wider and much more complex than usually supposed: when man 'loses his faith' in freedom, he is also at a loss with respect to the notions of virtue (or ethics in general) and god (or transcendence in general).

lives, there seems to be no doubt as to the limited scope of our freedom. Even if we were not born or live in shackles, in the common sense of the word, it seems fairly obvious that we cannot do whatever we want whenever we want to and that half of our choices are instinctive or involuntary. Indeed, hardly a day goes by without our being reminded of our impotence or faced with situations whose meaning eludes us and whose outcome escapes our control. Our entire existence amounts to a long and ceaseless compromise between what we could do and what is within our reach, or between what is conceivable and what is actually possible. And the exact nature of this compromise depends on how strictly we choose to define personal freedom. If the latter is construed in a tolerant manner, as the relative power to determine our choices and actions, there may still be considerable room for autonomy and independence. But if freedom is defined, *sensu strictiore*, as the absence of each and every kind of limitation, our lack of independence will appear even greater. For even our most modest wishes and projects will be found to depend on an endless series of constraints, related to the shape and strength of our body, to the physiology of our brain, to the spatial and temporal structure of the surrounding world, etc.

For every human being, freedom is confined to an interval located somewhere between complete powerlessness and omnipotence. The interval's upper end determines the actual extent of personal freedom, but its position is usually open to variation. Although most human choices are indeed very limited, not all limitations are equally essential or equally damning: while some are more easily endured, and sometimes even hardly noticed, others strike closer to home and represent more serious threats to the daily running of human life. In light of this variation, the different obstacles to personal freedom can be compared, in Stoic vein, to the walls of a citadel. At a more superficial level, the acknowledgement that one is not free to do anything one chooses is not particularly hard to accept and still leaves out a vast array of possible choices. Although human beings are unable to fly, for instance, or to breathe underwater, or to travel through time, they have learned to cope with these limitations and to live their lives accordingly. Further inside the citadel, however, freedom's interval can grow much smaller. If a man is prevented from speaking freely, or sent to jail, or sold into slavery, the balance between what he can and cannot do is significantly altered. Worse still, if he is denied the control of his body, by being gagged or shackled, his freedom is compromised

even further, with direct implications in almost every aspect of his life.

Yet these constraints, however serious, are never serious enough to invade one's inner citadel and to destroy one's implicit belief in personal freedom. Which brings us back to Schiller's claim: although the actual power to choose and to act freely may be wider or more restricted, depending on the number and the nature of the obstacles one is up against, it nonetheless *presupposes* a minimum of self-determination. A slave or a prisoner, despite all of their limitations, can still retreat into a more basic form of autonomy and regard their bondage as an accidental predicate, added to a more essential claim to personal freedom.

This basic sense of autonomy is the ground zero of human freedom and the foothold of every human choice. Whereas all the other forms of freedom may be denied and still leave room for some degree of independence, this one may not. It is the element that holds everything together, the inner core that grants meaning to freedom's outer layers. Therefore, if it should waver or prove unreliable, the entire interval of human freedom would be compromised. If we did not think of ourselves as fundamentally free agents, we would not experience the world as we normally do: the very notion of possibility would be transfigured and life would no longer appear to us as an open process, or as a task awaiting completion. In these circumstances, the ability to choose something over something else would no longer have any meaning because it would no longer be *ours*, but *someone else's*: even if all other impositions were suddenly lifted, and we were suddenly allowed to do whatever we please, this possibility would still amount to a mere deception. In like manner, if the most powerful man in the world were to cease believing he is the actual author of his choices, all of his others claims to freedom would be stripped of their value. Omnipotence itself would cease to be appealing or empowering: it would become a fake or borrowed power, and therefore no power at all.

In its most basic form, then, our belief in personal freedom amounts to the feeling that we *own* our life, that is, that we are the starting point of our thoughts and actions and not mere subjects of an alien power or mere links in a predetermined causal chain. This belief comes down to the inner certainty that each one of us is the ultimate master of his or her own destiny and that our life is not, as Horace puts it, the life of 'a wooden

puppet whose strings are manoeuvred by an alien power.'[6] This feeling of autonomy, embedded in the very notion of selfhood, is reiterated at each waking hour. Time and habit strengthen the conviction that our whole existence, despite its considerable share of contingency and impotence, is steered or manoeuvred by us; that we are its protagonists, as it were, and that its varying circumstances, undoubtedly real and effectively constraining, are ultimately but an outward setting, which can either favour or obstruct our natural sovereignty.

But this is not all. Although freedom is usually measured in terms of what we can and cannot do, it is not limited to this alternative. Apart from the power to choose and to act according to our will, personal freedom also requires a certain degree of insight into the actual import of our choices. When we choose something over something else, or act in a certain way rather than another, our decisions are supported both by an immediate feeling of independence and by a more or less ambitious belief in the ability to shape the course of events. In addition, however, we also tend to assume we have a correct grasp of what is at stake in each new context or situation. And this assumption entails a new set of presuppositions, which must also be identified and discussed.

Freedom is just as much about power as it is about knowledge. For our choices to be free choices, they must be meaningful, intelligible and coherent choices, and for that to be so we must be able to make sense of reality and to interpret, as best we can, the different circumstances we are faced with. If our decisions were not grounded in a specific understanding of what goes on around us, nor envisaged as part of a specific sequence of events, they would have no real value. If, when choosing or acting in a certain way, we did not have at least a faint idea of *what* we are doing, of *why* we are doing it and of *how* it affects us, our life, and the surrounding world, our choices would not be free choices and our freedom would amount to an empty word.

After examining freedom's concrete or actual requirements, indicated by the kinds of things we can and cannot do, let us now turn to freedom's cognitive requirements, and to the kinds of things we know and ignore. Once more, these requirements can be compared to different layers or walls, built around an inner core of cognitive self-

[6] *Satyrarum libri* II, 7, 82

confidence. Like personal power, personal knowledge is usually confined to an interval placed somewhere between complete ignorance and omniscience. The interval's upper end determines how knowledgeable we actually are, while the lower end indicates the minimal insight required for the exercise of personal freedom. Just like before, not all forms of ignorance are equally essential or equally constraining: whereas our lack of insight into quantum physics, for example, is liable to affect our power to make informed scientific decisions, the inability to read or write can prove far more damaging, as it is bound to condition our freedom of choice in a wider number of situations. And if we ignore our own identity, or the identity of our immediate surroundings, or the most basic principles of the physical world, we will be even less able to determine and direct the course of our lives.

Moving further down the interval, we arrive at a new core, or a new ground zero, whose suppression would entail the suppression of freedom itself. Once again, although our usual power may be wider or more restricted, depending on the number and the nature of the things we know, it presupposes a minimum of knowledge. At the very least, free choices must be grounded in a general assessment of what is *at stake* in our life: they require an identification of our life's main theme, as it were, of what matters most in it, of what motivates us to choose anything in the first place. And though this assessment may be vague, volatile and even unconscious, it is the foundation of a whole edifice of cognitive assumptions, whose inner composition is no less important to the definition of freedom. In particular, if we have no idea of who we are, where we stand or what goes on around us, our decisions can hardly be called such. For something to be freely decided, there has to be some kind of awareness of the subject, object and context in which the decision is made. Furthermore, apart from this basic knowledge, free choices also require the power to determine, in each situation, which and how many different alternatives are available to us. Again, if this power is lacking there can be no real freedom, for our choices are bound to reflect our ignorance, rather than our actual intentions. Finally, and just as importantly, personal freedom is also conditional on the ability to anticipate, at least up to a point, the consequences of our choices. Indeed, we may have a clear picture of the situation we are in and/or of the best possible course of action and yet be surprised by the actual result of our decisions. In this case, although we are free to choose, we are not able to guarantee the success of our choices, and may end up facing the opposite of

what we bargained for.

To appreciate the extent to which freedom is dependent on these basic requirements, as well as the extent to which they are dependent upon each other, let us look closer at what would happen if they were separately removed. If human beings, when faced with a given situation or choice, were able to identify the different alternatives available to them and to anticipate their consequences but utterly unaware of their own identity or context, the first two powers would be utterly worthless, for the freedom they afford would be grounded in ignorance and impotence. This kind of freedom would be as contradictory as that of a blind man allowed to see what he pleases. Alternatively, if human beings were able to make out their practical situation and to anticipate the effects or their choices but unable to determine which and how many options were open to them, the first two powers would once again be compromised. This combination would result in a new form of blindness, whether or not it were perceived as such. Thirdly, if we knew who we are, where we stand and what options we have before us but were utterly unable to foresee where they lead, our freedom would once again be illusory, for we might spend our whole life choosing and acting against our interests. In this case, our vision would amount, yet again, to a peculiar form of blindness, as it would lack the scope or reach necessary to reveal anything other than our own free will.

Since all of these clauses are essential to the actualization of freedom, the removal of one, two, or even all of them can only result in servitude, impotence and blindness. Such an exercise is typically enacted in Greek tragedies, where personal freedom is invariably countered by ignorance and self-deception. Oedipus' case is the most exemplary: although most of his decisions are freely made, they rest upon a globally mistaken diagnosis of the situation he is in; and though he is convinced he has a relatively clear knowledge of the different elements at play in his drama (viz. his own identity, the identity of his father, the identity of his mother, the identity of his wife, etc.), he soon finds out that nothing is what it seems (viz. that his adoptive parents are not his biological parents, that his mother is also his wife, that he himself is the killer of his father, etc.). Sophocles' tale is exemplary because of its unique focus on freedom's cognitive dimension. Oedipus embodies the extreme possibility of a man whose actual freedom to choose and to act is completely neutralized by his cognitive incompetence. Although he is able to do what he pleases, his choices are accompanied by a complete

lack of insight into his actual situation, a complete ignorance regarding the options that are actually open to him and a complete inability to anticipate their actual consequences.

The main point, however, is not simply that Oedipus knows not what he thinks he knows, but that his power, due precisely to this ignorance, is not as real as he takes it to be. Sophocles' sophisticated thought experiment highlights the fundamental gap between the power to choose and the power to translate one's choices into actual states of affairs. It aims to show that the ability to shape circumstances in one's favour is not immediate or automatic, but conditional on one's knowledge of reality. Indeed, one may be free to choose and yet remain as blind as Oedipus to the actual meaning of one's choices; and this blindness may lead one to misrepresent one's own interests, to mistake an unfavourable choice for a favourable one and to limit rather than expand one's freedom. In the play, these possibilities are conveyed by Sophocles' ingenious interweaving of freedom's concrete and epistemic claims. Whereas in the first scenes Oedipus' freedom appears to outshine everyone else's, his superiority is soon called into question. As the play unfolds, and his ignorance about himself becomes more apparent, his ability to shape circumstances in his favour is increasingly challenged, and eventually shown to be worthless. Oedipus' choices are grounded in a flawed diagnosis of his practical situation, which leads to an equally flawed prognosis of its future developments. And although his troubles are initially put down to outside forces or to unfavourable circumstances, they are ultimately due to an overestimation of his own cognitive powers.

This whole idea can be translated, once again, in Aristotelian terms, by the tragic notions of recognition (ἀναγνώρισις) and reversal (περιπέτεια).[7] Sophocles' play is designed in such a way as to maximize the impact of both of these phenomena. Firstly, Oedipus' tragedy does not consist merely in having to acknowledge that he is deluded about this or that particular aspect of his life. What is in store for him at the end of the play is not a partial or parochial ἀναγνώρισις, but a *total* one, whereby all of his beliefs are shown to be groundless and his entire life exposed as the very opposite of what it seemed to be. This recognition is all the more remarkable since Oedipus has already been informed by the oracle of what is going happen. He already knows the truth, but still

[7] See *Poetics* 1452a12ff.

does not see it, and his extreme situation is meant to illustrate just how complete human ignorance can be.[8] Secondly, Oedipus' peculiar situation does not lead to a simple change of fortune, but to a reversal of his entire fate. Again, what is at stake at the end of the play is not a partial or restricted περιπέτεια, but a *total* one, whose seriousness is meant to illustrate just how misleading human reason can be. Out of ignorance, Oedipus is led not only to misidentify all of the elements at play in his drama, but also to misrepresent the practical implications of all of his actions and resolutions. In the absence of a suitable diagnosis of reality, his alleged power veers round into impotence and servitude.

This idea can also be expressed with the aid of Hegel, who spoke of tragedy in general — and of Sophocles' play in particular — in relatively similar terms. Like Aristotle, Hegel locates the essence of Greek tragedy in the idea of self-deception. Although the tragic hero believes he knows what is at stake in his decisions, and can therefore choose and act in his best interest, this belief is usually groundless. As Hegel puts it, although he is convinced he knows reality's *true face*, what he sees is a mere *mask* of reality, wrongly perceived as an actual face,[9] and the purpose of tragedy is precisely to illustrate this antithesis: what the hero thinks he knows is 'immediately a not knowing'[10] and each dramatic event possesses therefore two opposite meanings, whose truth-value depends on whether one looks at it from the hero's perspective or from the spectator's. This dialectical contradiction is the motor behind the play's various developments and the source of its denouement. 'The action itself is this inversion of the *known* into its *opposite*',[11] the inexorable movement whereby all masks are finally dropped and every familiar truth is 'changed round' (*umgeschlagen*) into perplexity.[12]

[8] The tragedy — or the irony — of Sophocles' play does not lie simply in the fact that Oedipus ignores what he thinks he knows, and that he is led to suffer the consequences of his ignorance; what is worse, the very provisions he and the other characters make in order to avoid the fate prophesied by the oracle — namely, that Oedipus will kill his father and sleep with his mother — end up actively contributing to his downfall. For instance, Oedipus' departure from his home town in order to avoid fulfilling the oracle's prophecy leads him to his original birthplace and to his real parents (cf. lines 787ff.); Jocasta's surrender of her young son in order to escape the same prophecy unleashes the series of events which will lead to murder and incest, as well as to her own suicide (cf. lines 1014ff.), etc. In these and other situations, the characters' original intentions are not merely opposed to the actual outcome of their choices: they are the vehicle whereby that outcome is brought into being.

[9] See notably *HW* III, 541f.

[10] *HW* III, 537; *PS*, 446

[11] *HW* III, 538; *PS*, 447

[12] The most original aspect of Hegel's interpretation of tragedy is arguably its universal scope: instead

Yet despite Sophocles' eloquence — and the eloquence of Aristotle's and Hegel's interpretations — it may be argued that Oedipus' tale is overly dramatic or unrealistic, that Sophocles exaggerates the importance of freedom's cognitive dimension, or that his convoluted plot has no real equivalent in everyday life. Indeed, just as Oedipus reacts with incredulity and scepticism to Tiresias' predictions, so too most of us tend to reject the idea that our knowledge of reality amounts to an illusion, or that our freedom is liable to be overturned in such a complete and catastrophic way. Whereas Oedipus' whole existence is based on a series of mistakes, our lives are thought to be less uncertain, or at least less open to deception. And whereas Oedipus' freedom is made to depend entirely on what he knows or ignores, our own freedom is thought to hinge on more concrete circumstances, qualities and events. Interpretation and anticipation are generally regarded as relevant conditions for the actualization of freedom, but their importance is only fully acknowledged in connection with difficult or distant choices, with complex dilemmas or life-changing resolutions. And while the latter are certainly real and highly significant, they tend to represent but a small island in the wider sea of our daily concerns.

This reaction, however plausible, is really only a confirmation of the anatomical picture I have hitherto been sketching. It is precisely because none of us is ever indifferent to the issues of whether and to what extent we are free that Sophocles' tale seems so far-fetched. In our own case, we tend to assume we are knowledgeable enough to lead relatively free lives and to shape most situations according to our interests. Accustomed as we are to our usual cognitive perspective, we find it hard to believe that things might be completely different, and that our usual freedom might amount to an illusion. Yet Oedipus is just as attached to his own cognitive perspective and just as convinced of his own freedom. In this regard, his attitude is identical to ours, and

of the identity of specific persons (viz. Oedipus, his father, his mother, his wife, etc.) or the meaning of specific events (Oedipus' birth, the death of his father, the marriage to his mother, etc.), the *masks* Hegel refers to stand ultimately for all of the concepts and judgements employed throughout human life. Hegel borrows the tragedy's dialectical structure to illustrate a much wider cognitive phenomenon, related to the very definition of human cognition. In this respect, his analysis is more Platonic than it is Aristotelian: just like Plato, Hegel argues that the concepts and ideas we usually resort to are mere *versions* of true objects and events, immediately assumed to be true; and like Plato he insists on the need to question this assumption and to pull out reason's usual masks. I will return to this issue below, in section 9.2.

Sophocles' portrait is indeed an accurate illustration of human nature. For all of us, as for Oedipus, true freedom is the opposite of tragic freedom: it requires a kind of knowledge that cannot be reversed, or an insight into reality's true face. But just as Oedipus was eventually disabused of his cognitive self-confidence, so too may we be exposed to the possibility of such a refutation.

The former description, albeit very brief, already suggests a few important conclusions. Firstly, we have seen that personal freedom is not a simple, but a complex phenomenon. The usual assumption that we are free, or at least as free as we need to be in order to do what we happen to be doing, rests on two important sets of presuppositions, which correspond to several different forms of power. However, their differences are usually mingled into one blurry image. When I claim, for instance, that I am free to leave this room, what I am saying is that 'I have the power, if I so choose, to leave this room.' And this requires in turn a) that I am the primary source of my thoughts and actions, and not a puppet steered by an unknown master; b) that I can actually choose whether to stay or to leave this room; and c) that I am capable of carrying out my decision. Moreover, inasmuch as personal freedom is also dependent on a series of cognitive presuppositions, the power to leave this room also requires d) that I am aware of what is at stake in my own life; e) that I have a clear enough knowledge of my present circumstances; and f) that I can anticipate the consequences of my actually leaving this room.[13]

Secondly, although all of the powers usually entailed in the experience of personal freedom are equally important, they are not equally fundamental. Apart from the inner anatomy of our belief in personal freedom, we must also consider its inner physiology.

[13] At first sight, the awareness of what is stake in one's life (viz. of what matters most in life, of what motivates one to choose anything in the first place, etc.) does not seem required when deciding something as trivial as leaving this room. Whereas the determination of one's life goals is usually a complex issue, worthy of lengthy philosophical discussions, decisions such as leaving a room are usually considered much simpler and matter of fact. This attitude rests on a fundamental misconception, which stems from a natural tendency to separate what can only be thought of together. Although we usually fail to notice it, the decision to leave this room — like any other human decision — is indeed supported by a wider, if inexplicit appreciation of our whole life. This decision is never meaningful or desirable in itself, but always as an episode of a wider existential script, grounded in a million other decisions, assumptions and expectations. Therefore, the power to leave this room entails much more than the actual power to get up and move across the room: it is usually supported by all of the assumptions I have mentioned in points a) to f).

The different powers we usually take for granted are not merely coordinated with each other, but subordinated to one another, with the most elaborate presupposing the most simple. Considering the previous example, I am only free to leave this room because I have the power to anticipate the consequences of my choosing to do so. But this power is conditional on my ability to decide whether to stay or to leave, which in turn requires a minimal understanding of the practical situation I am faced with. Finally, none of this has any meaning if I am a mere pawn of fate or the mere instrument of an alien will. My belief in personal freedom requires, on the contrary, that I be the starting point of my choices and that I recognize myself as such.

All of these powers are at work, on different levels, in my usual conviction that I am free, if I so choose, to leave this room. And they are also present in most of the choices and decisions made throughout my life. However, not all of them are as certain or self-evident as they appear — which brings us to the third and most important point of our enquiry. The ordinary conception of personal freedom rests on a series of beliefs automatically held to be true. However, since their truthfulness is usually merely taken for granted, and not actually proven, it is open to various kinds of criticism. Considering what went before, this means not only that each of the powers we take for granted may turn out to be illusory, but also that *all* of them, inasmuch as they presuppose one another, may turn out to be illusory.

Let us start by considering freedom's concrete claims.[14] As regards the most basic assumption in my previous inventory, there are compelling philosophical arguments against the idea that we are the original source of our thoughts and actions. According to some brands of determinism, the feeling of self-control that runs like a *basso continuo* through our entire lives is not a sufficient proof of our independence. According to other, more radical brands of determinism, this feeling is downright deceitful, and we are indeed mere links in a predetermined causal chain, or mere puppets held by an unknown hand.

Furthermore, concerning the actual power to choose a specific course of action, we

[14] The following inventory is merely indicative of the main kinds of arguments that are usually levelled against the natural belief in personal freedom. As will be made clear, my aim is not to engage with all of them, but with one in particular, namely the critique of the cognitive assumptions entailed in our usual belief in personal freedom.

may reject the previous modes of determinism and still find that our independence is extraordinarily limited. Even if we are not an intermediate link in a predetermined causal chain, we may very well be the initial link of an equally inexorable causal chain, in which case our particular choices are not free choices, but mere reactions to previous events. Although we appear to be in control of our minds and bodies when deciding to move around, or to speak, or to pour a cup of coffee, we may simply be responding to highly complex arrangements of physical, biological, psychological and social stimuli. And if it were possible to isolate all of them, to determine their reciprocal interference and to account for all of their effects, we might very well find that our choices are indeed inexorable and that our belief in personal freedom is really a symptom of our ignorance.

Finally, to complicate matters even further, our belief in the ability to carry out our decisions and to achieve our projected goals is no less problematic. Even if we are indeed the original source of our thoughts and actions, and even if we are capable of choosing freely and acting upon our choices, there is ultimately no way of guaranteeing their success. For although we may be free to set in motion a determinate sequence of events, its final outcome is not exclusively up to us. Its evolution is dependent on a constellation of factors that cannot be fully anticipated or controlled, including the laws that govern the natural world, the choices made by other people or the dictates of chance.[15]

So much for freedom's concrete dimension. Let us return, now, to its cognitive dimension. As was seen earlier, although we are often faced with situations we do not understand, and forced to make decisions whose outcome is hard to anticipate, we are usually confident in our ability to interpret the meaning and implications of most daily events; even though we admit to being ignorant of a great number of things, our understanding of reality is usually deemed wide enough to inform most of our decisions; and even though we often recognize we are wrong or misinformed, our mistakes seem to

[15] If we decide to shoot an arrow in a determinate direction, the probability of hitting the target can be increased by practising our aim, by calculating how far the bow's string should be stretched, by determining the speed and the direction of the wind, by isolating the surrounding area so that no-one gets in our way, etc. However, despite all of these precautions, we will never be able to anticipate and control all of the variables liable to affect the success of our task. An earthquake might suddenly break out, a dead bird might fall from the sky, a momentary cramp might prevent us from stretching the string in the appropriate manner, etc. Due to these and a million other possibilities, the efficaciousness of our choices is always an open issue, and the actual extent of our freedom is thus much more limited than we are usually prepared to assume.

call for small and isolated corrections, and not for a global revision of our entire outlook on reality. But how solid are these assumptions? Are they liable to be refuted? And, if so, to what extent?

The main purpose of the present work is precisely to answer these questions. Throughout the previous pages, I have argued not only that our usual belief in personal freedom is complex and multi-layered, but also that most of the presuppositions whereupon it is grounded are liable to be contradicted. In what follows, my aim is not to take up the whole issue of personal freedom, nor to explore and discuss all of its underlying assumptions. I will focus rather on the relationship between personal freedom and personal knowledge, and on the assumption that we are knowledgeable enough to lead free lives. To test the solidity of this assumption, I will start by focusing on the nature, structure, and scope of our usual understanding of reality. I will argue that our cognitive sovereignty is not as certain as it appears, nor is the freedom it is thought to provide. Moreover, I will show that this is not merely a theoretical or epistemological problem, but one that has direct implications in every aspect of human life. Finally, I will argue that true freedom requires a radical revision of our usual standpoint and the pursuit of an entirely new cognitive perspective.

1.2. *Freedom and education*

The previous considerations have led to a conclusion that will remain at the centre of the following enquiry: namely, that there is no freedom without cognitive awareness. To be sure, our usual insight into the meaning and consequences of our choices is never fully adequate or fully complete. No matter how perceptive we take ourselves to be, our limited cognitive powers can hardly match the extraordinary complexity of human life. However, if we had no idea of who we are, where we stand, or what goes on around us, our choices would have no meaning. Living would be like walking through a dark room, with no familiar references to guide the way. The solution seems to lie, therefore, somewhere in the middle: although we do not expect our knowledge to be perfect, we do expect it to provide us with a sufficient degree of independence. In other words, we require a knowledge powerful enough to light up the darkened room of our lives and to

reveal, if only faintly and within a small perimeter, the shape and colour of the surrounding world, the direction in which we are heading and the main obstacles along the way.

The debate concerning the nature and the amount of knowledge required for the actualization of freedom is an old and difficult one. Historically, its origin can be traced back to the ancient Greek world, and particularly to the Greek contrast between human beings' primitive or natural standpoint and a civilized or educated standpoint. The Greeks were the first to insist on the idea that freedom is not a natural or immediate ontological asset, but the result of a specific form of education, or παιδεία, destined to release men from their natural state of ignorance. As the very word suggests, παιδεία was the transitional process whereby the παῖς, i.e. the child or youth, was elevated to the state of πολίτης, i.e. citizen, and admitted into the city's public and political life. But the word παῖς was also used for slaves and servants in general.[16] It referred to a kind of immaturity that was not exclusively tied with one's age, but also, and primarily, with one's power of judgement. Children and slaves — but also women and mad people, for example — were all included in this category because they were all deemed intellectually immature, and hence unaccountable for their judgements and actions. Lacking the emancipation provided by παιδεία, they were denied the right to free citizenship and destined to a life of servitude.

The history of Greek education is based on the antithesis between cognitive maturity (or majority, or adulthood) and cognitive immaturity (or minority, or infancy). From the more primitive to the more sophisticated versions of παιδεία, what distinguished the educated (παιδευτός) from the uneducated (ἀπαίδευτος) was the acquisition of the knowledge deemed appropriate to a free or independent life — the knowledge required to make one's own decisions, to shape circumstances in one's favour and to take part in the life of the community. With time, this formal antithesis was adapted to different contents and subordinated to different cognitive requirements. The kind of insights that were initially deemed sufficient to ensure one's independence were eventually found lacking,

[16] An ambiguity still found in the English word 'boy', or the French word 'garçon'. See, for ex., AESCHYLUS, *Choephoroi* 653; ARISTOPHANES, *Acharnians* 395, *Clouds* 132; or DEMOSTHENES, *Against Apatourius* 33, 8.

which prompted the need for more complete and sophisticated forms of education. As human life grew more complex, the emergence of new needs, habits and aspirations led to the redefinition of the frontier separating education and freedom from ignorance and servitude.

The first great step away from humankind's natural ignorance and towards a knowledgeable or educated form of existence is described by many Greek authors in mythological terms, as the result of a divine gift. According to Prometheus, in Aeschylus' (or Pseudo-Aeschylus') homonymous play, although primitive men 'had eyes to see, they saw to no avail; they had ears, but they did not understand; … just as shapes in dreams, throughout their length of days, without purpose they wrought all things in confusion'.[17] Only when they were handed down the gift of fire and the secret to a series of practical arts (τέχναι), skills (σοφίσματα), resources or contrivances (μηχαναί, μηχανήματα) — such as the arts of building, writing, fishing or mining, among many others — were they freed from their original blindness and 'endowed with reason'.[18] This same myth is evoked by Protagoras in Plato's homonymous dialogue, where Prometheus is also credited with having imparted to mankind the knowledge of the various arts.[19] When the gods created the living world, they provided all animals with speed, or strength, or a natural habitat or other means of survival and preservation. Men, however, were left 'naked, unshod, unbedded, unarmed'.[20] To compensate for this oversight, Prometheus offered them the gift of fire and the secret of all of Athena's arts, ensuring thereby their security and livelihood. A brief catalogue of these endowments, similar to the one offered by Aeschylus, can be found, moreover, in the first stasimon of Sophocles' *Antigone*, where the power of human reason is said to 'span the seas' and to 'wear away the earth', to 'turn the soil' and to 'snare birds and fish in the meshes of twisted nets', to 'master the beast who dwells in the wilds' and to 'flee the arrows of the inhospitable frost'.[21]

All of these accounts accentuate the extraordinary transformation brought about by the advent of reason and education. Whereas primitive men were ignorant and

[17] *Prometheus Bound* 447-450. Ed. cit., 493
[18] *Prometheus Bound* 444. Ed. cit., 491
[19] See *Pr* 320cff. Cf. *Phi* 16c
[20] *Pr* 321c5-6
[21] *Antigone* 332ff. Ed. cit., 35ff.

defenceless creatures, vulnerable to all sorts of dangers and constraints, educated men are stronger, cleverer, more industrious and more organized. They are able to challenge nature's tyrannical rule, to anticipate and influence the course of natural events, to harness the various resources of the physical world. As their original blindness is lifted, their power is strengthened and their freedom increased. At the same time, however, their new understanding of reality gives rise to more sophisticated cognitive demands. As their knowledge increases, so does the awareness of their own limitations, as well as the need for new and more complete forms of instruction.

In the ancient Greek world, this revolution was accompanied by another, closely related revolution, brought about by the so-called sophistic movement. With the sophists, the issue of education acquired a new importance and an even wider scope. Apart from reiterating the need for a transformation or an upgrade of natural reason, the sophists highlighted the wideness of the gap separating human beings' usual standpoint from a truly educated standpoint. Looking back at the Promethean leap from ignorance to knowledge and resourcefulness, they highlighted the possibility of a new and equally radical leap forward, towards an even higher cognitive stage. As in the previous transition, however, the result of this new transition is not accessible from its starting point, and can only be truly grasped at the point of arrival. This negative movement can be interpreted in light of what Hermann Fränkel called a 'Heraclitean thought pattern', that is, the analogical procedure whereby the contrast between a lower term, X, and a higher term, Y, is used as reference to envisage the identity an even higher term, Z, which relates to Y in the same way that the latter relates to X.[22] This kind of analogy is usually employed as a means of anticipating the content of ideas or realities that lie beyond one's cognitive reach: in this case, starting from the familiar contrast between ignorance and the basic resources available to most educated men (X and Y), one is led to envisage the unknown contrast between this basic set of resources and a new kind of knowledge (Y and Z), whose attainment would amount to an equally profound cognitive upgrade.

In other words, the sophists highlighted, with unprecedented vigour, the intrinsic *mobility*, or the intrinsic *versatility* of human cognition. They aimed to show that human

[22] FRÄNKEL, op. cit.

beings are not confined to a single cognitive perspective, or to a single outlook on reality, but capable of moving beyond their native standpoint and embracing new and previously inaccessible forms of knowledge. Παιδεία, in the sophistic sense, stands above all for a cognitive *transfiguration*: whereas education was previously construed in a chiefly quantitative or cumulative manner, as the possibility of expanding a basic repertoire of skills or resources, the sophists emphasized the qualitative gap separating ignorance from education. Challenging the immediate or universal accessibility of the knowledge handed over by the gods, they replaced the general distinction between skilful and unskilful individuals with the stricter distinction between experts and laymen.

This attitude is the starting point of a cognitive trend whose effects are still recognizable today. Although modern men and women tend to rely, in most daily situations, on the knowledge provided by experience and common sense, they are well aware of its limitations. What most people know about a certain subject or phenomenon is usually assumed to be different from what an expert or a scientist knows about it: his or her insight is not just any insight, but a specialized one, which exceeds human reason's basic cognitive repertoire. While most of us know, for example, the general difference between a headache and a toothache, not all of us are familiar with their exact causes or with what it takes to stop them. And although most of us know the difference between a house and a bridge, few of us know how to design either of them, or how to build them and make them stand. These insights are usually the prerogative of doctors, dentists, architects and engineers. They amount to artificial extensions of human reason, exclusively available to those who have studied and exercised themselves, sometimes for several years, in these specific activities.

In the sophist's vocabulary, doctors and architects who know their trades are in possession of a τέχνη, that is, an art or expertise enabling them to determine with accuracy what most people can only guess. Returning to the previous example, experienced doctors do not stare powerlessly at the suffering of their patients because they are able to see *beyond* or *behind* its immediate appearance. While disease and decay are for most of us, non-doctors, a mysterious affair, whose exact origin we believe to lie somewhere inside the physical body, a doctor's interpretation of the same phenomena is very different. Where we see a regular human body, a doctor sees a highly complex system of muscles, bones, organs, sinews and vessels. And where we stare alarmed at the

swelling of the skin or the onset of mysterious convulsions, a seasoned doctor instantly recognizes the usual symptoms of a lipoma, or the unmistakable signs of neuronal hyperexcitability.

Doctors, dentists, engineers and other τεχνῖται see the world differently. What is more, they can act upon it in different and more efficacious ways. By determining the causes of certain phenomena, they are free to manipulate them and to change their effects. Conversely, those who lack this kind of insight are less able to shape circumstances in their favour. They are doomed either to suffer the consequences of their ignorance or to subject themselves to the authority of others. This sophistic topos is evoked, for instance, in Plato's *Lysis*, when Socrates claims that 'with regard to matters in which we become intelligent […] we shall do as we please, and nobody will care to obstruct us. Nay, not only shall we ourselves be free and have control of others in these affairs, but they will also belong to us, since we shall derive advantage from them; whereas in all those for which we have failed to acquire intelligence, […] we on our part shall be subject to others.'[23]

On the political level, the sophistic parallel between knowledge and power led to an increase in the social importance of education, which prompted a redefinition of the duties and prerogatives associated with free citizenship. In his classic study on ancient Greek education, Werner Jaeger points out that the epithet 'free' (ἐλεύθερος), in the 5th and 4th centuries BC, usually meant that one is 'not a slave' (δοῦλος, παῖς), while 'the kindred word "liberal" (ἐλευθέριος) describes the conduct appropriate to a free citizen, whether in generous spending or in frank speaking (which would be improper in a slave), or in a gentlemanly way of life.' Accordingly, 'the "liberal" arts are those which belong to "liberal" education — and that is the paideia of the free citizen, as opposed to the uncultured vulgarity of the unfree, of the slave.'[24] Influenced by the sophistic model, the notion of παιδεία came to stand from an increasingly wide array of practical and theoretical subjects, ranging from gymnastics and military training to grammar, rhetoric, music or medicine. This curriculum reflected both an individual and a social commitment to freedom and autonomy: individually, it enabled young men to overcome

[23] *Ly* 210a9-c4
[24] *Paideia* II, ed. cit., 54f.

their natural weaknesses, by helping them improve their physical and cognitive powers; socially, it aimed to guarantee the security and welfare of the polis. Upon reaching adulthood, educated individuals were accorded the title of citizens and encouraged to become statesmen, to embrace a military career or to pursue a liberal profession. In all of these activities, they were expected to put their expertise at the service of the community and to help defend the city against natural and political threats, to increase its economic and military might and to enrich its cultural and intellectual life.

By putting a new premium on knowledge and education, the sophists raised the frontier separating cognitive minority from cognitive majority and placed higher demands on citizens and free men. They maintained, not without controversy, that the πολῖται were still, in a certain sense, παῖδες — that they were not as free or as insightful as they believed, but still in need of education. Yet although the sophistic revolution amounts to a radicalization of the primitive notion of παιδεία, its scope is still very restricted. Although the sophists highlighted the limitations of human beings' usual standpoint and the possibility of a new and more sophisticated cognitive outlook, their critical efforts were themselves targeted by a new and more extreme educational project, outlined and discussed throughout Plato's dialogues.

Plato's idea of education is also linked to the possibility of a radical cognitive reform, but its scope is very different from the one envisaged in the previous models. Like the sophists, Plato is very critical of human beings' usual cognitive perspective, but his criticism does not apply only to primitive or savage individuals, nor to educated Greek citizens. It also extends to the sophists themselves, as well as to artificers and experts of all kinds. To be sure, he also recognizes the difference between a layman's and an expert's outlook on a specific subject or situation. He does not regard all human knowledge as equally valid or equally profound, nor does he believe every human being to be equally cultivated or talented. Yet despite these variations, Plato highlights the possibility of a more general and more serious form of ignorance, shared by every man and woman — a form of ignorance that lies at the very heart of human cognition, affecting both the general knowledge of laymen and the specialized knowledge of doctors, engineers and other τεχνῖται. According to his suggestions, what human beings are missing is not simply an educated account of this or that aspect of reality, but of reality *as such*. In other words, their usual understanding of reality is not similar to a doctor's understanding of

the human body or to an engineer's understanding of a bridge or a house. Every human being is a layman as regards *life itself*.

The debates between Socrates and his interlocutors are centred, in most cases, on the definition of a familiar word or concept, whose meaning is usually taken for granted. As the dialogues unfold, this assumption is shown to be illusory and the first attempt at a definition gives way to a second, and to a third one, and often to many others. All the while, Socrates' task is simply to question his opponents, to compare their answers with one another and to point out their contradictions.

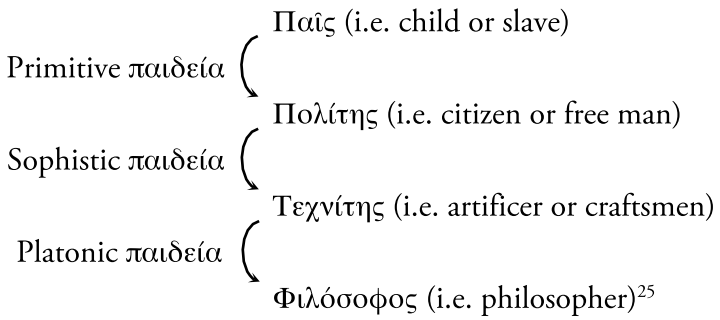
Initially, nothing seems to suggest that human reason is globally or fundamentally deficient. As the first inconsistencies start to emerge, Socrates' interlocutors are led to recognize they lack an adequate understanding of *certain* aspects of reality, but by no means *all* of them. If they are unable to offer a coherent definition of the concept in question, that simply means that the latter is more complex than usually assumed. To be properly understood, it must be broken down into simpler or more basic components. The exercise is thus repeated at a more elementary level: new concepts are brought up and new definitions are called for, but the results are equally discouraging. If the initial debate revolves around the notions of justice and injustice, for instance, or beauty and ugliness, the debaters soon find themselves grappling with much simpler determinations, such as 'good' and 'bad', 'big' and 'small', 'fast' and 'slow'. And here, once again, what seemed familiar and straightforward turns out to be neither. A new decomposition is called for. Eventually, the debaters are led to focus on notions as elementary as unity, identity or difference. But even here a consensus seems impossible. Just like before, their definitions prove ambiguous and unintelligible.

What emerges from this exercise is not merely the problematic nature of this or that human concept or set of concepts. Plato is making a much more extreme suggestion, namely that human reason is *entirely* or *globally* problematic. He suggests that the validity of every human belief is dependent on a series of concepts which are themselves merely assumed to be valid; and that the latter are supported, in turn, by a further subset of similar assumptions. Worse still, he suggests that the world's usual familiarity is grounded in a set of concepts which are themselves unfamiliar and ultimately unintelligible; and that the latter are grounded, in turn, in a further subset of similar assumptions. This extreme diagnosis amounts in many ways to a radicalized version of Oedipus' situation:

like Oedipus' knowledge, human knowledge is grounded in a web of assumptions that lend force to each other, but whose actual truth is far from certain. As in Sophocles' play, to question the validity of any one of these assumptions is to unveil a much wider cognitive problem, whose seriousness increases with the depth of one's enquiries.

The Platonic critique of human cognition will be discussed at length in the following chapters. For now, my main purpose is to highlight the specificity of Plato's conception of paideia. In the previous stages, the transition from a less educated to a more educated standpoint involved an element of self-criticism. The most primitive from of education reflected the acknowledgement of mankind's original impotence and the need for new and more sophisticated cognitive resources. In order to cope with a variety of practical demands and to attain the independence deemed worthy of free men, primitive individuals were led to overcome their limitations and to devise new forms of knowledge and understanding. With the advent of the so-called sophistic tradition, the acknowledgement of human ignorance was expanded and new forms of impotence were brought to the fore. Raising the stakes with regard to freedom and independence, the sophists highlighted the need for more sophisticated accounts of reality, irreducible to the general dictates of experience or common sense. Finally, Plato's educational reform can also be included in this global movement. Like the previous moments, it is also grounded in a critique of human cognition, but one that encompasses all of the previous educational achievements. Moving beyond the practical teachings of Prometheus and the 'technical' teachings of the sophists, Plato highlights the need for a *philosophical* reform of human cognition, liable to expose and reverse human beings' natural cognitive limitations.

This entire progression can be summed up in the following scheme:



Although Plato's philosophical project can be regarded as a new chapter in the history of education, it differs from the previous ones in at least two fundamental respects. The first one, already mentioned above, is the wideness of its scope. Plato's critique extends to the most basic elements of human cognition, and hence to every concept, idea or judgement relied upon in everyday life. His aim is not merely to identify and correct certain forms of ignorance, or to secure the knowledge required to achieve certain ends or aspirations. In his dialogues, human cognition emerges as a globally aporetic phenomenon, in need of a global rehabilitation.

What is at stake, for the first time, in Plato's enquiries is a truly universal form of education, directed at every human being as such. This aspect is in itself of great historical and cultural significance, as it introduces a cosmopolitan approach to human identity that is relatively new. In ancient Greek culture, human relations were not yet grounded in the modern biological notions of 'humanity' or the 'human species'. Ancient Greek men recognized others as their equals on the basis of their birthplace, language or social rank, and drew very clear and strict distinctions between citizens and non-citizens, fellow countrymen and foreigners, men and women, etc.[26] In Plato's dialogues, however, human ignorance is presented as an essential human trait, overarching racial and political distinctions. Although the latter are still at work in different and important ways, they are overshadowed by the universal *fact* of human ignorance. Education, rather than politics

[25] Although this scheme is useful, it is naturally simplistic. The different kinds of παιδεία under consideration are not pure or isolated phenomena, separated by fixed boundaries. Both from a conceptual and from a historical point of view, they amount to fluid and interrelated cultural events, which cannot be confined to a simple definition or to a precise historical period. Not only was human beings' 'primitive παιδεία' influenced, very early on, by the sophists' 'technical' model, but the latter was also instrumental to the development of Socrates' and Plato's philosophical model.

[26] See notably *St* 262c2ff.

or biology, is the true universal denominator, binding every man and woman to the same existential predicament.

The second main novelty of Plato's approach concerns the method employed in his dialogues. In the *Sophist*, for instance, when discussing the issue of education, Socrates explicitly evokes the Athenian conception of παιδεία, but only to reinterpret its conventional meaning: instead of defining education in an *acquisitive* manner, as the positive process whereby one is led to replace one's lack of knowledge with an increasingly specialized understanding of reality, Socrates defines education in a *subtractive* manner, as the negative effort to 'cast out the conceit of cleverness' (ἐκβολὴ τῆς δόξης)[27] underlying most human thoughts. He argues that the most grievous form of ignorance, 'separate from the rest, and as weighty as all other parts [of ignorance] put together', is 'thinking that one knows a thing when one does not know it'. 'Through this', he adds, 'all the mistakes of the mind are caused in all of us.'[28]

Accordingly, Plato's critique is a *negative* or *immanent* critique: its first aim is not to expand human knowledge, but to contract it, that is, to roll back the pretence of knowledge underlying most human ideas and to reduce what one *thinks* one knows to what one is *actually* certain of. Ignorance is shown thereby to be a widespread and involuntary phenomenon, whose first 'victims' are precisely those who judge themselves wisest or more cultivated. This unflattering conclusion, reiterated throughout most of Plato's dialogues, calls into question the validity of every traditional form of education, whether based on practical, political or technical standards. It suggests that every human being, however experienced or powerful, is still a mere παῖς, viz. a mere infant, or a mere slave.

The best and certainly the most famous illustration of this idea can be found in the beginning of the *Republic*'s seventh book, where human life, 'as regards both education and the lack of it' (παιδείας τε πέρι καὶ ἀπαιδευσίας)[29], is compared to a peculiar form of imprisonment. Socrates imagines a group of prisoners sitting in the bottom of an

[27] *So* 230b1. The word δόξα, used in many different contexts, can be translated as opinion, belief, judgement, expectation, notion, conjecture, etc. In Plato's works, however, it is usually associated with an *unwarranted* opinion, or belief, or judgement, etc. I will return to this issue below, in section 2.1.

[28] *So* 229c1-6

[29] *Re* 514a2

underground cave, forced to face forward and to watch the shadows of different objects projected on a wall. The prisoners' legs and necks are bound by shackles, which prevents them from seeing each other, their own bodies, or indeed anything other than the shadows on the wall. Moreover, since their imprisonment has lasted for as long as they can remember, their lack of freedom is not perceived or experienced as such. They do not know any other kind of life, and live therefore under the assumption that they are free, and that the wall in front of them encompasses the whole world.

According to Socrates, these prisoners are 'just like us'[30]: their ignorance resembles our own ignorance and their captivity illustrates the lack of freedom imposed on us by our usual understanding of reality. In other words, just as the prisoners are ignorant and unfree but fail to notice it, we too are ignorant and unfree but fail to notice it; and just as the prisoners, in order to acknowledge their captivity, must break their chains and look around, we too, in order to acknowledge our own captivity, must free ourselves from our usual pretence of knowledge. Once again, this image is grounded in a Heraclitean thought pattern, designed to convey a possibility whose actualization lies beyond our usual cognitive reach: starting from the familiar contrast between the prisoners' life and our own life, we are led to envisage the unknown contrast between our life and a new kind of life, whose embracement would amount to an equally profound cognitive transformation.

As Schopenhauer once wrote, 'every man regards the limits of his own vision as the limits of the world.'[31] The so-called allegory of the cave is an eloquent illustration of this phenomenon. Since the cavemen are unable to see the chains that bind them, they mistake the cave for the whole world. Therefore, their education seems to be, for the most part, a closed affair: although their learning is never really over, for there are always new shadows to see and new movements and meanings to observe and interpret, their overall conception of reality is not likely to be radically altered. In like manner, inasmuch as most human beings believe their usual understanding of reality to be fundamentally reliable, they do not bother to question its underlying principles. Indeed, although human reason is far from perfect, it is usually deemed accurate and empowering enough

[30] *Re* 515a5
[31] *Sämtliche Werke* V, ed. cit., 245

to allow for a relatively free life. And although one never ceases to learn new things, or to correct and refine one's previous knowledge, one does not expect one's usual standpoint to be radically refuted or transformed. According to Plato, however, this expectation may be entirely unwarranted. Human knowledge may amount, after all, to a complete illusion, and one's real education may not be a thing of the past, but a thing of the future.

1.3. *The origin and fate of a Platonic idea*

Socrates' enquiries usually lead to a damning conclusion: human beings are not as free or as knowledgeable as they tend to assume; their understanding of reality is not as solid or self-evident as it appears and the freedom it is thought to provide may well be illusory. This general idea was first articulated in Plato's dialogues, but it soon acquired a wider historical dimension. Throughout the centuries, following Plato's lead, different authors and doctrines highlighted the inherent fragility of human reason. Several explanations were offered as to how human ignorance limits human freedom and several methods were put forth to fix this problem. This long history is of course diverse and heterogeneous, but its different stages can nonetheless be regarded as different products of the same Platonic lineage, equally focused on the dialectical relationship between freedom, bondage and the pursuit of true knowledge.

In this context, it is worth mentioning, if only in passing, the Stoic tradition, and particularly the Stoic definition of freedom. Like Plato, authors such as Epictetus, Chrysippus, Marcus Aurelius or Seneca have highlighted the necessary co-implication of life's practical and intellectual dimensions. They have pointed out that living is just as much about choosing and acting as it is about interpreting reality and making sense of one's circumstances. Decisions and actions, however immediate or instinctive, are always dependent on a specific diagnosis of what things are, of how they stand, of how they are liable to evolve, etc. Human life is continually grounded in a vast web of cognitive claims, which guarantee not only the world's usual familiarity, but also the feeling that one is free to move around in it, to make one's own decisions and to pursue one's own interests. The Stoics' critique of human freedom, as it is usually conceived, is focused

precisely on this cognitive component: since freedom is dependent on a series of claims about reality, its actual value is conditional on the cogency of these claims. If one's diagnosis of reality is false or illusory, one's freedom will also be false and illusory. And, moreover, as long as this illusion is maintained, it will prevent one from acknowledging one's lack of freedom and embracing a truly free or independent life.

For everyday consciousness, reality is something given, a collection of facts endowed with fixed and self-evident meanings. Things are what they are and the subject's main task is to acknowledge their identity and to live and act accordingly. Against this cognitive attitude, the Stoics point out, in different ways, that reality is just as much a subjective phenomenon as it is an objective one. Human beings are never neutral or disinterested spectators, confronted with a pre-formed existential setting. Their usual account of reality is not selfless or impartial, but biased and self-centred. It is the reflection of a vital agenda, or a vital narrative, as it were, in which the self plays the leading role and everything else is accorded a secondary position. People, objects and events are never relevant *as such*, or *in themselves*, but as different elements of the same existential scenario. Everything that *is* is so because it affects one *in a certain way*, because it has *a certain degree of importance*, because it plays *a certain role in one's life*, etc. The world's identity emanates from one's own identity: it is a *function* of self-identity, and only as such is it intelligible and meaningful.

To borrow the Stoics' own vocabulary, life is not based on facts, but on ὑπολήψεις, that is, on subjective ascriptions of meaning grounded in a subjective account of reality.[32] In everyday life, however, this important detail is systematically overlooked. The self is led to forget its cognitive primacy and to regard its own interpretation of reality as reality itself. Ὑπολήψεις are mistaken for facts and one's entire existence is made to depend on judgements, events and circumstances whose meaning seems to allow for no alternative and whose relevance appears unquestionable. As a consequence, the Stoics argue, freedom itself is made to depend on a series of external or contingent 'facts', whose meaning and relevance is also taken for granted. For most people, freedom depends on material circumstances, on the satisfaction of bodily needs, on the possession of external goods,

[32] Cf., for ex., MARCUS AURELIUS, *Meditations* II, 15; IV, 3; VI, 52; IX, 32; XII, 8, 22, 26.

etc. Only by overcoming the 'factual' limitations of weakness, pain, disease or poverty are men and women able to lead free and independent lives. For the Stoics, on the contrary, this definition is inherently contradictory, as it places the source of freedom in the one place it can never be found, namely in the realm of contingency. If one's life is made to depend on fleeting goods or transient qualities, whose enjoyment can never be truly secured, one's freedom is also bound to be fleeting and transient, vulnerable to the sway of circumstances and to the random dictates of chance.

True freedom, as the Stoics understand it, cannot be achieved in this way. It requires a global cognitive reform, whereby the *hypoleptic* origin of every concept, fact and event is brought into view. Once consciousness acknowledges that 'everything is opinion, and that opinion is in [its] power'[33], reality acquires a new degree of plasticity. As soon as facts cease to be viewed as facts, their meaning is no longer univocal and their relevance is open to revision. Consequently, the vital narrative whereupon cognition is usually grounded is itself open to revision: what was formerly deemed essential may now be viewed as secondary and the life-goals placed at the top of one's usual list of priorities may now be dislodged and replaced by new ones. This corrective procedure is the central feature of the Stoic's paideia and the key to the Stoic redefinition of freedom. By reviewing and readjusting his vital priorities, the Stoic aims to reduce life's share of contingency. Instead of placing his hopes on what is not up to him, and falling prey to uncertainty and disillusionment, he shifts his attention to what is actually within his reach. Detaching himself from the bustle of worldly passions and concerns, the Stoic withdraws into the simple element of thought and looks therein for deliverance.

This ascetic release is the final goal of Stoicism. However, even if such a release is possible, and one is indeed able to review and readjust one's entire outlook on reality, this transformation can never be fully secured or taken for granted. The tendency to forget oneself, to rely on assumptions as though they were facts and to mistake thereby the contingent for the essential is a recurring aspect of human cognition, renewed and fortified with time and habit. Accordingly, each new experience is a potential source of delusion and the Stoic's task is never truly finished: it is a life-long exercise, or a life-long

[33] MARCUS AURELIUS, *Meditations* XII, 22: 'Ὅτι πάντα ὑπόληψις καὶ αὕτη ἐπὶ σοί.'

struggle, whose conquests are always provisional.

The idea that freedom requires a specific kind of education, argued by Plato and the Stoics, is not exclusive of these two traditions. It also played a central role in the evolution of Christian thought, from its primitive form to its modern developments. Like the Stoic outlook, the Christian outlook highlights the possibility of a complete transformation of human cognition, leading to a complete redefinition of the aims, concerns and priorities that shape human life. The starting point is once again the acknowledgement of everyday consciousness' limited cognitive perspective. Although for most of us the world is endowed with a fixed and self-evident meaning, open to a limited degree of variation, this attitude is the reflection of a specific existential scenario, grounded in a specific assessment of who we are, what matters most to us and what is at stake in our life. From the Christian point of view, however, this assessment is usually misguided: it tends to mistake what is secondary for what is essential and to misplace life's true cognitive foundation.

But though Stoicism and Christianity share the same general diagnosis, they prescribe different kinds of transformation. In the former case, the emphasis is placed on the contrast between subjectivity and objectivity, selfhood and otherness. The individual is led to forego his attachment to the objective world and to find freedom and satisfaction within himself. Stoicism's final purpose is a fully autarchic or self-sufficient life: by disregarding all external forms of authority, the Stoic aims to become his own master and to take his life in his own hands. In the case of Christianity, the emphasis is placed on the contrast between the self and something higher than the self, in which both subjectivity and objectivity are ultimately grounded. The deliverance promised by the Christian alternative is not about independence or self-sufficiency. On the contrary, it can only be attained through the recognition of a new and more essential form of servitude.

What is at stake in Christianity is the opposition between a *theocentric* outlook on reality — that is, a cognitive regime that has God as its centre and underlying condition — and a *non-theocentric*, or *anthropocentric* outlook on reality — that is, a cognitive regime that has the individual as its centre and underlying condition. From the Christian point of view, however, the second of these alternatives presupposes the first one: *deus* and *mundus* are not opposed to each other as two equivalent or symmetrical options are opposed to each other. The latter is contained within the former: it is a function of the

former, a specific angle or contraction of the former's universality. Therefore, though one may spend one's entire life in ignorance of, or in opposition to, the Christian alternative, one will still be bound to it in a primary or essential way. For the Christian believer, the non-believer's standpoint can only be maintained through self-ignorance or self-denial. Since cognition itself is the product of a divine will, to deny such a will is to deny one's own origin and underlying condition.

The Christian paideia amounts, therefore, to the widening of one's usual cognitive angle. As life's original source of meaning is transferred to a higher plane, the individual is led to see the world differently and to rethink his usual set of priorities. Satisfaction and self-fulfilment are no longer tied to concrete or material concerns and freedom itself can no longer be found within the worldly realm. It is now dependent on a global calling, or a global *vocatio*, to which everything else is subordinated. Unlike the objective freedom of ordinary consciousness, or the subjective freedom of Stoicism, Christian freedom is not about doing what one pleases, or avoiding external forms of compulsion. It consists rather in the possibility to break free from a life devoid of meaning and direction and to embrace one that is globally meaningful and rewarding.

Although this transformation entails a global disqualification of the earthly realm, Christianity does not prescribe, like Stoicism, a complete detachment from objectivity. In the Christian's eyes, the world is not eliminated, but transfigured. The whole universe is seen under a new light, and everything is revealed as the reflection of something else. Each objective content is both itself and not itself, a unique essence and the gateway to a higher cognitive order. The world is rendered meaningful *through God*, just as God himself resides in every worldly aspect, and this absolute reciprocity is traditionally embodied by the figure of Christ. By highlighting this co-implication — reiterated by the christological themes of incarnation, grace and salvation — Christian theology emphasizes the concrete nature of its spiritual demands: instead of a negative ideal, to be attained in another life, freedom is a positive ideal, to be pursued and cultivated in this one; it is the power to raise the worldly to the divine, and to find in this world the image of the next.

Both the Stoic and the Christian perspectives require specific forms of education. Furthermore, both can be read as historical developments of Plato's portrayal of human life as an imprisoned life. But despite their enormous influence, they are not the only heirs to Plato's founding idea. Stoicism and Christianity will play an important role in the following analysis, as will other forms of paideia, but its main focus will be placed elsewhere. I will pursue a wider historical thread and confront Plato's views with the work of two modern authors who have explicitly linked the notions of ignorance and truth to those of bondage and freedom — or slavery and mastery — and whose philosophical projects are also devoted to the double task of educating and liberating human consciousness. These authors are Georg Wilhelm Friedrich Hegel and Friedrich Nietzsche, and their ideas on the relationship between freedom and knowledge can be counted, in my view, among the most direct historical follow-ups of Plato's diagnosis.

The reasons for this choice can be anticipated by a brief outline of the main affinities shared by these three philosophical projects. For Hegel and Nietzsche, as well as for Plato, there is a fundamental disagreement between human beings' usual estimation of their cognitive powers and what the latter are actually worth. On the one hand, most of us believe we have a relatively good grasp of reality; we are convinced we can discern what is at stake in most ordinary situations and make informed and efficacious decisions. On the other hand, upon closer inspection, most of our judgements are found to rest on arbitrary and unintelligible assumptions. Our usual knowledge is a matter of belief, rather than truth, and promises more than it can deliver.

For Plato, Hegel and Nietzsche, this problem is not merely theoretical or epistemological. If we are indeed unable to vouch for the truthfulness of our judgements, our freedom may be seriously compromised. If our choices are based on unwarranted or unintelligible assumptions, their effects might prove unexpected, unfavourable or even harmful. In these situations, although we may be free to choose what we please, we are unable to achieve what we want, or to discern what is in our best interest. Our choices are not empowering, but debilitating, and our freedom amounts to an involuntary form of imprisonment.

Finally, Plato, Hegel and Nietzsche all agree that this lack of freedom calls for a global reform of human reason. For all of them, it is up to philosophy to correct this problem, either by raising our usual standpoint to the standpoint of truth and/or by

questioning the very standard of truth such a standpoint is supposed to meet. In either case, a philosophical reversal of human ignorance does not mean, in this context, a mere increase in the amount of things we know, nor the replacement of our usual beliefs about reality with different, but equally uncertain or arbitrary beliefs about reality. What is at stake in these three philosophical reforms is a *qualitative* transformation, aimed not only at the content, but also at the form of human cognition.

Despite these similarities, however, the Hegelian and the Nietzschean accounts of the relationship between power, knowledge and freedom are much more than mere reiterations of the Platonic account. In Plato's case, although Socrates stresses the need to break free from the chains of ignorance and opinion, it is at best unclear whether such release is actually possible. The dialogues are ambiguous with regard to their own purpose or outcome: all things considered, they can be read either as general descriptions of the imprisoned nature of human cognition or as actual attempts at liberation. In the first case, their task is a *negative* task. They are the starting point of a long philosophical tradition, whose *critical* role is still shared, in its essence, by Kant's epistemology. In the second case, their task can be envisaged in a *positive* way, as an effort to move beyond the critical stage, towards an actual deliverance from ignorance and bondage.

In Hegel's and Nietzsche's case, there is no room for such ambiguity. Both authors are explicitly set on overcoming, albeit in very different ways, the stage of criticism, and both claim to have found an actual solution to the Platonic predicament. As regards the relationship between freedom and knowledge, the temerity of their projects is precisely what sets them apart from previous philosophical undertakings. Hegel's and Nietzsche's aim is not simply to expose the limitations of human reason, nor to detail the form and scope of the imprisonment they produce. Rather, they have taken on the Promethean task of *setting humankind free*.

This introduction is not yet the place to consider how this transformation is supposed to come about. Nonetheless, I will start by offering a brief overview of Plato's, Hegel's and Nietzsche's standpoints, in order to bring out their main similarities and their main differences.

I. *Plato's scientific project*[34]

In Plato's dialogues, Socrates is usually engaged in the task of dissecting human reason and mapping out its inner structure. His aim is to bring to light what usually lies hidden, that is, to render explicit the methods and operations employed by the human mind in its daily activity. As a consequence of this procedure, what is habitually regarded as clear and self-evident turns out to be neither and one's usual account of reality emerges as a vast and complex labyrinth, filled with obstacles, blind spots and dead ends.

Socrates' characterization of human reason is based on the notion of εἶδος, or ἰδέα, usually translated as 'idea', 'form' or 'concept'. His first and most basic concern is to show that human beings' understanding of reality is entirely made up of εἴδη, or ἰδέαι. Ideas, forms or concepts are the basic currency of human cognition, the shape in which everything whatsoever comes to be known, and their variety is virtually unlimited: they may stand for actual things, for qualities of things, for quantities, for feelings, for actions, for logical functions, etc. This does not mean, however, that each act of cognition corresponds to a different and altogether unique εἶδος. If that were the case, human experience would be purely nominalistic and the world would be a permanent source of

[34] Like all general labels, the titles given to the three main parts of this work — namely 'Plato's scientific project', 'Hegel's phenomenological project' and 'Nietzsche's metacognitive project' — are merely indicative. The words 'scientific', 'phenomenological' and 'metacognitive' have all got a long history behind them and can all be understood in very different ways. In the present context, they will be used in connection with the problem of truth and true knowledge. More specifically, they will be used to characterize Plato's, Hegel's and Nietzsche's diverging views on the possibility of attaining a *true*, and hence *truly free* mode of cognition. Although the exact meaning of these words will only be fully understood in the course of the analysis, it can be anticipated up in a few brief notes. Firstly, in Plato's case, the term 'scientific' refers to the ideal of science, or ἐπιστήμη, that guides Socrates' dialectical enquiries. According to Plato, only when human consciousness' usual cognitive standpoint has been raised to a truly scientific or epistemic standpoint, no longer based on mere beliefs, or mere opinions, can human life be a free life. Secondly, in Hegel's case, the term 'phenomenological' refers to the peculiar philosophical undertaking carried out in the *Phenomenology of Spirit*. According to Hegel, only a 'phenomenological' critique of human cognition, in the terms we will later consider, can lead to the fulfilment of the scientific ideal pursued by Platonism and to the attainment of a truly free life. Finally, in Nietzsche's case, the term 'metacognitive' stands for a complete break from the scientific aspirations of the two previous models. For Nietzsche, human consciousness' usual lack of freedom stems not from the lack of truth or science, but from the very obligation of viewing reality in terms of true and untrue contents and events. His conception of freedom is dependent on the rejection of the very idea of truth and science that guides Plato's and Hegel's projects. According to Nietzsche, only a standpoint released from its *cognitive* pretensions — a *metacognitive* or *metagnoseological* standpoint, which has traded its usual concern with truth with a more primordial concern with *life* and *vitality* — can lead to a free or independent existence.

perplexity. On the contrary, in addition to maintaining that reason is entirely *eidetic*, Socrates also points out that its structure is originally *synthetic*. Instead of a different εἶδος for each cognitive content, human thoughts and judgements are grounded in a web of repeating and interchangeable εἴδη. In respect of 'all the ideas or forms, the same statement holds, that in itself each is one, but that by virtue of their communion with actions and bodies and with one another they present themselves everywhere, each as a multiplicity of aspects.'[35]

But not all of reason's eidetic combinations are equally legitimate or equally essential. As regards the first of these observations, Plato compares the multiplicity of human concepts to the letters of the alphabet: while some letters can be joined together to form words and sentences, others cannot; in like manner, while some εἴδη can be combined in intelligible and coherent ways, others cannot.[36] Human cognition is bound to a specific grammar, without which reason and meaning are impossible. Secondly, as regards the essentiality of reason's eidetic combinations, Plato points out that the alphabet of human cognition is a vertical alphabet, divided into different hierarchical stages. It is structured as a vast conceptual tree, where each particular concept is both supported by the combination of lower-order concepts and subordinated to higher orders of meaning. In the case of physical bodies, for instance, the terms 'pen' or 'pencil' refer to single, self-contained objects. Nonetheless, pens and pencils are only recognized as such because they occupy a specific position within a wider conceptual hierarchy. Looking downwards, they are the genera or classes (γένη) of which every particular pen or pencil partake; looking upwards, they themselves partake of higher-order γένη such as 'tool', 'object', or 'body', as well of universal γένη such as 'identity' and 'difference', 'rest' and

[35] *Re* 476a5-7. According to a popular exegetical trend, Plato's εἴδη are not ideas or concepts forged by consciousness, but the composing elements of *reality itself*, independent of one's cognition thereof. In other words, εἴδη are not so much ideas or concepts as the very things one's ideas and concepts attempt to convey. They lie *beyond* consciousness, and can only be grasped through the imperfect lens of human reason. This kind of reading, however, fails to appreciate the complexity and the seriousness of Plato's critique of human cognition. As we will see in the following chapters, the cognitive 'imprisonment' or 'enslavement' denounced by Socrates stems precisely from reason's *immanent* or *self-enclosed* structure: every act of cognition entails the simultaneous positing of a specific idea (or set of ideas) and the real thing(s) such an idea (or set ideas) is supposed to convey; ideas refer back to real things, which are themselves conveyed by ideas, and this entire cycle takes place within the eidetic realm — it is grounded in εἴδη, or in complex combinations of εἴδη.

[36] Cf. *So* 252e9ff. and *Phi* 17a8ff.

'movement', 'simultaneity' and 'duration', etc. To be sure, this complex architectural structure is not usually apparent when one looks at a pencil, or picks it up to jot something down. The point, however, is that the very idea of pencil *requires* the simultaneous, albeit unconscious acknowledgement of this entire conceptual apparatus. If a specific pencil were not regarded as an individual instantiation of the εἶδος or the γένος 'pencil', its appearance would be radically transfigured and its usual familiarity would be compromised. Likewise, if pencils were not automatically regarded as solid objects, endowed with a determinate length, subsisting through time, they would not be recognized as they usually are, nor used to jot things down.

In addition, a third and essential aspect of Plato's characterization relates to the practical or existential nature of reason's eidetic alphabet. As argued, for example, throughout the *Philebus* or the *Gorgias*, human beings' usual account of reality is not exclusively, nor indeed primarily motivated by theoretical or epistemological concerns. In this respect, Plato anticipates an essential element of the Stoic tradition, briefly mentioned earlier: since human beings are never impartial or disinterested spectators of reality, knowledge is rarely, if ever pursued *for its own sake*. Men and women wish to know things in order to achieve certain goals or to avoid certain obstacles. Their usual understanding of reality is guided by a global diagnosis of what is at stake in their life — of who they are and who they want to become, of what is in their best interest and what is not, of what is worth pursuing and what can be neglected, etc. In Plato's vocabulary, every human life is globally oriented towards what is perceived, at each given moment, as the *best* or the *greatest good* available (τὸ μέγιστον ἀγαθόν, τὸ βέλτιστον).[37] And this fundamental drive also extends to reason's predicamental scale. Again, when picking up a pen or a pencil, one is usually not absorbed by such difficult questions as who one wants to become or what is at stake in one's life. These interrogations are normally reserved to special, introspective moments, whose seriousness contrasts with the usual bustle of everyday life. However, Plato points out that one's usual conception of pens, pencils and indeed anything else presupposes a specific, albeit inexplicit or provisional answer to these fundamental questions. Pens and pencils are never perceived or defined on their

[37] Cf., for ex., *Go* 499e8-9, 468a-b; *Sy* 205e7-206a1; *Me* 77b-78b; *Eu* 282a2; *Phi* 20d.

own account, but as part of a wider existential narrative. Whereas for most people a pen is merely something to write with, or a means of fulfilling other, more important tasks, for pen manufacturers or pen collectors the εἶδος 'pen' is less generic. Since their life project is more directly tied to what a pen is or looks like, their interest in pens is more intense and more demanding. They tend to notice certain characteristics thereof which for most people tend to remain hidden, such as a pen's inner structure, its outward texture or its year or fabrication.

These and other issues will be addressed below in more detail. For now, it is important to sum up the main features of Plato's portrayal of human cognition. We have seen, so far, that human thoughts and judgements are supported by a universal, self-referential and stratified web of concepts. Each mental representation, however simple or immediate, is already endowed with a specific alphabet, a specific morphology and a specific syntax. Each is already committed to a highly complex constellation of εἴδη, subordinated to a specific assessment of the structure and meaning of human life. But apart from unveiling this complex cognitive structure, Plato's dialogues are also meant to highlight how far it usually is from most people's minds. Unlike Socrates, most human beings are usually blind to reason's inner workings. Their usual grasp of the value and scope of their knowledge is very superficial, and rarely called into question.

For everyday consciousness, cognition as such is not really an issue. Since one's account of reality is deemed clear and adequate enough, one's cognitive access to reality is not usually considered or discussed. This means, first of all, that most human concepts are not experienced as concepts, but as the very things whose meaning they are supposed to convey. One's conception of pens or pencils, for example, tends to coincide with pens or pencils *themselves* and the very idea of cognitive access is thereby disregarded. This also means, furthermore, that most human concepts are automatically regarded as clear and coherent. One usually knows what is meant by the terms 'pen' or 'pencil', for example, and has no difficulty employing them in everyday life. Pens and pencils are familiar props of a familiar setting, where one moves at ease.

These two assumptions are not only extensive to most human ideas, but also to all of the other cognitive properties listed above. The fact that most concepts are liable to partake in different conceptual arrangements is hardly construed as noteworthy or problematic. Likewise, the fact that most concepts can be decomposed in lower-order

concepts, or joined together to form higher-order concepts, is also accepted as something natural and self-evident. Everyday consciousness breezes through these difficulties, as it were, because it does not usually stop to consider them. Its relationship with concepts and conceptual arrangements is brief and inattentive, aimed at practical results, rather than epistemic accuracy. Socrates' enquiries, on the other hand, are designed to highlight and reverse this natural tendency. They are meant to show that the edifice of human reason, when examined more carefully, is anything but simple or self-evident.

Plato's method is to isolate a given concept or set of concepts and to challenge their adequacy and intelligibility. As regards the first of these issues, his usual strategy is to show that the meaning of human concepts is fully dependent on other concepts, whose definition is provided in turn by other concepts. The assumption that one's account of reality coincides with reality itself is therefore a fully *conceptual* assumption, backed by a constellation of concepts that are themselves supported by the same immediate assumption. Conceptual truth is usually a mere *presupposition* of truth, unable to guarantee is own validity. As regards the issue of intelligibility, Plato's argument is built along the same lines. Since the definition of a given concept is provided by a specific conceptual arrangement, pens and pencils are defined by their participation in continuity, solidity, length, colour, duration, etc. Yet all of these concepts are also used to define many other objects and events. What makes them specific features of *this pen* is their participation in the concept 'pen' — but the latter is what it is only insofar as it refers back to continuity, solidity or length! These and similar contradictions are bound to arise because the very ideas of participation, subordination and interchangeability, upon which the entire alphabet of human reason is founded, are ultimately confusing and unintelligible. Whenever they are singled out and dissected, consciousness is pushed to the wall and led to face its own ignorance, just like Socrates' interlocutors in the course of his enquiries.[38]

By encouraging natural consciousness to turn inwards and confront its own limitations, Socrates plants the seed of a terrible suspicion. If human cognition is based

[38] Cf. *So* 253b8ff.: what human beings usually lack is 'a clear perception of one form or idea extending entirely through many individuals each of which lies apart, and of many forms differing from one another but included in one greater form, and again of one form evolved by the union of many wholes, and of many forms entirely apart and separate.'

on a vast web of assumptions, whose truth-value rests on further, equally unverified assumptions, reason as we know it may well amount to one gigantic misunderstanding. Indeed, the scope of Plato's critique should not be minimized: it refers not to a specific domain of human reason, nor to a specific brand of ideas or concepts; his denouncement is primarily aimed at the *form* of human reason, and affects therefore its entire cognitive repertoire. The comparison between the human race and a community of prisoners, between human life and an involuntary dream,[39] between human consciousness and a diseased or deranged consciousness — all of these images refer to a global or pervasive problem. They suggest that the entire edifice of human ideas may be nothing more than a highly complex patchwork of beliefs, with no true foundation. Human knowledge, in its habitual form, may simply reflect 'the aberration of a soul that aims at truth, when the understanding passes beside the mark.'[40]

If Plato's suspicions are correct, this problem is liable to have serious practical consequences. If the concepts underlying our daily judgements are indeed unfounded and unintelligible, the vital narrative that guides our thoughts and actions is also open to attack. Apart from pens and pencils, or trees and cars, Plato's denouncement also extends to our usual idea of what is good or bad, pleasurable or painful, just or unjust, etc. What is ultimately being challenged, therefore, is the overall cogency of our usual diagnosis of who we are and where we stand, of what is in our best interest and what is not, of what is worth pursuing and what should be abandoned, etc. In short, what is at stake is the actual freedom to steer the course of our lives — the sovereignty we readily take for granted, but are usually unable to guarantee.

For Plato, the only way to deal with this problem is to do as Socrates does, that is, to

[39] The difference between dreaming and being awake is an important Platonic topos, usually conveyed by the contrast ὕπαρ / ὄναρ and other related forms. See notably *Re* 476c-d, 520c-d, 533b-c, 534c-d, 574c-e, 576b; *Phdr* 277d; *Phi* 36e, 65e; *Th* 158b; *Ti* 71e; *Cr* 439c-d; or *L* 969b, 533 b-c. I will return to this issue below, in section 11.1.

[40] *So* 228c10-d2. Plato describes this peculiar kind of delusion as a form of παραφροσύνη, that is, as a form of madness. The prefix παρα- is used in this context to convey both the idea of physical deflection or deviation — the soul passes *beside* truth just as an arrow passes *beside* the mark — and the idea of mental or spiritual deviation — the idea of madness or derangement.

The term παραφροσύνη is often used in the *corpus hippocraticum*, where it is usually contrasted with σωφροσύνη, that is, with the health or sanity required to see things *as they really are*. In the *Sophist*, as in many other Platonic texts, human reason's alleged health is exposed as a form of disease in disguise. Men and women do not know what they think they know, nor do they understand what they regard as familiar or self-evident.

bring reason's alphabet into focus and to try and make sense of its inner structure. In order to determine which ideas are in play in our usual standpoint, how they relate to one another and how reliable they really are, we need a specific science, and indeed 'the greatest of all sciences'[41], namely ἡ ἐπιστήμη διαλεκτική. In contrast to the usual captivity of natural consciousness, dialectics is 'the science that belongs to free men'[42] and the basic principle of a philosophical reform of human reason. But since most human beliefs are usually assumed to be true, the dialectical re-education of consciousness must start out as a negative undertaking. The philosopher's first task is to free men and women from their 'conceit of cleverness', to turn their self-confidence into perplexity and to clear the way for a non-presuppositive account of reality. Those who are willing to undergo this transformation 'will have to change the opinions which they had at first accepted, so that what was great will appear small and what was easy, difficult, and all the apparent truths in arguments will be turned topsy-turvy.'[43] In the end, if this reversal is successful, a new starting point will emerge. With all of their assumptions behind them, men and women will finally be open to the possibility of true knowledge and a truly free life.

II. *Hegel's phenomenological project*

The idea that human life, as regards one's usual powers of cognition, is not a free life has wielded great influence throughout the history of Western thought. Yet few authors have taken Plato's insistence on the need for a philosophical liberation of mankind as seriously or as far as Hegel. His critique of human reason represents in many ways an amplified and systematic response to Plato's dialectical critique, set on fulfilling the latter's main promise. As one commentator recently pointed out, Hegel's 'whole philosophical system, in all its incredible breadth and detail, can be understood as a single extended demonstration of the importance and meaning of freedom. Moreover, Hegel's philosophy is not only *about* freedom, but also claims to be productive *of* it: in the course

[41] *So* 253c4-5
[42] *So* 253c7-8
[43] *So* 234d6-e1

of his philosophical investigation of what it means to be free, Hegel arrives at the view that freedom depends upon the practice of philosophy.'[44]

Like Plato's views on freedom and philosophy, scattered throughout different dialogues and debates, Hegel's approach to this issue is not confined to a single text, to a specific angle or to a definitive line of approach. The concept of freedom is taken up at different junctures of Hegel's philosophical system, corresponding to different stages of self-awareness, or different 'shapes of Spirit'. Depending on the distance separating one's cognitive standpoint from a truly scientific or epistemic standpoint, freedom is liable to assume different meanings, implications and modes of expression.

Hegel's *Outlines of the Philosophy of Right*, devoted to the 'development of the idea of the absolutely free will'[45], is usually regarded as the most complete and direct exposition of his views on freedom[46] — from the more immediate notions of free will and free choice to the social and political institutions deemed necessary for the actualization of freedom. However, as Hegel points out in the preface, the dialectical progression developed throughout this work presupposes an acquaintance with the 'philosophical' or 'speculative' way of knowing already set forth in his logical writings, namely in the *Science of Logic*.[47] In the latter work, in turn, the reader is warned that the text's starting point does not amount to an absolute beginning, as it were, since the ensuing developments already take place at the level of scientific thought — that is, at a level which is essentially different from one's usual, 'pre-philosophical' or 'pre-scientific' standpoint.[48]

The first major step towards a philosophical release of ordinary consciousness must not be looked for in the higher stages of Hegel's system, but in an earlier propaedeutic work entitled *Phenomenology of Spirit*. Its overall aim is summed up in the opening lines of the ad published by Hegel on October 28th 1807 in the newsletter of Jena's *Allgemeine Literatur-Zeitung*: 'The *Phenomenology of Spirit* aims to replace the psychological

[44] DUDLEY, op. cit., 15

[45] *HW* VII, 87; *PR*, 50

[46] Along with the middle section of the *Encyclopaedia's Philosophy of Spirit*, entitled 'Objective Spirit', which amounts in many ways to a systematic reworking of the dialectical progression laid out in the *Philosophy of Right*.

[47] See *HW* VII, 12f.

[48] See *HW* V, 17f.

explanations and the abstract accounts of the foundation of knowledge. It considers the *preparation* to science from a point of view that makes it a new, interesting, and indeed the first science of philosophy.'[49]

The following analysis will be focused on Hegel's *Phenomenology*. The debate concerning Hegel's later assessment of this work, as well as its articulation with his mature philosophical output, will be bracketed in favour of a more direct approach to his phenomenological method. In keeping with my previous focus on the notion of paideia and its various historical declensions, Hegel's views will come up in connection with his unique understanding of the relationship between freedom, knowledge and education, laid out and tested throughout the *Phenomenology*. Like the previous educational projects — and unlike the Hegelian system's subsequent stages — the *Phenomenology* does not require a previous philosophical education. Its starting point is human consciousness in its native or natural form, with all the assumptions and beliefs that tend to dominate its usual outlook on reality.

Hegel's long dialectical progression is set in motion by the realization that the knowledge held by natural consciousness is not actual knowledge, but only the assumption or the *appearance* of knowledge.[50] Like Plato, Hegel aims to expose and overcome the spuriousness of natural ideas and concepts through an immanent critique of human cognition, designed to '[lead] the individual from his uneducated standpoint to [the standpoint of] knowledge.'[51] This critique is meant to precipitate the transition from a mode of consciousness that merely *asserts* the knowledge it possesses to be true to one that actually *knows* it to be such.

In addition, Hegel also believes that a higher or truer cognitive standpoint translates into a higher or more developed form of freedom. His phenomenological consciousness echoes the imprisoned consciousness of the *Republic*'s cavemen and his whole progression

[49] *HW* III, 593

[50] The ambiguity of Hegel's use of *Erscheinung* or *erscheinendes Wissen* is also present in the English words *appearance*, *to appear*, etc. According to Hegel, phenomenal knowledge is both the *knowledge of appearances* (viz. of that which immediately appears or comes into view — *Erscheinung*, φαινόμενον) and the *appearance of knowledge* (a mere semblance or pretence — *Schein*). As we shall see below, the transition from the first of these meanings to the second one is the *principium motionis* of the entire phenomenological progression.

[51] *HW* III, 31; *PS*, 16

can be compared to the long journey whereby the cave's darkness is progressively left behind and finally 'cut short' 'by a sunburst which, in one flash, illuminates the features of the new world.'[52] In particular, Hegel highlights and explores an aspect of Plato's account whose importance is often disregarded. As we will see later on,[53] the allegory of the cave is not about a simple contrast between ignorance and knowledge, or captivity and freedom. Socrates' narrative sets out an entire cognitive scale, or an entire scale of freedom, with many intermediate stages. In the beginning, as one of the prisoners is unchained, his ignorance gives way to a new and more sophisticated standpoint. His idea of reality is no longer reduced to the shadows on the wall and his life-long captivity is revealed as such. But this discovery is only the first in a long series of transformations.[54] Once the prisoner overcomes his *physical* imprisonment, he is led into a new form of imprisonment, namely an *optical* imprisonment, imposed by the fire's blinding glare. Afterwards, as he regains his eyesight, makes out the cave's ascending path and reaches the cave's entrance, the sun's blinding light prevents him once again from seeing ahead. To avoid this new obstacle, he turns his eyes to the shadows cast by objects, then to the reflections of objects in water, and only then — eventually — to real objects.

This long narrative brings out just how difficult it is to shed one's usual cognitive skin, as it were, and to embrace a fully epistemic or scientific standpoint. But although the prisoner's journey is long and varied, its different stages are merely repetitions of the same basic experience, and its obstacles mere reprises of the same basic imprisonment. This point is very important, and will be further detailed in due time. As the prisoner moves upwards, his freedom is not conquered gradually or in small doses. Each new imprisonment is a *complete* imprisonment and each new release is likewise a *complete* release — or so it seems at first sight. Accordingly, the prisoner does not regard the progression as a progression, but as a succession of unique and global transformations, whose final resolution is systematically postponed. The initial acknowledgement that the world is irreducible to the shadows on the wall leads to the discovery of a new world, hastily recognized as the real world. But just as the shadows turned out to be mere

[52] *HW* III, 18-19; *PS*, 7. Cf. PHILONENKO, op. cit., 331.

[53] I will return to this issue, as well as to the main parallels between Hegel's *Phenomenology* and Plato's allegory, in chapters 4 and 5.

[54] See *Re* 515ff.

shadows, this new world is eventually revealed as a mere cave. What seemed final is once again provisional and the progression's ending is shown to be a new beginning. The same illusion is then transferred to the shadows outside, and again to the reflections on water. Until the very ending, the prisoner is led from one world to the next, subjected to increasingly sophisticated forms of deception.

A very similar mechanism is deployed throughout the *Phenomenology of Spirit*. In Hegel's progression, the abyss separating ignorance from true knowledge is also filled by a long succession of intermediate stages. Starting from consciousness' most immediate cognitive outlook, Hegel draws up a catalogue of the different kinds of imprisonment or deception encountered and endured along the 'pathway to Science'. In each different stage, consciousness is led to confront its own ideas and assumptions and to test their solidity. This exercise leads to the discovery that the knowledge it takes for granted is inherently contradictory, and that the version of reality it relies upon is not real, but illusory. As in Plato's account, this cognitive volte-face entails the collapse of an entire worldview, but also the revelation of a new and more sophisticated one. Released from error and contradiction, consciousness transfers its former attachment to a new world and to a new set of principles. What is more, it endorses a new existential narrative, grounded in new aims, concerns and priorities. But this change is still only provisional: once these new ideas are subjected to a closer examination, they turn out to be mere beliefs, or mere shadows; and once the latter are considered and examined, they are also found lacking. This procedure is repeated over and over again, at increasingly higher cognitive stages, until no assumption or illusion is left to uncover.

The progression's final stage is no longer a mere version of reality, but 'the actual cognition of what truly is'[55]. The Platonic and Hegelian projects are both aimed at this final release, but their strategies are very different. Plato's approach is diverse and episodic. It combines a series of programmatic indications about the kind of reform required for the attainment of truth and a series of practical examples of how they should be applied. Overall, however, the dialogues amount to the pieces of an unsolved puzzle, which is up to the readers to interpret and complete. Hegel's approach, on the other

[55] *HW* III, 68; PS, 46

hand, is global and systematic. His work is not merely an introduction to philosophy or dialectics, but a complete and detailed account of how to go about transforming one's usual standpoint into an educated or scientific standpoint. While Plato points the way to freedom, the *Phenomenology* is itself a pathway of liberation, leading all the way from the immediacy of 'sense certainty' to the dialectical freedom of 'Spirit' and to the full deliverance of 'Absolute Knowledge'.[56]

The entire *Phenomenology* is built on the conviction that the realm of illusion inhabited by natural consciousness is not endless or chaotic, but liable to be translated into a coherent logical progression. In other words, irrationality itself, when closely inspected, is bound by a rational rule, which holds the key to its overcoming. Hegel's phenomenological catalogue is not a random catalogue, but a *systematic* catalogue, or a systematic chain of illusions, with each link necessarily bound to the next. Furthermore, since this chain can only be brought to an ending when every single illusion has been exposed and refuted, the *Phenomenology*'s catalogue is also a *global* catalogue, which includes every form of deception liable to be endorsed throughout human life. Hegel undertakes nothing less than the actual inventory of all possible worldviews available to human consciousness, as well as all of their combinations. His aim is to secure the attainment of truth by exhausting every form of untruth, and reaching thereby a truly presuppositionless standpoint. His progression is therefore an encyclopaedic progression, extending to every domain of existence: from logics to language, from science to religion, from ethics to aesthetics, from politics to world history.

[56] As noted earlier, Hegel's engagement with freedom is not confined to the *Phenomenology of Spirit*. It is also a central concern in his later systematic works. In light of this, the issue naturally arises as to whether the freedom attained at the end of the phenomenological progression is already a complete or absolute freedom — in which case Hegel's system is already directed at a free or educated consciousness — or whether it is still a new and more sophisticated mode of imprisonment — to be overcome at the end of the *Encyclopaedia*'s progression. This issue leads to the wider controversy regarding the exact role or function assigned by Hegel to the *Phenomenology*, which cannot be dealt with here. My present purpose is to contrast Plato's, Hegel's and Nietzsche's methodological approaches to the relationship between freedom and education. In Hegel's case, this theme is best illustrated, in my view, in the *Phenomenology*, which amounts after all to the 'detailed history of the education of consciousness itself to the standpoint of Science.' (*HW* III, 73; *PS*, 50)

III. *Nietzsche's metacognitive project*

The third part of the text will focus on Nietzsche's understanding of the relationship between freedom and knowledge. Just like Plato and Hegel, Nietzsche advocates a profound reform of human cognition. In line with the imagery employed by his predecessors, he compares human beings to 'involuntary slaves'[57], or to 'fettered spirits', 'chained … to [their] pillar and corner'[58]. Descriptions of this kind abound in his writings, as do prophecies of a 'great liberation', or appeals to the 'dangerous curiosity for an undiscovered world', ruled by 'the master's privilege of the free spirit'[59]. But although this kind of language is indeed evocative of Plato and Hegel, Nietzsche opposes the Platonic and Hegelian solutions to the problem of freedom and calls for an altogether different kind of release.

The key to understanding this difference lies once more in Plato's writings, and specifically in Plato's insistence on the fact that men and women are never indifferent to the truth-value of their mental representations. In the course of our lives, we usually assume our most basic beliefs to be true and tend to resist any suggestion to the contrary. Our ideas about ourselves and the surrounding world are not conjectural, but *assertive*, and indeed *dogmatic*. We take things to be actually as we believe them to be and not otherwise, and it is based on this certainty that we interpret what goes on around us, make plans and favour specific actions or attitudes. Moreover, even when we do think of something in conditional or conjectural terms, we think of it nonetheless as being in a certain way rather than another. Our understanding of it presupposes a specific version of events, even if only momentarily maintained.

In short, our usual representation of reality entails the simultaneous representation of its own truthfulness — which requires, in turn, the simultaneous representation of the *notion of truth*. In the introduction to the *Phenomenology of Spirit*, Hegel makes this point in the following terms: 'Consciousness simultaneously *distinguishes* itself from something

[57] *KSA* I, 74
[58] *KSA* II, 16; *HAH*, 6
[59] *KSA* II ff., 16; *HAH*, 7f.

and at the same time *relates* itself to it'[60]. More exactly, it distinguishes a *being-for-another*, *i.e.* its knowledge of reality, from a *being-in-itself*, *i.e.* reality as such. Hence, 'whatever is related to knowledge or knowing is also distinguished from it, and posited as existing outside of this relationship; this *being-in-itself* is called *truth*.'[61]

The notion of truth is the silent backdrop against which all human thoughts are measured. And this is the reason why, according to Plato, mankind as a whole, 'regarding both education and the lack of it', is trapped inside a cave. Indeed, it could just as well be the case that the truthfulness of our representation of reality were a matter of indifference to us, and that our lives were spent in a sort of logical relativism, which would render every representation as true as the next one, or all of them true in their own way. In this case, freedom and knowledge would no longer depend on each other and it would make no sense to speak of a cave, or of shackles, or of prisoners in the way Socrates does. But Plato's allegory is designed to show that this is never the case: human beings are imprisoned because their representation of reality falls short of reality itself. There is an original and potentially insurmountable opposition between the cave's darkness and the sun-lit world supposed to exist outside.

This fundamental tension must not be reduced, as it so often is, to a simple antithesis, in which truth as such is placed on one side, as an independent or preformed ideal, and ignorance is placed on the other. Plato's point is more subtle and much more alarming: he does not argue simply that truth exists in its own right, in some celestial kingdom of Ideas, and human beings are left outside of it; instead, he strives to show that all human beings, insofar as they think anything whatsoever, are bound to posit the notion of truth ahead of themselves, as it were, and to conceive their usual understanding of reality as either *truthful* or *falling outside of truth*. In this wider sense, truth is just as much a predicate of human thoughts and judgements as it is their underlying condition.

Now the main bone of contention between Nietzsche and the Platonic tradition is precisely the notion of truth, and particularly the referential status it usually enjoys. For Plato, Hegel and a vast lineage of thinkers, philosophy's starting point is the disagreement between human consciousness' usual account of reality and a true account

[60] *HW* III, 76; *PS*, 52
[61] *HW* III, 76; *PS*, 52f.

of reality. Our lack of freedom, our lack of power, our inability to attain real happiness and self-fulfilment are all secondary symptoms of a more basic cognitive condition, or a more basic cognitive discrepancy. Human life, down to its most essential aspects, is dependent on a global expectation of truthfulness, but the latter is continually thwarted by reason's inherent contradictions. Although we tend to believe otherwise, the knowledge we rely upon is usually just a pretence of knowledge, or the appearance of knowledge, and this error is liable to compromise the success of all of our decisions and practical achievements.

The specific approaches put forth in Plato's and Hegel's writings are grounded in this general diagnosis. Their differences concern the different methods employed to reverse this situation and the actual possibility of such a reversal. With Nietzsche, on the contrary, this common view no longer applies. To be sure, he also believes humankind to be imprisoned or enslaved, and in need of liberation; moreover, he also takes this lack of freedom to be primarily caused by consciousness' incapacity to meet a predetermined standard of truth. But whereas Plato and Hegel take truth as such for granted, and focus on how to attain it, Nietzsche takes a step back and questions the very need to think and choose in terms of truth and falsehood. While Plato described human life as being 'aimed at truth', but passing 'beside the mark', Nietzsche challenges the very existence of such a mark, as well the obligation to aim at it in the first place. In his eyes, the concern with truth and truthfulness is itself one of the dogmatic presuppositions consciousness hangs on to in everyday life. A serious philosophical critique, he argues, must extend to the very criterion on which every critique is usually grounded, namely the notion of truth, and unmask its uncritical or dogmatic nature.

For Nietzsche, life is not primarily about truth, but about power, vitality and volition. Men and women are indeed concerned with satisfaction and self-fulfilment, but these imperatives are not intrinsically cognitive or rational. Indeed, the contrary is true: reason itself is a symptom of an inner volitional core, or an inner volitional drive, whose immediacy predates the need to distinguish between true and false ideas, accurate and inaccurate perceptions, rational and irrational judgements, etc. Human will, according to Nietzsche, is originally free and unhindered, open to an endless repertoire of possibilities and modes of expression. In everyday life, however, it is reduced to a dogmatic concern with truth — a 'will to truth' (*Wille zur Wahrheit*, or *Trieb zur Wahrheit*) that outweighs

all other forms of valuation. This artificial exclusivity represents, for Nietzsche, a form of imprisonment. By sacrificing the richness of its resources to a unique cognitive standard, the will is limited to a diminished, anaemic, or slavish version of itself. Power and strength are tamed by reason and life itself is debased.

Nietzsche's exploration of this theme is not uniform. His critique of truth and rationality is grounded in diverse and sometimes conflicting arguments, borrowed from the domains of epistemology, biology, history, politics and aesthetics. Overall, however, his views on the relationship between truth and freedom may be associated with two main lines of enquiry. On the one hand, he seems to regard the 'will to truth' as a natural cognitive tendency, inseparable from the natural use of reason. If all human concepts are indeed supported by the notion of truth, human cognition, in its usual form, is globally imprisoned. In light of this acknowledgement, Nietzsche insists on the wideness of the gap separating natural consciousness from a truly free consciousness and highlights the depth of the transformation required to bridge this gap. The deliverance he advocates entails a full ontological revolution: the crossing of an abyss from which 'one returns *newborn*, having shed one's skin'[62], and the emergence of a wholly new kind of consciousness, with new ideas, concerns and priorities.

On other occasions, Nietzsche puts the emphasis on the cultural and historical nature of this transformation. He suggests that the 'will to truth' is not an essential feature of human consciousness, but the result of a long cultural development, leading to the universal dissemination of an artificial cognitive attitude — namely, the 'philosophical', or 'metaphysical', or 'theological' tendency to reduce reality to the dialectical opposition between truth and untruth, right and wrong, good and evil. Revisiting the history of human reason, Nietzsche contrasts the primitive wisdom of the ancient world (σοφία) with the scientific attitude of modern culture (ἐπιστήμη).[63] Primitive men and women, he argues, favoured life over cognition; they were free, like children or artists, to create their own values. Modern men and women, on the contrary, take truth and reason too seriously; forgetting themselves and their creative powers, they have come to place truth, in a more or less conscious manner, at the centre of their lives.

[62] *KSA* III, 351; *GS*, 7
[63] See notably *KSA* VII, 448.

The first great step was taken by Socrates, the promoter of 'a new, *unheard-of* esteem for knowledge and insight.'[64] According to Nietzsche, he is the founder of a new cultural type — the 'theoretical man', to whom freedom and happiness are essentially a matter of knowledge. Unlike other men before him, Socrates aims to explain away life's contradictions and to render the world a thoroughly rational place. To that end, he equates the True, the Good and the Beautiful, and turns truth simultaneously into a scientific, moral and aesthetic imperative. This attitude paved the way for a new and even greater form of imprisonment, brought about by the Jewish and Christian cultures. Judaism and Christianity inherited Socrates' and Plato's fixation with the divine value of truth and developed thereby a slavish mistrust of strength and power, as well as a morbid contempt for the concrete achievements and pleasures of earthly life. This 'denial of life', coupled with the epistemological and political misconceptions of modern and post-modern European philosophy, led to the nihilistic 'disease' which 'has spread unevenly across Europe' — a disease that 'appears greatest and most varied where the culture has been at home for the longest period of time.'[65]

Nietzsche's point is not only that these transformations have shaped the world's cultural and philosophical history, but also that they have penetrated consciousness' inner life and conditioned the way most men and women view themselves, each other and the surrounding world. Strongly opposed to this legacy, Nietzsche advocates a major cognitive change, or a 'transvaluation' of modern values, whose effects are personified by a series of mythological anti-heroes: Dionysos, the Greek god of music and drunkenness; Zarathustra, the rebellious hermit; and the Antichrist. All of these figures embody, in different ways, the recuperation of a primitive or pre-modern ideal, averse to the laws of cognition and rationality. But since this primordial worldview has grown unintelligible and out of reach, these figures also personify an entirely new standpoint, no longer trapped within the dialectical opposition between truth and untruth, and no longer prey to Christian-Socratic morals.

Although Nietzsche's project raises a number of difficult questions, namely with

[64] *KSA* I, 89; *BT*, 65. Emphasis added. The original adjective is *unerhört*, which also conveys the stronger meaning of a *shocking* or *outrageous* esteem for knowledge and understanding.

[65] *KSA* V, 139; *BGE*, 101

regard to its scope, meaning and viability, its contrast with the two previous projects is now more apparent. Following Plato and Hegel, Nietzsche regards the problem of freedom as 'the problem of science itself.'[66] Nonetheless, his definition of freedom is incompatible with the scientific ideal espoused by Plato and Hegel. For Nietzsche, what keeps humanity in chains is not the lack of a scientific mode of cognition, but the yearning towards it. Freedom is not achieved by overcoming ignorance, but by overcoming the very need to do so.

1.4. *Historical affinities*

Finally, a serious study of the relationship between the different philosophical projects under analysis requires a closer consideration of their historical significance. More precisely, if one is to grasp the specificity of Hegel's and Nietzsche's reinterpretation of Plato's original idea, one must consider their own understanding of the history of philosophy as regards human freedom and the lack of it.

Although Hegel and Nietzsche disagree on their assessment of the philosophical merits of Plato's scientific project, both acknowledge its lasting historical influence. In addition, both emphasize the transitional nature of their own historical age and the importance of the role played by their own philosophical standpoint in the global narrative initiated by Plato. Indeed, both Hegel and Nietzsche regard their solution to the 'problem of science' as the necessary culmination of a long historical progression, whose different stages are regarded as either positive or negative achievements, depending on whether they lead human consciousness towards, or away from, the kind of deliverance they prescribe.

Firstly, let us return briefly to Hegel. In his account, freedom is dependent on science and the latter is the result of a global cognitive progression, leading from an immediate to an educated outlook on reality. This progression is not just a logical or epistemological progression: it is meant to capture the extraordinary complexity of

[66] *KSA* I, 13; *BT*, 4

human life, in its multiple angles and dimensions. What is at stake in each stage of development is a global worldview, with logical and epistemological, but also psychological, moral and political implications. Moreover, these different scenarios are not purely fictional or imaginary. They also reflect, albeit in an indirect or idealized way, consciousness' concrete historical evolution. The whole *Phenomenology* is built on the idea, central to Hegel's entire philosophy, that there is a necessary link between reason's immanent or dialectical development and its actual historical development — in other words, that the history of human thought, despite its share of contingency and randomness, amounts to a global movement away from ignorance and towards more demanding or sophisticated forms of rationality. Accordingly, in Hegel's eyes, modern reason is the necessary, albeit indirect outcome of Greek science; and Plato's scientific project is but the starting point of a long and complex philosophical saga, whose culminating chapter — prior to Hegel's own contribution — corresponds to German idealism, and particularly to the philosophies of Kant, Fichte and Schelling.[67]

For Hegel, Kant's critical project amounts in many ways to a systematic reform of Plato's critique of human cognition. In his own words, Kant 'turns back to the standpoint of Socrates', but his philosophical quest is 'placed on a higher level.'[68] Nevertheless, although his account of human cognition is somehow closer to the epistemic standpoint craved by Plato and his successors, his conclusions are far from reassuring. Instead of solving the 'problem of science', Kant's criticism seems to aggravate it by revealing just how problematic it really is. His insistence on the fundamental discrepancy between human cognition and the actual things it is meant to cognize — that is, between the appearance of reality and reality itself[69] — is not liberating, in the Platonic or Hegelian sense of the word. Rather, it offers a new and even more damning image of human beings' natural servitude.

The issues of whether and how Kant's cognitive dualism can be overcome are a constant theme in the writings leading up to the *Phenomenology of Spirit*, and constitute

[67] This historical-philosophical saga is expounded throughout the *Lectures on the History of Philosophy*, where Hegel presents his own philosophical standpoint as the last link in the 'sacred chain' leading from Greek philosophy to German idealism. See *HW* XX, 456.

[68] *HW* XX, 329f.

[69] See, for ex., *KrV* A114, B311f.

the immediate philosophical ground whereupon Hegel's phenomenological project is erected. Whereas Kant does highlight the reciprocal implication of reason and reality, or subjectivity and objectivity, his philosophy remains trapped within this fundamental antithesis: 'it makes the identity of the opposites into the absolute terminus of philosophy, the pure boundary which is nothing but the negation of philosophy.'[70] For Hegel, however, 'the sole Idea that has reality and true objectivity for philosophy is the absolute suspendedness of the antithesis. This absolute identity is not a universal subjective postulate never to be realized. It is the only authentic reality. Nor is the cognition of it a faith, that is, something beyond all knowledge.'[71] Hence, if philosophy is to become more than a mere critique of human cognition, it must be able to unite both sides of Kant's equation, and provide human consciousness with a standpoint that is both subjectively valid and objectively true.

Fichte's and Schelling's efforts to bring about such a unification amount, in Hegel's eyes, to new and more sophisticated stages in the path towards science. Yet both of their solutions replicate Kant's dualism by replacing his 'universal subjective postulate' with their own 'universal subjective postulates'. In other words, by professing the unity of thought and reality, and attempting thereby to seal the breach between self-consciousness and objective consciousness, Fichte's *Ich* and Schelling's *intelektuelle Anschauung* are ultimately just as undialectical as Kant's critical approach. Whereas the former postulates the formal antithesis between the self and its object, the latter postulates the formal unity of these two instances. In so doing, Hegel argues, both solutions adopt as their starting point that which can only be known as a *result*, and maintain the presuppositive element that a true philosophical education, if it is to be actually liberating, must be able to overcome.[72]

As Hegel writes in the preface to the *Phenomenology*, truth is not 'an *original* or *immediate* unity as such', but a 'self-*moving* sameness' (*die sich bewegende*

[70] *HW* II, 302; *FK*, 67

[71] *HW* II, 302; *FK*, 68

[72] Hegel's engagement with Kant and post-Kantianism is much more complex and nuanced. My sole purpose, for now, is to highlight the structural affinities between Hegel's and Nietzsche's accounts of the historical evolution of Plato's scientific project — both of which assign a central role to Kant's critical philosophy.

Sichselbstgleichheit), or a 'reflection within itself' (*die Reflexion in sich selbst*).[73] Therefore, if philosophy is to provide human consciousness with a truly scientific standpoint, it must cease to envisage the opposition between truth and appearance as either inclusive or exclusive. It must move beyond the immediate contrast between self-sameness and difference and acknowledge contradiction itself as the driving force behind reason and reality. In this new conceptual framework, subjectivity and objectivity no longer stand for self-identical or self-excluding instances. They are defined, instead, in truly dynamic terms, as the simultaneous movement whereby each gives way to the other and becomes thereby what it is.

This dynamic or dialectical mode of identity — 'the identity of identity and non-identity'[74] — is the founding principle of Hegel's logic and the cornerstone of his entire philosophical system. Accordingly, the revolution it sets into motion represents the culmination of a long conceptual journey and the beginning of a brand new philosophical era. Whereas Kant and post-Kantianism are still burdened with the critical exposition of human beings' natural imprisonment, the onset of Hegelian dialectics signals the true dawn of human freedom, dramatically announced in the *Phenomenology*'s preface:

> Ours is a birth-time and a period of transition to a new era. Spirit has broken with the world it has hitherto inhabited and imagined, and is of a mind to submerge it in the past, and in the labour of its own transformation.[75]

Interestingly, Nietzsche also associates the idea of liberation with a major historical change. Half a century later, he also exalts his own age as a time of transformation. And like Hegel, he also points out the conceptual continuity between Plato, Kant and post-Kantianism. Yet unlike Hegel, Nietzsche does not regard his own standpoint as the final stage in the historical and philosophical ascent towards science. On the contrary, insofar as the Western philosophical canon amounts to the long chronicle of Platonism's historical evolution, Nietzsche's intended liberation is also a break with history itself. Its

[73] *HW* III, 23; *PS*, 10. Translation altered.
[74] *HW* II, 96
[75] *HW* III, 18; *PS*, 6

aim is to revive a pre-historical or pre-scientific mode of cognition, in which the Socratic dream of absolute knowledge plays as little part as possible.

In order to situate this transformation, it is useful to compare the historical roles assigned by Hegel and Nietzsche to Kant's *Critique of Pure Reason*. In the first case, though Hegel acknowledges the specificity of the impasse revealed by Kant's critical diagnosis, he regards it as the logical outcome of the 'Socratic standpoint', and hence as a necessary stage in the global progression towards science and true knowledge. Kant's critical dilemma represents the last and most decisive philosophical antinomy, to be overcome by a truly dialectical reform of human reason. For Nietzsche, on the other hand, Kant's critique is not a prelude or a preparation to science, but the first major evidence of the scientific project's global failure. It is a premonition of the approaching break with the 'old order of things' and the restoration of an altogether different historical-philosophical paradigm, announced in a no less dramatic manner:

> The catastrophe slumbering in the womb of theoretical culture is gradually beginning to frighten modern man; in other words, he is beginning to suspect the consequences of his own existence; he therefore dips into his store of experiences for some means of warding off the danger, although he does not really believe in them. Meanwhile great natures with a bent for general problems have applied the tools of science itself, with incredible deliberation, to prove that all understanding, by its very nature, is limited and conditional, thereby rejecting decisively the claim of science to universal validity and universal goals.[76]

The 'great natures' Nietzsche refers to here are Kant and Schopenhauer, whose scientific criticism of science has led to the crisis of modern reason and to the need for a radical change of perspective. Kant's critical philosophy, in particular, has dealt a serious blow to the belief in the possibility of an actual communion between the realms of subjectivity and objectivity — a belief underlying both the *natural* conviction that one's thoughts provide one with an adequate perception of reality and the *philosophical* conviction that one's usual standpoint, albeit lacking such perception, may come to acquire it by means of education. It must be noted, however, that Nietzsche's enthusiasm

[76] *KSA* I, 117f.; *BT*, 87

for this kind of criticism was short-lived. As with many other of his philosophical infatuations, he soon came to regard Kant's and Schopenhauer's sceptical approaches — as indeed the whole tradition of modern German philosophy — as a continuation, rather than actual break with the scientific tradition. By exposing the intrinsic limitations of human knowledge, Kant and his successors did not abandon, after all, the Platonic canon. They merely uncovered, as Nietzsche puts it, 'a hidden path to the old ideal'.[77] Once 'reality was made into mere appearance'[78], a 'real world' was placed beyond it and accorded new and different names: a 'Thing in itself', the 'Absolute', 'Spirit', or 'Will'. But the operating principle in all of these concepts is once again the old standard of truth. In Nietzsche's later writings, these different perspectives are increasingly treated as symptoms of the same fundamental problem, which he refers to, almost interchangeably, as the 'idealist', 'metaphysical', 'moral', 'religious' or 'theological' fallacy of modern reason. The idea of an immanent (and imminent) collapse of scientific reason, vaticinated in the *Birth of Tragedy*, is replaced by an increasingly severe condemnation of philosophy in practically all of its forms. Nietzsche's efforts are devoted to anticipating and bringing about a new, post-scientific and post-moral world, which remains for the most part unexplored. And although he does recognize, in past historical ages, a few promising echoes of his own standpoint, he regards the latter as the first real stage in the path towards freedom.

The following dissertation, divided into three main parts, will document and compare Plato's, Hegel's and Nietzsche's views on the philosophical liberation of natural consciousness. As we have seen in the previous sections, this issue is not confined to the work of these three authors. A detailed genealogy of the relationship between power, freedom and knowledge would have to include, among others, the Stoic, the Christian and the Kantian traditions, as well as all of their historical and cultural offspring. In what follows, however, my aim is not to offer an exhaustive account of this issue, nor to chart its historical or chronological evolution, but to determine its full conceptual scope. To do so, I will focus on three philosophical projects which represent, in my view, the most

[77] *KSA* VI, 176; *A*, 9
[78] *KSA* VI, 176f.; *A*, 9

extreme approaches to the problem of freedom, and constitute the conceptual boundaries within which all other theories and doctrines operate.

Plato raised the issue. His dialogues highlight human beings' native imprisonment and the need for a full cognitive reform, or a full cognitive release, extending to every aspect of human life. This radical diagnosis gave rise to a long critical lineage, whose different members reiterated and amplified reason's intrinsic limitations. But whereas Plato was the starting point, Hegel and Nietzsche are, in a way, the ending — or rather, two alternative endings, pointing to two very different directions. Hegel's project builds on Plato's standpoint and sets out to fulfil its main challenge. His aim is to overcome the 'problem of science' by raising consciousness to the element of science, and inaugurating thereby a new, post-critical brand of philosophy. Nietzsche's project rejects Plato's standpoint, and sets out to reverse its philosophical legacy. His aim is to overcome the 'problem of science' by challenging the very idea of science, and inaugurating thereby a new, post-rational or post-cognitive brand of philosophy.

This illustrious triangle encloses an entire range of possibilities, concerning one of philosophy's most important debates. Indeed, these three projects may be regarded as conceptual limits not only for the classical, medieval and modern understanding of freedom and science, but also for post-modern and contemporary approaches to this relationship. Insofar as the works of Husserl, Heidegger or Wittgenstein, for example, highlight the cognitive dimension of freedom and power, they may also be located within this set of coordinates and interpreted in light of its different alternatives.

Part I

Plato's Scientific Project

2.

PLATO'S IDEA OF TRUTH

Die Wahrheit! Schwärmerischer Wahn eines Gottes!
Was geht die Menschen die Wahrheit an![79]

In Plato's dialogues, the subject of freedom is often presented in the form of a paradox. Although human life is usually assumed to be free, and although human beings usually regard themselves as independent and powerful beings, these beliefs are uncertain and open to refutation. According to Socrates, things might be the other way around: human life may in fact be an imprisoned life and human beings may in fact be powerless or enslaved beings, unaware of their true condition.

This idea is famously illustrated in the *Republic*'s seventh book, in the so-called allegory of the cave. But this well-known image is not an isolated case. In other texts, Plato argues, for example, that human beings are usually compelled to regard reality 'as through prison bars'[80], or that men and women are usually 'in the condition of common slaves.' [81] Imprisonment, servitude and impotence play a central role in Plato's characterization of human life and their scope is much wider than the one we tend to acknowledge. The lack of freedom highlighted by Socrates does not refer simply to the incapacity of achieving certain goals or overcoming certain obstacles or difficulties. It

[79] *KSA* I, 759
[80] *Phd* 82e2-4
[81] *Sy* 215e6-7

relates to the way we cognize reality and to the meaning we ascribe to objects and events. More specifically, it stems from our usual lack of a true and intelligible account of reality, on which to ground our judgments, choices and decisions.

In order to determine the exact implications of Plato's argument, let us start by considering his main views on the structure and scope of human cognition, scattered throughout his dialogues.

2.1. *Truth, opinion and doxa*

Plato's dialectical enquiries are usually focused on the lack of coincidence between what most of us think we know about reality and what we actually know about it. Their aim, however, is not simply to point out the finite or limited nature of human cognition. To be sure, all men and women are ignorant about a great number of things, and mistaken or deluded about a great many more. Human knowledge, even in the best cases, is always but a very small island amid a vast and potentially infinite sea of ignorance and uncertainty. Yet the primary focus of Plato's critique is not this generic form of ignorance. In the dialogues, Socrates' main concern is rather to show that human beings are ignorant and deluded about the very things they claim to know best.

Unlike other men and women, Socrates 'recognizes that he is … of no account in respect to wisdom.'[82] While his interlocutors are usually confident about their cognitive powers, or at least sufficiently convinced of their ability to grasp the meaning of what goes on around them, Socrates is not so optimistic. Firmly convinced of his own ignorance, he sets out to put other people's knowledge to the test, and his painstaking enquiries invariably lead to the conclusion that most people do not know what they claim to know — that human wisdom, even as regards the simplest or most familiar notions, 'is of little or no value'[83].

For Socrates, as for Plato, the primary source of human ignorance lies in human

[82] *Ap* 23a6-7
[83] *Ap* 23b2-5

consciousness itself. That is to say, consciousness' usual incomprehension of the most basic aspects of reality is primarily due to a basic incomprehension of itself. Though all men and women assume they know a great deal of things, and though they live their whole lives under this assumption, they are usually ignorant about what knowing actually means. Focused as they are on specific objects and events, projects and purposes, they rarely take the time to look at how all of these things come to be cognized. The world's meaning and intelligibility are taken for granted and reason's efforts are usually saved for more concrete problems. In Plato's dialogues, on the contrary, Socrates goes back to a more elementary form of enquiry. He turns his attention to the backstage of human cognition, so to speak, and looks for the cognitive structures and procedures whereupon it is grounded.

Usually, whenever consciousness comes into contact with a given object or event, it tends to regard its own role as that of a mere spectator. In everyday life, objects and events are usually perceived as mere facts or occurrences, to be witnessed and taken note of. Their identity, however vague or unclear, is thought of as something fixed and fully determined. The circumstance of their being known does not alter in any way what they already are. Against this attitude, Socrates points out that things are never quite this simple. In the course of his enquiries, he strives to show that human cognition does not amount to the direct contact with some primordial or pre-conceptual core of reality, but to the manipulation of a basic supply of concepts, or εἴδη, through which the world is rendered intelligible and meaningful. Moreover, he argues that these εἴδη are never independent or isolated. As we have seen earlier, they are always part of a complex cognitive web, liable to be organized in different ways and to convey a potentially infinite array of meanings, contents and shapes. Just as the identity of each object or event is the result of a complex combination of εἴδη, so too is the latter part of a wider conceptual web, whose reach extends to the entire cognitive realm. Every thought or judgement corresponds to a specific configuration of this global system, where different concepts take up different hierarchical positions and assume different degrees of importance.

As Plato unearths this complex cognitive system, he calls attention to its global arbitrariness. And his critique is directed both at the system's alleged *validity* and at its alleged *intelligibility*. In what regards the first of these aspects, Socrates points out that human reason is entirely dependent on the idea of truth. Indeed, it could be the case that

we were somehow indifferent to the truth-value of our ideas and concepts and went about our business not caring whether things are in fact what we perceive them to be. If that were so, the different configurations assumed by our global web of concepts would be wholly interchangeable. But this is never the case. Our thoughts and judgements are always accompanied by an *expectation* of truthfulness: we expect our usual account of reality to be true, rather than false; and we expect it to reveal not incomplete or distorted images of things, but things *as they really are.*

This does not mean, of course, that we wonder ceaselessly about the truth-value of all our thoughts, nor that we are seized by endless philosophical doubts when performing the most basic everyday tasks. This is not the case because our usual account of reality is not *interrogative*, but *assertive*: given any object or event, our representation of it is never reduced to the representation of a mere object or a mere event; it also includes, simultaneously, the representation of its truth-value. In other words, whenever we perceive an object or an event, we perceive it to be *so* or *so — and not otherwise.* The conceptual web underlying each of our representations is automatically accompanied by a thesis regarding its own validity, without which objects and events would amount to pure possibilities, and not to actual things or states of affairs.

Secondly, according to Plato, the conceptual web underlying our usual account of reality is also accompanied by a thesis regarding its own intelligibility, whose origin is no less arbitrary. Objects and events are usually familiar to us because their meaning rests on familiar concepts and ideas, combined in different ways. Once again, however, just as these concepts are merely asserted to be true, they are also merely assumed to be clear or intelligible. Once their meaning is seriously questioned, what emerges is a peculiar form of indetermination. Since each different concept, or combination of concepts, is defined by its position within a global conceptual hierarchy, where all concepts are likewise defined by other concepts, every meaning lies outside of itself, as it were, and cannot be reached.

In Plato's dialogues, this double critique is usually conducted in one of two different ways. In some cases, Socrates' charges are triggered by the consideration of a specific notion or concept, taken up at random or at the request of his interlocutors. The enquiry moves from a particular case to an increasingly wider conceptual debate, where reason's inner contradictions are gradually unveiled. In other occasions, Socrates' analysis is

directly grounded in a global description of human cognition, where the same contradictions are discussed in a more systematic way. In what follows, for practical reasons, my own exposition of Plato's argument will be focused on the second of these approaches.

With the previous considerations, we have arrived at what is arguably Platonism's central theme: the recurrent claim, phrased and illustrated in multiple ways, that reason is not a mere *taking note* of reality, as though the latter were immediately accessible or self-evident; human beings' usual mode of cognition consists rather in a series of *theses* about reality, automatically assumed to be true. In Plato's vocabulary, these theses, grounded in more or less complex constellations of concepts, are called δόξαι, and their exact definition is crucial to understanding the essence and scope of the Platonic reform of human cognition.

The first step towards grasping what Plato usually means by the terms δόξα and δοξάζειν is to avoid the false friends associated with the translation of these terms as 'opinion' (and 'to have an opinion' about something), 'view' (and 'to hold a view' about something), 'judgement' (and 'to judge' something to be so or so), etc. Through these and similar formulas, we tend to contaminate Plato's notion of δόξα with our own modern notions of opinion or judgement, and risk missing thereby his entire point. This contamination stems above all from three main properties we assume to belong to all opinions or judgments, but which are absent from Plato's specific definition — namely, what might be called a) the *non-binding* nature of opinions; b) the *episodic* nature of opinions; and c) the *intelligible* nature of opinions.

The first of these properties relates to the degree of certainty enjoyed by opinions or judgements. Usually, what defines an opinion is the lack of proof supporting it, and hence its exposure to the possibility of error. When one holds a specific view about something, one is claiming that, in what regards the object or state of affairs in question, *A* or *B* is the case — and not *C*, or *D*, or *E*, etc. An opinion, however uncertain, is by definition the determination of a specific version of things. But precisely because it is an opinion, and nothing more than an opinion, it is unable to exclude the possibility that things might be otherwise. Unlike factual or apodictic truths, opinions convey a provisional kind of truth, whose full confirmation requires further evidence.

Depending on the context, Plato resorts to what might be called a common use of the term δοξάζειν (and all related words or expressions) and a technical use of this term (and all related words or expressions). The first one, frequently found in many ancient authors, can indeed be translated as 'having a determinate opinion' or 'holding a determinate view' about something. In this case, δόξαι amount to beliefs about reality whose degree of certainty or self-assurance can vary considerably. In the second sense, though — the Platonic sense *par excellence* — the term δοξάζειν refers to a much stronger kind of opinion. More precisely, it refers to an opinion that is not seen or lived as an opinion, but as *the actual truth*. In this second case, δόξαι amount to beliefs exposed to the possibility of error, but which have nonetheless been raised to the status of factual or apodictic truths.[84]

The second way in which Plato's δόξαι differ from simple opinions concerns their overall scope. For most of us, opinions are frequent, but nonetheless isolated events. They arise in connection with a specific object or state of affairs, amid a wider universe of objects which elicit no opinion on our part — either because we already know what they are, and do not need to hazard an opinion about them, or because we are indifferent to them. In either case, opinions are usually construed as explicit cognitive events, aimed at those segments of reality that interest or puzzle us, of which we have no definite knowledge. Once again, however, what is highlighted in the Platonic notions of δόξα and δοξάζειν is precisely the opposite, namely, the fact that human opinions are not episodic or isolated, but *pervasive*. By insisting on these notions, Plato is not alluding to a regional phenomenon, or to a restricted aspect of human cognition. He is pointing out that human knowledge *is from top to bottom a matter of δοξάζειν*. Whenever human beings think they know something, they really only have an opinion about it — and one whose appearance of truth belies a disconcerting lack of evidence.

Thirdly, Plato's δόξαι also differ from simple opinions in that they are ultimately

[84] This second sense, however, can also be subdivided into two different variants: in Plato's dialogues, the words δόξα and δοξάζειν may be used either in a neutral or in a pejorative way. In the first case, they are meant to emphasize a given judgement's categorical nature (as opposed to διστάζειν, which conveys the idea of doubt or indecision), regardless of its actual truth-value. In the second case, they apply to malformed or unreliable judgements, wrongly perceived as sound or solid. Nevertheless, as we will see below, this semantic distinction is ultimately irrelevant, since every δόξα, according to Plato, is by definition a malformed judgement.

unintelligible. When most people state an opinion or make a judgement, they are usually convinced they know what they are talking about. The content of opinions or judgements may be uncertain or controversial, but the terms in which they are expressed are usually not. The opinions themselves, regardless of their truth-value, take the form of intelligible words or sentences, whose meaning is generally familiar to those who state them and to those who listen. In Plato's definition of δόξα, on the other hand, things are not that simple. By calling attention to the doxastic nature of human knowledge, Plato is also highlighting the potential unintelligibility of the cognitive alphabet employed by human consciousness.

To appreciate how these three differences come into play in the Platonic conception of δοξάζειν, let us turn to an important passage from the *Philebus*, where Socrates and Protarchus discuss the origin of human opinions.[85] In order to explain how a given δόξα comes into being, Socrates resorts to the image of a man who sees a distant object and tries to make out its identity.

> SOCRATES: Would you say that often when a man sees things at a distance and not very clearly, he wishes to distinguish between the things which he sees?
> PROTARCHUS: Yes, I should say so.
> SOCRATES: Next, then, would he not ask himself —
> PROTARCHUS: What?
> SOCRATES: 'What is that which is visible standing beside the rock under a tree?' Do you not think a man might ask himself such a question if he saw such objects presented to his view?
> PROTARCHUS: To be sure.
> SOCRATES: And after that our gazer might reply to himself correctly 'It is a man'?
> PROTARCHUS: Certainly.
> SOCRATES: Or, again, perhaps he might be misled into the belief that it was a work of some shepherds, and then he would call the thing which he saw an image.
> PROTARCHUS: Yes, indeed.[86]

This whole passage calls attention to a procedure that is usually invisible or automatic. Whereas most of our beliefs about reality take shape instantaneously, with

[85] See *Phi* 38c-39c.
[86] *Phi* 38c5-d11. Cf. *Phi* 42a, *So* 234b, *Th* 191b, or *Pr* 356c.

little or no hesitation, the observer imagined by Socrates is purposefully held in a state of indecision, tempted by a series of compelling alternatives. Unlike what happens in everyday life, what is here highlighted is not the end result of reason's representational process, but the process itself. What is brought to the fore is the very act of δοξάζειν, prior to the closing in on a determinate δοξαζόμενον.

Every human belief is born from a βούλεσθαι κρινεῖν, that is, from the urge to distinguish between different accounts of what something is or means. In the lines leading up to this passage, Socrates argues that this urge is triggered ἐκ μνήμης τε καὶ αἰσθήσεως, by the combined action of sense perception (αἴσθησις) and memory (μνήμη), whereby reality is made available or cognizable in the first place.[87] But this latter claim must be read with care. At first sight, it might be taken to mean that sense perception and memory are the original sources of human cognition, which would contradict my previous insistence on the *pervasive* nature of δοξάζειν. Indeed, if there is something prior to the doxastic plane — a more primitive mode of representation, on which δοξάζειν itself is grounded — it is only natural to assume that human cognition is *not* from top to bottom a matter of δοξάζειν. And, moreover, that it is not *primarily* a matter of δοξάζειν. If human opinions are indeed born from sense perception and stored by memory, the stage of δοξάζειν is but a secondary (or tertiary) stage, added to an already formed nucleus of sensory and mnemonic data. And if that is the case, it is also natural to assume that δοξάζειν, albeit responsible for validating and bringing into relation the correlates of sensation and memory (the various αἰσθήματα and μνῆμαι), possesses but a limited range of variation. For surely sensations and memories, insofar as they are not originally produced by δοξάζειν, convey a kind of truth that is not as open to being questioned as the wavering truth of human opinions. And surely the brightness of a determinate colour, for example, or the pitch of a determinate sound, or the smell of a determinate flower, etc., are not matters of opinion in the same way that the recognition of an object seen from a distance might be. Given the priority of sensation and memory in human consciousness' cognitive hierarchy, there seems to be an *immediate* or *factual* core of information that is independent from δοξάζειν and

[87] *Phi* 38b12

unaffected by it.

Once more, however, this is precisely the kind of misconception Plato aims to expose through his insistence on the notions of δόξα and δοξάζειν. To be sure, human beliefs presuppose the combined action of αἴσθησις and μνήμη. If men and women were not endowed with the ability to see, or to hear, or to touch, nor with the capacity to remember what they saw, heard, or touched, they would be unable to cognize reality. Plato's point, however, is that everything one is able to represent as *being something*, or as *having a determinate meaning* — and *a fortiori* one's representation of sense perception and memory themselves, as well as one's sensations and memories — is represented *from the viewpoint of δοξάζειν*. Although memory and sensation are presumed to underlie every human belief, their role can only be recognized and appreciated through the ubiquitous and all-embracing filter of δοξάζειν. And this is so because human beings' usual standpoint is not a sensuous or mnemonic standpoint, but a *doxastic* standpoint.

In the *Philebus*, Socrates places human beings' usual standpoint within a cognitive progression ranging from a minimum of cognitive insight, corresponding to a complete absence of cognition (ἀναισθησία), to a maximum of cognitive insight, corresponding to an epistemic or scientific outlook on reality (ἐπιστήμη).[88] Its structure can be summed up as follows:

↓
Ἀναισθησία
Αἴσθησις (viz. αἰσθάνεσθαι)
Μνήμη (viz. μνημονεύειν)
Δοξάζειν
᾿Επιστήμη (viz. ἐπίστασθαι)

The first aspect to consider, if one is to get to the core of Plato's conception of δοξάζειν, is the qualitative discontinuity of this progression. Inasmuch as every stage stands for a different cognitive regime, the transition from each one to the next entails a radical change of perspective. If a given consciousness were situated at the stage of μνήμη, for example, it would be incapable of conceiving either the cognitive perspective of a

[88] *Phi* 33c-34c, but also 21b-d. Cf. *Th* 163c-164b.

consciousness situated at the stage of αἴσθησις or the cognitive perspective of a consciousness situated at the stage of δοξάζειν. In the first case, a mnemonic standpoint would be incapable of envisaging a purely sensuous account of reality because memory is something entirely different from the mere succession of αἰσθήματα. In order to remember something, one must be familiar with the idea of duration and capable of distinguishing between past, present and future events. These two features are utterly inaccessible to a purely sensuous standpoint, to which reality is reduced to infinitely brief and unrelated flashes of information.[89] Likewise, a doxastic standpoint is incapable of envisaging what a purely mnemonic standpoint might amount to, because to hold a view about something is an altogether different thing from storing and combining series of sensations. To claim that something is *this* or *that* (is *this way* or *that way*, has *this nature* or *that nature*, etc.), one must be able to represent the very being of things, and thereby distinguish between what *is* and what *is not*, between truth and error, between adequate and inadequate portrayals of things, and so on. All of this is inconceivable to a purely mnemonic standpoint, to which memories are nothing more than combinations of sensuous units, utterly alien to the notions of truth and untruth.[90]

When confronted with Plato's cognitive scale, every reader is led to form a mental image of what each of its stages stands for. Encouraged by the seemingly boundless plasticity of our usual standpoint, we are led to believe we can shed our cognitive skin, as it were, and put on radically different cognitive costumes. But this belief is, of course, illusory. As Plato points out, we are anchored all the while to the stage of δοξάζειν and fated to represent reality according to its parameters. When we consider the inferior stages of ἀναισθησία, αἴσθησις and μνήμη, we are really only considering our own doxastic standpoint *minus* the ability to distinguish between truth and falsehood (in the case of a purely mnemonic standpoint), the ability to retain previously felt sensations (in the case of a purely sensuous standpoint) and the ability to feel *as such* (in the case of a

[89] Indeed, the cognitive perspective of a purely sensuous standpoint cannot be translated into words: the very adjectives 'brief' and 'unrelated' refer to ideas which such a standpoint is utterly unable to grasp (namely duration and multiplicity). To such a standpoint, the world is reduced to something like an 'absolute' or 'blind instant', devoid of duration, difference, existence or appearance of any kind.

[90] Again, the cognitive perspective of a purely mnemonic standpoint can only be described in purely negative terms as a 'pure collection' of sensations, devoid of identity, truth or normativity of any kind.

purely senseless or 'anaesthetic' standpoint). In like manner, when we consider the superior stage of ἐπιστήμη, we are really only considering our own doxastic standpoint *plus* the ability to vouch for the truthfulness of our usual account of reality. In all of these situations, our standpoint remains as doxastic as ever, albeit *disguised* as a senseless, sensuous, mnemonic or epistemic outlook on reality.

It might be stated, in Kantian terms, that sensations and memories are the *matter* of human thoughts, whereas beliefs, or δόξαι, are its *form*. Let us focus, firstly, on the first half of this statement. Sensations and memories are the matter of one's thoughts because without them cognition as such would not be possible. Without αἰσθήματα and μνῆμαι there would be no fuel for δοξάζειν, and hence nothing whatsoever to assert or deny. In the absence of sensuous and mnemonic information, the mind would have nothing to work with, no flesh to put on its opinions and beliefs. On the other hand, αἰσθήματα and μνῆμαι are by definition incapable of producing a determinate thought-content, whatever it may be. In the *Theaetetus*, for instance, when discussing the qualitative difference between sensation and knowledge, Socrates points out that the sight of a determinate colour or the audition of a determinate sound are not, by themselves, knowable thought-contents.[91] In order for a colour or a sound to be thinkable, one must be able to acknowledge that both of them exist, that each is different from the other and the same as itself, that both together are two and each separately is one, etc., etc. Unlike what is sometimes assumed, none of these acknowledgements are included in, or conveyed by, the sheer perception of colours or sounds. They are neither sensory nor mnemonic, and can never be acquired within the stages of αἴσθησις or μνήμη. They belong in a higher cognitive stage, which presupposes the silent action of memory and sensation, but has no direct access to either of them.

Which brings us to the second half of my previous statement, namely that human beliefs, or δόξαι, are the *form* of human cognition. Since sensations and memories are silent as regards the being and truth-value of what they represent, they need to be brought to life or *made to speak* by the animating power of δοξάζειν. And sure enough, returning to the *Philebus*' example, the man who sees an object from a distance and tries

[91] See *Th* 185a-e.

to discern what it is starts by asking questions to himself. Plato's use of direct speech to convey the observer's inner doubts reiterates the idea that the knowledge derived from the senses (or indeed from anywhere else) can only become knowable as a discourse. Only when what appears to the observer (the φαντασθέντα) takes the form a question, and becomes afterwards a declaration or thesis, can it come alive as a determinate thought-content. In the *Theaetetus*, following this same idea, Socrates argues that 'the soul, … when it thinks, is merely conversing with itself, asking itself questions and answering, affirming and denying.' And so 'when it has arrived at a decision, whether slowly or with a sudden bound, and is at last agreed, and is not in doubt, we call that its opinion [δόξαν]; and so I define forming opinion [δοξάζειν] as talking and opinion [δόξαν] as talk which has been held, not with someone else, nor yet aloud, but in silence with oneself.'[92] In the *Sophist*, moreover, thought is also defined as 'a silent inner conversation of the soul with itself.'[93]

The reason why Plato is so attached to this idea is by now easy to make out. At first glance, it seems fairly obvious that in order to perceive something one needs to be conscious that one perceives something, whether or not this acknowledgement takes the form of an 'inner discourse', or an 'inner conversation'. But the point, once again, is that this acknowledgement is not the mere translation or the mere clarification of a previously held knowledge about reality. This acknowledgement is the very source of human cognition: the 'big bang' which sets knowledge into motion, the 'ground zero' on which every impression and idea are made to stand. In the *Philebus*' image, the *formal* priority of beliefs over sensations and memories is made clear by the fact that the object seen from a distance never presents itself as a pure appearance, or as a cluster of αἰσθήματα and μνῆμαι. The unknown shape is perceived from the very beginning as one of different contending δόξαι. Accordingly, the observer is not required to take note of a sensuous impression, but to decide between different theses or versions of reality. His aim is simply

[92] *Th* 190a2-6

[93] *So* 263e-264b. This 'inner conversation' is not usually an *actual* or *explicit* conversation. That is to say, one does not usually stop to consider every option or alternative available before arriving at a given judgement or decision. In everyday life, the vast majority of judgements are *automatic* — or taken 'with a sudden bound'. They are based on 'conversations' that have already taken place, and whose results have been recorded and incorporated into consciousness. This does not mean, however, that one's judgements are any less doxastic. It only speaks to the complex, inexplicit and enduring nature of δοξάζειν.

to gather enough information to be able to settle for a determinate version of things — be it a coloured shape, a man or a statue sculpted by shepherds.

All of these elements lead to the same central conclusion: it is within the realm of δοξάζειν — and nowhere else — that human knowledge is born. And it is due to δοξάζειν — and to nothing else — that there is, indeed, a world to be seen and heard, and objects to be distinguished and known. But if human cognition is indeed entirely made up of theses about reality, all human thoughts and ideas are exposed to the possibility of error. And herein lies the true motivation for Socrates' concerns. If everything is what it is simply because it is asserted to be so, every single truth is liable, at least in principle, to be proven wrong. A cognitive system based on such shaky foundations is far from reliable — for despite all of our expectations, there may be, after all, no actual connection between our representation of reality and what it is meant represent.

In light of this extreme possibility, the metaphorical reach of the *Philebus*' image appears even greater. Socrates is not stating simply that objects seen from a distance are hard to identify, but that every one of our representations, insofar as it is based on the same doxastic procedure, is cursed with the same kind of difficulty. In other words, whenever we consider any object whatsoever, we are bound to see it *at a distance.* Or rather, borrowing Goethe's later formula, 'everything we possess, we see as if from afar.'[94]

But if Plato's image is indeed intended as an allegory of human cognition *in general*, the conclusions it leads to are at best perplexing. If the situation imagined by Socrates is taken to represent a specific kind of cognitive experience — namely, the kind where the information retrieved by the senses is not sufficient or not clear enough to trigger the adoption of a determinate δόξα — the observer's hesitation is quite understandable. This kind of predicament, not infrequent in everyday life, can be easily solved by an increase in the amount of information provided by the senses. In Plato's example, if the distant shape approaches the observer, or vice-versa, its visual appearance will become more detailed and easier to identify. Moreover, it seems absurd to think that someone looking

[94] *Sämtliche Werke* 6.1., ed. cit., 536:

Was ich besitze seh' ich wie im Weiten,
Und was verschwand wird mir zu Wirklichkeiten.

at another person standing only two feet away, for example, would mistake him or her for an image or a statue. In this and similar cases, where there is no actual distance separating the observer from what is observed, the exposure to the possibility of error seems much smaller. And if one does get things wrong, one's mistake cannot be very significant. For if a determinate δόξα turns out to be false, it can only be so regarding a small detail, and not the big picture.

It would seem, then, that there must be at least a hierarchical division between those beliefs that are more exposed to the possibility of error and those that are less. Even if one accepts that *all* of them are ultimately grounded in a mere assumption of truth, the belief in the existence of God, or the foretelling of a future event, for example, are surely more uncertain than one's belief that the sky is blue, or that one has a body.

In the *Phaedrus*, Plato mentions the formal difference between controversial or 'abstract' concepts and more straightforward or 'concrete' ones. In fact, he himself draws a distinction between the concepts whose meaning we all agree upon, like 'iron' or 'silver', and those whose meaning lead us to 'part company, and disagree with each other and with ourselves' (ἀμφισβητοῦμεν ἀλλήλοις τε καὶ ἡμῖν αὐτοῖς)[95], such as 'justice', 'goodness', etc. Furthermore, it is a well-known fact that many of Plato's dialogues are explicitly centred on the discussion of 'abstract' or openly controversial concepts, such as 'justice', 'beauty', 'pleasure', 'piety', etc. Nevertheless, an attentive reading of any of the dialogues is bound to show that the distinction between self-evident and controversial concepts is a spurious one. In the course of Socrates' enquiries, the explicit ἀμφισβήτησις surrounding the notions of justice, beauty, piety, etc. is invariably revealed as the keyhole from whence to peek at a much wider and much more alarming form of ἀμφισβήτησις.

Usually, Plato resorts to what might be called a zooming device: starting from openly controversial notions, he proceeds to zoom in on their basic components. His aim is to eliminate the controversial nature of these concepts by isolating simpler concepts, devoid of ἀμφισβήτησις, and to use them as the ground on which to 're-construct' or 're-synthesize' the controversial concepts. However, his hopes of finding such a ground are habitually shattered. The concepts of justice, beauty and the like are shown to rest upon

[95] *Phdr* 263a2-b2

simpler concepts that are themselves controversial, and these in turn are supported by yet another set of controversial concepts. In his attempts to zoom in on the most basic components of human cognition, what Socrates finds is not an original core of simple or self-evident concepts, but a seemingly endless succession of controversial concepts, with no firm ground in sight.

By calling attention to the potential disputability of human consciousness' entire cognitive vocabulary, Socrates is not simply suggesting that we are unable to vouch for the truthfulness of our concepts. He is also suggesting that most of us, when employing either 'abstract' concepts like 'justice' or 'concrete' concepts like 'chair', are in fact unable to grasp their actual meaning. When we state, for example, that 'justice is the noblest of virtues'; and when we say, for example, that a 'chair is green', we usually assume that we know what we mean — viz. that although admitting we do not know everything there is to know about justice or chairs, we know enough to speak about them in a clear and fairly unproblematic way. Yet when forced to explain what justice is, or what a chair is, we immediately resort to a chain of interconnected theses whose ultimate meaning is as unclear as that of justice or chair. In the first case, we say, for instance, that justice is to 'accord everyone his share', and that one's share is 'that which best suits his needs'; and that the best is the 'highest degree of good'; and that good is 'the opposite of bad', and so on. In the second case, we say, for instance, that a chair is a 'solid object', and that solidity is the 'degree of resistance to motion', and that motion is the 'change of location over time', and that time is 'a measure of physical change', etc. In both examples, the definitions offered amount to mere δόξαι, which are in turn defined by more δόξαι. Their meaning is caught up in an immanent, self-enclosed and ultimately unintelligible maze of δόξαι, with no clear exit in sight.

In view of this, the amount of beliefs that may lead us to 'part company, and disagree with each other and with ourselves' is liable to be much greater than we tend to acknowledge. Plato's zooming device suggests that we are ultimately unable to reach a consensus not only regarding the meaning of openly controversial concepts, but also regarding the meaning of much more straightforward concepts, such as 'iron', 'silver', 'big', 'small', 'table', 'chair', 'green', 'red', etc. And this inability brings us back to the perplexity caused by the *Philebus*' allegory: whereas we have no difficulty in accepting that the former set of concepts is indeed controversial, and open to interpretation, we are

not usually as open-minded when it comes to the latter. This is so, Plato argues, because our usual standpoint is not used to zooming in on its own concepts. Its habitual perspective is distant and fuzzy, just like that of the *Philebus*' observer. It deals with concepts without actually seeing them or pausing to examine what they really mean.

In short, Plato's image does not refer simply to distant shapes or to distant objects, but also to the very things placed in front of us. To borrow further from Goethe, 'the hardest thing of all is that which seems easiest: to see what lies before our eyes.'[96] The difficulty exposed in Plato's dialogues can be described in similar terms. It does not depend, ultimately, on the kinds of objects we perceive, or on the concrete conditions in which we perceive them. It is not derived from mere ignorance, or from a lack of experience, or from some other contingent factor added to an otherwise sound cognitive standpoint. What Plato keeps highlighting is a structural problem, which stems from the very nature of our usual mode of cognition and affects all of our ideas, judgements and decisions.

2.2. *Self-confidence and self-interest*

In Plato's dialogues, Socrates' usual task is to expose the doxastic foundation of human cognition. Throughout his enquiries, what is usually perceived as natural and self-evident is shown to be neither and the whole world is turned into a problem. For Socrates, things are never what they seem, but only claimed to be such. Their meaning, when subjected to a closer examination, is reduced to a string of contradictions, with no clear solution in sight. Moreover, the dialogues also document the resistance put up by most of us to this kind of criticism. If reality as we know it is grounded in a system of theses, or δόξαι, it is certainly not lived or experienced as such. On the contrary, human life is continually supported by the conviction that reality is indeed what it appears to be. And this conviction is so natural and deep-rooted that any suggestion to the contrary is usually

[96] *Sämtliche Werke* 4.1., ed. cit., 694:

Was ist das schwerste von allem? Was dir das leichteste dünket,
Mit den Augen zu sehn, was vor den Augen dir liegt.

regarded as a theoretical fiction. We may entertain the possibility of a life entirely grounded in theses, and hence entirely open to revision, but we are usually convinced of the contrary. In real life, most δόξαι translate as facts, certainties and apodictic truths.

It might be argued, against this general diagnosis, that human beings are not usually as confident about their cognitive powers. Indeed, as a quick survey of anyone's past experience will show, human thoughts are not immune to uncertainty and error. In everyday life, there seems to be much more room for error and illusion than there is for certainty, and although we are indeed certain of a good number of things, we often admit to being wrong or deluded about many others. Furthermore, we are usually ready to concede that what we *do* know is very little when compared with everything there is *to* know. Our limited grasp of reality, while enabling the performance of most of the practical and intellectual tasks required of us, is usually regarded as a small cognitive island amid a vast sea of ignorance.

At first glance, this description seems accurate. Our usual cognitive perspective, confined to a small region of space, bound to a single moment in time, is indeed very limited. On closer inspection, however, things are not quite that simple. Even though our usual knowledge is perceived as an island, or an enclave, what lies outside of it is never purely unknown. If that were the case, the identity of distant objects and realities would be thoroughly indeterminate, open to an infinite range of possibilities: they might exist or not exist, they might be identical to the objects and realties we are familiar with, they might be radically different from them, etc. These possibilities, however, are never truly undecided. Although we have no experience of what lies outside our immediate cognitive horizon, we have an *idea* of it. And this idea is much less vague than we tend to acknowledge.

In everyday life, what we do not know or have no access to is never regarded as a pure mystery, or a pure riddle. On the contrary, it is always perceived as an extension of what we have already experienced. Our usual standpoint is guided by a general presumption of *homogeneity*, that is, by the general assumption that all empirical contents, whatever they may be, are bound to the same fundamental principles and endowed with the same fundamental properties. When we think about a distant country, for example, or a distant planet, we base our idea of them in the rules and qualities that define our own country and our own planet. In like manner, when we picture the inside

of our body, we do so based on its outer appearance. The basic characteristics of the things we are familiar with are automatically projected into all other regions of reality and transformed into global phenomenal principles. Distant or inaccessible objects, regardless of their actual nature, are never thought of as being colourless, for instance, or shapeless, or weightless, or devoid of texture, etc. Whatever they might be, they are always represented as having a determinate colour, a determinate shape, a determinate weight, or a determinate texture.

Simultaneously, however, our usual standpoint is also guided by a general presumption of *heterogeneity*, that is, by the general assumption that whatever lies outside our immediate cognitive horizon is different from what is contained in it. And, moreover, by the general assumption that the various things located outside our empirical horizon are as different from each other as the things we are familiar with. If we were walking downtown, for example, and would suddenly realize that every street is exactly the same, we would not fail to be amazed. Likewise, if we were to find out that every single house, in the same street, is perfectly identical to the next, we would also be perplexed. Even though we usually assume all things to follow the same fundamental principles and to share the same fundamental qualities, we also expect them to be different and varied.[97]

But these two assumptions are not incompatible. The differences we expect to find outside our empirical horizon are not radical or *substantive* differences, but limited or *adjectival* ones. And this is so because our usual presumption of heterogeneity is subordinated to our usual presumption of homogeneity, and not the other way around. Whereas the latter refers to basic properties or categories (colour, shape, texture, etc.), the former refers to specific declensions of these basic categories (different colours, different

[97] These two principles can be compared to what Leibniz named the principle of uniformity, or 'Harlequin's principle' (*le principe d'Arlequin*) and the principle of variation, or 'Tasso's principle' (*le principe du Tasse*). Cf. *Philosophische Schriften*, ed. cit. II, 75; III, 339, 343ff., 497, 500, 530; V, 64f., 463s.; VI, 545f., 548, 604; VII, 392. While the first of these principles states that everything we know or experience is fundamentally similar to everything that exists (*c'est tout comme ici partout et toujours*), the second one states that everyhing is nonetheless different from everything else (*che per variar natura è bella*). Nature is both uniform and varied because every particular difference amounts to a different expression of the same totality. But whereas for Leibniz these two principles qualify the universe as it really is or exists, they are being used here to characterize human consciousness' usual *assumptions* about reality. In this case, homogeneity and heterogeneity are not metaphysical, but *doxastic* principles, whose legitimacy is merely presupposed and not actually demonstrated.

shapes, different textures, etc.).[98] What we expect to find beyond our empirical horizon are different variations of the same categorial repertoire. And since these variations are inspired, once again, by those we are used to finding within our empirical horizon, the principle of heterogeneity is really only another instantiation of the principle of homogeneity.

This brief description is enough to disprove our initial hypothesis. Instead of an island surrounded by a sea of ignorance, our usual knowledge resembles a wide continent, or a wide cognitive web, extending in every possible direction. What we tend to regard as a purely uncharted territory is in fact already peopled with all sorts of δόξαι, regarding every aspect of reality. Indeed, instead of projecting *some* of the principles and qualities that define our empirical horizon into the distant or inaccessible regions of space, we usually do so with regard to *all of them*. And, moreover, instead of projecting these principles and qualities into *some* regions of space, leaving others unexplored or uncharted, we do so into *all of them*. Every reality we can think of is covered with δόξαι — highly complex bundles of δόξαι, concerning the world's colour and shape, the laws of motion and causality, the unity and hardness of physical objects, the continuity of space, and so on. This does not mean, of course, that we have as clear a picture of what goes on before us as we do of what goes on in distant countries, or of what goes on inside our body. But it does mean that these realities hold a very limited degree of surprise: what we do not know about them is only how the different δόξαι we ascribe to everything else are combined and declined.

However, there is still much more to this preliminary picture. For one thing, these considerations do not apply only to the element of space, but also to the element of time. Just as we tend to project our understanding of nearby objects and realities into other regions of space, we also tend to project our past and present knowledge into the future. When we consider any object, we automatically assume it will go on existing after the present moment. If a book lying in front of us would suddenly vanish into thin air, we would not fail to be amazed. Moreover, when we consider any object, we also assume its

[98] Leibniz draws a similar picture: [Quoique ces deux principes] paraissent se contrarier, [il faut les] concilier en entendant l'un *du fond des choses* [sc. le principle d'Arlequin], l'autre *des manières et des apparences* [sc. le principe du Tasse].' (*Philosophische Schriften*, ed. cit. III, 348. Emphasis added.)

colour will remain the same after the present moment, as will its shape, size and texture. If the same book were to suddenly change colour, or lose its usual rigidity, or shrink to half its size, we would not be any less surprised. Finally, just as these things are not expected to happen to the book in front of us, neither are they expected with regard to any other book, anywhere else in the world. Here, once again, our doxastic web extends in every possible direction. It is globally prescriptive, or globally proleptic, setting the ground for every future insight or experience.

The same holds, in addition, with regard to the representation of our own identity. Throughout our lives, no temporal moment is lived as an isolated moment, or as the final moment there is. Besides assuming that the world will go on existing, in the same moulds and with the same general appearance, we also assume *we* will go on living and experiencing reality as we have always done. What is more, the temporal horizon we place in front of us is not finite, but infinite. Even though we know we will eventually die, this knowledge does not prevent us from assuming, at each given moment, that a new moment will come, and a new one after that. Our thoughts and judgements, albeit conditioned by the idea of death, are always supported by the expectation of an endless temporal horizon, where all of the δόξαι we have previously considered are maintained and multiplied.

A full description of our standpoint's inner structure would require a much longer analysis, but the previous remarks are already indicative of the extraordinary complexity of our usual web of assumptions. Once again, however, all of the above is only a small part of an even wider doxastic picture. Apart from all of our presuppositions regarding the world's spacio-temporal structure, a second and equally important set of presuppositions is also at work within us. Although reality as we know it is usually filled with assumptions about the appearance and duration of objects and events, it might be the case that all of this were merely witnessed by us, in a neutral or disinterested way. If that were so, every object and event would be equally important and equally worthy of our attention. But this, of course, is never the case. We are never mere spectators of reality, but active and self-interested participants, to whom there is always something at stake. Our life has the form of a task, or rather of a highly complex set of tasks, whose fulfilment requires a vast series of judgements and decisions. Accordingly, our usual priority is not the acquisition of an epistemically pure or accurate version of things, but

one that is good enough to guide our actions and to satisfy our practical demands. The main criterion whereby we determine things to be so or so, to be true or false, to have this or that nature, etc., is not objective or epistemic, but practical and self-centred.

In accordance with this new set of judgements, the world is usually divided into things that matter and things that don't, things that are worth knowing and doing and things that are not, things that are beneficial and things that are harmful, necessary and superfluous, pleasant and unpleasant, secure and dangerous, etc. And these distinctions shed a new and important light on the principles of homogeneity and heterogeneity we have previously considered. Even though our doxastic web extends to every corner of the universe, our idea of what goes on in distant planets, for example, is usually not as detailed or as demanding as our idea of what goes on in our street, or in our house, or in our living room. And the reason for this is not only that our senses are too limited to reach such distant realities. This cognitive asymmetry is due, above all, to our lack of interest in them. More specifically, it is due to the assumption that what goes on in other planets has no direct influence on the tasks and projects we usually focus on. Nothing of what exists there is liable to alter our immediate perception of what goes on around us, or to solve the concrete problems we are faced with everyday. Then again, all of this might change if our practical horizon were different: if we were astronomers, for example, and earned a living from studying other planets, their interest for us would increase; likewise, if we were forced to abandon the Earth, due to some natural catastrophe, and were forced to move to another planet, our perception of it would also be radically altered.

Our usual account of reality, albeit universal in scope, is centred on a specific cognitive perimeter. It is here that our life plays out, that our daily tasks are decided and executed and that out vital goals are liable to be reached. Everything falling outside of this perimeter is deemed irrelevant or taken care of.[99] It is part of the background against which our life unfolds, and treated accordingly. However, as suggested in the previous example, this perimeter is not fixed or final. What lies in the background can always be

[99] This perimeter does not correspond necessarily to a single geographical location. It is not merely a spatial perimeter, but a *practical* or *functional* one, reflecting one's vital projects and aspirations. In the case of an astronomer or an astronaut, for example, it can encompass both a suburban house and the surface of the Moon, insofar as these two locations are recognized as part of the stage wherein their life goals are liable to be reached.

brought to the fore, and vice-versa. This cognitive balance is dependent on a highly complex constellation of δόξαι, whose inner composition is permanently updated, in response to new contexts and situations.

In Plato's dialogues, Socrates calls attention to both of the main features I have briefly described. He emphasizes both the doxastic foundation of our spatial and temporal representations and their practical or self-interested nature. But his characterization of human cognition includes a third fundamental element, on which the other two are grounded. And this too contradicts our habitual understanding of our own cognitive perspective. Usually, we tend to conceive the different tasks envisaged and carried out every day as separate events, oriented towards different purposes. But this is never truly the case. All human δόξαι, despite their extraordinary diversity, are ultimately directed towards a single end. Although human life is diverse and incoherent, torn between different and often conflicting aims and projects; and although these aims and projects keep changing with time, with circumstances, sometimes with mere chance, all of them share a commitment with what is perceived, at each given moment, as the best or most advantageous course of action. Everything is chosen in view of what is deemed best — either directly, by aiming straight at what one wants, or indirectly, by aiming at something else, judged the best vehicle to attain what is best. This, according, to Plato, is the real source of human beings' constitutive self-interest.

The doxastic web that supports human life is organized in the form of a global scale, or a global pyramid, whose highest term is 'the good' (τὸ ἀγαθόν) or 'the best', or 'the most excellent' (τὸ μέγιστον ἀγαθόν, τὸ βέλτιστον) of all contents or determinations.[100] This crowning element is that in view of which everything else is thought of, judged, chosen, envisaged, decided, attempted, risked or pursued. The greatest good does not refer only to big decisions or to life-changing events. All daily actions and resolutions, from the most insignificant to the most important, are grounded in a specific version of it. 'It is in pursuit of the good that we walk, when we walk, conceiving it to be better; or on the contrary, stand, when we stand, for the sake of the same thing, the good.'[101] Τὸ μέγιστον ἀγαθόν is the ultimate telos of human existence, the driving force behind all

[100] See above fn. 37.
[101] *Go* 468b

forms of cognition, action and enjoyment. Furthermore, this crowning element is not an isolated conquest or achievement. We are not interested in what is best only for a while, or only in a limited way. We long to enjoy it *for as long as possible*, and *in the highest degree possible*.[102]

But this is still only one half of Plato's argument. Apart from bringing out the absolute priority of τὸ ἀγαθόν over everything else, he also points out that 'the good' is by definition a *formal* notion, liable to be translated into a multitude of different contents. On the one hand, 'nothing is more certain … than that every intelligent being pursues [the good], desires it, wishes to catch and get possession of it, and has no interest in anything in which the good is not included.'[103] On the other hand, since the good can be made to stand for almost anything, its pursuit can also assume very different forms. What drives human life, therefore, are different versions of τὸ ἀγαθόν, whose content depends on each person's circumstances, interests and aspirations. Or rather, what drives human life is a global doxastic web entirely subordinated to what is assumed by each person, at each given moment, to be the best possible life.

By insisting on the formal nature of τὸ ἀγαθόν, Plato is arguing that the crowning element of our doxastic edifice is itself a doxastic judgement, prone to the same contradictions that affect all other δόξαι. At the same time, he is also pointing out that this problem is not usually apparent to us, since our definition of τὸ ἀγαθόν, like that of all other δόξαι, is automatic and taken for granted. In everyday life, instead of wondering ceaselessly about what the greatest good is, or how to obtain it, we simply go on living, focusing on specific projects and tasks. This does not mean, of course, that we do not have doubts about what to do or how to act, nor that we never wonder about the global meaning of our lives. These occasions, however, are not the rule, but the exception. And even when they do represent real moments of indecision, the latter usually affects only a very small part of our cognitive outlook. We may have trouble placing a specific δόξα, or set of δόξαι, in our global cognitive edifice, but all of its other elements are left untouched. All other δόξαι remain committed, in one way or another, to the same version of τὸ ἀγαθόν.

[102] Cf. *Sy* 206a.
[103] *Phi* 20d8-10

Plato's point, however, is not simply that the best can be made to stand for different things. Insisting on the formal nature of τὸ ἀγαθόν, he is also underlining its perpetual or self-renewing nature. Indeed, precisely because the best is always *better* than anything else, no concrete content is good enough to exhaust its meaning. One may dedicate one's entire life to saving up a million dollars, for example, or buying a specific kind of car, or having a specific number of children, but all of these goals, once achieved, are immediately surpassed by the idea of something even better, still left to accomplish. Since life's ultimate goal is formal, rather than concrete, it is doomed to this permanent dissatisfaction. Every particular experience propels one the next, as it were, and every moment is a prelude to something else.

In the *Philebus*' thought experiment, the identification of a distant shape illustrates the complex cognitive structure we have hitherto considered. To begin with, the observer imagined by Socrates is not baffled by the absolute novelty of what he sees. The object whose meaning he tries to make out is surrounded by a multitude of familiar shapes, whose meaning has already been decided. The theses hazarded about the shape's identity are already framed by a myriad of other related theses, concerning not only the observer himself, his surroundings and his past experience, but also the very structure of the objective world. Regardless of what the shape turns out to be, it is already perceived as a solid body, with a specific colour and texture, taking up a specific amount of space, subject to the laws of gravity and motion, etc. The shape's entrance into the observer's empirical horizon is not a pure φαίνεσθαι, but the addition of a new element to a preformed version of things. What the observer sees is automatically judged to be something determinate and intelligible, but not yet recognized — a man, or a statue, or something else.

This general idea is made clearer in the dialogue's immediate continuation, where Socrates compares human cognition to the joint work of a scribe (γραμματεύς) and a painter (ζώγραφος) active within the human soul.[104] After suggesting that the 'matter' of our representations are sensations and memories, and their 'form' beliefs or δόξαι, in the

[104] *Phi* 38e-39e

terms I have previously discussed, he goes on to compare the 'inner discourse' that accompanies the endorsement of different δόξαι to the double task of 'writing words in our souls'[105] and painting therein 'the images of what was established and said' (τὰς τῶν δοξασθέντων καὶ λεχθέντων εἰκόνας).[106] In both of these metaphors, consciousness is likened to a sort of book or document (βιβλίον) in which the various δόξαι are recorded and combined.

At this point, Socrates adds a new and important element to his description. He observes that human cognition is not concerned only with the past and the present, but also with the future — and hence that the words and images written and painted in our souls concern not only things that *are* or *were*, but also things that *will be*.

It is important to bear in mind Plato's exact wording:

> SOCRATES: Do the writings and pictures, then, which we imagined a little while ago to exist within us, relate to the past and present, but not to the future?
> PROTARCHUS: To the future, especially.
> SOCRATES: Do you say 'to the future especially' because they are all hopes [ἐλπίδες] relating to the future and we are always filled with hopes all our lives?
> PROTARCHUS: Precisely.[107]

The vital book Socrates uses as metaphor is not a mere depository of facts, past and present, as they are witnessed and taken note of. The λόγοι and the εἰκόνες recorded by consciousness refer to memories and direct observations, but also to future situations and events. In order to grasp Plato's meaning, it is important to note that the word ἐλπίς stands not only for hopes or prospects — that is, for the desire for things to happen or evolve in a certain way — but also for expectations, grounded in presuppositions — that is, for the assumption that things will remain or become *so* or *so*, and *not otherwise*. The ἐλπίδες written and painted in our cognitive book concern not only how things *should turn out*, but above all how things *are supposed to turn out*, given their past and present condition.

Returning to the *Philebus*' initial example, the vision of a distant shape would not be

[105] *Phi* 39a3. Writing words, sentences or whole discourses — the word is λόγος.
[106] *Phi* 39c6-7
[107] *Phi* 39d7-e7

possible (or, at least, would take a radically different form) without a series of contextual presuppositions relating not only to past and present, but also to future events. As the observer struggles to make out what he sees, he assumes, for example, that he will go on living after the present moment, that the world around him will go on existing, that his immediate surroundings will retain their colour and shape, and so on and so forth. All of these beliefs, along with all the others mentioned above, make up the setting in which the distant shape makes its appearance, or the ground on which its meaning is made to rest. They provide the cognitive text, or the cognitive painting, as it were, in which the new shape is to be included.

By comparing human consciousness to a book, or a document, and by calling attention to the complexity of its content, Socrates is highlighting an aspect of human cognition that is usually misunderstood. In his image, the scribe and the painter are continually at work, recording every new idea, object and event. However, their registry does not consist in a mere collection of ideas, objects and events, taken down in the order in which they are experienced. If this were so, the book of consciousness would amount to a journal of human experience, starting at the beginning of one's conscious life and ending in one's death. Whereas the first chapters would be comparatively empty, in accordance with the empirical poverty of infancy and childhood, the last chapters would be full of words and images, reflecting the empirical wealth of a long and seasoned life.

What Plato seems to be describing, however, is something very different. The word he uses is not βίβλος, i.e. a book or a roll of papyrus, but a diminutive form, βιβλίον, which evokes the image of a single paper, tablet, or canvas. This difference is crucial, as it suggests an entirely different form of registry. The scribe's task is not to write at length, recording one sentence after another, but to *rewrite*, again and again, the same original text. Likewise, the painter's task is not to paint different images, one after the other, but to *repaint*, again and again, the same original one. Unlike the words and the images of a regular book or journal, the λόγοι and the εἰκόνες mentioned by Plato are not definitive. While some of them are maintained or repeated, confirmed by the course of experience, others are erased and corrected, in order to accommodate new versions of the same ideas and insights.

This paper or tablet is best described, therefore, as a cognitive *palimpsest*, whose changing content reflects consciousness' permanent cognitive updates. By highlighting

the idea of a *complete* cognitive text, or a *complete* cognitive painting, continually reiterated and redefined, Plato is emphasizing the *complete* or *global* nature of our usual account of reality. The βιβλίον of human consciousness is already, from the very start, a global map, or a 'vital atlas', which charts the territory in which our life usually unfolds. Instead of a partial or incomplete account of reality, it offers a complete, albeit provisional picture of everything there is, was and ever will be.

But the δόξαι written and painted in our cognitive tablet do not refer simply to objects and events. As we have seen earlier, they also document our usual interests, aims and aspirations, as well as our overall assessment of what is best or most important in life. And this practical dimension can also be illustrated with the aid of the *Philebus*' initial image. In the beginning, the shape seen from a distance is only a blurred outline, open to different interpretations. Once it moves closer, however, its interference in the observer's practical horizon changes, and so does the kind of definition it elicits. A series of aspects that were previously irrelevant, like its actual size, its exact colour, or its ability to move or speak, become increasingly noteworthy, leading to the emergence of new δόξαι and to the recycling of old ones. Conversely, as the shape moves away, or ceases to intrigue the observer, the composition of his 'vital atlas' is once again transformed. As his priorities change, his attention is caught by other, more pressing concerns, and the same cognitive totality is reorganized. New objects and events are brought to the fore, while the old ones are gradually demoted. The shape's definition becomes less detailed and less important, until it finally fades into the background.

2.3. *Body and soul*

Plato's characterization of human cognition is remarkably modern. The idea that one's knowledge is determined by a continual urge for practical orientation, grounded in a global appreciation of what is at stake in one's life, seems to lead as far as the phenomenological and the existential traditions. And even though Plato does not speak of self-interest and practical reason in the explicit way Kant, Husserl or Heidegger would later do, he does seem to address the same general issues by means of a more intuitive language. In many of his dialogues, the distinction between the doxastic nature of human

cognition and the possibility of a true or epistemic account of reality is often framed by another, equally important distinction: the contrast between σῶμα and ψυχή. Once again, however, to understand what Plato means by these words, and how they relate to one another, it is important to dispel a few common misconceptions.

In the dialogues, as in most classical Greek texts, σῶμα and ψυχή are usually translated as 'body' and 'soul'. But although this translation is undoubtedly correct, it does not get us very far — not so much because the words 'body' and 'soul' are in themselves incapable of conveying the true meaning of this conceptual opposition, but because they are tainted by a series of spurious and misleading assumptions. For modern readers, σῶμα usually refers to the human body with its members, organs, and concrete shape, whereas ψυχή is thought to evoke some obscure and ineffable entity lodged somewhere inside it. While the body can be seen and touched, the soul deals in ideas and feelings. But body and soul, however different they may be, are both regarded as *things*, that is, as the basic ingredients of a collection. When thrown together, they somehow blend into one another, giving rise to a mysterious construct called life.

For Plato, as for most Greek authors, σῶμα and ψυχή are not different things, but different ways of looking at reality, or different forms of interpreting human life. To reduce these two terms to the elements of a collection is the surest way to lose sight of their true meaning, and this is so for two main sets of reasons. Firstly, this kind of description relies on the possibility of a neutral or disinterested observer, placed outside the opposition and capable of a clear and impartial understanding of both terms. Plato's point, however, as we will see below in more detail, is precisely that such an observer does not exist. Since all of us are naturally endowed with a soul and a body, we are doomed to see reality from within this opposition. Therefore, our understanding of the soul and the body is itself conditioned by an original mixture of bodily and 'psychic' (ψυχικά) concerns. We do not know what a soul really is, or what a body really is, because we lack the distance required for such an assessment. A pure body and a pure soul can only be known indirectly, as the different limits to which our usual standpoint can tend.

Secondly, by construing σῶμα and ψυχή as the elements of a collection, we fail to realize just how different they really are. There can be no symmetry between σῶμα and ψυχή because their relationship is not one of coordination, but one of subordination — and, moreover, one of *absolute* subordination. To understand this, let us recall Socrates'

recurrent claim that the soul rules the body, and not the other way around.[108] If this is taken to mean that the soul rules the body as a thing rules another thing, we are led back to our initial dualism. But if this claim is read under a different light, and Plato's insistence on the qualitative difference between soul and body taken more seriously, what emerges is a new and more radical idea: instead of an obscure entity lodged inside the body, or merged with it, the soul is the very condition for the body to be thought of as such; nay, it is the source of meaning *in general*, the founding principle of human consciousness.

By insisting on the ruling nature of ψυχή, Socrates is calling attention to the soul's absolute cognitive priority. Or rather, borrowing a term unknown to Plato and coined long after his lifetime, he is calling attention to the soul's *transcendental* nature. In Plato's own language, he is arguing that human cognition is by definition a 'psychic' event, or that reality is revealed, primarily, *through the soul*, or *with the eyes of the soul* (διὰ τῆς ψυχῆς). The soul's limits are not those of the body, or of any other finite thing. They coincide with the limits of reality and extend as far as cognition itself. Accordingly, the original meaning of ψυχή can also be conveyed by modern expressions such as consciousness, cognitive access or cognitive awareness. Ψυχή is not a part of reality, but the window, as it were, from whence to look out on it.

Yet this is not all. By insisting on the ruling nature of ψυχή, Socrates is usually also saying something about the specific quality of the cognitive access it provides. In its pure or most essential form, ψυχή is not simply the ability to cognize reality, but to do so in a thoroughly clear and unhindered way. A purely 'psychic' standpoint amounts to something like a global and absolutely revealing glance on the whole of reality. It represents a superlative form of cognitive access, to which nothing whatsoever remains hidden. Such a standpoint does not need to favour a specific thought-content over another, or to look away from a given region of reality to focus on another. Likewise, the distinctions between truth and falsehood, clarity and obscurity, certainty and uncertainty, etc., are of no use here. Its cognitive perspective reveals reality as it is in itself, with no omissions, distortions or ambiguities.

[108] See, for ex., *Phd* 94b, *Phdr* 246b or *Re* 353d.

All of this applies to ψυχή in its pure or original form. But things are very different in the concrete element of human life. Here, one's cognitive access is never as clear or as revealing. Unlike a purely 'psychic' standpoint, human beings' usual standpoint amounts to a permanent struggle. It is continually forced to decide whether reality is so or so, to distinguish between true and false versions of things, to interpret ambiguous meanings and to make sense of obscure objects and events. Furthermore, it is usually focused on very small portions of reality and forced to make up for its ignorance with an endless series of assumptions, oversimplifications and inaccuracies and untruths. Its ideas are usually mere beliefs and its ignorance, apart from all of the things that are difficult or distant, also affects the ones that lay close at hand.

According to Plato, our usual standpoint is never a purely 'psychic' standpoint because our soul is always accompanied by a body. Each human being is a 'body and a soul put together' (ψυχὴ καὶ σῶμα παγέν)[109] and each human being's outlook on reality is determined by the specific balance of forces struck by these two components. Unlike the soul, however, the body is devoid of cognitive insight. Whereas the soul is a source of life and knowledge, the body is a mere burden or a dead weight. It is that *despite which* the soul attains a clear understanding of reality. And yet its sheer presence is enough to alter the soul's outlook in profound and significant ways.

If the soul were to suffer no interference from the body, its vision would reveal the world as it really is, in all its scope, meaning and detail. When in contact with the body, its universal gaze is reduced to a regional, unclear and unreliable form of access. The soul is forced to cognize reality through a restricted set of bodily organs, whose account of reality is limited to a restricted cognitive horizon and bound to a specific kind of cognitive access (viz. the cognitive access enabled by sight, smell, hearing, touch or taste). Moreover, the soul has to attend to a series of concrete and pressing bodily needs, ranging from the need to breathe, for example, to the specific challenges raised by physical motion, hunger or sexual drive. As a result of all this, the soul's original cognitive condition, whose only motivation was the pure contemplation of reality, is replaced by a self-interested concern with the body: with its preservation and welfare, with its practical

[109] *Phdr* 246c5

orientation, with its enjoyment and satisfaction, etc. In contrast to the soul's original asceticism, this embodied standpoint is restless and passionate. And instead of the soul's original immobility, it is continually led from pleasure to pain and from pain to pleasure.

Since the soul is usually confined to the body, it must view the world *with the eyes of the body* (διὰ τοῦ σώματος) and give in to the body's cognitive demands. In Plato's dialogues, this situation is often portrayed as a form of imprisonment, captivity or servitude. In the *Cratylus*, for instance, Socrates associates the noun σῶμα with the verb σώζειν (i.e. to 'preserve' or 'keep safe') and defines the physical body as 'an enclosure to keep [the soul] safe, like a prison.'[110] In the *Phaedrus*, he claims that the soul is 'imprisoned' in the body 'like an oyster in its shell.'[111] In the *Phaedo*, a pure soul is said to be 'freed from the body as from fetters'[112], and later on a regular soul is described as 'entirely fastened and welded to the body and … compelled to regard realities through the body as through prison bars, not with its own unhindered vision.'[113] As a result, 'each pleasure or pain nails it as with a nail to the body and rivets it on and makes it corporeal, so that it fancies the things are true which the body says are true.'[114]

Human life, insofar as it depends on the human body, is not a free life. And this lack of freedom is grounded, once again, in a peculiar cognitive impairment. Yet Plato's contrast between the soul's intrinsic clairvoyance and the body's intrinsic ignorance does not amount to a simple antithesis. He is drawing up an entire cognitive scale, based on the negative idea of *hindrance* or *interference*. More precisely, he is placing human life inside a cognitive interval whose upper limit corresponds to the soul's originally unhindered vision (viz. to the complete independence of ψυχή from σῶμα) and whose lower limit corresponds to an absolutely hindered, and therefore utterly blind outlook on reality (viz. to the complete subordination of ψυχή to σῶμα). Human life, in its habitual form, is confined to the space left open between these two standpoints. It is a mixture of 'psychic' and non-'psychic' (or somatic) elements, whose inner composition is open to variation.

[110] *Cra* 400c6-7
[111] *Phdr* 250c5-6
[112] *Phd* 67d1-2
[113] *Phd* 82e1-4
[114] *Phd* 83d4-6

Ψυχή and σῶμα do not stand, therefore, for different cognitive attitudes, but for different degrees of the same 'psychic' element. Whereas the soul represents a plenum of cognitive awareness, or a zero of cognitive hindrance, the body represents a zero of cognitive awareness, or a plenum of cognitive hindrance. But despite the formal protagonism of the 'psychic' element, its true or authentic meaning remains out of sight. Although Plato's description of the body-soul compound is often grounded in a description of the soul in its pure form, counterposed to the impurity of its *fallen* or *embodied* form, this kind of account is necessarily allegorical. Since none of us is or was ever a 'pure soul', our understanding of the scale's highest term is always speculative. Not only do we lack the kind of insight ascribed to a pure soul, but we are also incapable of anticipating what it would be like to have it. Whenever we think of a 'pure soul', or a purely 'psychic' standpoint, we do so negatively, by mentally detaching ourselves from our body and trying to imagine what would be left once our bodily needs, contingencies and limitations were entirely suspended.

This kind of limitation also applies, for similar reasons, to the other end of Plato's cognitive scale. Since none of us is or was ever only a body, our understanding of the scale's lowest term is also thoroughly speculative. Our cognitive access to a life devoid of cognitive access entails a performative contradiction, which we usually bypass by means of a mental fiction. Whenever we think of a 'pure body', we do so negatively, by mentally ridding ourselves of our soul and trying to imagine what we would be left with. This is the kind of procedure we might resort to, for instance, when picturing our own death. In these occasions, we propose to step outside of ourselves, as it were, and place our own dead self in front of us. But this experiment is only possible, of course, because we remain conscious of ourselves throughout the whole process, no less alive than we were before.

In short, all of our cognitive representations — including, *a fortiori*, the representation of a purely psychic and a purely somatic standpoint — are both psychic *and* somatic. In the body-soul compound, the compound itself has priority over what is compounded, so that the identification of its ingredients is always a negative or immanent procedure. But Plato's aim is not simply to bring out the constitutive nature of this synthesis. He also points out that this compound is not a peaceful or harmonious one, but a source of tension and unease. In its usual, hindered form, the soul is the target of two different forces, pulling in opposite directions. On the one hand, it is naturally

attached to its pure or unhindered form. On the other hand, it is 'nailed' to the physical body, subject to its needs and desires. Due to this contradiction, the soul is not at home in its skin. Imprisoned and denatured, expelled from its original element, it still holds on to the *memory* of it. In the physical world, the soul's primitive grandeur lives on in the form a faint, but inextinguishable echo, which corresponds precisely to the notion of truth.

We have thus arrived at the central element of Plato's description. The soul, deprived of its immediate communion with truth, is bound to look for it in the bodily realm. And herein lies the primary source of human beings' concern for the truthfulness of their beliefs. Insofar as every human idea or judgement is born within the soul, it is automatically accompanied by a natural attachment to truth. But whereas for a purely 'psychic' standpoint reality and true reality are the same thing, for a composite standpoint such as ours the notion of truth is a formal notion, lacking an immediate or self-evident content. There is, in our usual standpoint, an *excess* or a *disproportion* with regard to truth, which reduces every idea and judgement to a possible *version* of truth, or to a mere *applicant* to the status of truth. In other words, truth is not automatic or transparent, but bound to be translated into specific theses about reality, which may or may not be true.

Moreover, this attachment to truth is not merely logical or gnoseological, but practical and self-interested. Whereas a purely 'psychic' standpoint knows everything, including what is best or most advantageous, for a composite standpoint such as ours the greatest good is not immediately apparent. It is also the object of a thesis, or of several different theses, whose content is liable to change with time and circumstances. The expectation of truthfulness underlying all of our judgements and decisions is also accompanied by a belief in their ability to forward our interests and aspirations. And just as the former may turn out to be groundless, so too is the latter open to refutation.

This general description reiterates the main aspects of Plato's conception of δόξα and δοξάζειν, discussed in the previous sections. Since the soul is always attached to the notion of truth, every cognitive act amounts to the establishment of a specific version of things, usually assumed to be true. Every judgement or decision is already a doxastic judgement, or a doxastic decision, and this general rule is valid even in the lowest stages of cognitive awareness. Although none of us are ever purely 'psychic' or intellectual

beings, our bodily needs and desires are never so strong as to obliterate our natural concern with truth. Even when our soul is faced with the most violent passions, or the most pressing physical demands, our enjoyment and pain are still dependent on the validation of a specific version of things. In these situations, the soul is still called upon to assent to 'the things … which the body says are true.'[115] And although this cognitive approval is often automatic and unconscious, it is still a necessary condition for pain and enjoyment to be perceived or felt as such.

Out of the various passages one might choose from to illustrate the Platonic contrast between σῶμα and ψυχή, the *Phaedrus*' depiction of the soul as a chariot led by winged horses is one of the most revealing. In this well-known allegory, Plato sums up most of the issues mentioned above in a poetic, but nonetheless remarkably accurate manner.[116] Although this is not the place for a thorough examination of this text, let us recall the main elements of Socrates' description. He starts by likening the soul's inner structure to a chariot guided by a charioteer and pulled by a pair of winged horses. This chariot is similar to the one that carries the gods across the heavens, but with a few important differences: whereas the charioteer and the horses of the gods 'are all good and of good descent (τε ἀγαθοὶ καὶ ἐξ ἀγαθῶν)'[117], those of the human soul are mixed. One of the horses, 'noble and of noble breed (καλός τε καὶ ἀγαθός)'[118], pulls the chariot upwards, towards the dwelling-place of the gods; the other horse, 'quite the opposite in breed and character (ἐξ ἐναντίων τε καὶ ἐναντίος)'[119], persists in dragging it downwards, towards the bustle of concrete pleasures and mundane affairs. Therefore, whereas the journey of the gods is safe and easy, the journey of human life is difficult and troublesome, prone to accidents and changes of course.

Socrates' description suggests, from the very beginning, three important ideas, which is useful to bear in mind. Firstly, the human soul, in its usual form, is a decayed or downgraded soul. Although it is naturally akin to the world inhabited by the gods, it fails

[115] *Phd* 83d6
[116] See *Phdr* 246a-257a.
[117] *Phdr* 246a7-8
[118] *Phdr* 246b2-3
[119] *Phdr* 246b3

to attain its degree of perfection. It is modelled after an image or a standard it is unable to meet. Secondly, the human soul is not simple, but complex and self-divided. It is the battling ground of two opposite forces, represented by the horses' different breedings. Thirdly, due to this self-division, the soul's nature is not fixed, but variable. Depending on the skill of the charioteer, as well as on the strength and speed of each of the horses, its journey can be very different, and lead to very different places. These three ideas are all echoed, in a more or less direct manner, in Plato's definition of ψυχή, summed up in the previous pages. Firstly, the soul's usual form is indeed a long way from its essential or potential perfection. Instead of a true and transparent glance on the whole of reality, it is normally confined to a restricted cognitive perspective, grounded in a complex system of beliefs. Secondly, the soul is indeed self-divided, torn between an original urge for truth and clairvoyance and a practical urge for self-preservation, self-orientation and enjoyment. Thirdly, depending on how strongly each of these tendencies is felt, the soul is indeed liable to change. Its nature can vary according to a cognitive scale ranging from a very small degree to a very large degree of somatic hindrance.

At first sight, the opposition between the noble and the rebel horse might be interpreted as the simple contrast between the rational element in human life, which leads men to favour wisdom and prudence over pleasure and gratification, and a more immediate or rash attitude, leading to the opposite behaviour. But although this kind of reading is plausible, it fails to capture the various subtleties of Socrates' image. Recall, firstly, that his whole description is about the soul and not about the body. Each element of the allegory represents, therefore, a specific part or function of the soul. The charioteer, as the ruling element, symbolizes the soul's most primary function, which is that of granting access to reality. The charioteer represents not only cognition as such, but also the possibility of choosing different cognitive contents, as well as the need to settle for one in particular.[120] The horses, on the other hand, symbolize the psychic and

[120] Socrates' comparison of the soul to a chariot led by a charioteer amounts to a variation of a wider Platonic topos, namely the description of human life as a peculiar journey, which requires a peculiar form of navigation. In this context, human beings are often compared to inexperienced pilots (κυβερνῆται), in need of navigation skills (a τέχνη κυβερνητική). Cf. esp. *Al* 135a; *Cl* 498b; *Go* 508cff.; *Phd* 85c, 99d; *Re* 488aff., 551cff.; or *Thg* 123bff.

the somatic tendencies within the soul, that is, the soul's choice between freeing itself from the interference of the body or 'nailing' itself further to it.

Regardless of which is dominant, both alternatives entail the endorsement of specific versions of reality, whose truth-value is determined by the entire soul. Once more, Socrates is intent on showing that every act of cognition, however immediate or instinctive, presupposes a specific set of beliefs assumed to be true. In the allegory, this is made clear by the fact that the unruly, or 'somatic' horse, albeit unwilling to cooperate with the charioteer and the noble horse, never acts on its own. He is bound to the chariot and forced to drag it along towards the things he prefers, so that the entire soul is led to regard those things as truthful and desirable. If the unruly horse proves stronger or more obstinate than the noble one, and if its jerks become too violent or impossible to restrain, the charioteer and the noble horse 'go forward with him, yielding and agreeing to do his bidding'[121]. In other words, even amid the most primitive bestiality or the most unbridled physical lust, when bodily needs gain the upper hand and all of one's thoughts and actions seem oblivious to truth and true knowledge — even then it is the soul that renders pleasure pleasurable, pain painful, enjoyment enjoyable, etc.

The soul, not the body, is the ultimate source of the body's impressions, feelings or affections. The body only 'seems to be self-moving, because of the power of the soul within it'[122]. This conclusion plays a major role in Plato's analyses of pleasure and desire, developed at length in dialogues such as the *Philebus* or the *Gorgias*. In these texts, Socrates attempts to discredit the idea that pleasures are good in themselves or that pains are evil in themselves. And he does so precisely by arguing that pleasure and pain are never automatic or self-evident, but always dependent on the specific course taken by the soul's chariot. Anticipating one of Stoicism's main themes, Socrates aims to show that our usual judgements concerning pleasure and pain are more flexible than we tend to suppose, allowing for a significant degree of variation.

This brings us to a second and crucial element of Plato's allegory, namely the contrast between the ascending and the descending movement of the soul's chariot. Instead of picturing a vehicle travelling on wheels, along a road or a track, Socrates

[121] *Phdr* 254b2-3
[122] *Phdr* 246c4

evokes the image of an Olympian carriage, pulled by flying horses. This peculiarity conditions the whole allegorical setting, subordinating all of its elements to the alternative between moving forward and/or upwards, propelled by the horse's wings, or falling down, due to the force of gravity.[123]

To begin with, this alternative reiterates the idea of a soul confined to a downgraded version of itself, burdened by a calling it cannot fulfil. Not only is the soul's chariot modelled on a divine chariot, meant for perfect or omniscient beings, but its horses are also special horses, morphologically similar to those of the gods. Their wings suggest that their natural habitat is not the earth down below, but the sky up above — that they are destined to move upwards, rather than downwards. And this suggestion is made clearer in Socrates' subsequent description of the soul's wings, and of how they are liable either to shrink into mere stumps or develop into full-fledged organs.[124] What is outlined in this poetical (and comical) passage is once more an entire cognitive scale, measuring the degree of hindrance or impurity endured by the soul in its embodied form. In Socrates' own words, 'the natural function of the wing is to soar upwards and carry that which is heavy up to the place where dwells the race of the gods.'[125] Therefore, when a soul 'is perfect and fully winged, it mounts upward and governs the whole world; but the soul which has lost its wings is borne along until it gets hold of something solid, when it settles down, taking upon itself an earthly body.'[126] Yet although the embodied soul is fated to 'lose its wings' (πτεροῤῥυεῖν), it is never *fully* wingless (ἄπτερος): its wings are always present, even if only as mere stumps, because the soul is permanently affected by the memory of its divine provenance; or again, because whatever its position within Plato's cognitive scale, it already entails some kind of relationship, however remote, with the scale's highest term.

[123] The contrast between an upward and a downward movement (viz. a 'vital' ἐπάνοδος and a 'vital' κάθοδος, or a 'vital' ἀναβαίνειν and a 'vital' καταβαίνειν, etc.) is also a structuring theme in the allegory of the cave, for example, or in the *Phaedo*'s final cosmological myth. Along with the *horizontal* contrast between physical closeness and physical distance, this *vertical* contrast is one of Plato's preferred devices to convey human life's *ambiguous* or *equivocal* condition. See below fn. 221.

[124] See *Phd* 246b6ff.

[125] *Phdr* 246d6-7

[126] *Phdr* 246b7-c4

This point is crucial, and must be sufficiently brought out: however weighty or earth-bound, the soul is always haunted by a perfect, or perfectly winged version of itself. But this haunting does not refer simply to a better version of itself, or to the best version it is able to achieve. The *best* stands, here, once again, for the *absolute best*, the *best there is*, the *best thing conceivable*, etc. And this superlative does not indicate merely a plenum of cognitive insight, but also a plenum of satisfaction or self-fulfilment. Since the soul, in its purest or highest form, knows everything there is to know, it also knows what is best, and leads the best possible life. Its original dwelling place is nothing less than ὁ ὑπερουράνιος τόπος, that is, the very top of the celestial vault, where the gods are said to dwell. But this image may once again be misleading: when we picture this divine realm, we tend to place it somewhere above the clouds, or above the sky, or above the universe — at a great, but finite distance from the human realm. However, this ontological summit, precisely because it stands for the *absolute best*, can never be placed at a finite distance. Although it is always somehow 'close at hand', in the form of a longing or a haunting, it is by definition *thoroughly inaccessible* or *thoroughly out of reach*. If the ὑπερουράνιος τόπος were to correspond to a definite place, with definite qualities and a definite appearance, it would cease to be the best thing conceivable. The soul would be able to conceive something *even better*, even if only in a purely negative or formal way. Accordingly, the superlative determination which the soul aims at is in fact the *superlativized form of a comparative determination*, continually open to higher forms of meaning. It is not simply τὸ βέλτιστον, or τὸ κάλλιστον, or τὸ ἀληθέστατον, but a superlative form of βέλτιον, a superlative form of κάλλιον, or a superlative form of ἀληθέστερον.

Finally, let us turn our gaze downwards. Apart from the soul's wings, and their varying degrees of development, the chariot's movement is also conditioned by the force of gravity. And this simple fact is also of great importance. In order keep the chariot floating, and to elevate it even further, the horses must be able to counter its natural tendency to fall. But whereas the horse of noble breed, determined to pull the chariot upwards, is forced to exert an active effort, the unruly horse is helped downwards by the force of its own weight. The heavens above may well be the soul's primary aim, but the body's weight reverses its ideal course, separating what is *truly* essential from what is only *immediately* so.

The automatic nature of gravity, as opposed to the artificial nature of the soul's ascending flight, illustrates the radical difference in appeal or persuasiveness enjoyed by the 'psychic' and the somatic components of human life. While the somatic mode relates to concrete facts and decisions, the 'psychic' mode amounts to a much less evident possibility, usually perceived as vague and metaphysical, more suited to philosophical enquiries than to the practical concerns of everyday life. Bodily truth is the truth we (literally) fall back on, the one we most readily recognize and accept. It is the cognitive standard we adhere to by default. 'Psychic' truth, by contrast, requires a voluntary shift of perspective — a 'soaring upwards', away from the body and its usual concerns.

Plato's allegory reveals a very peculiar crossroads, whose alternative routes are not truly equivalent, nor equally visible, nor equally comfortable or easy to tread. Yet both of them condition, in varying degrees, the kind of life we lead, the sort of things we value and the aims and projects we favour and pursue. They represent different attitudes towards life, knowledge, and happiness, whose different existential translations will be considered in more detail in the following sections.

3.

MASTERS, SLAVES AND PHILOSOPHERS

Throughout Plato's dialogues, Socrates usually confronts two different ways of life, based on two different kinds of knowledge. In the first case, human beings' main concern is the practical efficiency of their views and judgements. Their interest in truth is superficial and the standard whereby they accept or reject a given idea or belief is not very demanding. In the second case, this standard is raised and truth is valued more highly. The priority of efficiency over accuracy is reversed and one's beliefs are subjected to a closer examination.

In the previous chapter, the first of these attitudes was described with the aid of the *Philebus*. I have pointed out, following Plato, that the observer imagined by Socrates is not a mere spectator, but a self-interested agent, committed to a complex system of assumptions, aims and aspirations. When faced with a distant object, his main concern is not to get to the bottom of what he sees, but simply to gather as much information as necessary in order to settle for a determinate version of things. Since the object's identity is determined by the observer's practical situation, its meaning is always provisional. As the observer's situation changes, and as his aims and aspirations are redefined, so too is the object's definition liable to be revised.

This example illustrates what goes on with most of us throughout most of our lives. Just as with Plato's observer, our account of reality is grounded in a vast array of practical concerns, whose meaning and importance is continually revised and adjusted. But although we do not value truth as such, and seldom wonder about the absolute truthfulness or falsity of our judgements, our cognitive standpoint is nonetheless

dependent on the notion of truth. And indeed just as much so as a purely philosophical or epistemic standpoint. To be sure, philosophical debates about truth and error are usually far from our everyday thoughts. Our life is centred on concrete goals and actions, on things to do and things to avoid. But even though we do not often think about truth and error, we do wish our actions to be successful and our goals to be achieved. And for this to happen our judgements about reality must be reliable — they must reveal the world as it truly is.

Socrates' dialectical charges are usually directed at this very demand. By pointing out the unreliability of our ideas and judgements, he means to highlight our usual incapacity to secure our own interests. If the efficiency we strive for, and usually take for granted, is dependent on the accuracy of our beliefs, and if the latter is open to criticism, our own priorities and aspirations are likewise open to criticism. For Plato, the solution to this problem requires the adoption a new mode of cognition, guided by a new set of priorities and a different relationship with truth.

3.1. *Rhetors and philosophers*

Halfway through the *Theaetetus*, as Socrates and Theodorus discuss the exact definitions of δόξα and ἐπιστήμη, their debate is momentarily side-tracked by a secondary, albeit closely related discussion. From 172c on, Socrates compares two different kinds of life, based on two different attitudes towards knowledge and truth. The first is ascribed to lawyers and public speakers, and the second one to philosophers, or 'lovers' of knowledge.

Let us focus on the first kind of life described by Socrates. Although the passage in question is somewhat long, it is important to bear in mind Plato's exact wording:

> SOCRATES: Those who have knocked about in courts and the like from their youth up seem to me, when compared with those who have been brought up in philosophy and similar pursuits, to be as slaves in breeding compared with freemen.
>
> THEODORUS: In what way is this the case?
>
> SOCRATES: In this way: the latter always have that which you just spoke of, leisure, and they talk at their leisure in peace; just as we are now taking up argument after argument, already beginning a third, so can they, if as in our case, the new one

> pleases them better than that in which they are engaged; and they do not care at all whether their talk is long or short, if only they attain the truth. But the men of the other sort are always in a hurry — for the water flowing through the water-clock urges them on — and the other party in the suit does not permit them to talk about anything they please, but stands over them exercising the law's compulsion by reading the brief, from which no deviation is allowed (this is called the affidavit); and their discourse is always about a fellow slave and is addressed to a master who sits there holding some case or other in his hands; and the contests never run an indefinite course, but are always directed to the point at issue, and often the race is for the defendant's life.

At first glance, Socrates' description is rather specific: his mention of courts of law (δικαστήρια), public contests (ἀγῶνες), an opposing party (ἀντίδικος) and a sworn statement, or 'affidavit' (ἀντωμοσία), leaves no doubt as to the legal nature of his account. The rhetors or public speakers he refers to are lawyers or litigators, engaged in a judicial hearing. What is being highlighted, therefore, is the difficulty of arriving at true judgements during a legal dispute, inasmuch as public litigations do not allow lawyers the time or the detachment needed to reach fair decisions. Since they are always in a hurry, and since their decisions can result in the life or death of their defendants, they can hardly remain impartial and serene.

This permanent tension is bound to take its toll on the litigators' character and to shape it in negative ways:

> As a result of all this, the speakers become tense and shrewd; they know how to wheedle their master with words and gain his favour by acts; but in their souls they become small and warped. For they have been deprived of growth and straightforwardness and independence by the slavery they have endured from their youth up, for this forces them to do crooked acts by putting a great burden of fears and dangers upon their souls while these are still tender; and since they cannot bear this burden with uprightness and truth, they turn forthwith to deceit and to requiting wrong with wrong, so that they become greatly bent and stunted. Consequently they pass from youth to manhood with no soundness of mind in them, but they think they have become clever and wise.[127]

[127] *Th* 172c8-173b3

In addition, Socrates points out that the aim of legal arguments is not so much to be truthful as to be efficient, and that the job of litigators is not so much an art of truth as an art of persuasion. In order to dominate their opponents and win over their audience, litigators are led to rely on artifice, on demagoguery and on increasingly sophisticated forms of deceit. Hence, as they become better at their job, 'their souls become small and warped' and 'they pass from youth to manhood with no soundness of mind.'

By contrasting the life of public speakers and litigators with the life of 'those who have been brought up in philosophy and similar pursuits', Socrates introduces a well-known Platonic theme. Despite his focus on legal justice, and on the specific proceedings of an Athenian court of law, his description turns up in connection with a debate concerning the teachings of Protagoras of Abdera, the most famous and the most conceited of all Greek sophists. Just like Protagoras — as well as Gorgias, Prodicus, Hippias, Thrasymachus or Tisias, among many others — the rhetors mentioned in the *Theaetetus* consider themselves 'clever and wise' and 'know how to wheedle their master with words and gain his favour by acts.' Their distinguishing feature is not their wisdom or their virtue, but their ability to persuade others that they are wise and virtuous.

As in so many other of Plato's dialogues, Socrates is alluding to the fundamental distinction between philosophy's primary concern with wisdom, or σοφία, and rhetoric's primary concern with persuasion and public approval. He is pointing out that the eloquence exhibited by most professional speakers is not indicative of the truthfulness or cogency of their arguments, but merely of the ability to pass them off as true and cogent. Whereas philosophers are concerned with the clarity and consistency of each idea or argument, rhetors and sophists are merely focused on the 'rough and tumble of arguments' (ἡ ἐν τοῖς λόγοις χυλίνδησις)[128]. Like the litigators mentioned above, 'they do not care what the truth is in the matters they are discussing, but are eager only to make their own views seem true to their hearers.'[129]

At first glance, then, Socrates' digression in *Th* 172cff. can be read as yet another

[128] *So* 268a2
[129] *Phd* 91a3-6

charge against the widespread popularity of rhetoric and sophistry.[130] And this is all the more plausible since the role of public speakers in Athenian society during Socrates' and Plato's lifetime was not confined to philosophical or epistemological debates. Their political and social influence was enormous, and their services were requested by many citizens at numerous junctures of personal and public life. Rhetors and sophists acted as intermediaries in job applications, matrimonies and lawsuits[131], and it is the latter capacity — highly sought after in Athens, the city of quarrels and legal disputes[132] — that Plato highlights in the aforementioned passage.

However, this reading only reveals a small part of what goes on at *Th* 172cff. Regardless of this classical topos, Socrates' distinction between 'those who have knocked about in courts and the like from their youth up' and 'those who have been brought up in philosophy and similar pursuits' must not be interpreted too strictly. This contrast also speaks to a wider and more profound cleavage, separating two radically different *modes of being*. Socrates' condemnation of the hasty and self-interested attitude that inspires most legal decisions is not aimed solely at litigators or rhetors, nor at a specific kind of elocution. He is suggesting, moreover, that all human beings, as regards their usual cognitive perspective, are rhetors or litigators in their own right.

In order to appreciate the true reach of Socrates' description, let us look closer at its main elements. He mentions, firstly, the problem of time, and the fact that litigators cannot reflect at leisure, but are forced to make sudden and often unsubstantiated decisions. Whereas philosophers can go over several arguments, study each one in detail and avoid drawing a final conclusion until they are absolutely satisfied, the rhetors

[130] Although the words ῥήτωρ and σοφιστής may be used as synonyms, they differ in a few important respects. The former, used by Plato in the *Theaetetus*, applies generally to a public speaker or orator, in all of his different capacities. The latter, on the other hand, is derived from the adjective σοφός ('wise', 'skilled', etc.) and the verb σοφίζομαι ('to make wise', 'to instruct', etc.), and refers originally to someone who excels in, or is adept at, a determinate craft or faculty. From the late 5th century B.C., however, the word *sophist* acquired a more specific meaning, related to the rise of the so-called sophistical school or movement. The sophists of Plato's day were itinerant teachers of grammar, rhetoric, politics and mathematics. Their discourses (public or private) were highly praised and their lessons were often very expensive. Simultaneously, however, the word *sophist* also acquired a negative tinge, widely explored in Plato's writings and still present in the modern terms *sophism* and *sophistry*. For a brief discussion of these different meanings, cf. GOMPERZ, *Griechische Denker* I, ed. cit., 345ff.; GOMPERZ, *Sophistik und Rhetorik*, ed. cit., esp. 35-48 and 279-290; or GUTHRIE, op. cit., 27-34.

[131] See PHILONENKO, op. cit., 81.

[132] GOMPERZ, *Griechische Denker* I, ed. cit., 347

described by Socrates are not as free. The strict rules that govern legal proceedings prevent them from being thorough in their assessments and the pressing nature of the problems they are faced with calls for clear and speedy resolutions. In court, the clock is always ticking. The timing of a decision may be as important as its content, and a good call made too late can be worse than a bad one made on time.

Furthermore, Socrates also highlights the self-interested nature of legal decisions. Whereas philosophers take a step back to observe and to judge, looking at each issue from different angles, rhetors and litigators are not as circumspect. Given their interest in their own success and in that of their clients, as well as their usual longing for the crowd's approval, their decisions are never selfless or neutral, but always 'tense' and 'shrewd'. For rhetors and litigators, it is not indifferent to decide or not to decide, to decide this way or that way, to decide sooner rather than later, etc. Unlike mere observers, they are continually forced to take a stand, and their decisions usually translate into real and inescapable consequences.

This diagnosis is now becoming more familiar. Recalling our previous discussion, the decisions now linked to the activity of speakers and litigators are very similar to those Plato calls δόξαι. And their self-interested nature, grounded in a permanent urge for practical success and orientation, is also evocative of Plato's characterization of δοξάζειν. Socrates' criticism of the deliberative method employed by legal orators can be read, therefore, as a criticism of human cognition in general. The ῥήτορες mentioned in the text stand for all of us, in most occasions of everyday life: just like litigators, we are always wrapped up in our own vital cause; the clock urges us on, as it does all litigators, forcing us to choose between different judgements, deeds and life paths; just like litigators, we are used to sacrificing truth and accuracy to practical efficiency; and hence, just like litigators, we are continually led to rely on rash and unsubstantiated decisions.

All of this would not be as serious were we aware of it. But Plato's point is precisely that we are not. To be sure, there is a conscious dimension to the deceitful nature of rhetoric. Rhetors and litigators are not naïve creatures, but cunning and shrewd ones, aware of the tricks and ruses they resort to in their speeches. Plato's attack on rhetoric is thus also an attack on the rhetors' usual dishonesty — on their 'crooked acts', their recourse to flattery and their tendency to 'requite wrong with wrong'. But Socrates' criticism also refers to a more profound cognitive tendency, of which the rhetors'

'conscious' dishonesty is only the most visible example. He accuses litigators (and ourselves) of having 'small and warped souls', as well as 'no soundness of mind', due to their (and our) tendency tend to overlook the inherent limitations of their (our) decisions and to mistake the practical convenience of a given judgement or argument for its actual truth. In this cognitive regime, each new decision reinforces the litigator's (and our) illusion of having 'become clever and wise', when in fact it has only helped to establish a provisional and ultimately unwarranted version of things.[133]

We find here yet another instance of the Socratic procedure *par excellence*: in spite of focusing on the activity of orators and litigators, Socrates is making the much more general point that all human beings usually think they know what they do not, and base their entire lives on 'facts' which are in truth mere beliefs. This general misconception affects all human ideas and judgements, and it also extends to the tricks and ruses favoured by rhetoric. Since rhetors are also prey to this peculiar form of ignorance, the tricks with which they deceive others are ultimately also a form of self-deception. Whenever they lead others into thinking something is true, or good, or advantageous, they do so based on their own judgement of what is true, or good or advantageous. But since the latter is just as doxastic as every other judgement, its truth-value is just as open to refutation.

After this brief description, Socrates goes on to characterize the 'band' or 'troop' (χορός) of philosophers, whose life is very different. The first difference between philosophers and rhetors, already mentioned earlier, is their relationship with time: whereas the latter are 'always in a hurry', with the water-clock 'urging them on', philosophers 'talk at their leisure in peace' and 'do not care at all whether their talk is long or short, if only they attain the truth.' Secondly, whereas rhetors are professional

[133] Although in some contexts the modern meaning of the word 'rhetorical' might suggest the idea of abstraction, and thus be used to stress the overly theoretical nature of a given discourse, or the remoteness of its practical implications — as when someone says, for instance, that a question is 'merely rhetorical', or that an argument is 'nothing but rhetoric' — what Plato highlights in his critique of rhetoric is exactly the opposite. Rhetors are above all practical men, concerned with practical truths, and although they are capable of discussing highly abstract epistemological issues, their ultimate goal is to succeed in contradicting their opponents and winning over their audience. It is based on this practical definition that our own lives can be deemed 'rhetorical': like Plato's ῥήτορες, we are primarily focused on practical, rather than theoretical concerns, and usually willing to sacrifice the epistemic accuracy of our beliefs to their practical efficiency.

orators and litigators, masters of eloquence and elocution, philosophers 'appear ridiculous when they enter the courts of law as speakers'. Their quest for truth leads them away from the concrete concerns of everyday life and into the world of thought, making them ignorant of all things practical and political. To illustrate this peculiar kind of alienation, Socrates tells the story of Thales of Miletus, who is said to have fallen into a pit while looking upwards to study the stars. Tradition has it that a Thracian servant girl, having watched the incident, made fun of the philosopher on account of his being 'so eager to know the things in the sky that he could not see what was there before him at his very feet.'[134]

The story of the Thracian girl is a well-known anecdote, whose profound philosophical meaning cast a long shadow in the history of Western thought.[135] Thales can be credited as the first great incarnation of a specific and enduring cultural type: the 'philosopher' or 'sage', clumsy and aloof, constantly absorbed by complex theoretical problems and oblivious to the most immediate or concrete aspects of human life. According to Socrates, what sets Thales and most philosophers apart from ordinary men is not a lack of interest in human life — quite the contrary: precisely because they are unusually drawn to human nature, and inordinately intrigued with what it means for human beings to think or act the way they do, they tend to abstract from life's contingent aspects and to focus on its true essence. Although Thales, the archetypal philosopher, is 'ignorant of what he is doing' and 'hardly knows whether he is a human being or other kind of creature', he exerts himself to find out 'what a human being is and what is proper for such a nature to bear different from any other.'[136]

The list of Thales' descendants is long and varied. His closest successor is Archimedes, whose excitement over the discovery of a fundamental law of physics is said to have made him jump out of his bath and take to the streets of Syracuse naked, eager to communicate his findings.[137] His lack of common sense is also brought out in a well-known anecdote involving his death, recounted by Plutarch: during the Roman siege of

[134] *Th* 174a6-8

[135] For a discussion of its historical and philosophical implications, see BLUMENBERG, op. cit.

[136] *Th* 174b

[137] This well-known anecdote, of questionable historical accuracy, first appeared in Vitruvius' *De Architectura* IX.

Syracuse, Archimedes was absorbed in the contemplation of a mathematical problem and did not notice the military manoeuvres unfolding around him. When a soldier came up to him, with orders to take him before General Marcellus, commander of the Roman troops, Archimedes refused to break off his reflections until he had found a solution to his problem, and ended up being killed.[138]

But whereas Archimedes' death is tragic, the usual remoteness of philosophers has also been explored to comic effect. One of the best examples of this can be found in Aristophanes' *Clouds*, where Socrates and his pupils are depicted as the fanatic enthusiasts of a ludicrous philosophical sect, whose main practices consist in walking on air, contemplating the sun and finding out the truth about all sorts of 'celestial matters' (μετέωρα πράγματα)[139]. The satire caricatures both the aloofness of philosophers and their utter disinterest in (and disdain for) immediate daily concerns. The objects of their studies range from minute biological details, such as the physiology of small insects, to fantastic cosmological theories, concerning the origin and structure of the universe. In all of these cases, however, their conclusions turn out to be of little or no practical value. Their alleged wisdom is humorously contrasted with their complete lack of common sense.[140]

[138] See *Plutarch's Lives* XIX, §4. There are different versions of this story, also reported by Plutarch. While some insist on Archimedes' indifference to the events surrounding him, others attribute his death either to his stubbornness or to a mere accident.

In Plato's dialogues, Socrates' keen interest in theoretical matters also leads him to forget himself and others and to become unaware of his immediate surroundings. Cf., for ex., *Sy* 174d4-175b3 and 220c1-d5. The latter passage, which also takes place during a military campaign, is especially evocative of Archimedes' story: '"… The valiant deed our strong-souled hero dared" on service [in Potidaea] one day, is well worth hearing. Immersed in some problem at dawn, [Socrates] stood in the same spot considering it; and when he found it a tough one, he would not give it up but stood there trying. The time drew on to midday, and the men began to notice him, and said to one another in wonder: "Socrates has been standing there in a study ever since dawn!" The end of it was that in the evening some of the Ionians after they had supped … brought out their mattresses and rugs and took their sleep in the cool; thus they waited to see if he would go on standing all night too. He stood till dawn came and the sun rose; then walked away, after offering a prayer to the Sun.'

[139] *Clouds*, ed. cit., 37

[140] Similar charges can be found, for ex., in ARISTOPHANES, fr. 672 K. I 557; *Birds* 1553ff. and 1282 ff.; or *Frogs* 1491ff. In the *Clouds*, however, as in other writings, Aristophanes does not focus solely on the philosophers' distractedness or lack of common sense. His tone, as Nietzsche puts it, is 'half outraged, half scornful' (*KSA* I, 88; *BT*, 65), and he does not hesitate to accuse Socrates — through the mouth of Strepsiades — of being 'all affectation and fuss' and 'assuming the grand air' (ed. cit., 299). Interestingly, Aristophanes' charges echo Socrates' own charges against the rhetors and sophists of his day. Socratic philosophy is portrayed not only as useless or devoid of practical value, but also as downright fraudulent. The Aristophanic Socrates is therefore immune to the *natural* naïvety of Archimedes or Thales. He is aware

The definition of philosophy as a form of alienation is a recurrent theme in Plato's writings, where Socrates is also portrayed as an outsider, whose ideas, concerns and priorities are very different from those of regular men. On the one hand, his enquiries into the structure and scope of human cognition call into question what most people take for granted. His insistence on reason's intrinsic contradictions leads him to challenge not only the social, religious and political tenets of Athenian society, but also the very foundations on which human existence is usually grounded. As the dialogues unfold, philosophy is increasingly revealed as an *absolutist* enterprise, whose universal vocation is incompatible with the guiding principles of everyday life. On the other hand, however, Plato seeks to show that the alienation produced by philosophy is only perceived or felt as such from a *non-philosophical* point of view. Socrates appears eccentric and unreasonable precisely because his ideas and priorities are judged from a standpoint he himself deems unreasonable. But if human reason were to assume a different form, and to favour different ideas and concerns, this judgement might be radically revised. The possibility of such a revision is the main theme of Plato's dialogues and the main focus of Socrates' arguments and disputes.

3.2. *Socrates and Callicles*

In the second half of the *Gorgias*, Socrates engages in a heated debate with Callicles about the merits of rhetoric and philosophy. During this long exchange, Callicles underlines the eccentricity of philosophers in much the same way as Socrates does in the *Theaetetus*. But whereas Socrates' account is followed by a redeeming moment, in which philosophy is presented as a real alternative — and indeed as the only true alternative — to the life of rhetors and litigators, Callicles' description is limited to its negative aspects. In his eyes, the practice of philosophy must only be encouraged at an early age, and even then only to a limited extent. In adult life, philosophy is not only useless, but also harmful, and even

of the shaky logic underlying his teachings, and comes across as a conceited charlatan, who seeks to profit from his supposed wisdom.

dangerous:

> However well endowed one may be, if one philosophizes far on into life, one must needs find oneself ignorant of everything that ought to be familiar to the man who would be a thorough gentleman and make a good figure in the world. For such people are shown to be ignorant of the laws of their city, and of the terms which have to be used in negotiating agreements with their fellows in private or in public affairs, and of human pleasures and desires; and, in short, to be utterly inexperienced in men's characters. So when they enter upon any private or public business they make themselves ridiculous.[141]

In his long and scornful speech, Callicles goes on to concede that 'it is a fine thing to partake of philosophy just for the sake of education', and hence that 'it is no disgrace for a lad to follow it.'[142] Philosophy is a suitable occupation for young men, and the mark of a free or liberal mind (ἐλεύθερος).[143] But when schooling gives way to real life, and one's decisions become serious and binding, philosophy is no longer beneficial. In its most innocent version, it becomes a laughing matter and a source of ridicule for those who pursue it. Inexperienced in practical matters and public affairs, they are bound to be made fun of — in the same way that Thales is made fun of by the Thracian servant girl, or that Socrates is mocked by Aristophanes. In its most dangerous version, however, philosophy may develop into a serious vulnerability and pose a threat to the welfare of its followers. Since philosophers are usually naïve creatures, unable or unwilling to take an active part in social and political life, they are also less able to defend themselves against other people's malice. Whenever their words and deeds are judged incommodious to others, they are liable 'to be stripped by [their] enemies of all [their] substance, and to live in [their] city as absolute outcast[s].'[144]

With these two worries in mind, Callicles alternates between a patronizing attitude, full of conceit and mockery, and a more ominous, quasi-menacing tone. At first, he describes Socrates' life-long engagement with philosophy as ridiculous and unmanly,

[141] *Go* 484c7-d7
[142] *Go* 485a
[143] *Go* 485c
[144] *Go* 486b7-c2

'more befitting a slave' than a gentleman.[145] He likens his obstinacy to a 'boyish travesty', unworthy of his noble nature and rendered grotesque by his advanced age. Afterwards, considering Socrates' usual insolence and the animosity liable to arise from his enquiries, Callicles' mockery gives way to a more serious admonition, complete with proleptic allusions to Socrates' forthcoming trial and execution.[146]

Callicles' criticism of philosophy and philosophers stands out for its vehemence, but it is by no means an exception within the Platonic corpus. Depending on the interlocutor and the context, Socrates' views are often met with different kinds of resistance: they are the target of perplexity, whenever they seem too absurd or too contrary to common sense; ill will, whenever, they hurt his opponent's pride or self-esteem; indifference, whenever they are deemed irrelevant for the successful running of everyday life; or reproach, whenever they are perceived as intrusive or dangerous.[147]

Again, what is interesting about this catalogue is not simply what these reactions say about Socrates' views, but what they tell us about ourselves. When reading Plato's dialogues, we usually make two automatic assumptions, which condition our entire understanding of the text: firstly, we tend to assume that the views held by Socrates and his interlocutors are equally clear or intelligible; secondly, we tend to liken our own standpoint as readers to that of a spectator, or a referee, whose role is to weigh the arguments on both sides and to decide on their merits.[148] These assumptions, however, are usually unwarranted and tend to conceal the true complexity of Plato's habitual strategy. To begin with, the dialogues are never merely two-way conversations, with Socrates on one side and his interlocutors on the other, but *three-way* conversations, where the reader also plays a part. Plato is continually bringing us into the game, as it

[145] *Go* 485a

[146] The tragic effect achieved in *Go* 486a-b is worthy of Sophocles or Euripides: '[CALLICLES, addressing SOCRATES:] ... if somebody should seize hold of you or anyone else at all of your sort, and drag you off to prison, asserting that you were guilty of a wrong you had never done, you know you would be at a loss what to do with yourself, and would be all dizzy and agape without a word to say; and when you came up in court, though your accuser might be ever so paltry a rascal, you would have to die if he chose to claim death as your penalty.' Cf. also *Go* 521b-522e.

[147] For good examples of each of these attitudes, see, for ex., *Go* 481b6-c4, *Pr* 333e-334a; *Eu* 11b4-5 and *Go* 485e1-496d1.

[148] This is only valid, of course, in the case of dialogues where there is in fact a disagreement between Socrates and his interlocutor(s), and where there are indeed different views under debate. In some dialogues, Socrates has no opposition, as it were, and his interlocutor's role is simply to provide the cue for his ideas and reasonings.

were, and playing on our expectations, on our assumptions and on our past experience. His arguments are often built in such a way as to address a specific prejudice of ours, to elicit a specific reaction on our part, to catch us off guard on a specific subject, etc. Furthermore, the dialogues are never simple contests between symmetrical or equivalent standpoints. Although we tend to assume otherwise, and to put ourselves in the shoes of both contenders, their worldviews are not equally familiar or accessible to us. And this is so, once again, because we are never mere spectators or referees: when reading the dialogues, we are already committed to a specific outlook on reality, and our understanding of each argument is therefore just as biased and self-interested as our understanding of anything else.

From where we stand, Plato's dialogues amount to the confrontation between a standpoint that is familiar to us, embodied by most of Socrates' opponents, and one that is not, embodied by Socrates himself. Although we may empathize with Socrates' views, and even share some of them, they are nonetheless the expression of a foreign world, made of new and strange possibilities. Confined as we are to our usual doxastic perspective, we tend to interpret Socrates' suggestions in a doxastic way, either by downplaying their subversive power or by construing them in a negative way, as a bridge or a pathway to a cognitive beyond. Accordingly, the different ways in which Socrates' interlocutors react to his ideas reflect our own natural ways of reacting to them. Firstly, we are often perplexed at Socrates' suggestions, and tend to regard them as overly radical or extreme. We may accept that some of our beliefs and judgements about reality are uncertain and unintelligible, but we are hardly prepared to accept that *all of them* are open to revision. This idea seems too fantastic to be believed, and too contrary to sound common sense. Secondly, Socrates' suggestion also represents a hard blow to our cognitive self-esteem, and tends to be resisted in direct proportion to its degree of severity. Again, our cognitive powers may not be perfect, but they are good enough to provide us with a reasonable understanding of what goes on around us. Thirdly, even if the two previous assumptions lose their force, and we come to take Socrates' suggestions more seriously, we nonetheless tend to minimize their effects by downplaying their importance. In other words, Socrates may be right on a theoretical level, but his denouncements do not usually prevent us from living our lives, from achieving our usual aims or from making sense of the things that really matter to us. Finally, even when this

last assumption is called into question, and we catch a glimpse of the actual seriousness of Socrates' charges, we are still led to protect ourselves against them by looking the other way. There may be, indeed, a constitutive problem with the way we perceive reality, but why choose to dwell on it, at the risk of sacrificing our usual self-confidence? Why not simply carry on as usual, however unphilosophical that may be?

All of these reactions, instinctively shared by Socrates' opponents and ourselves, speak to the tenacity of the bond that binds us to our usual cognitive situation. Like Callicles, the litigators mentioned in the *Theaetetus* and the vast majority of people, we are naturally convinced that our usual perspective, despite all of its imperfections, can nonetheless reveal what is actually important in life, whereas philosophy's abstract discourses cannot. Unlike Socrates, we are concerned with *concrete* things: that is, with things that count, with things whose actual existence cannot be disputed and whose importance every reasonable man will allow. 'Reasonable men' are of course modelled after ourselves: they too share our preference for concrete things and our distrust of philosophy's obscure and counter-intuitive reasonings.

Now what is at stake in Plato's writings is precisely the possibility of reversing this state of affairs. If the dialogues are read as the opposition between a familiar standpoint and an unfamiliar one, the reader remains fundamentally unchanged. With a familiar perspective to hang on to, consciousness is allowed to take refuge in its usual web of assumptions and to remain within the doxastic realm. But Plato's strategy is to prevent this by turning what is familiar into a problem. As our beliefs and aims are called into question, we are gradually dislodged from our usual cognitive element and led into a new and strange territory, where the previous rules no longer apply. Whenever this operation is successful, our immediate identification with Socrates' opponents loses its force and our attitude towards his ideas is transformed. Instead of mere eccentricities, they are envisaged for the first time as serious possibilities, as concrete and as important as our usual ideas and concerns.

This, then, is Socrates' primary aim: not merely to fight arguments with other arguments, but to induce his listeners to shed their usual skin, as it were, and to contemplate the possibility of a new cognitive order. This peculiar metamorphosis is compared by Alcibiades, near the end of the *Symposium*, to the 'ravishment' produced by the music of Marsyas, the satyr-god, or by the song with which the Sirens tried to seduce

Odysseus.[149] In a passage that is in many ways evocative of Callicles' indictment of philosophy, Alcibiades points out the unique power of Socrates' words. Whereas other orators, however talented, are often unable to capture their listeners' attention, Socrates is always sure to astound and entrance. As an example of this, Alcibiades recounts his own experience and the 'strange effects' he often felt in connection with Socrates' discourses:

> When I listened to Pericles and other skilled orators I thought them eloquent, but I never felt anything like this; my spirit was not left in a tumult and had not to complain of my being in the condition of a common slave: whereas the influence of our Marsyas here has often thrown me into such a state that I thought my life not worth living on these terms. [...] For he compels me to admit that, sorely deficient as I am, I neglect myself while I attend to the affairs of Athens. So I withhold my ears perforce as from the Sirens, and make off as fast as I can, for fear I should go on sitting beside him till old age was upon me.[150]

In this telling passage, philosophy is once again deemed incompatible with the concrete demands of social and political life. Like Callicles, Alcibiades accuses Socrates of diverting attention from the duties a responsible citizen 'ought to mind'. Whenever one is led into the maze of philosophy, what was important and respectable no longer appears so and one's position in the world is suddenly at risk. By Alcibiades' own admission, only by fleeing Socrates and his charms is he able to think straight, value what he used to value and cherish the 'favours of the many' (τῶν πολλῶν).[151] Although he knows Socrates' teachings to be important, and feels ashamed to neglect them, he is instinctively drawn to 'take a runaway's leave'[152] and preserve what is left of his social and practical life.

This passage also brings up a theme which was already present both in Callicles' speech and in the *Theaetetus*' legal description: namely, the association of the alternative

[149] The verbs used are ἐκπλήσσεσθαι, which conveys the idea of being *driven out of oneself* with amazement or astoundment, and κατέχεσθαι, which conveys the idea of being *gripped* or *held hostage* by something or someone else. Alcibiades describes the experience of *disempowerment* brought about by philosophy's criticism of human cognition — the discovery that one's supposed power over one's choices and actions is spurious and illusory. As we will see below, this feeling of disempowerment is a necessary condition, according to Plato and Socrates, for the actual attainment of power. Only by renouncing one's pretence of knowledge and independence can one embrace a truly knowledgeable and independent life.

[150] *Sy* 215e4-216a8

[151] *Sy* 216b5

[152] Ibid.

lives proffered by Socrates and his opponents to the contrast between freedom and slavery. While Callicles considers the pursuit of philosophy 'befitting a slave' and urges Socrates to break his dishonourable habit, the latter characterizes the rushed lives of litigators as a form of slavery, and points to philosophy as a form of release. In like manner, Alcibiades recounts how his own life was revealed by Socrates to be that of a 'common slave', and how he had to flee his presence in order to regain his sense of freedom.

This general motif brings us back to the correlation between freedom, power and knowledge, which lies at the centre of Plato's philosophy. Callicles' position is that of most 'reasonable men', and the one we naturally tend to endorse: from our usual point of view, philosophy is enslaving because it compromises the attainment of the very things we regard as liberating and empowering. If we are unwilling to trust our usual account of reality, or to meet the concrete challenges of everyday life, or to attain the 'means, repute and other good things' praised by Callicles,[153] our freedom will indeed be very limited. Our philosophical qualms, if given free reign, will end up preventing us from getting ahead in life, from controlling our fate and from sharing in life's pleasures and rewards. Too naïve or conscientious to raise our voices and fight for our interests, we will be left at the mercy of others, exposed to all sorts of abuses.

For Callicles, for Plato's litigators and for all of us, freedom is dependent on the ability to get what we want. And what we want, as we have seen earlier, is nothing less than *the best there is* — or, at any rate, the best we are able to get at each given moment of our lives. Yet both for Callicles and for us the identification of what is best only allows for a limited degree of variation. Although our specific life-goals may differ, we all wish to have 'means, repute and other good things in plenty'. And although the value or desirability of certain things may be judged differently by each of us, there are other, more general things that no 'reasonable man' would refuse. Indeed, life's main difficulties are not usually about deciding what is best or worth having, but about the actual possibility of obtaining it. Most of our dilemmas are practical, rather than theoretical or philosophical: they relate to the efficacy of our choices and not to their actual meaning.

[153] *Go* 486d1

Since we usually know what we want, or what we would like to have if given the opportunity, philosophy is of limited interest to us. At the end of the day, we would all trade its abstract 'truths' for the concrete possibility of having 'means, repute and other good things in plenty.'

For Plato, however, this attitude is already the product of an assumption. And like all assumptions, it is liable to be proven wrong. If our entire belief system is exposed to the possibility of error, its highest-order determination is also liable to be false or unintelligible. If the self-interested drive that fuels our choices and actions is based on an automatic definition of the greatest good, all of our aims and priorities may be in need of revision. And therefore, although we usually believe we are choosing and acting in our best interest, we might be doing just the opposite: the 'good things' we set our aim on, along with everyone else, may not be so good after all, and the freedom they are thought to provide may not be as liberating as we tend to suppose.

3.3. *Masters and Slaves*

In the *Gorgias*, Socrates' debate with Callicles about the merits of philosophy is based on a fundamental question, 'which has the highest conceivable claims to the serious interest even of a person who has but little intelligence — namely, what course of life is best.'[154] On one side, Socrates maintains that the concrete definition of the best kind of life (ὁ βέλτιστος βίος) is not immediate or self-evident, and calls for a thorough investigation. On the other side, Callicles, Gorgias and Polus are all convinced that the best things in life are not that ambiguous or difficult to recognize. To be sure, each person's definition of the best life is not only personal, but liable to change with time and circumstances: while one man may choose to spend his entire days eating, another may regard hunting as the most vital of occupations; and while the former may develop, with time, an equally strong passion for hunting, the latter may succumb, in turn, to a life of gluttony. But although the variety of life choices available to human beings is indeed endless, and

[154] *Go* 500c1-4

continually open to revision, Socrates' opponents are nonetheless convinced that there are more general definitions of ὁ βέλτιστος βίος which every man can agree on.

The first of these definitions is offered by Gorgias, who links the attainment of the best possible life with the art of rhetoric, (ἡ τέχνη ῥητορική).[155] In his eyes, rhetoric is superior to all the other τέχναι because its scope is virtually unlimited. Although an experienced doctor, for example, possesses a special insight into the inner workings of the human body, and is thus more able to avert disease and suffering, he is usually as ignorant as most people with regard to the art of hunting. And although a hunter, in turn, is more skilled than other people in the hunting ground, and thus less vulnerable to the dangers of malnutrition and death, he is usually as ignorant as most people with regard to the art of medicine. And the same holds for all the other τέχναι: each of them grants a privileged insight into a specific aspect of reality, but leaves the others more or less untouched; each τέχνη amounts to a ray of light cast upon a small region of reality, surrounded by obscurity and ignorance.

According to Gorgias, a skilled rhetor differs from a skilled doctor or a skilled hunter in that he is not limited to a specific kind of expertise. And the reason for this is twofold. Firstly, since his eloquence enables him to talk doctors, hunters and other τεχνῖται into doing what he wants, he is able to reap the benefits granted by other τέχναι without having to learn them. The actual knowledge about medicine and hunting is still a prerogative of doctors and hunters, but its empowering or liberating effect, which derives from the actual ability to interpret and change the course of events, is appropriated by the rhetor and used in his favour. Secondly, skilled rhetors are more persuasive than doctors, hunters and other τεχνῖται in matters pertaining to the latter's fields of expertise. Indeed, 'there is no subject on which the rhetorician could not speak more persuasively than a member of any other profession whatsoever, before a multitude.'[156] For these two reasons, rhetoric may be said to 'comprise in itself practically all powers at once'[157], and to be 'a cause not merely of freedom to mankind at large, but also of dominion to single

[155] See esp. *Go* 452d1ff.
[156] *Go* 456c4-6
[157] *Go* 456a7-8

persons in their several cities.'[158] Rhetors are the most powerful citizens in the state and the chief consultants in all public matters. If they so wish, they can have doctors as their slaves, along with hunters, trainers, money-getters and everyone else liable to advance their personal interests.[159]

But the fact remains that rhetors, inasmuch as they lack the actual knowledge provided by the different τέχναι, are ignorant of the matters they persuade others about. Socrates' first charge against Gorgias' definition of rhetoric consists in highlighting the fundamental difference between knowing, or *having learnt* that something is true (or false) and believing, or *having been persuaded* that something is true (or false).[160] Since rhetors are not τεχνῖται in the classical sense of the word, the kind of persuasion they produce generates belief without knowledge, and is thus unable to instruct others about true and false beliefs, right and wrong decisions, aims to be pursued and to be avoided, etc. Furthermore, since rhetoric is really only able to produce the *appearance* of knowledge, rhetors can only persuade those who are as ignorant as themselves. Whilst doctors and hunters can be talked by rhetors into acting in their favour, and putting their expertise at their disposal, rhetors cannot persuade doctors or hunters about medical or hunting matters. For the beliefs generated by rhetoric are never as persuasive as those derived from true knowledge.

What, then, is the origin of rhetoric's power? Although rhetors do not know the things doctors know about medicine, nor the things hunters know about hunting, they do possess *some kind of knowledge*. But instead of a real τέχνη, this knowledge amounts to a mere ἐμπειρία, or 'habit'[161], whose real aim is not to produce knowledge, but pleasure and gratification. Socrates proceeds to illustrate this claim by selecting different branches of knowledge and dividing them into two separate lists: the first one includes the arts

[158] *Go* 452d6-8. Cf. 210a9-c4.

[159] *Go* 452e

[160] See *Go* 454c.

[161] Cf. *Phdr* 270b5-6, where Socrates draws a methodological distinction between τέχνη and τριβῇ καὶ ἐμπειρίᾳ, i.e. 'practice and routine'. In both cases, the word ἐμπειρία can also be translated as 'experience', and associated with empiricism and empirical knowledge. In the *Gorgias*, Plato is pointing out that the kind of knowledge conveyed by rhetoric is not epistemic, but empirical. Unlike doctors and other experts, rhetors are ignorant of the truths acquired through a serious study of medicine, or hunting, or economy. However, they do know, out of observation and practice, what other people usually consider true or false, good or bad, desirable or undesirable, etc., and this knowledge allows them to ingratiate themselves with their hearers and to influence their opinions.

related to the welfare of the body (gymnastic and medicine) and the soul (justice and legislation); the second one includes different activities also related to the body (cookery and self-adornment) and the soul (rhetoric and sophistic).[162] The main difference between these two sets lies in the final aim of each activity: whereas gymnastic, medicine, justice and legislation aim to determine what is best for the body and the soul, cookery, self-adornment, rhetoric and sophistic aim to please the body and to flatter the soul. In the case of medicine, for instance, patients rely on doctors to determine the cause of their illnesses and to select the best course of treatment. However, an effective treatment might entail an unpleasant or painful therapy, whose benefits are not immediately apparent. In the case of cookery, the situation is reversed: although the best dishes are the most immediately appetising, they might turn out, in the long run, to be physically harming.

The doctor is in possession of a τέχνη, while the cook cultivates a mere habit. Yet both medicine and cookery claim to give the body what is best — τὸ βέλτιστον. The disagreement stems, therefore, from a different understanding of what the best might be: whereas doctors consider the best to be 'the healthiest', τὸ ὑγιέστατον, cooks regard the best as 'the pleasantest', τὸ ἥδιστον; and while doctors base their opinions on their knowledge of human anatomy and physiology, cooks rely on their knowledge of the human palate. In the case of cookery, the good, τὸ ἀγαθόν, and the pleasant, τὸ ἡδύ, coincide, because the pleasure conveyed by well-cooked dishes is instantaneous. In the case of medicine, however, τὸ ἀγαθόν and τὸ ἡδύ often contradict each other, and only truly coincide when the patient is healthy and the medical art no longer needed. The reasons for this disagreement are once again twofold. Firstly, a state of disease is not always noticeable, and one may lead the pleasantest of lives and suffer, at the same time, from the worst of diseases. Secondly, there is nothing pleasant about taking a bitter medicine, or adopting a restrictive diet, or submitting to a painful therapy. Needless to say, these actions are not usually carried out in view of themselves, but as a means of securing a final and greater good, namely the decrease of pain or the cure of one's illness. However, precisely because the medical road towards the good is often long and unappealing, the healthiest options are often neglected in favour of more palpable or

[162] See *Go* 464b-466a

immediate pleasures.[163]

The parallel between medicine and cookery is offered by Socrates as an illustration of the parallel between justice and rhetoric. Inasmuch as the rhetor's eloquence enables him to steer the will of the crowd, his voice carries a lot of weight with the state and he is often invited to give advice on important public matters. In particular, he is called upon to help decide which laws are just and unjust, which courses of action are advantageous for the state and which should be avoided, which men should be rewarded and which should be punished, etc. Like the cook, however, the rhetor is not a real expert, but 'a shrewd, gallant spirit which has a natural bent for clever dealing with mankind.'[164] Consequently, he is not in possession of a sure method whereby to distinguish justice from injustice. What he does know is human nature and public opinion. His talents are limited to the ability of determining which decisions are likely to please or flatter the crowd and which will be refused or met with distrust. Just like the cook, he prescribes what is pleasant, but not necessarily what is best. And his decisions, albeit flattering, may come to harm the citizens and the state.

Pressed by Socrates, Gorgias ends up conceding that a good rhetor, if he is to be more than a mere demagogue, must learn what is just and act accordingly. But Polus, who takes over from Gorgias, rejects this idea. In his eyes, the fact that most rhetors are not concerned with justice, but merely with their own self-interest, does not restrict the power of rhetoric. On the contrary: if rhetors are not bound by moral or political scruples, their versatility is even greater, and so is their eloquence. They are free to talk anyone into saying or doing anything they please, and the extent of their power is potentially infinite.

From this point on, the dialogue enters its true subject matter, which is not confined to the merits of rhetoric. The debate between Polus and Socrates turns into a debate about power, its nature and the kind of freedom it affords. More specifically, it turns to the issue of whether power is *intrinsically* advantageous and liberating, or whether it must be *rendered so* through the acquisition of knowledge. Polus likens the rhetor's power to that of a king or a despot. Although his intentions are not always just, his authority is

[163] Cf. *Re* 358a-d.
[164] *Go* 463a7-8

usually strong enough to compensate for his iniquities and to protect him from harm. More importantly, Polus argues that the despot's self-interested regard for justice echoes an attitude that is instinctively, albeit secretly shared by most of his subjects. Even though most citizens condemn crime and injustice, they do so out of respect for public order and fear of the law. But all of them would be glad, if given the chance, to swap places with a powerful and unjust ruler. Who would not lay aside his or her moral qualms in exchange for 'the liberty of doing anything one thinks fit in one's city', including 'putting people to death and expelling them and doing everything at one's own discretion'?[165] If the fear of punishment were eliminated, and wrongdoing were rewarded with success and impunity, most men would choose power over justice, and sacrifice rectitude to the actual advantages of a privileged life.

For Socrates, on the contrary, rhetors and despots are not only powerless, but indeed the least powerful men in the whole city (466b-e). And to the dismay of Polus and Callicles, this claim is grounded in an even more 'shocking' (σχετλίη) or 'monstrous' (ὑπερφυής) one: to do wrong is the greatest of evils — greater still than to suffer wrong, however great one's suffering might be (469b-472e).[166] Rhetors and despots, insofar as they do wrong, are not to be envied, but to be pitied. And their alleged freedom is no freedom at all.

At first sight, these paradoxical claims might be read as a mere defence of justice and honesty against Polus' Machiavellian standpoint. Whereas Socrates sides with goodness and rectitude, Polus sides with evil and iniquity, and while the former is a truly honest man, the latter is an opportunist. But this dichotomical reading fails to appreciate the true depth of Socrates' argument. Instead of a mere contest between good and evil, or justice and injustice, the debate between Socrates and Polus is a debate about knowledge and the extent to which the things one knows and ignores are brought to bear in the definition of freedom, power and happiness.

According to Socrates, what unites rhetors, despots and common men is their incessant search for τὸ βέλτιστον: every human being, whether rich or poor, powerful or miserable, eloquent or tongue-tied, wants to lead the best possible life. But although all

[165] *Go* 469c5-7. Cf. *Re* 344b-d and 359b-360d.
[166] Cf. *Phdr* 274a8-b1: 'It is noble to strive after noble objects, *no matter what happens to us*.'

of us choose and act in view of what is best, most of us lack a solid criterion whereby to determine what the best really is. The reason for this, Socrates argues, lies in the difference between the things we *wish* and the things we *know*. Since we all wish what is best, we all do what we think is best; but since what we think is best is not necessarily so, we do not always do what we wish. And herein lies the contradiction underlying rhetoric's and despotism's supposedly unlimited power. For Polus, there is no doubt that everyone would welcome 'the liberty of doing anything one thinks fit in one's city'. In his eyes, such a power is *intrinsically* or *automatically* good, and no one in his or her right mind would fail to realize it. For Socrates, such a power is the result of a determinate *thesis* about the greatest good. And like all theses, it requires demonstration. The power hailed by Polus can only be shown to be beneficial if it is shown to be good, and not merely claimed to be so.[167]

To grant Socrates' point, one needs simply to return to the comparison between doctors and cooks, and to the difference between those actions that are good in themselves and those that are carried out in view of a greater and more or less distant good. When a sick man takes a bitter medicine, he does something unpleasant in view of something pleasanter, and thus *better*. But when a diabetic eats too much sugar, he does something that is immediately pleasant, but might turn out to be unpleasant, and thus *worse*. In the case of rhetors and despots, 'the liberty of doing anything one thinks fit in one's city' — and particularly that of 'putting people to death and expelling them and doing everything at one's own discretion' — might turn out to be harmful for a number of reasons. The despot, confident in his own power, might be led to overestimate the extent to which he is immune to retaliation and end up being harmed or killed by those he has harmed. Alternatively, he might be free of danger but become prisoner of the constant need to reassert his authority and to protect himself from rivals and conspirators.[168] Finally, even if he does manage to escape suffering and requital during his

[167] As Hegel will later point out, the novelty in Socrates' argument consists in showing that 'this good, which has to be esteemed by me as a substantial end, *must be known by me*; with this the infinite subjectivity, the freedom of self-consciousness in Socrates breaks out.' (*HW* XVIII, 442. Emphasis added.)

[168] The despot's 'imprisonment' and the constant need to reassert his power are stressed by Plato in the *Republic*, when describing the onset and development of tyranny (cf. *Re* 467d-569c). Overall, the image of the paranoid ruler, dominated by suspicion and mistrust, is a fertile literary motif, famously explored, for example, in Xenophon's *Hiero*, in Suetonius' biographies of Roman emperors or in Shakespeare's *Macbeth*.

earthly life, he is still exposed to the possibility of facing an unearthly tribunal, with which his earthly authority carries no weight.

But Socrates is not a consequentialist. He does not believe despotism to be evil merely because the despot is not immune to being harmed. He claims, rather, that doing wrong is *itself* the greatest of evils — even greater than being harmed. And what is more, he argues that it is impossible for an unjust man to be happy, regardless of whether he is punished or comes to harm. The wrongdoer, *because he is unjust*, 'is wretched anyhow.'[169]

To prove his point, Socrates starts by having Polus admit that doing wrong is nonetheless 'uglier' or 'fouler' (αἴσχιον) than suffering it. This admission sets the ground for the demonstration of his previous and more controversial claim that 'doing wrong is worse than suffering wrong, and even worse when the wrongdoer is left unpunished.' As regards the first half of this statement, Socrates' argument can be divided into three main steps: he argues a) that a 'fair' thing (καλόν) is usually deemed so because it is either pleasant (ἡδύ) and/or beneficial (ὠφέλιμον); b) that doing wrong can therefore only be

[169] *Go* 472e5. This point is highly ambiguous. By arguing that doing wrong is worse than suffering it, and even more so when the wrongdoer is left unpunished, Socrates is doing more than simply denouncing the groundlessness and/or unintelligibility of Polus' standpoint. In contrast with the *negative,* or *critical* procedure that usually guides his criticism of δοξάζειν, Socrates is advancing two *positive* or *categorical* claims about justice and wretchedness — which is at best atypical for someone whose usual concern is precisely to show that no positive or categorical claim is ever truly substantiated. Since the *Gorgias* fails to offer, up until the very end, an explicit justification of Socrates' claims, there seems to be no way of guaranteeing that they are not themselves doxastic, and liable to be refuted in the same way that Polus' are.

This apparent incongruity can be explained by various different arguments, which at present can only be briefly stated: (1) It may be put down to the alleged difference, pointed out by many commentators, between Plato's initial (Socratic) dialogues and his later (Platonic) dialogues. Apart from noticeable changes in style and composition, Plato's 'Socratic' writings are often regarded as *negative* or *elenctic* dialogues, aimed solely at exposing the inherently faulty or contradictory nature of human concepts and beliefs, whereas Plato's later writings are deemed irreducible to this negative form; (2) Socrates' affirmative stance, in the *Gorgias* and elsewhere, may be due to pedagogical and political concerns. In his brief introduction to the dialogue (ed. cit., 249), W. Lamb points out that Plato's 'design of attracting the attention of the ordinary man of some culture is evident'; (3) Socrates' double standard can be interpreted along the lines of Pascal's famous *pari*. That is to say, by refuting Polus' doxastic defence of injustice and by favouring his own doxastic defence of justice, Socrates is choosing between the lesser of two evils. Whereas doing wrong amounts to openly embracing a mistaken standpoint, doing good amounts to embracing a standpoint that has not yet been refuted, and can still be proven right. Morally speaking, the first of these options is much more hazardous than the second one… (4) It might be argued, finally, that the whole dialogue is written not from a doxastic standpoint (viz. not from our usual, uneducated standpoint), but from a philosophical or epistemic standpoint. If Socrates is already in possession of such a standpoint, he may be *assuming* the epistemic validity of his defence of justice without bothering to prove it. In this case, the dialogue's aim would simply be to refute Polus' and Callicles' mistaken theses, while a proper demonstration of the correct alternative would be left for another occasion.

I will return to this issue below, in section 12.2., when discussing Nietzsche's views on Socrates' position.

deemed 'fouler' than suffering wrong if it is either less pleasant and/or less beneficial than suffering wrong; and c) that doing wrong, since it is not less pleasant than suffering it, must be less beneficial — which leads to the conclusion that doing wrong is indeed worse than suffering it, *quod erat demonstrandum*. As regards the second half of Socrates' claim, his argument can once again be divided into three main steps: he argues a) that a fair action is usually deemed so because it is either pleasant and/or beneficial; b) that punishing someone justly can therefore only be deemed fair if it is either pleasant and/or beneficial; and c) that a just punishment, since it is not pleasant, must be less beneficial — which leads to the conclusion that doing wrong is indeed worse when the wrongdoer is left unpunished, *quod erat demonstrandum*.

This demonstration prompts the intervention of Callicles, which opens the third and last section of the dialogue. Like Polus, Callicles starts by insisting on the counter-intuitive nature of Socrates' claims: if every wrongdoer is indeed wretched, and even more so if he escapes punishment or requital, then 'the life of us human beings [must] have been turned upside down' and we must be doing 'quite the opposite, it seems, of what we ought to do'[170]. Unlike Polus, however, Callicles is not willing to accept Socrates' association between τὸ καλόν and τὸ ἀγαθόν (or τὸ αἰσχρόν and τὸ κακόν). According to Callicles, Polus was not able to dodge Socrates' charges because he was not as bold or as brazen as he should have been. Instead of minding about 'fairness' and 'foulness', he should have dared to look for the truth about justice and injustice beyond the realm of social conventions.

Callicles' views on the merits and inconveniences of a just life echo most of the key-points addressed by Socrates, Thrasymachus and Glaucon in the *Republic*'s first and second books. What is discussed in both dialogues is whether justice and a just life are desirable in themselves, as Socrates argues, or whether they are pursued out of necessity, but against one's innermost will, as argued by Polus, Callicles and Thrasymachus. Just as Polus had maintained that no one in his right mind would choose a life of virtue over the powerful, albeit unjust life of a despot, Glaucon argues that for most people justice 'belongs to the toilsome class of things that must be practised for the sake of rewards and

[170] *Go* 481c3-4

repute due to opinion, but that in itself is to be shunned as an affliction.'[171] If human beings created laws that favour justice over injustice, they did not do so out of the acknowledgement that the former is superior to the latter, but due to 'a compromise between the best, which is to do wrong with impunity, and the worst, which is to be wronged and be impotent to get one's revenge.' [172] The respect for justice is thus artificially ensured by a 'social compact', devised by the weak and impotent to protect themselves against the strong:

> When men do wrong and are wronged by one another and taste of both, those who lack the power to avoid the one and take the other determine that it is for their profit to make a compact with one another neither to commit nor to suffer injustice.[173]

The aim of this compact is to curb human beings' natural appetite for power through an artificially imposed respect for equality and righteousness. According to Callicles, as well as to Thrasymachus and Glaucon, this idea is at the origin of the distinction between what is true or just *according to law*, or *custom* (κατὰ νόμον), and what is true or just *according to nature* (κατὰ φύσιν).[174] Although a just life is socially commendable and practically necessary, within certain boundaries, lest one be marginalized or punished, it is by no means natural. For nature is by definition indifferent to human conventions, guided only by the immediate laws of life, growth and strength.[175] There is, therefore, a fundamental incompatibility between what is *naturally* good or evil and what is *morally* or *legally* so:

[171] *Re* 358a4-6

[172] *Re* 359a5-7

[173] *Re* 358e5-359a2

[174] Cf. Go 482e-484c and *Re* 359c, but also *Pro* 337d, *Phdr* 272d, *La* 715a and 890a. Though the debates regarding justice and morality carried out in the *Gorgias* and in the *Republic*'s first books are very similar, they also differ in important respects. For a brief comparative analysis, see, for ex., BARNEY, op. cit., or BROZE, op. cit.

[175] The translation of νόμος as 'custom', 'convention', etc., may be misleading, and requires some clarification. These English words usually refer to a set of codes and practices considered acceptable to the members of a community. However, when obeying these rules, one is often aware that they are artificial or arbitrary — that they are mere *customs* or *conventions*. What Callicles and Thrasymachus have in mind, however, is a stronger and more binding form of obedience. They are referring precisely to νομοί that are not lived as νόμιμοι, but as as φυσικοί. And herein lies the force of their claim: they are denouncing what appears natural and unquestionable as an 'ideological' fabrication, designed to benefit only a few members of the community.

> By nature everything is fouler that is more evil, such as suffering wrong: doing it is fouler only by convention. Indeed this endurance of wrong done is not a man's part at all, but a poor slave's, for whom it is better to be dead than alive.[176]

Callicles' eulogy to natural justice evokes once again the theme of mastery and slavery. On his view, these two notions do not correspond merely to social or political roles, but to natural or ontological categories, which may or may not be reflected in the actual fabric of social and political relations. 'Natural' masters, on the one hand, are slaves to no one. They are free to do what they please whenever they please, to indulge their every impulse or desire without fearing any kind of requital. This, of course, is not allowed to the many, but only to the 'better type of mankind'[177], to whom 'luxury and licentiousness and liberty', since they 'have the support of force, are virtue and happiness.' [178] On the other hand, the lot of 'natural' slaves includes all of those who lack the natural power or the natural strength to enjoy such freedom. But since these amount to the vast majority of people, the 'natural' opposition between masters and slaves is overturned and replaced by the 'social compact' mentioned above: most human beings, in order to disguise their own powerlessness, decry luxury and liberty as shameful and unjust, and manage thereby to restrain other people's power. 'Being unable themselves to procure the achievement of their pleasures they praise the temperance and justice by reason of their own unmanliness'[179], 'thus … enslaving the better type of mankind'.[180]

Fit in dominatu servitus, in servitute dominatus.[181] Although sovereignty and slavery amount to natural categories, determined by one's natural endowments, the 'unnatural covenants of mankind' have led masters and slaves to switch places. Opposed to this role reversal, Callicles prophesies the elimination of social conventions and the restoration of

[176] *Go* 483a7-b2. The distinction between νόμος and φύσις and the debate concerning the origin of human justice and human laws are central themes of the sophistical tradition. On this subject, see notably HEINIMANN, op. cit.; MENZEL, op. cit., esp. 1-75; GUTHRIE, op. cit., 55-134; or KERFERD, op. cit., 111-132.

[177] *Go* 492a6-7

[178] *Go* 492c4-6

[179] *Go* 492a7-b1

[180] *Go* 492a6-7

[181] CICERO, *Oratio Pro Rege Deiotaro* XI, 30: 'He is sometimes slave who should be master; and sometimes master who should be slave.'

humankind's natural order:

> When some man arises with a nature of sufficient force, he shakes off all that we have taught him, bursts his bonds, and breaks free; he tramples underfoot our codes and juggleries, our charms and "laws", which are all against nature; our slave rises in revolt and shows himself our master, and there dawns the full light of natural justice.[182]

Like Callicles, Socrates also speaks of mastery and slavery, but with a very different aim in mind. In his eyes, freedom is not primarily based on power or strength, but on true knowledge. To refute Callicles' position, he subjects his conception of freedom to his usual zooming method, in order to bring out its unintelligibility. He starts by focusing on Callicles' distinction between a 'better' and a 'baser part of mankind': when asked what he means by 'superior' (κρείττων) and 'inferior' (ἥσσων) sorts of men, Callicles offers a series of different, but equally inconsistent definitions, whose contradictions are systematically exposed by Socrates. At 488d he claims that some men are superior to others due to their physical strength (ἰσχυρότης); at 489e he associates superiority with wisdom (or soundness of mind, φρονιμότης); at 491b with courage and manhood (ἀνδρεία); and at 491e with fullness and gratification (πλεονεξία).[183] In this last definition, the essence of Callicles' standpoint is finally brought to the surface. His idea of freedom comes down to an extreme form of hedonism, where power is measured by the ability to 'let [one's] desires be as strong as possible, and not chasten them', and 'to minister to them when they are at their height', and 'to satisfy each appetite with what it desires.'[184]

This definition leads back to Polus' standpoint, to Socrates' parallel between medicine and cookery and to the distinction between τὸ ἀγαθόν and τὸ ἡδύ. The dialogue's final debate amounts, therefore, to a development of Plato's initial argument, concerning the relationship between freedom and knowledge: since a) what is pleasant or desirable is always pursued in view of what is best, and not otherwise, ergo b) the ability to do what one wills is only liberating insofar as it is actually beneficial. In other words,

[182] *Go* 484a2-b1

[183] The word πλεονεξία only comes up later, at 508a7, but Socrates uses it to sum up this passage and its subsequent developments.

[184] *Go* 491e8-492a2

true freedom rests on the ability to distinguish good or beneficial pleasures from bad or harmful ones. And such ability presupposes a specific knowledge, provided by a specific τέχνη. For Socrates, the name of this τέχνη is none other than philosophy, 'a certain practice or preparation' for the acquisition of the good.[185] Following this new definition, the initial dispute between Socrates and Gorgias concerning the merits of rhetoric is recast under a new light: whereas the rhetor's art consists in the procurement of pleasure or gratification, naïvely understood as naturally or immediately beneficial, the philosopher's role is to expose the ambivalent nature of pleasure and to distinguish good pleasures from false or illusory ones; moreover, whereas Gorgias had defined rhetoric as the finest of all arts, 'comprising in itself … all powers at once', this epithet is now shown to belong by right to philosophy. Unlike the freedom provided by medicine, hunting, or any other particular τέχνη, the kind of freedom pursued by philosophy is not limited to a specific domain of human existence. By attempting to define what is best, and to select thereby the best possible way of life, the philosopher's aim is to lay out the very ground on which freedom of any kind can flourish.

We are now in a better position to understand the Platonic reversal of Callicles' master-slave dialectic. When distinguishing between masters and slaves, or free and imprisoned men, Plato does not refer merely to the political meaning of these words, as we often do, or to their biological meaning, as Callicles does. He is suggesting, rather, that *all of us*, insofar as we lack a true knowledge of what is best, are naturally enslaved, or imprisoned.[186] Although we are usually convinced we are free to choose what we want, we may in fact be doomed to choose what we do *not* want — or what is not, after all, in our best interest. We find here the essence of the contradiction that will reappear in the allegory of the cave and in other Platonic texts: the greatest obstacle to human freedom is not to be found outside human consciousness, where it is usually placed, but in the very idea of freedom. The chains that bind us are the chains of δοξάζειν, and the only chance

[185] *Go* 500d8-9

[186] Plato's philosophical understanding of mastery and slavery is indeed much wider than the social ans political meanings associated with these words. Ironically, though, Plato himself was once sold into slavery (that is, actual slavery), and eventually liberated for the price of twenty minas. According to Diogenes Laertius, he was released when someone informed his owner that he was a philosopher — the implication being that philosophers, from a practical point of view, are useless creatures, with a very low market value. See RIGINOS, op. cit., 86-92.

of release lies in a philosophical reform of human cognition.

Plato's conception of freedom entails a radical transformation of our usual cognitive perspective, affecting all of our ideas and beliefs. The universal scope of this transformation, suggested throughout the *Gorgias*, is also highlighted in the *Theaetetus*, following Plato's critique of rhetors and litigators. When describing the 'troop' of philosophers, Socrates points out their usual lack of common sense and their ignorance in public and political matters: whenever a philosopher meets with other people, or is made to speak in court and elsewhere, he soon becomes 'a laughing-stock not only to Thracian girls but to the multitude in general, for he falls into pits and all sorts of perplexities through inexperience.'[187] The philosopher's 'inexperience', however, refers to the things deemed important by the 'multitude in general', and his 'perplexity' is directed at the things the multitude claims to know and understand. If these things are neither inherently important nor inherently clear, but merely assumed to be so, inexperience and perplexity change sides. The alleged superiority of 'reasonable men' disappears and the alleged servitude of philosophers is revealed as the only real path towards human freedom:

> When, my friend, [the philosopher] draws a man upwards and the other is willing to rise with him […] to the investigation of … human happiness and wretchedness in general, to see what the nature of each is and in what way man is naturally fitted to gain the one and escape the other — when that man of small and sharp and pettifogging mind is compelled in his turn to give an account of all these things, then the tables are turned; dizzied by the new experience of hanging at such a height, he gazes downward from the air in dismay and perplexity; he stammers and becomes ridiculous, not in the eyes of Thracian girls or other uneducated persons, … but in those of all men who have been brought up as free men, not as slaves.[188]

[187] *Th* 174c3-5
[188] *Th* 175b9-d7

4.

CHAINS AND SHADOWS

In Plato's dialogues, men and women are often compared to prisoners or slaves. In the two previous chapters, we have seen some of the central ways in which Plato articulates this idea: firstly, with regard to the notions of truth, certainty and self-interest, I have highlighted Plato's insistence on human beings' usual tendency for self-deception, illustrated by the soul's usual confinement to the body; secondly, with regard to Plato's definition of philosophy and rhetoric, I have discussed the different meanings ascribed by Socrates and his adversaries to the notions of power, freedom, mastery and slavery.

In the next chapter, I will explore this Platonic theme by returning to its canonical formulation, found in the so-called allegory of the cave. This well-known image, featured in the *Republic*'s seventh book, is summed up by Socrates in a few brief sentences. In the bottom of a dark, subterranean cave, a group of prisoners is forced to face a stone wall. Their backs are turned to the entrance of the cave, located at the end of a long upward path, to a fire burning somewhere along this path and to a low wall resembling a screen used in marionette performances, placed between the fire and the prisoners. Various objects are held above the screen and their shadows are projected on the wall at the bottom of the cave. The prisoners watch the movement of the shadows and hear the echoes of voices but are unaware of the marionette performance taking place behind

them.[189]

Before focusing on the different elements of this description, it is important to consider Plato's overall strategy and what it aims to achieve. As we have started seeing earlier, the allegory of the cave is not a simple image, with a simple or straightforward meaning. Socrates' description sets up a complex allegorical situation, which can be looked at from very different angles. At the beginning the allegory, two different standpoints are at work. The first, and most basic one, is the prisoners' outlook on reality, limited to the shadows projected on the wall. Since the prisoners' bodies are bound by chains, and have always been so, their state of imprisonment is not lived or acknowledged as such. Moreover, since the prisoners' necks are also bound by chains, which prevent them from looking in any other direction, the wall and the shadows are all they know. From where they stand, the shadows are not simply shadows, but real things, and the limits of the wall coincide with the limits of the world.

The second standpoint at work in the allegory is our own perspective as readers and interpreters of its meaning. Unlike the prisoners, we have access to a wider portion of reality. We know that the wall is not the whole world and that the images on it are mere shadows of real objects, physically placed elsewhere. The prisoners' life amounts, therefore, to a limited or downgraded version of our own life: their freedom is more limited than our usual freedom and their cognitive perspective is more limited than our usual cognitive perspective.

Everything changes, however, when Socrates claims that the prisoners are *like us*, and that their situation is *our situation*.[190] This declaration introduces a turning point in the narrative, shifting the focus from the prisoners' cognitive limitations to our own cognitive limitations. Whereas the allegory started out as the illustration of a cognitive scale in which we occupy the highest position, and the prisoners an inferior one, this turning point introduces the possibility of a wider cognitive scale, in which our perspective and the prisoners' are both limited or downgraded versions of a higher cognitive perspective. Plato's allegorical procedure amounts to I have called earlier, borrowing Fränkel's

[189] Cf. *Re* 514a-515c
[190] *Re* 515a5

formulation, a Heraclitean thought pattern:[191] Socrates resorts to a comparison whose terms we *can* understand as a means of introducing a second comparison, of which only one term is familiar to us. Starting from the analogy between our own life, usually supported by an immediate belief in freedom and clairvoyance, and that of a group of prisoners who are equally convinced of their freedom and clairvoyance, Socrates establishes a new analogy, in which we are accorded the prisoners' former position and our own position is shifted forward: just like the prisoners are fettered and fail to notice it, we too may be fettered and fail to notice it; and just like the prisoners inhabit a cave, which they mistake for the whole world, we too may be trapped inside a cave which we mistake for the whole world. But this new scenario, precisely because it alludes to a possibility that lies beyond our usual cognitive horizon, can only be envisaged in a negative way. To understand Plato's suggestion, we must project the difference between the prisoners and ourselves into an unknown cognitive terrain, and to posit a new human being ahead of us — one who is somehow freer and more knowledgeable than we are.[192]

[191] See above section 1.2.

[192] The difference between the allegory's *illustrans* — viz. a community of prisoners trapped inside a cave — and the allegory's *illustratum* — viz. the prisoners as symbols of ourselves and our own life — can be further clarified by looking at the structural parallels between the allegory of the cave and the so-called allegory of the divided line, featured in the *Republic*'s sixth book (esp. 509d–511e). Although this is not the place to dwell on the similarities and differences between these two allegories, a few brief remarks are nonetheless in order. As is generally known, the line imagined by Socrates at *Re* 509d, intended to illustrate human cognition's different hierarchical stages, is divided into two main sections, referring to the visible world (τὸ τοῦ ὁρωμένου) and to the invisible world (τὸ τοῦ νοουμένου). These two sections are each divided into two subsections, referring, respectively, to the cognition of images or shadows of real things (εἰκόνες), to the cognition of real objects and living beings (ὁρατά), to the cognition of abstract or intellectual notions, and finally to the cognition of true ideas. Insofar as the prisoners inside the cave are forced to see shadows of real things, their cognitive perspective is initially placed in the line's first subsection. And since the allegory documents, as we will see below, the transition from this world of shadows to the real world, located outside the cave, the prisoners' final destination is the line's second subsection. But this correspondence refers only to the allegory's *illustrans*, and not yet to its *illustratum*. Since *we*, unlike the prisoners, already inhabit the real world, *our* starting point is the line's second subsection, and *our* destination is the line's second main section. In *our* case, the allegory documents the possibility of adopting a higher cognitive standard, liable to reveal what we usually take to be real and intelligible as unreal and unintelligible. The transition from the third to the fourth subsection corresponds, therefore, to the basic experience re-enacted in all of Plato's dialogues — and the allegory of the cave is no exception: starting from our usual cognitive perspective, Socrates calls attention to its inner contradictions and to the need to search for a new and more solid outlook on reality.

4.1. *Inside the cave*

Socrates' initial description refers not to a simple imprisonment, but to a double one. Inside the cave, the prisoners are subjected both to a *physical* confinement, due to the chains that bind them, and to an *optical* confinement, due to their having to see shadows, instead of real objects. Both of these restrictions reflect specific aspects of human beings' usual cognitive perspective, which must now be briefly considered.

The first form of imprisonment is grounded in a paradox that is highlighted in different ways throughout Plato's dialogues. As already noted, since the shackles fastened round the cavemen's legs and necks are in place ἐκ παίδων — that is, since early childhood, or since they are old enough to remember[193] — their captivity is the only kind of freedom they know. Unaware of their own bodies, or the other prisoners, or the surrounding space, they are incapable of grasping their real situation or appreciating the real extent of their freedom. The very notions of freedom and captivity, along with every other idea or belief, are directly conveyed to them by the shadows on the wall. But although the prisoners are liable to spend their entire lives in this state of delusion, their situation is nonetheless an artificial or unnatural situation, maintained by force. If the cavemen were depicted in such a way that their anatomy would somehow prevent them from moving their heads and looking in different directions, then their cognitive limitations would be part of their original identity. In this case, there would be no real tension between what the cavemen see and what lies behind them, because their limited visual scope would coincide with their natural visual horizon. But Plato's cavemen, just like us, are *potentially* able to look in different directions and to see different things. Their captivity is not natural, but imposed: it prevents them from achieving their full cognitive potential and reduces them to a lesser or diminished version of themselves. In other words, the knowledge they are able to attain exceeds the knowledge they actually possess, or claim to possess, and this idea leads back to the paradox we have encountered when discussing the Platonic contrast between body and soul. As suggested in the *Phaedrus*, for

[193] *Re* 514a5

instance, through the image of the flying chariot, there is an original disproportion between the soul's natural cognitive vocation and its actual cognitive condition — a yearning for truth and clarity that is usually quenched by a vast web of assumptions.

Which brings us to my second point. The captivity described by Socrates is not due only to the chains that bind the prisoners, forcing them to see shadows, but also to the shadows themselves, and to their peculiar mode of identity. In the allegory, as in real life, shadows amount to incomplete or imperfect representations of real things, produced by an object or set of objects interposed between a luminous source and a given surface. Depending on the relative positions of each object, the luminous source, the surface of projection and the observer, shadows are liable to have different degrees of sharpness — ranging from clear-cut black projections to diffuse greyish stains — and different degrees of distortion — ranging from familiar silhouettes to abstruse and unrecognizable shapes. In Plato's allegory, however, the most important thing about shadows is their *borrowed* or *second hand* nature. They amount to a particular instantiation of what Plato often refers to simply as 'images' or 'copies' of real things (εἰκόνες, εἴδωλα) — as when he argues, for example, that our usual account of reality is only an *image* of reality, or that our usual ideas are only *images* of true ideas.

Unlike real things, or real objects, shadows are not direct sources of meaning. When considered in isolation, disconnected from the objects they refer to, they cease being shadows and become abstract and meaningless shapes. Their definition is intrinsically tied to the object that produces them, whether or not that object is known, and their identity is thus inherently transitive or dialectical. Shadows amount to visual signposts, as it were, or to visual reminders of real objects. What defines them is precisely their lack of self-identity: they are what they are by pointing to something other than themselves, which is no less a part of their identity. Accordingly, whenever we see a shadow, we do not see a mere shape, but the image of a specific object or a specific person. And even if we do not see or cannot identify that object or person, we nonetheless assume that it must exist somewhere. To identify an unknown shadow is not usually to say what it is *in itself*, but what it is *a shadow of*.

The identification of shadows consists therefore in a comparative procedure, whose success depends on the scope and clarity of one's visual horizon. In Plato's cave, this procedure applies more directly to the reader's standpoint, which has access both to the

objects paraded above the screen and to the shadows projected on the wall. Unlike the prisoners, we readers know that the shadows are mere shadows and that their true identity must be looked for behind the prisoners' backs. From where we stand, the whole situation comes down to a simple lack of correspondence: if the prisoners were allowed to turn around, they would find out what we already know, and would link the shadows on the wall to the objects above the screen. From the prisoners' perspective, however, things are very different. Since they have never seen anything other than the shadows on the wall, they are not aware of their shadow-like quality. In their eyes, the shadows amount to real, self-identical objects, whose meaning is to be looked for in themselves. The world perceived by the prisoners, entirely made up of shadows, is radically different from the real world, but since the very standard with which the prisoners measure what is real is provided by shadows, this difference usually goes unnoticed.

If the prisoners' perspective were presented in purely conjectural terms, as the peculiar outlook of a peculiar sort of men, its interest would be merely theoretical. But with Socrates' 'Heraclitean' move, everything is transformed: the readers are led to abandon their omniscient standpoint, to step into the prisoners' shoes and to take their predicament more seriously. By suggesting that the cavemen are 'just like us', Socrates is implying that our own world — the world we see, touch and hear every day — is entirely made up of shadows. As extraordinary as this suggestion might seem, its scope and gravity must not be minimized: Socrates is implying, indeed, that none of the objects we see or come into contact with are in any way real or reliable. They are all mere shadows, wrongly perceived as real objects.

But this is not all. Socrates is implying, furthermore, that we are surrounded by shadows produced by objects *we have never seen or experienced*. And this suggestion is even less intuitive than the previous one, as it challenges the very idea of shadow we are familiar with. Since shadows are what they are by referring to something else, their definition is necessarily dependent on the position (*Setzung*) of an object, whatever it may be. Usually, then, *objects precede shadows*: they grant them their meaning and purpose, provide the backdrop against which shadows can be thought of as such. But by reducing our standpoint to that of the prisoners, Socrates is turning this basic model upside down. In human life, he argues, *shadows precede objects*. Our account of reality is primarily based on shadows, not on objects, which means not only a) that all of the

things we conceive and experience are mere signposts pointing *somewhere else*, but also b) that all of the things we conceive and experience amount to *empty* signposts, or to signposts pointing *in an utterly unknown direction*.

This description is still somewhat vague. How does Plato's suggestion translate into real life? And why, if at all, should it worry us? To answer these questions, let us abandon for a moment the Platonic cave and turn our attention to two Platonic notions explicitly brought up in other dialogues, but no less present in the *Republic*'s allegory. The first one is the notion of μέθεξις, i.e. 'inherence' or 'participation'. Its meaning can be briefly clarified by turning to the *Phaedo* and to the widely quoted passage in which Socrates discusses the definition of beauty (or fairness, or nobility — τὸ καλόν):

> I think that if anything is beautiful besides beauty itself [αὐτὸ τὸ καλόν] it is beautiful for no other reason than because it *partakes* [μετέχει] of beauty itself; *and this applies to everything.*[194]

Shortly thereafter, Socrates adds:

> If anyone tells me that what makes a thing beautiful is its lovely colour, or its shape or anything else of the sort, I let all that go, for all those things confuse me, and I hold simply and plainly and perhaps foolishly to this, that nothing else makes it beautiful but the presence or communion (call it which you please) of beauty itself, however it may have been gained; […] I do insist that beautiful things are made beautiful by beauty. For I think this is the safest answer I can give to myself or to others, and if I cleave fast to this, I think I shall never be overthrown.[195]

Whenever we ask, as Socrates often does, what it means for something to be 'beautiful', or 'coloured', or 'one', or 'solid', etc., we are inevitably led back to a set of general categories, or εἴδη. Something is beautiful because it *participates* or *inheres* in the notion of beauty, just as something is white because it *partakes* of whiteness, and so on.

[194] *Phd* 100c4-6. Translation altered; emphases added.

[195] *Phd* 100c10-e1. The universality of this principle is also highlighted, for ex., in *Phi* 15d-e: 'We say that one and many [viz. abstract notions and their particular instantiations] are identified by reason, and always, both now and in the past, circulate everywhere in every thought that is uttered. This is no new thing and will never cease; it is, in my opinion, a quality within us which will never die or grow old, and which belongs to reason itself as such.'

When we say 'white', we 'cannot help thinking' of 'whiteness'. And moreover, according to Socrates, to say that something is white because it partakes of whiteness is 'the safest answer we can give'.

Before delving deeper into Socrates' argument, two brief clarifications are in order. Firstly, although many authors translate Plato's 'αὐτὸ τὸ *x*' as 'absolute *x*', this solution is ambiguous and tends to misrepresent, in my view, Plato's entire argument. In this and similar passages, Socrates is not qualifying beauty in terms of its quantity or intensity, as though the expression αὐτὸ τὸ καλόν were meant to single out a 'concentrated' or 'saturated' form of beauty, as opposed to a 'faded' or 'diluted' one. Such a reading usually leads to the simple dualism found in most accounts of Plato's so-called 'Theory of Forms' or 'Ideas' — viz. to the simple contrast between human beings' earthly existence and an 'enhanced' and 'supernatural' existence, where concepts are somehow livelier or purer than the ones employed in real life. But what is at stake in this passage, as in all related ones, is not a *quantitative* distinction between different degrees of beauty, or fairness, or whiteness, or anything else. If that were the case, our usual standpoint would already be in possession of an adequate conception of beauty, for example, and would simply have to look for a clearer or livelier one. What is at stake in Socrates' argument is rather the *qualitative* opposition between *true* beauty, or beauty *in itself*, and a mere *version* of beauty, wrongly perceived as the real thing.[196] Socrates' point, here and elsewhere, is that our usual conception of beauty does not coincide with 'beauty itself' because it is only a mere δόξα, or a mere image, or a mere shadow of beauty. Although we speak of beauty as though we knew what it means, we are usually dealing with a *transitive* or *second hand* determination, whose meaning is missing and nowhere to be found.

The second clarification, which must also be stated very briefly, relates to the scope of Socrates' argument. In the *Phaedo* and in other dialogues, Socrates' discussion of μέθεξις is centred on elementary concepts such as beauty, goodness, unity or difference.[197] These εἴδη are the most basic ingredients of human cognition, and play a therefore a privileged role in Socrates' debates. But this selection does not mean that the

[196] As Hegel will later point out, Plato's ideal realm 'is not [located] beyond concrete reality, in Heaven or in some such place; … It is not metaphysical, but rather *reality itself brought closer.*' (*HW* XIX, 39. Emphasis added.)

[197] See, for ex., *Phd* 100cff.; *So* 256b, 259a; or *Pa* passim.

rule of μέθεξις only affects a small part of our cognitive horizon. As Socrates points out, it 'applies to everything', from the most general to the most specific human beliefs. It also applies, indirectly, to every combination of beliefs, and indeed to the whole web of assumptions underlying our usual outlook on reality. Every complex mental representation partakes of simpler conceptual units, which partake in turn of a series of concepts or ideas 'in themselves'. The whole edifice of human reason is grounded in a thoroughly transitive web of beliefs, intrinsically related to concepts or ideas 'in themselves'.[198]

According to Socrates, 'there is no other way by which anything can come into existence than by participating in the proper essence of each thing in which it participates' (τῆς ἰδίας οὐσίας ἑκάστου οὗ ἂν μετάσχῃ).[199] Therefore, to say that something is beautiful, fair, or white because it partakes of beauty, fairness, or whiteness is 'the safest answer we can give'. But it is by no means a good answer. For when asked about what beauty or whiteness *are*, we can only provide examples of beautiful or white things. While every specific belief leads to back to its εἶδος, or to a more or less complex combination of εἴδη, the latter can only be defined in relative terms, as *that in which each belief inheres or participates*. And this circularity — the circularity of μετέχειν, or μέθεξις — is precisely what is highlighted in Plato's cave. The prisoners can only see shadows because they can only know things as particular instantiations of real objects. But these objects, in turn, are themselves posited through their shadows, that is, through their particular instantiations. The relationship between shadows and real objects, or between beliefs and true knowledge, is not disjunctive, but dialectical. Both terms are essential moments of the relationship and neither can be thought of without the other. Yet precisely because of their intrinsically transitive nature, neither shadows nor real objects provide a stable ground in which to anchor a true account of reality. Whether we cling to

[198] In *Pa* 130c-e, Parmenides argues that abstract ideas are not confined to notions such as the beautiful, the good, and the like. Even the most 'ridiculous' or 'worthless' things, such as hair, mud or dirt, are endowed with specific εἴδη. In the other dialogues, however, Plato does tend to speak of εἴδη in connection with basic or elementary ideas, as opposed to his use of δόξα, which is more generic. But this difference does not indicate a qualitative distinction between different kinds of ideas or beliefs. Inasmuch as human cognition is supported by a global web of ideas, combined in different ways, it is entirely *eidetic*; furthermore, inasmuch as this whole web is automatically assumed to valid and intelligible, it is entirely *doxastic*.

[199] *Phd* 101c2-4

our beliefs or invoke the authority of a concept 'in itself', we are bound to deal with shadows, *and nothing else.*

Another Platonic notion that is directly, albeit inexplicitly linked to Socrates' allegorical description, is the notion of ἀνάμνησις, or 'recollection', along with the well-known Platonic claim that 'learning is nothing else than recollection.'[200] In the *Phaedo*, once again, when discussing the idea of equality, Socrates distinguishes between equality 'in itself', or 'in the abstract', i.e. αὐτὸ τὸ ἴσον, and the equality predicated of different things or objects. Although the notion of equality is not usually considered in isolation, but in connection with concrete things — equal pieces of wood, for example, or equal pieces of stone — it is nonetheless irreducible its particular instantiations. 'Abstract' equality is not the same as 'concrete' equality because equal pieces of wood or stone are never perfectly equal to each other: however similar they might be, they are always liable to be looked at from different angles, placed in different positions, lighted in different ways and perceived in accordance to different cognitive sensibilities, priorities or concerns. In short, although stones and pieces of wood may be considered equal, they 'fall short of being like abstract equality.'[201] Or rather, as Socrates also puts it, their equality *yearns* or *longs* (ὀρέγεται, προθυμῆται, βούλεσθαι) to be like abstract equality but falls short of it.[202]

Our usual knowledge of equality is irreducible, therefore, to the knowledge of equality conveyed by our encounters with equal things. Abstract equality is not a product of experience, but a *condition* thereof: not the result of an empirical observation, but the very standard whereby we observe and measure the equality of the things we encounter. Plato writes:

> When anyone on seeing a thing thinks, "This thing that I see aims at being like some other thing that exists, but falls short and is unable to be like that thing", [—] he who thinks this must necessarily have previous knowledge of the thing which he says the other resembles but falls short of.[203]

[200] Cf. notably *Phd* 72e-73b and *Me*, 80dff.
[201] *Phd* 74d6-7
[202] *Phd* 75af.
[203] *Phd* 74d9-e2

Since our knowledge of equality transcends the realm of experience, it must have been acquired before we began to see, or hear, or use any of the senses. In other words, it must have been acquired before we were born. Simultaneously, however, since our knowledge of equality is not perfect or immediate, but reduced to the concrete instantiations of equality conveyed by our encounters with equal things, it must have been forgotten at birth, or soon after we were born. These two conclusions are the two main principles of Plato's doctrine of recollection, devised in order to account for the peculiar nature of our usual understanding not only of 'the equal', but also of 'the greater and the less', and of 'all such abstractions.'[204] In everyday life, we neither know what equality is nor ignore it completely: we learn it by recollecting what we once knew, but lost touch with. To learn is therefore to bring to mind a knowledge that is somehow already present within us, but must be recovered or re-activated.[205] On the one hand, if we did not carry with us the idea of equality, we would not be able to recognize it in equal stones or pieces of wood; on the other hand, if we had not forgotten it, or lost touch with it, we would not be reduced to the imperfect versions of equality conveyed by equal stones and pieces of wood. As it is, we are trapped between knowledge and ignorance, divided between our *idea* of reality and our *experience* of it. Although equality, greatness and whiteness in themselves are irreducible to concrete equality, or concrete greatness, or concrete whiteness, they can only be known in a negative or indirect way, as *that which concrete concepts are recollections of*. The ideal and the empirical moments are

[204] *Phd* 75c9-10

[205] By tracing our knowledge of the different εἴδη back to the soul's disembodied existence, Plato uses a *temporal* image to describe what the Modern idealist tradition would later call *transcendental* or *a priori* knowledge. This parallel was highlighted by the neo-Kantian school: cf. notably HARTMANN, op. cit., 180f.: 'An dieser Theorie des Lernens durch *ἀνάμνησις* wird der Gedanke durchaus klar, daß es ein Wissen geben muß, das einen anderen Ursprung hat als den im Einzeldasein der Dinge. Und in diesem Wissen soll die wahre Erkenntnis gegründet werden. [...] Denn sie hat ihren gesicherten Grund (*αἰτία*) in dem, was das Bewußtsein aus sich selbst schöpft, was Plato mythisch „Wiedererinnerung" nennt, was wir wohl in modernerer Terminologie als a priori bezeichnen dürfen'; and NATORP, op. cit., 41: 'Das ist die bedenkliche, der großen Errungenschaft des MENO: der Entdeckung des *a priori*. Die Erkenntnis [...] muß Erkenntnis *a priori* sein, d. h., sie muß im Selbstbewußtsein wurzeln.' Yet the most clear formulation of the distinction between the temporal and the transcendental understandings of the concept of recollection can be found directly in Kant, in the classical definition of *a priori* knowledge offered at the beginning of the introduction to the 1787 edition of the *Critique of Pure Reason*: '*Der Zeit nach* geht also keine Erkenntnis in uns vor der Erfahrung vorher, und mit dieser fängt alle an. Wenn aber gleich alle unsere Erkenntnis *mit* der Erfahrung anhebt, so entspringt sie darum doch nicht eben alle *aus* der Erfahrung.' (*KrV* B1. Emphases added.)

intrinsically bound to one another and cannot be thought of in isolation.

Once again, this general description requires two brief clarifications. The first one relates to the tendency, already noted earlier, to interpret Plato's αὐτὸ τὸ ἴσον, or αὐτὸ τὸ ἔλαττον, or αὐτὸ τὸ καλόν, etc., as quantitatively enhanced versions of familiar concepts, located in a 'super-world' of Ideas or Forms. This kind of reading is often coupled with a similar reading of Plato's doctrine of recollection, which tends to reduce Socrates' argument to a simple contrast between the soul's earthly existence and its 'spiritual' or 'supernatural' form. In this and similar interpretations, the *immanent* circularity of human cognition is explained away in a *transcendent* manner: like the reader's omniscient outlook on Plato's cave, the interpreter places before him both the natural and the supernatural realms and decrees their incommunicability. In doing so, however, he fails to appreciate the intrinsically dialectical nature of the recollection process. He construes the object of the soul's recollection as an enhanced or intensified version of a regular object, placed beyond the confines of physical life, and avoids thereby the true implications of Socrates' criticism.

This simplistic interpretation of Plato's ἀνάμνησις is also associated with an equally simplistic interpretation of the role played by mythology within the *corpus platonicum*. To be sure, Plato's myths concerning the pre-existence and re-incarnation of the soul are essential components of his dialogues. On several occasions, Socrates refers to εἴδη as notions learned in another life and to ἀνάμνησις as a recollection of what went on before our physical existence.[206] But the meaning of these images tends to be reduced to one of two limited interpretations. The first one consists in reading Plato's myths as literal descriptions, inspired by religious or mystical beliefs, and/or by the poor scientific resources available during Plato's lifetime. This condescending approach reduces the mythological aspects of Plato's writings to naïve cosmological descriptions, surpassed by modern scientific accounts of human cognition. The second main line of interpretation consists in reading Plato's myths as purely metaphorical or poetical descriptions, put down to Socrates' fertile imagination and separated from the more 'serious' or properly 'philosophical' sections of the dialogues. This reading, albeit in a way more sophisticated

[206] See, for instance, *Phd* 76d-e and *Phdr* 249e-250c, among many other examples.

than the previous one, also fails to appreciate the myths' intrinsic philosophical significance, treating them as secondary or downright superfluous elements.[207]

In my view, the importance of mythology in Plato's dialogues can only be truly grasped by focusing not on the myths themselves (viz. on what they tell us about eternal life, or metempsychosis, or the supra-celestial realm), but on the issue of human cognition (viz. on what they tell us about ourselves and the way we cognize reality). Although the myths as such, prior to their allegorical function, are indeed fanciful tales, which may or may not be taken seriously, the myths as explanations designed to capture the essence of human cognition are as full of significance as the most sophisticated philosophical theories. And this general principle is corroborated by Plato's usual philosophical method. Instead of speaking about migrating souls and the outer surface of heaven as of things existing in their own right, and afterwards trying to determine how they affect human life, Socrates usually does the opposite: having identified a given contradiction or inconsistency at the heart of human cognition, he then comes up with stories designed to bring it out in the clearest and liveliest ways.

Since Plato's myths amount to *ad hoc* explanations, their accuracy must not be measured by their degree of verisimilitude, but by the extent to which they succeed in capturing the peculiarity of our usual cognitive situation.[208] Nevertheless, when suggesting that every particular determination presupposes a set of transcendental εἴδη *as though* they existed in some remote region above the vault of heaven; or when suggesting that human beings possess a latent knowledge of the various εἴδη *as though* they were learned in some previous life, Plato is not resorting to mere metaphors, in the conventional sense of the word. For a metaphor consists in replacing a direct meaning with an indirect one, which presupposes in turn that one has access to the direct meaning

[207] Whereas Leibniz, for example, refers to the 'error of pre-existence' in Plato's otherwise 'excellent' doctrine of anamnesis (cf. *Discours de Métaphysique*, ed. cit., 241, and *Nouveaux Essais*, ed. cit., 40), Natorp regards the 'whole mythical-mystical covering of the doctrine of recollection' as a mere 'poetical addition' (op. cit., 35).

[208] This is more or less what Socrates means, in the *Phaedrus*, when reproaching his interlocutor's concern for the reliability of a myth's sources: 'They used to say, my friend, that the words of the oak in the holy place of Zeus at Dodona were the prophetic utterances. The people of that time, not being so wise as you young folks, were content in their simplicity to hear an oak or a rock, *provided only it spoke the truth*; but to you, perhaps, it makes a difference who the speaker is and where he comes from, *for you do not consider only whether his words are true or not.*' (*Phdr* 275b4-c2. Emphases added.)

that is being replaced. This presupposition, however, is precisely what Plato and Socrates keep calling into question. The peculiar nature of our usual understanding of reality must be conveyed by means of images and myths because what we do not know or understand cannot be accessed in any direct or immediate way. Or again, because all of our ideas and beliefs are themselves mere images or metaphors, whose original meaning is nowhere to be found.

This first clarification leads to a second one, related to the scope of Plato's conception of ἀνάμνησις. According to Socrates, the notion of recollection concerns not only the equal, the great or the less, but also beauty in itself, the good in itself, the just, the holy, 'and, in short, … all those things which we stamp with the seal of "being in itself"' (αὐτὸ ὃ ἔστι).[209] But this does not mean that the notion of ἀνάμνησις refers only to a small part of our cognitive horizon, leaving out other, less abstract or basic concepts. Once more, these specific εἴδη are singled out because they are the basic ingredients of human beliefs, of the various combinations of human beliefs, and indeed of the whole web of assumptions underlying human cognition. Every mental representation, whatever its nature or content, entails a movement of recollection; every idea or belief 'yearns to be something else but falls short of it.'

Even though Socrates, in the allegory of the cave, does not refer explicitly to the notion of ἀνάμνησις, the prisoners' situation is no less representative of this peculiar phenomenon. Just as the soul forgot what it knew upon entering the body, at the moment of birth, so too the prisoners are chained to their seats 'since childhood'; and just as each of the soul's ideas is a mere image or echo of a true idea, so too the prisoners are forced to see images or shadows of real things. In both cases, consciousness is burdened with a trace, or a shadow, or a memory of true knowledge, but denied its actual content.

[209] *Phd* 75d1-2. Translation altered.

4.2. *Towards the light*

After describing the prisoners' situation, Socrates goes on to imagine what would happen if one of them were released from his chains and made to stand up and look around.[210] Seeing his own body and the surrounding space, he would come into contact with a whole new reality, radically different from the one he was used to. His cognitive horizon would no longer be reduced to the wall in front of him and the shadows would eventually be perceived as shadows and lose their referential value. However, as the narrative immediately makes clear, this release would not be a full release. By getting rid of his chains, the prisoner would indeed overcome the first and most immediate form of imprisonment, but only to find himself restrained in new and equally severe ways. Instead of a binary scheme, designed to illustrate a simple contrast between captivity and freedom, or ignorance and clairvoyance, Plato's allegory depicts an entire cognitive scale, made up of different and potentially endless layers of imprisonment, ignorance and servitude.

To begin with, Socrates emphasizes the ideas of distance and remoteness. His cave is a subterranean cave, far removed from the world above. Its entrance is separated from the place where the prisoners are settled by a long, rough and steep uphill path, accessible only to the strongest and most persistent climbers. He points out, moreover, that this subterranean prison is very poorly lit. Its source of light is not the cave's entrance, located too far away, but a fire burning behind the prisoners, which is itself located at a great distance. Given these constraints, a prisoner who would manage to break his chains would still be faced with a number of obstacles. Having succeeded in abandoning his seat, starting up the cave's path and approaching the fire, his previous paralysis would be replaced by a new form of paralysis, caused by the fire's glare. Used to the cave's obscurity and to the shadows paraded on the wall, he would be blinded by the light and prevented from moving further. Although his new standpoint would be 'closer to reality',

[210] See *Re* 515c4ff.

and 'turned more towards real things'[211], the prisoner would be led to think otherwise. Confused and uncomfortable, with sore eyes and a spinning head, he would 'turn and run back to what he can see, thinking that this was really clearer than what he is now being shown.'[212]

With time, however, the prisoner might grow accustomed to the light and to his new surroundings. He would then manage to proceed further along the cave's path and approach the objects whose shadows are projected on the wall. But this new discovery, instead of a full release, would amount to a more sophisticated re-enactment of the allegory's initial deception. Instead of shadows, the prisoner would now be faced with 'props of all kinds', including 'statues and other creatures made of wood and stone', 'fashioned in all kinds of ways.'[213] These objects, held by 'puppeteers' above a screen, modelled after the things that exist in the world up above, are no longer mere projections or ghostly visions, but real bodies, endowed with a specific size, a specific colour and a specific texture. Once again, however, they are not yet real things, in the proper sense, but mere images or copies of real things. Although the prisoner's initial ignorance has been transformed, it has not been eliminated. His idea of reality is still based on a mere representation of reality, passed off as reality itself.[214]

This second stage leads to a third allegorical situation, which takes place at an even higher altitude. If the prisoner, overcoming the first two obstacles, were forced up the cave's ascending path, leaving behind its darkened depths, the screen and the fire, he might end up reaching the cave's exit. But if he were led outside and made to look around, he would once again be paralyzed. The light of the sun, much stronger than that of the fire inside the cave, would dazzle him altogether, preventing him from seeing anything whatsoever. Once again, he would feel dizzy, confused, and suspicious of his

[211] *Re* 515d2-3

[212] *Re* 515e2-4. Translation altered.

[213] *Re* 514b8-515a1

[214] As the prisoner moves beyond the fire and comes across the screen and the objects held above it, he is unaware of the puppeteers hidden behind it. Therefore, when he hears their voices, he is persuaded that the objects themselves are speaking. As the allegory progresses, however, this illusion is not explicitly addressed or cleared up. Although the prisoner eventually realizes that the world is not reduced to the objects held above the screen, he does not actually encounter or confront the puppeteers. Socrates' narrative is henceforth focused on the growing tension between light and darkness, which will culminate in the discovery of the outside world.

new discovery. The outside world would appear less real than the cave's semi-obscurity and the prisoner's first impulse would be to run back inside. With time, however, he might adapt to his new address and learn to endure the world's brightness. But even then, he would not be immediately able to face the objects and shapes in front of him. He would start by focusing on the shadows of real objects. Afterwards, as his eyes grew less sore, he would move on to the reflections of people and other things in water, making out their colour, shape and size. Finally, fully accustomed to his new dwelling place, he would come to stare directly at people and things, discerning at last their actual qualities.

One of the most striking aspects of Socrates' description is indeed the length of the prisoner's journey and the amount of obstacles he has to face. From his initial standpoint to his final destination, he is confined to obscurity and dazzled by strong lights, led up a steep pathway and subjected to multiple forms of deception. Moreover, he is made to do these things against his will, pushed back all the while by a natural aversion to novelty and change. Yet despite the complexity of Socrates' description, what is at stake in the allegory is not a succession of different challenges, nor the acquisition of different kinds of knowledge. Throughout his entire journey, the prisoner yearns for — and is repeatedly denied — *the exact same thing*, i.e. access to reality *as it actually is*. With which new trial, he is led to realize that the knowledge he relies on is a mere illusion, and that the things he takes to be real are in fact mere shadows, or mere copies, or mere reflections of real things, whose true meaning lies elsewhere. The diversity of situations imagined by Socrates is not meant to illustrate the diversity of human knowledge, but the diversity of human ignorance. It is meant to emphasize the plasticity of our usual standpoint and the variety of costumes and masks it is willing to put on in order to conceal its usual groundlessness.

In other words, Plato's allegory is not so much about the alternative between ignorance and knowledge as it is about the ubiquitousness of ignorance and the *possibility* of actual knowledge. Whenever the prisoner believes he has overcome the realm of illusion and found reality 'in itself', he is eventually led to recognize that the latter is merely a new and more sophisticated variation of the former. Not even when reality seems more real, as it were, it is above suspicion, and this continual exposure to error is the central theme of the allegory and the main source of the imprisonment highlighted

by Plato. Since our usual ignorance is characterized precisely by its appearance of truth, true knowledge cannot be attained in a simple or direct way. A new account of reality, dogmatically deemed truer or more real than the previous one, is still as open to criticism as the previous one, and still as liable to be revealed as a form of ignorance in disguise. True knowledge can only be found, if at all, when all forms of deception have been tested and left behind. It can only emerge at the end of a long pathway of ignorance, in which every intermediate stop appears at first as the finishing line. The progression outlined by Socrates is thus a *negative* progression, whose highest term is both present and absent in all of its stages. The truth supposed to lie at the end of the prisoner's ascending path is already known from the very beginning, but in a negative form — as a mere assumption, or a mere signpost, or a mere yearning. The dialectical movement illustrated throughout the allegory is once again the movement of δοξάζειν, or μετέχειν, or ἀναμιμνήσκεσθαι.

This point is extremely important, and often misunderstood. When reading Plato's allegory, we tend to adopt the standpoint of a mere observer and to look at the whole situation as if from above. Unlike the prisoners, we know from the very beginning that the shadows are mere shadows, that the cave is a mere cave and that the real world can only be found outside, at the end of the cave's ascending path. Seeing the whole picture at once, we already know that the prisoner's different stops are the intermediates stages of a global progression, and we also know how it ends. Therefore, when Socrates declares that these prisoners are 'just like us', we picture ourselves at the lowest end of a similar progression, starting with our current standpoint and leading all the way to truth. In doing so, however, we tend to conceive this new progression in a *positive* or *cumulative* way, as a succession of cognitive upgrades culminating in a state of absolute wisdom or clairvoyance. Whereas our usual standpoint is limited and flawed, confined to a very small portion of reality and exposed to a multitude of errors and inaccuracies, the insight to be gained by leaving our imaginary cave is nothing other than a perfect or omniscient account of reality.

This kind of reading agrees with most interpretations of the so-called 'Theory of Ideas' or 'Theory of Forms'. Just as Plato's εἴδη are usually construed as 'enhanced' or 'concentrated' versions of familiar concepts, so too is the allegory of the cave usually read as the illustration of an ascent towards a total or absolute cognitive perspective. As argued earlier, however, this kind of reading tends to underestimate the seriousness of Plato's

criticism. By comparing human life to the life of the cavemen, Socrates is not simply pointing out that our usual cognitive perspective is limited or flawed, that it needs expanding or perfecting. He is suggesting, rather, that none of the things we claim to know or understand are actually known or understood by us. In other words, he is suggesting that our entire outlook on reality, down to our most simple ideas or insights, is grounded in ignorance and indetermination. And he is highlighting, therefore, the need for a complete reform of human cognition.

From the prisoner's perspective — and hence, according to Socrates, from our own perspective — the different versions of reality endorsed throughout the allegory are not the intermediate stages of a wider cognitive ascent. Each new version is perceived as the end of the path, that is, as *reality itself*. Likewise, when each version is called into question and shown to be unreal, the prisoner's disillusionment is not partial, but complete. What he realizes is not that *some* of the things he believed in are false or unclear, nor that *most* of the things he believed in are false or unclear, but that *none* of them are what they seemed to be. With each transformation, the prisoner renounces his whole world and enters an entirely new one, which he takes once again to be real in itself.

This dialectical process is illustrated by two complementary themes, namely the visual contrast between light and shadow (or between clarity and obscurity, σαφήνεια and ἀσάφεια) and the physical contrast between the bottom of the cave and the world up above (or between the prisoner's ascending and descending movements, ἀνάβασις and κατάβασις). As regards the first theme, Socrates distinguishes between 'two ways in which the eyes become confused', namely when one moves 'from darkness into light' and 'from light into darkness.'[215] Let us start by focusing on the former movement, corresponding to the liberated prisoner's ascending journey. In the bottom of the cave, the lights are at their lowest. The fire is reduced to a distant glow and the cavemen can hardly make out the shadows on the wall. However, darkness is never absolute because the prisoners are never completely devoid of cognitive insight. Even at the lowest of cognitive stages, their outlook on reality is already a mixture of knowledge and ignorance, or vision and blindness, in the peculiar sense we have seen earlier. Since each prisoner is

[215] *Re* 518a2-3

already burdened with a recollection of true knowledge, what he sees is already a version of reality, but one that falls sort of reality itself.

As one of the prisoners breaks free and moves towards the fire, the shadows recede and the light increases. Paradoxically, however, this change ends up compromising his ability to see the objects in front of him. Although his previous standpoint was less revealing than his new one, his vision is now blocked by the fire's glare. The cave's brightness (φώς, φανότης) is no longer synonymous with cognitive clarity (σαφήνεια) and the prisoner's first instinct is to return to the cave's obscurity (σκότος, σκοτεινότης), which he regards as clearer (σαφέστερος), and hence truer or more real (αληθέστερος).[216] This situation echoes the cognitive challenge usually faced by Socrates' interlocutors throughout his enquiries: as their beliefs are subjected to a closer examination and found to be uncertain and unintelligible, they are also left in a state of perplexity and disorientation; as Socrates dismantles the web of assumptions that supports their usual cognitive perspective, their first instinct is to 'take a runaway's leave'[217] and regain their usual self-confidence.

The second source of visual confusion, experienced when one moves 'from light into darkness', is illustrated by Socrates through an inversion of the allegory's initial situation. After imagining what would happen if a prisoner were to break his chains, move on upwards and exit the cave, Socrates pictures that same prisoner returning to the bottom of the cave and settling down in his initial dwelling place. This time around, his eyes would be accustomed to the brightness of the sun, and 'filled with darkness' upon entering the cave.[218] Just as he was once a foreigner in the outside world, new to the appearance and meaning of real objects, so too would he now be a foreigner inside the cave, incapable of making out the meaning of the shadows on the wall. Before his eyes had had the time to adjust back to the cave's darkness, he would be criticized and made fun of by the other prisoners. His visual impairment would be regarded as a symptom of ignorance and foolishness, and his praise of the upper world, of its brightness and clarity, would be greeted with scepticism and distrust. Furthermore, if the liberated prisoner

[216] Cf. *Re* 515d-e
[217] *Sy* 216b5
[218] *Re* 516e4-6

were to try and free his companions and convince them to follow him upwards, their initial reserve might evolve into open disdain and hostility. Angered by his provocations, the others might even decide to 'get their hands on him' and 'put [him] to death.'[219]

This description is once again evocative of the behaviour adopted by Socrates' interlocutors throughout his enquiries. Whereas the visual impairment experienced during the prisoner's ascent referred to human consciousness' habitual standpoint, and to its usual blindness with regard to its own contradictions, this second form of impairment is characteristic of a philosophical standpoint. In Plato's dialogues, it is usually illustrated by Socrates' habitual detachment from the social and political world, by his unwillingness to engage in mundane affairs, by his disinterest in earthly pleasures or possessions and by his overall devaluation of the things that are cherished and pursued by most people. From his point of view, human beings' usual aspirations are grounded in mere assumptions, or mere shadows, and belong therefore in the bottom of the cave. For regular men and women, on the other hand, Socrates' concerns are themselves shady and obscure. His obsession with clarity and coherence is not only ridiculous and impractical, but also a nuisance to other people and the ultimate cause of his downfall. Like Callicles' warnings against the practical dangers of philosophy, Socrates' emphasis on the hostility of the cavemen towards the liberated prisoner is a direct allusion to his own tragic fate.[220]

Finally, let us look more closely at the second of the two aforementioned allegorical themes, namely the cave's topography and the specific significance of the variations in altitude considered by Socrates. Here, once again, it is important to highlight two central aspects of his description. Firstly, as in the *Phaedrus*' allegory of the horses and the chariot, there is a fundamental asymmetry between the 'way up' (ἄνω) and the 'way down' (κάτω), which conditions the whole allegorical setting. On the one hand, just as the soul's winged horses were naturally destined to fly upwards, so too are the prisoners destined to lead a free life and to inhabit the real world. As things stand, they are settled outside of their natural element and prevented from fulfilling their true cognitive vocation. On the other hand, just as the soul's chariot was naturally pulled down by the force of gravity, so too are the prisoners put down by the steepness and by the roughness

[219] Cf. *Re* 517a; *Go* 486a-c, 521b-522e; *Ap* 31a.
[220] *Re* 517a4-6

of the cave's ascending path. It is much easier for them to stay seated than to venture upwards, towards an unknown and remote destination.

In both allegories, Plato emphasizes the contrast between consciousness' 'aerial' calling and its 'earthbound' or 'sedentary' nature. Even though our usual standpoint, however limited or flawed, is always aimed at true knowledge, and used to regarding what it sees as the real world, our concern for clarity and coherence is not very high. In other words, although we are naturally interested in truth, we have a tendency to 'sit back' (θᾶκος) and take the truthfulness of what we see for granted.[221]

But the theme of altitude is also revealing of a second and equally fundamental idea. In line with many other Platonic texts, the allegory of the cave illustrates the contrast between human consciousness' usual cognitive perspective, based on a web of beliefs automatically assumed to be true, and the possibility of a new and radically different cognitive perspective, based on a true or epistemic account of reality. Unlike other texts, however, the allegory is particularly emphatic about the length of the journey leading up to this transformation and the variety of pitfalls found along the way. Given the remarkable versatility of our usual standpoint, none of the cognitive conquests made inside the cave can be taken for granted. All of them are provisional, and end up revealing new and more sophisticated forms of imprisonment. Nevertheless, not all of

[221] The allegorical contrast between an ascending and a descending journey, as well as the effort required to rise above one's usual cognitive position, are also central themes in the cosmological fable recounted by Socrates near the end of the *Phaedo* (109a-114c). In a description that is very similar, in many respects, to the allegory of the cave, Socrates imagines a group of creatures dwelling in the bottom of the sea, but convinced they inhabit the upper surface of the earth. Having lived their entire lives underwater, with no idea of the world up above, these creatures would see the sun and the stars through the water and would be led to think that the see was the sky. By means of a Heraclitean thought pattern identical to the one employed in the *Republic*'s allegory, Socrates compares these creatures to us and to our own cognitive situation: like them, 'we dwell in a hollow of the earth and think we dwell on its upper surface; and the air we call the heaven, and think that is the heaven in which the stars move. … By reason of feebleness and sluggishness, we are unable to attain the upper surface of the air; for if anyone should come to the top of the air or should get wings and fly up, he could lift his head above it and see, as fishes lift their heads out of the water and see the things in our world, so he would see things in that upper world; and, if his nature were strong enough to bear the sight, he would recognise that that is the real heaven and the real light and the real earth.' (*Phd* 109d6-110a1) In the *Phaedrus*' allegory, human consciousness' natural 'sluggishness' was illustrated by the chariot's natural tendency to fall, due to the force of gravity; in the allegory of the cave, this same difficulty was conveyed by the steepness of the cave's ascending path and by the prisoners' inability to endure the brightness of the sun. In the *Phaedo*, by comparing human beings to fish, Socrates underlines once again their natural aversion to a higher or more demanding cognitive standpoint. Although human beings, like fish, can rise above water, they quickly lose their breath. Their natural tendency is thus to regain the depths of the sea, where they can breathe at ease.

these stages are alike. The cave's path is an ascending path, starting at the bottom and ending at the top. Socrates' image corresponds, therefore, to a global hierarchy of ignorance, with each new conquest leading the prisoner 'a little closer to reality' and 'more toward real things'. Although each disillusionment is a complete disillusionment, each new starting point is superior to the previous one and more distant from the beginning. Moreover, although the journey is very long, the prisoner is bound for a specific destination, located at a finite distance. According to Socrates' image, his fate is not to go on climbing indefinitely, as though the cave were indeed the whole world, but to reach the sunlight and the real world.

Here, however, a word of caution is required. As we have seen at the beginning of this chapter, Socrates' allegorical device is a negative or 'Heraclitean' device[222]: he focuses on the transition from the prisoners' standpoint to our own standpoint in order to convey the possibility of a transition from the latter to a new and even higher standpoint, located beyond our usual cognitive horizon. Accordingly, when the prisoners reach the end of their journey, their cognitive perspective coincides with our usual cognitive perspective, prior to the journey we ourselves are called to undertake. But what characterizes our usual outlook on reality, as Socrates repeatedly points out, is precisely the conviction that no such journey is needed — that the world we see is the real world, made of real people and real objects. Therefore, if this conviction is unwarranted, the prisoners' final release may also be misleading. Just as the shadows on the wall, the props above the screen, the shadows outside and the reflections on water were all revealed as images devoid of content, so too is the sun-lit world liable to be part of a new and even more elaborate cognitive hoax. When the prisoners reach the end of their journey, they may be entering a second cave, lit by a new and more potent fire. And this new cave may lead in turn into a third one, contained within a fourth one, and so indefinitely. Plato's story of captivity and release may indeed be an endless story, in which every conclusion is a new beginning and every reality the image of something else.

[222] See above section 1.2.

4.3. *The self as shadow*

Before abandoning Plato's cave, one last remark concerning shadows and real objects is in order. At the beginning of the allegory, Socrates compares our usual standpoint to that of prisoners bound by shackles and forced to see shadows projected on a wall. But although these shadows are paraded in front of us, like images on a theatre screen, the allegory does not refer merely to an external problem. The prisoners, the shadows and the cave are meant to illustrate not only what we are able to see or know about the world around us, but also *who we are* and *who we claim to be*. Plato's description is just as much a diagnosis of human knowledge as it is a diagnosis of human identity and a reflection about the possibility of a radical ontological transformation.

According to Socrates, our usual account of reality is based on shadows, rather than real things. But the shadows we rely upon, unlike regular shadows, are unable to reveal the objects and events they are shadows of. Their transitivity is compromised because the things they refer to are missing, or were never known in the first place. And this problem affects not only some of our ideas and beliefs, nor even most of our ideas and beliefs, but all of them. As we have seen earlier, Socrates' indictment is directed at the entire edifice of human cognition: from the simplest to the most complex insights, from the way we represent the surrounding world to the way we represent ourselves, and even to the way we represent our representation of the world and ourselves. Every one of our ideas amounts to a shadow, or to the shadow of a shadow, or to an even more complex combination of shadows. No matter how far or how deep we look, the actual source of our beliefs remains hidden and out of reach.

Usually, as the dialogues demonstrate, most of us tend to reject this diagnosis or to underestimate its seriousness. Like the prisoners inside the cave, we are naturally attached to our usual standpoint and averse to radical change. Moreover, even when we do consider Plato's views, and admit to being *confronted* with shadows, or *in possession* of shadow-like representations, or anything of this kind, we are led to construe the problem in a strictly objective manner. As the world's identity is called into question, we take refuge in our own identity, as though it were immune to Socrates' charges. By reducing the issue to a disagreement between selfhood and otherness, we hang on to a primary

core of self-identity, unaffected by the circularity that contaminates everything else.

This expedient, however convenient, is still grounded in a mere assumption. Although our own identity appears to be more immediate or more familiar than the identity of objects and events, it is no less bound by the rule of μέθεξις or ἀνάμνησις. Indeed, even though the idea of selfhood is undoubtedly different from the idea of a chair, or of justice, or of the greatest good, it is nonetheless mediated by the same cognitive alphabet employed to envisage all of these things. And since this alphabet, according to Plato, is entirely made up of shadows, self-identity is itself made up of shadows. Ultimately, then, by comparing our life to that of the prisoners, Socrates is not suggesting simply that we *see* shadows, or that we are *faced* with shadows, but also that we *are* shadows — and *nothing more than* shadows. On closer inspection, the problem of not *knowing* what we think we know is also, and above all, the problem of not *being* who we think we are.

The *Republic*'s allegory draws on an ancient literary topos, frequently found in Greek poetry and drama. In Pindar's *Pythian Odes*, for instance, man is said to be 'the dream of a shadow' (σκιᾶς ὄναρ), set alight with 'the brilliance given by Zeus'[223]; In Sophocles' *Ajax*, Odysseus declares in similar vein that 'all we who live are nothing more than phantoms or fleeting shadows' (εἴδωλ᾽ ὅσοιπερ ἢ κούφην σκιάν)[224]. These and like images anticipate Plato's own characterization of human beings as *shadows* of real beings, and of human existence as a *shadow-like* existence. In the allegory, he suggests not only that every human belief 'yearns to be something else and falls short of it', but also that every human being 'yearns to be something else and falls short of it.' He suggests, in other words, that human life is globally burdened with the recollection of a different kind of life — a *real* life, where shadows are not mere shadows, but real objects, and beliefs are not mere beliefs, but true ideas.

Once again, however, this difference should not be interpreted in positive or

[223] *Pythian Odes* VIII, 95

[224] *Ajax* 126 (ed. cit., 44). The word εἴδωλον may refer to the phantom or to the ghost of a living being (viz. the εἴδωλα of the dead wandering through Hades), but also to a mental image (cf., for ex., *Phd* 66c or *Th* 150c), or to an image reflected in a mirror or in water (cf., for ex., *So* 266b or *Re* 516a). In all of these cases, the defining trait of εἴδωλα is their *negative* or *defective* nature: viz. the fact that they are not what they appear to be, or fall short of their alleged meaning or nature. Εἴδωλα amount to φαυλότερα ὄντα, that is, to 'paltry' or 'second-rate' versions of real beings.

cumulative terms. This new life is not simply a better or enhanced version of our current life. In fact, strictly speaking, it is not new at all. The life we fall short of is the life we already claim to know and lead every day. The ontological tension Plato keeps highlighting, by means of different images and allegories, refers not to the contrast between who we are and who we could become, in some ideal or transcendent world, but to the immanent contrast between who we are and who we take ourselves to be. To 'know ourselves', as Socrates recommends, is thus to acknowledge that we have yet to become who we think we are and to arrive at the place where we claim to be.

The idea that man is a mere shadow (σκία), or a mere phantom (εἴδωλον), or a mere image (εἶκον) of himself — evoked in different contexts by Pindar, Sophocles, and Plato — is also related to another important Greek topos, explicitly referenced in the *Republic*'s allegory. As Socrates highlights the incompatibility of the life led by the cavemen and the one that is eventually embraced by the liberated prisoner, he observes that the latter, having grown accustomed to the outside world, would lose all interest in the honours and rewards cherished and fought for inside the cave.[225] To illustrate this point, Socrates quotes from Homer's *Odyssey*, and more precisely from the episode where Odysseus, in search of Tiresias, descends into Hades, 'where dwell the unheeding dead, the phantoms of men outworn. (βροτῶν εἴδωλα καμόντων)'[226] Moving past dead servants, heroes, kings and relatives, all reduced to mere images of their former selves, Odysseus comes across the shadow of Achilles, who complains to him about his current condition. Although Odysseus tries to comfort him by claiming that even in death he is a great prince, Achilles retorts that he would rather be a poor slave to a poor master 'and live on earth' than 'to be lord over all the dead that have perished.'[227] These same words are used by Socrates to describe the liberated prisoner's reluctance to return to his previous life: like Achilles, he would eagerly wish to be a poor slave to a poor master and 'undergo anything rather than what he thought about down there and live in the way he did down there.'[228]

With this comparison, Socrates renders explicit a parallel that had already been

[225] See *Re* 516c-e
[226] *Odyssey* XI, 475-476
[227] *Odyssey* XI, 489-491
[228] *Re* 516d6-7

suggested ever since the beginning of the allegory: the contrast between shadows and real things mirrors the contrast between the world of the dead and the world of the living, and the cave evokes the subterranean depths of Hades. The implication, of course, is that the cavemen, reduced as they are to images or shadows of themselves, are not truly alive. Their peculiar kind of existence is grounded in the peculiar kind of transitivity we have already encountered at different junctures of the allegory: although the prisoners are not truly alive, they are not truly dead either; their life is permeated by death (or by darkness, or by deception), just as their death permeated by life (or by light, or by truth). And something very similar happens in Homer's description: although the dead have kept their former appearance, and therefore *seem* to be alive, they were divested of their former strength, power, and freedom, and reduced to mere shadows; everything in them is a sign of something else, which no longer means what it used to. They are pure *evocations*, or pure *recollections* of real beings, whose original no longer exists.

Like Homer's phantoms, the cavemen's power is reduced to the appearance of power and their freedom is reduced to the appearance of freedom. They are not real people, in the true sense of the word, but mere shadows or visions of real people. However, since we readers are the allegory's true protagonists, it is our own power, our own freedom, and indeed our whole life that is being called into question: just as the cavemen may be dead and not know it, *we* too may be dead and not know it; and just as the cave itself may be the world of the dead, *our own world* may be the world of the dead. This possibility is explicitly mentioned in the *Gorgias*, during Socrates' debate with Callicles: quoting from one of Euripides' lost plays, Socrates wonders whether 'to live is to be dead, and to be dead, to live.'[229] This brief formula encapsulates the *Republic*'s entire allegorical situation and aims to highlight, once again, a possibility that lies beyond our usual cognitive horizon. If life as we know it is indeed a form of death, the contrast between life and death is transposed to a higher plane, in which both terms acquire a radically new significance: just as our own life is qualitatively superior to death, in the traditional sense of the word, so too a real or genuine kind of life, as Socrates understands it, is qualitatively superior to the life we lead. This new life is at first merely the negative image

[229] *Go* 492e. Socrates quotes a verse from Euripides' play *Phrixus* (fragment 833).

of our current life, and can only be rendered positive through a complete reform of our usual cognitive perspective.

5.

PHILOSOPHY AND DELIVERANCE

In the previous chapters, following Plato, I have highlighted the universal scope of human beings' cognitive imprisonment: although personal freedom is limited in a variety of external or contingent ways, the kind of captivity described by Socrates in the *Republic* and elsewhere is not external or contingent, but an essential feature of human identity, underlying every human thought, belief or decision; likewise, although personal freedom may be conquered or increased in different ways, the deliverance envisaged by Socrates is not a regional or partial deliverance. It entails a complete cognitive transformation, bound to affect every single aspect of human life. In the following sections, I will discuss the nature, relevance and possibility of this transformation, in an attempt to determine whether Plato's philosophical project is bound to remain a critical project, focused on the negative effort of uncovering human reason's inner contradictions, or whether it can lead to the actual attainment of a true or epistemic account of reality.

In either case, Plato is well aware of the obstacles faced by any serious attempt to reform human cognition. Most human beings, in order to perform the intellectual and practical tasks usually required of them, tend to embrace an automatic, non-philosophical standpoint, and to stand by it. For a good illustration of this, one must simply recall the reactions of Socrates' interlocutors throughout many of his enquiries. At first, the discovery that the most basic human concepts, when isolated and looked at more carefully, are in fact dogmatic and unclear is a natural cause of perplexity. But as the dialogues unfold, this initial perplexity gives way to exasperation, and Socrates' charges are resisted with increasing vehemence. It rapidly becomes clear that what is

taking place is not a disinterested discussion about knowledge and truth, but a struggle between two conflicting modes of existence. Psychologically, by accusing his interlocutors of not knowing what they think they know (and sometimes claim to be an authority on), Socrates strikes a blow at their self-esteem; philosophically, though, his attack is much more serious, for it aims at the very core of their usual cognitive perspective, challenging their ability to think, choose and act the way they do.

To borrow another of Plato's images, human life, in its habitual form, is a 'life of slumber', and Socrates' role is to wake human beings from their usual 'nap'[230]. Like most sleepers, however, men and women are comfortably dreaming and do not wish to be disturbed. Moreover, even when they do consent to being woken, their alertness is usually short-lived. This natural incapacity to stay awake — like the prisoner's incapacity to endure the brightness of the sun, in the allegory of the cave, or the fishes' incapacity to breathe outside water, in the *Phaedo*'s cosmological myth[231] — speaks to the paradox underlying Socrates' critical efforts: despite the seriousness and the importance of his criticism, its grip upon us is less strong and less enduring than the grip of our usual standpoint, and although it is possible to live our entire life 'slumbering' and dreaming, the same does not seem to apply to Socrates' alternative.

In Socrates' own words, once we agree to join him in his enquiries, 'one of two things will happen: either we shall find what we are after, or we shall be less inclined to think we know what we do not know at all; and surely even that would be a recompense not to be despised.'[232] But it is precisely this latter claim that we tend to disagree with. If the contradictions underlying our usual outlook on reality were perceived or lived as an inconvenience, the acknowledgement of our ignorance would indeed be a valuable recompense. This is not the case, however, because our usual commitment to clarity and truth is not very serious. When faced with Socrates' dilemma, we tend to choose a third and less complicated alternative: if we can spend our entire life dreaming, and if doing so is more comfortable and more practical than to try and wake up, why is philosophy, as Plato understands it, really necessary?

[230] *Ap* 31a
[231] See above fn. 221
[232] *Th* 187c1-3

The debate concerning the need for a philosophical reform of human cognition leads to a second and even thornier debate, concerning the actual possibility of such a reform: if the doxastic foundation of human beliefs is really as ubiquitous as Plato suggests, there seems to be no way of guaranteeing, beyond a reasonable doubt, that philosophy's critical charges are not themselves doxastic, or that the transformation advocated by Socrates is not simply the more sophisticated dream of a more sophisticated slumberer. This suspicion calls into question not only the epistemic ideal pursued by philosophy, but also the freedom it is supposed to provide: if it turned out to be justified, the deliverance envisaged in Plato's dialogues would amount to a new form of imprisonment, open to new forms of criticism.

5.1. *The need for deliverance*

The issues of the need for a philosophical reform of human cognition and the possibility of such a reform have haunted, in a more or less explicit fashion, every serious post-Platonic attempt at a definition of philosophy. In modern times, David Hume addressed these issues with remarkable directness, in an effort to determine the true role played by philosophy in human life.[233] Just like Plato, Hume highlighted the basic contradiction between the cognitive self-confidence enjoyed by every human being and the severe limitations of his or her outlook on reality. He pointed out, like Socrates, that the usual success of one's practical achievements is dependent on a considerable degree of ignorance and inattention, which is hardly ever identified or called into question.

Generally speaking, Hume's philosophical method is not very different from Plato's. In both cases, philosophy's first task is to probe the edifice of human reason, to isolate its most basic components and to question their meaning and intelligibility. Zooming in on elementary notions such as identity, difference, space, time, cause or effect, the philosopher brings out their intrinsic contradictions and challenges thereby their usual familiarity. Moreover, insofar as these simple notions provide the basis for all other ideas

[233] See notably *A Treatise of Human Nature*, ed. cit., 171ff.

and judgements, the philosopher's criticism ends up contaminating the whole fabric of human reason, as well as every dimension of human life.

Hume's description of the practical and intellectual turmoil generated by this kind of criticism is reminiscent of Alcibiades' description of the effects produced by Socrates' discourses:

> The *intense* view of these manifold contradictions and imperfections in human reason has so wrought upon me, and heated my brain, that I am ready to reject all belief and reasoning, and can look upon no opinion even as more probable or likely than another. Where am I, or what? From what causes do I derive my existence, and to what condition shall I return? Whose favour shall I court, and whose anger must I dread? What beings surround me? and on whom have I any influence, or who have any influence on me? I am confounded with all these questions, and begin to fancy myself in the most deplorable condition imaginable, inviron'd with the deepest darkness, and utterly depriv'd of the use of every member and faculty.[234]

This confusion, however, is only temporary. After emphasizing the destructive or paralyzing power of philosophy, Hume acknowledges the lack of weight carried by philosophical propositions when compared to the automatic assumptions of natural reason. Although human consciousness is capable of questioning its own methods and composition in a seemingly uncompromising manner, these 'very refin'd and metaphysical [reflections] have little or no influence upon us.'[235] The peculiar kind of awareness promoted by philosophy, even when pursued in a serious and sustained manner, is never the rule, but always the exception.

In addition, Hume suggests that this indifference, which translates into the usual tendency to withdraw into natural reason, amounts to a defensive mechanism, which protects us from the disorientation bound to arise from a close scrutiny of our usual cognitive limitations:

[234] *Treatise,* ed. cit., 175. Cf. *Sy* 215e-216a.
[235] Ibid.

> Most fortunately it happens, that since reason is incapable of dispelling these clouds, nature herself suffices to that purpose, and cures me of this philosophical melancholy and delirium, either by relaxing this bent of mind, or by some avocation, and lively impression of my senses, which obliterate all these chimeras.[236]

In response to this paradox, Hume discusses the issue of the need for philosophy by asking whether one should oppose reason's natural 'slumber', and reject 'all the trivial suggestions of the fancy'[237], or whether it is possible to quash one's philosophical aspirations by determining, once and for all, that 'no refined or elaborate reasoning is ever to be received.'[238] Both of these strategies, however, turn out to be problematic. The first one, if pursued in a serious manner, might lead from 'melancholy' and 'delirium' to a complete existential paralysis: by criticizing and rejecting reason's basic cognitive alphabet, one would be left with a radical and self-destructive scepticism, which might end up compromising one's very ability to judge and act. The second strategy, albeit more compatible with the practical demands of everyday life, might also prove untenable. If pursued in a serious manner, it might lead to a complete break with objectivity and coherence, and to the endorsement of a complete logical relativism. By 'cutting off entirely all science and philosophy', one would 'proceed upon one singular quality of the imagination, and by a parity of reason … embrace all of them.'[239] In the absence of a fixed criterion with which to distinguish true beliefs from false ones, reason would be reduced to self-confidence and caprice, and even the wildest absurdities could come to be regarded as true or justified.

We are left, then, with a 'very dangerous dilemma, which-ever way we answer it.'[240] The choice between ignoring the contradictions unveiled by philosophy and accepting their consequences comes down, in Hume's words, to the choice between 'a false reason and none at all.'[241] These alternatives are very different, but both entail a dangerous or

[236] Ibid. Hume's definition of philosophy as a form of *delirium* echoes Plato's definition of philosophy as a form of μανία. In both cases, what is at stake is a deviation from consciousness' *normal* or *standard* condition, prescribed by 'sound common sense'.

[237] *Treatise,* ed. cit., 174

[238] Ibid.

[239] Ibid.

[240] Ibid.

[241] Ibid.

debilitating form of ignorance, liable to compromise the viability of our usual mode of cognition. This danger, however, is only truly real when these attitudes are pursued for their own sake, in a radical or unrelenting manner. But this is hardly ever the case: human consciousness is always divided between objectivity and self-interest, bound to heed and to harmonize both of these concerns. With regard to the pursuit of philosophy and science, Hume highlights the laziness and short-sightedness of our usual standpoint, always ready to dispel the 'clouds' of self-criticism in favour of more efficient or rewarding alternatives. With regard to the possibility of a complete emancipation from philosophy and science, he highlights the intrinsically normative nature of our usual standpoint. If our concern for coherence and objectivity were entirely ignored, the very distinction between truth and falsehood would end up losing its meaning, along with all other forms of judgement and valuation. And even if this attitude were somehow tenable, it would nonetheless entail a fundamental contradiction, for the decision to break free from philosophy and science would itself be the result of a 'refined reasoning', of the very kind consciousness would be set on repudiating. In this respect, Hume is also thoroughly Platonic: on his view, truthfulness and coherence are defining features of even the least scientific or demanding cognitive perspective; even the crudest or least scientific beliefs are still in some sense 'refined', albeit sometimes in the lowest of degrees.

This basic correlation is already conveyed, with equal force, by Plato's usual definition of philosophy. In his writings, the words φιλοσοφία and φιλοσοφεῖν stand for an attachment (φιλία) to 'wisdom' (σοφία), that is, to the truthfulness and clarity of one's ideas and judgements. Generally speaking, to be σοφός is to be able to cognize reality as it really is, in a truly clear and transparent way, with no distortions or limitations. But although our idea of wisdom is relatively clear, our attachment to it is liable to assume very different forms. On the one hand, as we have seen in the previous chapters, none of us are indifferent to the truthfulness and to the clarity of our judgements because none of us can live or act without a specific version of how things are (and were, and will be). Whether guided by the most demanding or by the most liberal cognitive standard, we are always attached to a complex system of assumptions, whose content and articulation conditions every aspect of our life. In this general sense, all of us

are already, by definition, 'followers of wisdom' (φιλοσοφοῦντες).[242] On the other hand, however, we are usually only philosophers *in principle*. As we have also seen, our attachment to the truthfulness of our beliefs, albeit philosophical in kind, is usually hasty and opportunistic. Truth and wisdom *as such* are not usually a direct concern of ours, and only become so in the face of practical obstacles or limitations. Whenever our web of assumptions falters, or is called into question, our concern with truth is rendered explicit and intensified. Faced with an unknown or unclear object or situation, we are forced to look for its meaning, or to question our previous idea of it. But this moment of indecision is usually brief and superficial: it only lasts until a new set of beliefs is brought in, like a band-aid, to correct the problem and re-establish the system's integrity.[243]

Given this complex picture, Socrates' enquiries are never about the simple contrast between knowledge and ignorance, or truth and error. Through his arguments, myths and allegories, he sets forth an entire cognitive scale, ranging from the lowest degree of philosophical awareness, or attachment — the lowest degree of φιλεῖν, or φιλοσοφεῖν — to the highest degree of philosophical awareness, or attachment — the highest degree of φιλεῖν, or φιλοσοφεῖν. The interval between these two limits is the native realm of human life, and Socrates' habitual task is to explore its different regions, stages or altitudes.

As the dialogues usually suggest, our position within this interval is usually low and stationary. Since most of us are 'at ease' or 'at home' in our daily environment, focused on concrete tasks and achievements, the philosophical reassessment of our ideas and judgements feels unnecessary and inconvenient. However, as Socrates repeatedly points out, the co-implication of the scale's upper and lower limits is valid throughout the entire scale. Wherever we happen to be placed, our life is always burdened with a 'recollection' of true knowledge: it is a 'winged' life, an arrow 'aimed at truth', a 'heavenly plant'

[242] *Sy* 204a8ff.

[243] At *Phi* 52a-b, when Socrates includes the pleasures of knowledge (μαθημάτων ἡδοναί) in the category of pure pleasures — that is, pleasures which do not result from the (conscious) fulfilment of a need or longing — he points out, rather uphemistically, that they 'are felt only by very few.' Whereas human beings usually long for knowledge or truth as a means of making up for a specific shortcoming (usually aroused by, or translated into, a practical encumbrance of some kind), the search for knowledge as such is much rarer, if not downright illusory.

growing upside down, with its roots in the heights and its branches down on earth.[244] Accordingly, to ask about the need for a philosophical reform of human life is not to ask about the need for something entirely new or unknown to us. Philosophy stands for an attachment to truth that is already present, albeit in an incipient or undeveloped way, in all of our beliefs, choices and decisions. We are already situated — whether we want to or not — within a global scale whose summit is true knowledge, which means we are already implicated — whether we want to or not — in a vital quest for true knowledge. The real choice, then, is either to stop halfway and settle for spurious versions of true ideas and judgements or to embark in the ascending journey depicted in Plato's dialogues.

The first of these alternatives is certainly the easiest one. However, it may also prove the riskiest one, or even the most contrary to our own interests. Inasmuch as our usual ignorance extends to all of our ideas and beliefs, it also affects our most important choices and decisions. Therefore, by settling for a spurious or doxastic account of reality, we may in fact be placing our whole life, our whole happiness and our whole future on an equally spurious or doxastic foundation. It is for this reason, more than any other, that Plato insists on the vital importance of philosophy and science. In his eyes, the pursuit of a true or scientific account of reality is not a choice among others, but the real vocation of human life. All other activities and projects presuppose a specific assessment of what is best, or most beneficial, or most important in life, but are unable to vouch for its actual truth. Grounded in mere presuppositions, automatically assumed to be true, they are inherently exposed to the possibility of error. Philosophy, on the other hand, precisely because it rejects all presuppositions, is the only possible means of arriving at the truth about what really matters in life. And although this acknowledgement is usually not compelling enough to turn all of us into full-time philosophers, it is certainly worth taking into account.

[244] *Ti* 90a7-b2

5.2. *The possibility of deliverance*

Philosophy, in its strict or stronger sense, is singled out in Plato's dialogues as the highest and most important of human activities. Only through philosophy, it is said, are men and women able to recognize the misconceptions that plague their usual outlook on reality. And only through philosophy can they hope to leave their ignorance behind and find the truth about themselves and the surrounding world. But although the former claim is demonstrated, time and again, during Socrates' enquiries, the latter claim is much less obvious. For it is one thing to acknowledge that one is ignorant or imprisoned and quite another to break free from ignorance and attain a true or scientific account of reality.

According to Socrates, when men and women 'take refuge in philosophy', they 'escape from their former selves by becoming different.'[245] This transformation, however, is at first merely a negative transformation. Philosophy's first task, as Plato understands it, is not to replace one's usual beliefs with different beliefs, or one's usual understanding of reality with a new understanding of reality. Socrates' method is an immanent method, whose primary aim is simply to free human consciousness from its usual pretence of knowledge.

In the *Sophist*, as Socrates addresses the issue of παιδεία and compares different educational methods, he refers to this negative procedure as a kind of 'purge' or 'purification' (καθαρμός), designed to liberate the soul from ignorance and self-conceit.[246] Whereas more traditional forms of education consist in punishing or admonishing the ignorant for their mistaken beliefs and replacing them with different and truer ones, imparted by the teacher or the educator, the 'cathartic' method (καθαρτική) is based on a different strategy. The traditional method, Socrates argues, is not only authoritative, but also ineffective, since 'he who thinks himself wise would never

[245] *Th* 168a5-7
[246] See *So* 229a-230e

be willing to learn any of those things in which he believes he is clever.'[247] Although 'traditional' pupils may be led to revise some of their views and judgements about reality, their most in-grained assumptions are bound to prevail over the teacher's corrections. To counter this problem, the 'cathartic' method sets about to 'cast out the conceit of cleverness [ἐκβολὴ τῆς δόξης] in another way.'[248] Its proponents start by questioning a man about the things he thinks he knows or has experience of. In doing so, 'they easily discover that his opinions [τὰς δόξας] are like those of men who wander, and in their discussions they collect those opinions and compare them with one another, and by the comparison they show that they contradict one another about the same things, in relation to the same things and in respect to the same things.'[249] Through this immanent procedure, which is the one usually employed by Socrates, the ignorant are led to recognize by themselves the inconsistency of their beliefs. As a result, their 'conceit of cleverness' is abandoned and their mind is free to consider new ideas and beliefs.[250]

In the *Republic*'s seventh book, following the allegory of the cave, a different version of the same method is also put forth during a debate about education.[251] This time around, Socrates argues that teaching cannot amount simply to inserting knowledge into an empty soul, as if 'putting sight into the eyes of the blind.'[252] This method does not work because the soul is never truly empty or blind, but full of ideas and beliefs it is strongly attached to and unwilling to sacrifice. The teacher must focus, rather, on 'turning the soul around' and 'devising a way of making the organ see, which already has vision, but is not properly orientated or looking where it should.'[253] The definition of this new strategy leads Socrates and Glaucon into a longer discussion about the role of education within the state and the kinds of subjects that should be taught to its leaders. Yet after examining the relative merits of arithmetic, geometry, astronomy and music, Socrates returns to his initial argument: all of these subjects are mere preludes or

[247] *So* 230a6-8
[248] *So* 230b1-2
[249] *So* 168b5-8
[250] In doing so, 'they grow angry with themselves and gentle towards others, and this is the way in which they are freed from their high and obstinate opinions about themselves.' (*So* 230b8-c2) Note the striking contrast between the docility of these pupils and the murderous hostility of Plato's cavemen.
[251] See esp. *Re* 518a-519b and 521c-535a
[252] *Re* 518c1-2
[253] *Re* 518d5-7

preparations to the true and final task of education, which is to enable the soul to 'grasp methodically in every case what each individual thing is in and of itself.'[254] And this final and most basic of insights can only be attained, if at all, through a negative or immanent educational method, now properly called the dialectical method (ἡ διαλεκτική μέθοδος).

Unlike other forms of education, designed to replace ignorance with knowledge, dialectics aims to replace one's alleged knowledge with a higher form of ignorance, and thereby clear the way for the acquisition of real knowledge. Whereas all other τέχναι and μαθήματα — including, in varying degrees, arithmetic, geometry, astronomy and music — build their conclusions on a pre-existing substratum of beliefs, concerns, and desires, dialectics means to question the entire edifice of human cognition, down to its most profound sources. The dialectical method leads consciousness to turn inwards, to reason's internal structure and to its intrinsic fragility. The dialectician isolates criticizes, and removes one by one the different 'hypotheses' (ὑποθέσεις) on which one's habitual perspective is made to rest, and 'proceeds in this way to the actual first principle [ἐπ' αὐτὴν τὴν ἀρχὴν] in order to be securely based.'[255] To describe this procedure, Socrates evokes an Orphic image that is once again linked to the theme of altitude and elevation: whereas the 'eye of the soul' is usually 'buried in a kind of barbaric filth' (βόρβορος βαρβαρικός), dialectics 'quietly draws and leads it upwards.'[256]

One of the key elements at play in this passage is the notion of ὑπόθεσις, whose definition, in this context, is very different from that of the modern term 'hypothesis'. Instead of a thesis hazarded about something, admitting to a certain degree of doubt or uncertainty, the word ὑπόθεσις, just like the word δόξα, refers to a thesis that is not lived or experienced as such, but as a sure fact.[257] Its meaning is thus closer to that of the

[254] *Re* 533b2-3

[255] *Re* 533c7-d1

[256] *Re* 533d1-3. The same image comes up in *Re* 363d7-8.

[257] Although the word ὑπόθεσις is often used, throughout Plato's corpus, in the modern sense of 'hypothesis', or 'thesis', or 'supposition', etc. (see, for ex., *Eu* 11c5; *Go* 454c4; *La* 743c5, 812a4; *Me* 86e, 87a2, 87d3, 89c3; *Pa* 127d7, 128d5, 136a, 142b-c; *Phd* 92d6; *So* 244c4; *Th* 183b3), its meaning in the *Republic*'s sixth and seventh books is closer to that of 'presupposition', or δόξα, in the Platonic sense of the word. For further reading on the Platonic usage of ὑπόθεσις, see, for ex., CARVALHO, *Μέθοδος e ὑπόθεσις*, ed. cit., esp. 39-52; WOLFSDORF, *The Method ἐξ ὑποθέσεως at 'Meno' 86e1-87d8*, *ed. cit.*, esp. 39-52; or IDEM, *Trials of Reason*, ed. cit., 157-161.

words 'assumption' or 'presupposition'[258]. Moreover, like Plato's δόξαι, ὑποθέσεις are not episodic, but universal. The soul does not have a few presuppositions about a few specific objects or events: it is 'buried', from top to bottom, in assumptions. Here and elsewhere, Plato suggests that human cognition, prior to the cathartic intervention of philosophy, is entirely 'hypothetical', i.e. entirely presuppositive or doxastic. Finally, just like δόξαι, ὑποθέσεις are not clear or self-evident, but ambiguous and opaque. Although they are usually perceived as self-evident, the aim of dialectics is precisely to reveal their inner contradictions.

The philosophical re-education proposed by Socrates consists in recognizing one's cognitive assumptions *as* assumptions; and what is more, as part of a ubiquitous and ultimately unintelligible web of assumptions. The dialectical method's main goal is thus to identify and suspend all of reason's ὑποθέσεις, in search of a truly *unhypothetical* principle, on which to ground a new and truly scientific account of reality. This achievement, however, is far from simple or self-evident, not only because human consciousness, in its usual condition, is utterly unable to anticipate what the discovery of such a principle would amount to, but also because the journey leading up to it is anything but brief and straightforward.

Let us start from the ending. Assuming such a journey were possible, and all of one's assumptions were indeed exposed and suspended, what would await one on the other side? Different scenarios spring to mind. Firstly, if reason's doxastic or hypothetic web really is ubiquitous, Plato's τὸ ἀνυπόθετον is probably nothing more than an empty place. If philosophy's role is purely negative, it is fair to assume that its ultimate truth is not a concrete or positive truth, but merely the *untruth* of human reason. If so, its ultimate reward is not pure knowledge, but pure perplexity, or pure wonder.

Secondly, this absolute negativity might somehow bring into view a new and *positive* reality. In this case, the *emptiness* unveiled by the dialectician would be instantly revealed as *fullness*, and indeed as the *absolute fullness* of ἐπιστήμη. Since what 'blocks' or hinders the 'eye of the soul' is the 'filth' of its own assumptions, their removal would allow for a clear vision of reality, and an instant apprehension of 'what each individual thing is in

[258] A translation suggested by the word's etymology, and preserved, for example, in the German word *Voraussetzung*.

and of itself.'

Thirdly, as a variation of this second alternative, it might also be argued that Plato's dialectical method is but one half — the negative half — of his scientific project. Accordingly, once the philosophical or dialectical chapter were over, a new, properly epistemic chapter would ensue, in which cognition would no longer be filtered by the distorting lens of δοξάζειν. This new world would still have to be known and explored, but reality would now be a fully intelligible place, where one's beliefs would immediately coincide with things in themselves. Following the image offered in the allegory of the cave, this would mean that the prisoners' underground ascent would be followed by a new ascent, this time in broad daylight. Instead of an underground path, leading to the world they already believed to inhabit, they would now have to climb a mountain or a mountainous path, leading to a new and truly positive summit.

Which of these solutions, if any, is the right one? What is the real outcome of Socrates' dialectical method? Instead of offering a direct answer to this question, Plato highlights, in different ways, the self-contradiction inherent to all of the previous scenarios. Since the aim of Socrates' method is a complete transformation of human cognition, its results cannot be known prior to the transformation itself. In the absence of a fully fledged philosophical standpoint, the anticipation of what it would amount to can only be entertained ἐξ ὑποθέσεως, that is, by means of the very assumptions consciousness is supposed to be leaving behind.[259]

The final answer to Socrates' riddle can only be revealed, therefore, at the very edge of δοξάζειν, when all of one's assumptions have been successfully identified and called into question. But before such a turning point can even be considered, a closer and more concrete problem must be grappled with. The uncertainty regarding philosophy's ultimate destination is preceded by the uncertainty regarding the actual possibility of getting there — or, indeed, of ever coming close. Since our usual standpoint is

[259] Indeed, not even Socrates, the philosopher κατ' ἐξοχήν, seems to have reached a fully fledged philosophical standpoint: on the day of his death, with a whole life of philosophy behind him, he still claims to be 'pursuing' or 'practising' (ἐπιτηδεύειν) philosophy (see *Phd* 64a6). Like a regular student, he reviews his doctrines and tests the soundness of his arguments, debates and refutes as though there were still much more to learn, correct and improve. His understanding of τὸ ἀνυπόθετον is still a thoroughly negative understanding, filtered by the distorting lens of δοξάζειν.

continually grounded in a global account of reality, the number of concepts and beliefs that come into play in each act of cognition is virtually endless. Although this multiplicity is not apparent when we consider a given object or a given person, for example, the very idea of an object or a person is intrinsically dependent on a whole army of silent assumptions, without which objects and persons would not be thinkable. And these assumptions relate not only to basic notions such as existence and inexistence, selfsameness and difference, truth and falsehood, number and size, shape and colour, etc., but also to the definition of our interests and life-goals, of what is good and evil, beneficial and harmful, worth pursuing and avoiding, etc. Consequently, a complete dialectical reversal of human cognition would require a complete survey of all of these silent elements. And even if such a titanic undertaking were somehow possible, the global 'purge' envisaged by Socrates would also require a sure knowledge of which and how many assumptions are implicated in each of our thoughts, a clear understanding of how they change over time and a clear awareness of how they relate to one another.

As if this were not enough, a second major difficulty must also be considered. As Plato recurrently points out, one of the main peculiarities of human beliefs or theses about reality is their habit of concealing their thetical nature and posing as actual facts. In the allegory of the cave, this point is particularly well illustrated: as the prisoner breaks his chains and starts his ascent, he believes to have found the knowledge he was initially denied. However, the objects paraded above the screen, like the shadows projected on the wall, are eventually revealed as mere images of real things, whose actual content lies elsewhere. This disillusionment is then repeated when the prisoner steps outside the cave and looks at the shadows of real things, and again when he turns to the reflections on water. In all of these situations, what seemed real and reliable turned out to be false and deceptive.

The same kind of deception applies to Socrates' dialectical purge. Given the inherently doxastic nature of human cognition, there seems to be no way of guaranteeing, beyond a reasonable doubt, that the assumptions one has singled out were indeed suspended, and not simply replaced by new assumptions. For even when a given belief is revealed to be spurious and unintelligible, this very revelation is supported by a conviction regarding its own truthfulness and intelligibility, which is itself open to being denounced as spurious and unintelligible. And this denouncement is liable to be

redirected, again and again, at its own result, without ever reaching a final conclusion.

Once again, Hume also called attention to this problem. In line with Plato's own suspicions, he suggested that there is ultimately no solid criterion with which to distinguish between a successful reform of one's usual standpoint, capable of exposing and eliminating all of the silent presuppositions on which it is grounded, and the covert action of yet another subset of silent presuppositions:

> Can I be sure, that in leaving all established opinions I am following truth; and by what criterion shall I distinguish her, even if fortune should at last guide me on her footsteps? After the most accurate and exact of my reasonings, I can give no reason why I should assent to it; and feel nothing but a *strong* propensity to consider objects *strongly* in that view, under which they appear to me. […] Without this quality, by which the mind enlivens some ideas beyond others (which seemingly is so trivial, and so little founded on reason) we could never assent to any argument.[260]

To be sure, the effort of questioning the cogency of our ideas and beliefs is very different from our usual willingness to accept them. Nevertheless, Hume argues, the conviction with which we undertake the former task is not substantially different from the one presiding over the latter. Even though we may commit ourselves to Socrates' dialectical method, and attempt to expose the groundlessness of all of our assumptions, we might be doing so as the result of a further assumption, whose content was left uncriticized. Hume is stating, in different words, what Plato had already suggested in his subterranean image: however vigorously we try to rid ourselves of the presuppositions underlying our usual cognitive perspective, there might still be a presuppositive nucleus at work within us, whose action cannot be suspended.

Finally, adding to the two previous difficulties, one last problem is worthy of mention. Apart from the sheer number and complexity of our cognitive assumptions, and from the intrinsically deceptive nature of our usual standpoint, our philosophical efforts are also at odds with the ephemeral nature of their results. Even if we were able to single out and suspend all of the 'secrets judgments of common reason'[261], such a task would

[260] *Treatise*, ed. cit., 172
[261] KANT, *AA* XV, 180

immediately prove out-dated. And for two main reasons: firstly, because the internal composition of our usual cognitive perspective is never the same — with time and circumstances, our account of reality is continually revised and readjusted, in response to new priorities and demands; secondly, because even those assumptions we have successfully isolated and suspended are never permanently gone — insofar as the circumstances that led us to adopt them in the first place go on being felt (namely, the need for practical orientation, the sway of bodily needs and desires, the pursuit of certain aspirations and life goals, etc.), they are bound to slip back into consciousness and win back their former influence.

Considering all of these difficulties, the possibility of a complete deliverance from error and contradiction cannot be taken for granted. Given the peculiar nature of our usual understanding of reality, philosophy can only be construed as an open process, with no clear end in sight. The philosopher's labour, as Plato understands it, is similar to the labour of Penelope: just as her web was constantly unwoven and started anew, so too philosophy's efforts tend to lose their force and must be revived and reasserted.[262] The nature of philosophical conquests is such that none is ever truly guaranteed: they are not fixed or final, but part of a wider on-going movement, made of advances and retreats, of victories and setbacks. As Seneca once wrote, 'when we can never prove whether we really know a thing, we must always be learning it.'[263] And that is precisely what Plato and Socrates set out to do. Rather than replacing human ignorance with some version of true knowledge, they set out to denounce and resist its permanent renewal, through an equally persistent and renewed educational effort.[264]

[262] See *Phd* 84a. Cf. *Re* 532d2-5.

[263] *Epistles*, ed. cit., 191

[264] Of course there is a different and (presumably) final solution to the philosopher's predicament, which is the deliverance brought about by physical death. Since human blindness, according to Plato, consists precisely in a pretence of lucidity, an absolutely lucid standpoint can only emerge once our usual belief in the clarity of our thoughts has been completely abandoned. And that can only happen once we cease to have a body, and bodily needs, and practical concerns, and the need for practical orientation, and the desire to go on living, and so on — in other words, *when we cease to be alive.* 'When we are dead we are likely to possess the wisdom which we desire and claim to be enamoured of, but not while we live. For, if pure knowledge is impossible while the body is with us, one of two things must follow, either it cannot be acquired at all or only when we are dead.' (*Phd* 66e2-4)

5.3. *Plato and Hegel*

All things considered, Plato's philosophy is about questions, rather than answers. His writings are the very opposite of neatly constructed treatises, designed to offer final solutions to strictly defined issues. Although their main theme is truth and true knowledge, their main focus is on doubt, illusion and contradiction. Throughout the dialogues, human life is described as the meeting point of two contrary and seemingly irreconcilable movements. This basic conflict is conveyed by the contrast between freedom and slavery, but also soul and body, light and darkness, dream and wakefulness or life and death. In all of these examples, Plato does not simply favour one term in detriment of the other, nor does he simply argue that one is right and the other wrong. He points out, rather, that both terms are essential and usually found together. Human life *is itself* this contradiction and human beings are what they are by partaking simultaneously of both of these elements. Men and women are originally hybrid creatures, caught halfway between reason and unreason, and their time on earth — to borrow the *Republic*'s haunting formulation — resembles a 'day shrouded in darkness' (νυκτερινὴ ἡμέρα).[265]

To be sure, Plato advocates a transformation of this state of affairs. Philosophy, as he understands it, is charged precisely with 'bringing the soul round' from its 'nocturnal day' to a 'true day' (ἀληθινὴ ἡμέρα)[266], or with elevating human cognition from the realm of δοξάζειν to the realm of ἐπίστασθαι, or ἐπιστήμη. But although Plato's philosophical project is directed at the attainment of a true or scientific standpoint, the emphasis is placed on the possibility of such an attainment, rather than on its actual concretization. Plato's dialogues are not mere guides or recipes for the acquisition of true knowledge, but reflections about whether and to what extent true knowledge can be reached. And although Socrates anticipates and exemplifies the kind of criticism required for a complete dialectical reversal of human cognition, his enquiries are only mere samples, illustrations, isolated forays into a wider and wholly unknown territory. They

[265] *Re* 521c6
[266] *Re* 521c6-7

reveal the different ways in which human knowledge falls short of true knowledge, but they also highlight the difficulty of changing this situation.

Historically, Plato's critique of human cognition is the starting point of a long philosophical journey, extending all the way to the Modern and contemporary traditions. His focus on the intrinsic limitations of human concepts and judgements opened the door to a variety of critical projects, centred on different premises, aims and methodologies. Among these projects, Hegel's idea of a dialectical re-education of natural consciousness can be read, in my view, as a direct response to the philosophical challenge issued throughout Socrates' enquiries.

Even though Plato and Hegel belong to very different worlds and philosophical traditions, their works are focused on the same general task, namely the elevation of human consciousness from an uneducated, or doxastic standpoint, to an educated, or scientific standpoint. In both cases, the plea for such an elevation follows from the acknowledgement of a basic disagreement between consciousness' usual understanding of reality and reality itself. For Plato and Hegel, our usual standpoint is grounded in a fundamental contradiction: although we are thoroughly ignorant or uneducated beings, we usually believe otherwise; and, moreover, although our ignorance limits our freedom to choose and act according to our interests, it is not usually lived or felt as a limitation. The assumption that we are more knowledgeable than we actually are entails the assumption that we are freer, more independent and more powerful than we actually are.

Apart from sharing the same diagnosis, and taking on the same general task, Plato and Hegel adopt a similar methodology. Since consciousness is never a *tabula rasa*, but a vast and potentially boundless repository of beliefs, and since it is usually attached to these beliefs and unwilling to let them go, the elevation to the standpoint of science cannot be achieved simply by replacing them with new beliefs. The change must come about in an immanent way, through the explicit acknowledgement of one's usual ignorance. Only when self-confidence gives way to self-criticism and perplexity is consciousness free to embrace the possibility of a truly different outlook on reality.

In Plato's case, this feat is to be accomplished, if at all, through the dialectical method discussed in the *Republic*, the *Sophist* and elsewhere. By collecting and comparing

one's usual ideas, and by realizing they 'contradict one another about the same things, in relation to the same things and in respect to the same things'[267], one is led to abandon one's usual pretence of knowledge and to embrace a new, 'enlightened' ignorance, which may lead in turn to the attainment of true knowledge. In Hegel's case, the elevation to science also rests on a dialectical enquiry, whereby consciousness is also led to identify, confront and overcome its own cognitive limitations. This basic methodological approach is featured, in varying forms, in most of Hegel's works, from his early essays to his later systematic writings, but its first thorough enactment is found in the *Phenomenology of Spirit*. Unlike the later chapters of Hegel's so-called system of philosophy, outlined in the *Encyclopaedia of the Philosophical Sciences*, the *Phenomenology* is a propaedeutic work, charged with an explicitly educational task. It is targeted at everyday consciousness and at the immediate beliefs and presuppositions that shape everyday life.

Like Socrates' dialectical method, Hegel's phenomenological method is designed to reveal what appears close, self-evident and familiar as distant, obscure and strange. In the preface to the *Phenomenology*, Hegel conveys this idea in thoroughly Platonic terms:

> Quite generally, the familiar, just because it is *familiar* [*bekannt*], is not *cognitively understood* [*erkannt*]. The commonest way in which we deceive ourselves or others about understanding is by assuming something as familiar, and accepting it on that account; with all its pros and cons, such knowing never gets anywhere, and it knows not why. Subject and object, God, Nature, Understanding, sensibility, and so on, are uncritically taken for granted as familiar, established as valid, and made into fixed points for starting and stopping. While these remain unmoved, the knowing activity goes back and forth between them, thus moving only on their surface.[268]

If reason is to be more than a superficial and unfounded habit, consciousness must turn inwards and reassess the solidity of its usual account of reality. Moreover, this procedure must be extended to every one of its beliefs, so as to clear the way for the future acquisition of true knowledge. Only 'the scepticism that is directed against the

[267] *So* 230b7-8

[268] *HW* III, 35; *PS*, 18. Cf. *HW* XVIII, 39: 'Es ist … das Eigentümliche der Philosophie, das zu untersuchen, was man sonst für bekannt hält. Was man unbesehen handhabt und gebraucht, womit man sich im Leben herumhilft, ist gerade das Unbekannte, wenn man nicht philosophisch gebildet ist.'

whole range of phenomenal consciousness … renders the spirit for the first time competent to examine what truth is. For it brings about a state of despair about all the so-called natural ideas, thoughts and opinions … with which the consciousness that sets about the examination [of truth] straight away is still filled and hampered.'[269]

However, this task is by no means an easy task. Like Plato, Hegel is well aware of consciousness' inner complexity, of its usual dogmatism and unique aptitude for self-deception. Therefore, like Plato's dialectical method, Hegel's dialectical method is not limited to the simple contrast between truth and error, or reality and illusion. The *Phenomenology* sets forth an entire cognitive scale, ranging from natural knowledge to scientific knowledge, and documents consciousness' lengthy and laborious journey from the scale's lower limit all the way to the top. This long ascent, divided into a series of intermediate stages, evokes the prisoner's ascent in the allegory of the cave, equally divided into different intermediate stages.[270] What is at stake in both progressions is not the quantitative transition from a minimum to a maximum of knowledge, as though consciousness would merely accumulate new ideas and insights as it moves along. Both progressions consist in a succession of increasingly sophisticated *versions* of true knowledge, wrongly perceived as such. At each new stage, the prisoner and the phenomenological subject believe they have arrived at the end of their journey, but each new transition ends up proving them wrong: the acknowledgement that their new account of reality is merely an image or an appearance of reality gives rise to a new 'state of despair' and to the need to go on climbing.

The most Platonic aspect of the whole phenomenological progression is arguably Hegel's extreme, almost obsessive insistence on the tenacity of consciousness'

[269] *HW* III, 73; *PS*, 50. Regarding the translation of *erscheinendes Bewußtsein* as 'phenomenal knowledge', see above fn. 50. As for the 'despair' brought about by scepticism, the English translation is unable to capture Hegel's wordplay: *Verzweiflung* means despair, but it is also derived from the word *Zweifel*, which means 'doubt'. The term is also meant to evoke, therefore, the active movement of 'being led into a state of doubt'. I will return to this theme in the next section.

[270] See PHILONENKO, op. cit., 331: 'Platon suppose qu'"on" libère un prisonnier; ce dernier voudra s'élever à la lumière, mais plus il se rapprochera de la lumière, plus ses yeux habitués aux ombres le feront souffrir. En un sens, il est question d'une véritable *Phénoménologie de l'esprit*, car enfin ce "on" qui déchaîne l'âme doit lui-même être libre, et sa position par rapport au feu est juste l'inverse de celle du prisonnier.' The structure of Hegel's *Phenomenology* and the meaning of its different standpoints, or 'inner voices' — viz. the standpoint of the 'narrator' ('on', viz. 'we') versus the standpoint of natural consciousness ('es', or 'it') — will be discussed below, in section 6.3.

presuppositive core. In the *Phenomenology*'s own terms, as in Socrates' allegory, there can be no real knowledge as long as there remains within us but a single unverified assumption. Truth can only be envisaged when reason's entire repertoire of presuppositions — or shadows, or images, or ὑποθέσεις — has been successfully neutralized, and the *Phenomenology*'s lengthy progression is the detailed history of this struggle, extended to every corner of human cognition and to every dimension of human life. But here, where Plato and Hegel come closest to each other, is also where they part company. Although the phenomenological progression is grounded in an originally Platonic idea, Hegel's encyclopaedic fervour contrasts with Socrates' tentative, almost gestural enquiries. Whereas Plato's scientific project is above all an *interrogative* or *inconclusive* project, continually haunted by the shadows of failure and unfulfilment, Hegel's phenomenological project is an *affirmative* or *conclusive* project, directed towards a concrete and final destination.

'The development of philosophical science as science', Hegel writes, 'begins with Plato.'[271] But although the latter's standpoint is 'clearly defined and necessary, one cannot remain in it', for 'Reason makes higher demands.'[272] The entire history of philosophy can be interpreted, in Hegelian terms, as the long and fruitful development of Plato's critical diagnosis, culminating in Kant's 'return to the standpoint of Socrates', albeit 'on a higher level.'[273] Departing from this long *negative* lineage, Hegel takes on the *positive* task of solving Plato's riddle and freeing human consciousness from the fetters of ignorance and self-deceit. His work is infused with a unique and characteristic philosophical hybris, whose closest match is perhaps Nietzsche's own (anti-)Platonic attempt to reform human cognition. But whereas Nietzsche takes issue with the starting point of Plato's philosophy, and with dialectics in general, Hegel embraces the dialectical method, criticizing its ultimate lack of direction. In his eyes, the actual release from consciousness' usual cognitive bondage can only come about once dialectics is raised to a systematic form. In other words, if Plato's journey is to be successful, the different stops along the way cannot be regarded merely as successive or contingently related events. They must be

[271] *HW* XIX, 11
[272] *HW* XIX, 20
[273] *HW* XX, 329f.

revealed as the interconnected stages of a global dialectical progression, whose goal is as necessarily fixed for consciousness as the path leading towards it.

Part II

Hegel's Phenomenological Project

6.

HEGEL'S PHENOMENOLOGY OF SPIRIT

„Wie komm ich am besten den Berg hinan?"
Steig nur hinauf und denk nicht dran![274]

In line with Plato and the Platonic tradition, Hegel's conception of freedom is part of a wider philosophical debate concerning the nature, scope and truthfulness of human cognition. To be free is to be in control of one's life, which can only happen if one is able to identify one's usual existential setting, to interpret the objects and events one is faced with and to assess their value and their degree of importance. In other words, true freedom requires a true account of reality, whereupon one's judgements and choices can be safely grounded. According to Hegel, however, this is precisely what human beings usually lack. Human consciousness, in it habitual form, is ignorant and self-deluded. It is usually guided by a series of immediate assumptions, whose truthfulness is hardly ever called into question. In like manner, human life, in its habitual form, is not a free or independent life. True freedom can only be achieved through a global re-education of natural consciousness, or through the replacement of one's usual cognitive perspective with a true or scientific perspective.

This is the task laid out in the *Phenomenology of Spirit*. Hegel's first major opus is a highly ambitious work, with a unique structure and a wide thematic scope. It consists in

[274] *KSA* III, 356

a long succession of cognitive stages, leading from the most immediate form of cognition, which Hegel names 'sense certainty' (*sinnliche Gewißheit*), all the way to the standpoint of science, or 'absolute knowing' (*absolutes Wissen*). Throughout this succession, one is led to consider different standpoints, to embrace different versions of reality and to adopt different sets of concepts and ideas. Moreover, one is led to examine each different standpoint, to assess its reliability and to acknowledge its inner contradictions. Although each version of reality appears to coincide, at first, with reality itself, this coincidence is invariably proven wrong. As a result, each standpoint is replaced by a new one, and this procedure is repeated again and again, leading to new and increasingly sophisticated accounts of reality. Hegel delineates a complete educational programme, covering an extraordinarily vast array of concepts, subjects and situations. The complexity of his survey reflects the complexity of human life and the various cognitive directions available to human consciousness.

Despite this diversity, however, the progression follows a very specific plan. All of its stages are connected to each other and their meaning is determined by their position within the progression. Therefore, unlike many other philosophical texts, the *Phenomenology* cannot be broken down into a collection of arguments about a specific subject or set of subjects. Although its different sections are sometimes considered separately, in connection with concrete philosophical, scientific, political or aesthetic debates, their meaning can only be truly grasped in its original context. Every fractional reading, centred on this or that particular element, must start by situating it within the work's global itinerary and determining in what way and to what extent it is meant to forward Hegel's overall argument.

By the same token, the philosophical merits of any of the *Phenomenology*'s particular theses can only be fully appreciated in light of the work's global structure. Since each phenomenological stage is but an intermediate stop along a wider argumentative path, its specific conclusions are necessarily provisional. They are neither true conclusions, nor completely false ones: as the progression unfolds, each different stage is criticized, abandoned and replaced by a new one; at the same time, however, each is revealed as the necessary vehicle for attainment of new stages and new conclusions. Ultimately, then, their truth must not be looked for in themselves, but in the progression's final

destination. It hinges not on the solidity of a particular thesis or idea, but on the solidity of Hegel's entire methodological approach.

In what follows I will offer a brief overview of Hegel's phenomenological progression. The analysis will be divided into two main parts: throughout the present chapter, I will deal with the *Phenomenology* as a whole and discuss its aim, scope and methodology; in the following chapters, I will delve into the actual progression and examine some of its central moments. My aim, however, is not to offer a detailed reading of the entire *Phenomenology*, nor to sum up its main arguments, but to focus on the issues of freedom, power and knowledge. More specifically, I will attempt to highlight the different ways in which Hegel's text echoes Plato's critique of human freedom, the new elements it adds to the previous debate and the kind of deliverance it proposes.

6.1. *Truth and appearance*

The *Phenomenology of Spirit*, regardless of its unique methodology, is deeply indebted to the Platonic tradition. Like Plato, Hegel challenges the overall legitimacy of natural reason and highlights the need for a global reform of human cognition. He argues that the ideas and judgements underlying our usual understanding of reality are not accurate or straightforward, but inaccurate and unclear. Furthermore, like Plato, Hegel points out that this problem is not usually seen or experienced as such. Although we recognize the limited nature of our cognitive powers, we nonetheless assume our daily grasp of reality to be fundamentally adequate and reliable. In other words, although we admit to being ignorant of, or mistaken about, an endless variety of subjects, we consider ourselves competent enough to know most things as they really are.

The unwillingness or incapacity to question our beliefs in any serious or radical way is tied, once again, to the universal scope of our usual cognitive perspective. Like Plato, Hegel highlights the fact that human cognition is never about a specific part or region of reality, but always about the whole of it. Our knowledge is not confined to a series of islands, or enclaves, interspersed by gaps of complete ignorance. We are guided, at any given moment, by a global version of things, presupposed and reiterated by every new act of cognition. And what is more, this global account is not rigid, but plastic and versatile.

It can be expanded and detailed, adapted to new situations and requirements. Therefore, when faced with a new object or event, we are hardly ever lost in pure wonder. New ideas and contents are never entirely new because they are always perceived as extensions of a familiar cognitive landscape and adjusted to a familiar set of rules and assumptions. Their novelty is an internal variation of the same cognitive repertoire, whose global legitimacy is never truly questioned.

To borrow another Platonic image, our usual standpoint speaks a specific language, or employs a specific alphabet. Its apparent virtuosity derives from the ability to combine letters and syllables in a seemingly infinite number of ways. Yet all of our words and sentences derive from a finite group of letters. And these letters, when subjected to a closer inspection, are not as clear as we tend to assume. For Hegel, as for Plato, a true reform of human cognition depends on the acknowledgement of this double paradox. Even though our usual standpoint seems remarkably free and malleable, ready to cope with all kinds of situations, it is really self-enclosed and dogmatic, averse to difference and change. Once this contradiction is recognized, the universal enclosure of natural reason can be finally breached: if the world we are used to is not *the whole world*, but only a specific version of reality, it is liable to be replaced by other versions of reality; indeed, it might just be that our current cognitive language is not the language spoken by reality, or that our current cognitive alphabet is not the alphabet in which reality is actually written.

These preliminary remarks provide a sufficient basis for dispelling a misconception that is commonly associated with the *Phenomenology.* Contrary to what is sometimes assumed, Hegel's progression is not a quantitative or cumulative progression, leading from a minimum to a maximum of cognitive insight. The elevation of natural consciousness from the stage of 'sense certainty' to the stage of 'absolute knowing' does not amount to a simple increase in the broadness or the scope of one's cognitive outlook, as though the phenomenological subject would start by knowing very little about reality and would end up by knowing very much. Hegel's progression is not about completing an incomplete cognitive picture, or about adding cognitive layers to the same cognitive core. It documents a different and more profound transformation, or series of transformations, extending to the very notions of reality, identity and cognition.

Likewise, the elevation from the stage of 'sense certainty' to the stage of 'absolute knowing' does not amount to a simple increase in the clarity of one's cognitive outlook,

as though the phenomenological subject would start by seeing things in a vague or unclear way and would end up by seeing them very clearly. Once again, Hegel's progression is not about sharpening an obscure version of reality, or about adding detail to the same cognitive image. The transition from each stage to the next is fuelled by a different kind of transformation, with wider and more profound implications.

Both of these readings follow the aforementioned tendency, characteristic of our usual cognitive perspective, to construe difference and change in an immanent way. Using our standpoint as reference, we imagine a series of enhanced or amplified versions of our current knowledge, leading to a universal and absolutely transparent vision of reality. Yet this vision, however wide and illuminating, is still but a variation of our initial perspective, bound to the same cognitive idiom, based on the same cognitive alphabet. The 'absolute knowledge' it is thought to convey is merely the absolute version of a specific kind of knowledge, which is still as regional and dogmatic as before, and still affected by the structural limitations Hegel criticizes and aims to overcome.

As Marcel Proust famously wrote, 'the only true voyage … would be not to visit new landscapes, but to possess other eyes, to see the universe with the eyes of an other'[275]. This is the kind of voyage set out in the *Phenomenology*. With each new stage, what changes is not merely the scope or the clarity of what is seen, but the very act of seeing. Consciousness comes to see different things, but it comes to see them *differently*, with the aid of new eyes and a new cognitive alphabet. Instead of exploring and expanding its initial world, it is led into a new world, or a succession of new worlds, where nothing whatsoever is what it used to be. However, precisely because these new worlds are in no way similar to the initial one, they cannot be envisaged using the latter's cognitive resources. As long as consciousness holds on to its usual eyesight, or to its usual alphabet, it is incapable of actual change. It can only anticipate the kind of leap recommended by Hegel in purely negative terms, as a blind man anticipates the world's visual appearance.

But this is only part of what is at stake in Hegel's approach. Although he insists on the radical difference between each phenomenological perspective, he is equally adamant about the progression's global continuity. Indeed, apart from emphasizing the fact that

[275] Op. cit., III, 258

each different stage amounts to a whole world, or to a whole universe, shaped by its own concepts, rules and priorities, he also claims that all stages are nonetheless connected to each other — that they are all somehow linked to a single overarching principle, which lives on in every one of them and holds the key to their mutual articulation. The *Phenomenology*, he argues, is not a mere collection of worldviews, but an orderly and organic progression, brought about 'through one and the same principle having spontaneously assumed different shapes.'[276]

Following this claim, Hegel also argues that each phenomenological stage, because it is the product or expression of the same generating principle, is already somehow revealing of all the other stages. It already contains within itself, in an indirect or inexplicit manner, the entire progression, as well as its final result. And this also applies to the *Phenomenology*'s starting point: in Hegel's words, true knowledge is 'the circle that presupposes its end as its goal, having its end also as its beginning.'[277] Although natural consciousness, in its most basic form, is confined to a specific cognitive alphabet, it is nonetheless bound, as it were, by a higher cognitive calling; although it is unable, from its limited vantage point, to envisage other, more sophisticated modes of cognition, it is already burdened, *in posse*, with all of its future insights.

At first glance, these two sets of claims appear irreconcilable. How can each phenomenological perspective be *at the same time* a whole and a part, self-sufficient and dependent on something else? Both of these qualifications appear to entail the exclusion of the other. And yet if either one of them is favoured over the other, the whole progression, as Hegel understands it, is liable to be compromised. On the one hand, if each standpoint is truly independent or self-sufficient, consciousness is apt to remain eternally tied to the same cognitive alphabet. If every form of otherness is but a variation of the same cognitive repertoire, one's native standpoint is indeed inescapable, and no real progression can ever take place. On the other hand, if each stage is not a whole world, but a mere region of a wider cognitive standpoint, one is led back to the quantitative model criticized earlier. In this case, consciousness may indeed discover new landscapes, but it is bound to do so with the same pair of eyes. Just like before, difference

[276] *HW* III, 21; *PS*, 8
[277] *HW* III, 23; *PS*, 10

and otherness are mere variations of the same familiar alphabet, and no real progression can ever take place.

In the introduction, Hegel defines the *Phenomenology*'s entire progression as a *Realisierung des Begriffs*[278], that is, as the complete 'realization' of a single unifying 'notion'. As he proceeds to make clear, consciousness is always burdened with an idea or 'notion' of reality — nay, it *is itself* this idea or notion, which constitutes the primary source of human cognition. But whereas up until now I have used the terms 'standpoint', 'perspective', or 'outlook' interchangeably, in order to characterize consciousness' usual *version of reality*, there is an important difference between the latter and Hegel's specific definition of *Begriff*. Throughout the *Phenomenology*, he argues, consciousness is led to endorse different versions of reality — but not, strictly speaking, different 'notions', or *Begriffe*. In the beginning, in the end and in every intermediate stage, consciousness is guided by *the same basic notion*. Its understanding of it, however, is not always the same: it depends on its *degree of realization*, which is radically altered with each new transition.[279]

At the beginning of the progression, this overarching notion is fully 'unrealized'. In its most basic or immediate form, true knowledge is automatically taken for granted. One's ideas and judgements are thought to convey things *as they really are*, and one's version of reality is not lived or experienced as such, but as *reality itself*. This initial agreement, however, is only apparent: once consciousness is brought to the stand, so to speak, and confronted with its own ideas and judgements, it '[shows] itself to be only the Notion of knowledge, or in other words, not to be real knowledge.'[280] Contrary to what is assumed, then, our usual account of reality does not coincide with reality itself, nor is it

[278] *HW* III, 72; PS, 49

[279] Throughout the introduction, Hegel's use of the word *Begriff*, or notion, is twofold. In a more general sense, which is the one that I will favour to in the following pages, *Begriff* is the guiding principle of the entire phenomenological progression, and that which every different stage seeks (and fails) to capture. Accordingly, the *Phenomenology* as whole amounts to a *Realisierung des Begriffs*, i.e. to the elevation of natural consciousness to the standpoint of science. At the same time, however, Hegel also uses the term *Begriff* to designate the specific notions adopted by consciousness at each phenomenological stage. In this second sense, there is not only one *Begriff*, but different *Begriffe*, corresponding to different objects and different cognitive criteria. These two uses are nonetheless closely connected: the different notions adopted by consciousness amount to different *versions* of the same founding notion. They are pseudo-notions, as it were, whose ultimate truth can only be found at the end of the progression.

[280] Ibid.

able to convey things as they really are. But this realization is at first merely negative. Since our cognitive perspective is a global perspective, which mediates all of our thoughts and judgements, it extends *simultaneously* to our idea of reality and to reality itself. Therefore, we cannot replace our initial version of reality with a truer one without falling back on the same cognitive alphabet, and repeating thereby the very mistake we are trying to avoid.

In other words, we are trapped within a standpoint we know to be untrue, and unable to escape its contradictions. Aware of this problem, Plato devised a negative form of education, grounded in an immanent critique of human cognition. He argued that true knowledge can only be achieved through a complete refutation of one's current knowledge. Only by releasing oneself from every unwarranted assumption is one able to see reality with new eyes, and free to learn its true language. For Hegel, this negative approach is the starting point of a long critical tradition, which links Platonism to Stoicism and Scepticism, and whose chief methodological premise is still at work in Modern philosophy, and notably in Kant's critical philosophy. In the *Critique of Pure Reason* and elsewhere, Kant is also focused on the disagreement between consciousness' usual account of reality and reality itself. Moreover, he also devises an immanent form of criticism, equally motivated by the universal (or transcendental) scope of natural reason: since natural consciousness is confined to a global account of reality, it cannot step outside of it and assess the degree to which it coincides with reality itself. Things *in themselves* (*viz.* things as they are independently of one's coming to know them) can only be cognized indirectly, as that which remains after the acknowledgement and the (hypothetical) suspension of reason's global lens. This negative procedure is made possible through a detailed exposition of consciousness' internal composition, including all of the 'secrets judgments of common sense'. Once this exposition is completed, the critic can then subtract it to his or her usual account of reality, and deduce thereby reality's true nature.

This whole strategy can be clarified with the aid of a simple formula:

$$R + C = R'$$

in which R stands for *reality in itself*, C for the *ability to cognize reality* and R' for the joint result of these two variables, i.e. *cognized reality*.

Kant's critical project is grounded in an inversion of the way we tend to conceive the relationship between knowledge and reality: rather than assuming, as we normally do, that reality exists in its own right and cognition is somehow added to it, Kant points out that all reality is by definition *cognized reality*, and hence that the very idea of an object prior to any act of cognition is itself the correlate of an act of cognition. In other words, he highlights the transcendental scope of natural reason and the impossibility of attaining truth or science in any immediate or direct way. Instead, he proposes to reverse the operation set out in the first equation. Taking cognized reality as its starting point, the critical project seeks to isolate and cancel out the distorting effects of human cognition:

$$R' - C = R$$

According to Hegel, although Kant correctly emphasized the universality of natural consciousness and the need for an immanent critique of human cognition, his critical strategy is not critical enough. Despite his so-called 'Copernican revolution', he still regards cognition and reality as two separate things, opposed to one another. His entire critique is grounded in a dichotomical approach to human cognition, which presupposes the very difference it aims to overcome.

This point is made in the opening paragraphs of the introduction, where Hegel contrasts this critical approach with his own critical approach.[281] It should be noted, however, that this initial charge is not aimed solely at Kant or the critical tradition. Hegel blames critical philosophy for having failed to criticize a 'natural assumption' that is already at work in human beings' usual cognitive perspective. Namely, both the critical standpoint and our ordinary standpoint tend to view reason uncritically as either an instrument (*Werkzeug*) for 'getting hold of the truth' or the medium (*Mittel*) whereby the 'light of truth' reaches us. Whereas natural consciousness regards this medium as an absolutely transparent filter (R + C = R), the critical project acknowledges its distorting

[281] See *HW* III, 68f.

power, and proposes to correct it (R' - C = R). But although the critical model is undoubtedly more sophisticated than the natural one, both presuppose a fundamental divorce between reason and reality. In both cases, the true is defined as that which lies *beyond* cognition, and which the latter is supposed to reveal.

This approach, Hegel argues, is inherently contradictory. In the first case, if cognition is indeed an *instrument* used on reality, its use can either bring reality closer to us without altering its original nature — this is consciousness' ordinary or *uncritical* view — or reveal the world at the price of altering its shape — this is the philosopher's *critical* view. We may try to overcome this dilemma, as the philosopher does, by getting acquainted with the way in which this instrument works, but this effort is also doomed to failure: by removing 'from a reshaped thing what the instrument has done to it, the thing … becomes for us exactly what it was before.'[282]

In the second case, if cognition is the *medium* whereby reality is brought into view, it can either offer a fully transparent image of reality — this is consciousness' *uncritical* view — or reveal the world at the price of distorting its original appearance — this is the philosopher's *critical* view. Again, we may try to overcome this dilemma by determining how this medium filters the 'light of truth', but this effort is equally vain. As Hegel puts it, 'if by testing cognition, which we conceive of as a *medium*, we get to know the law of its refraction, it is again useless to subtract this from the end result. For it is not the refraction of the ray, but the ray itself whereby truth reaches us, that is cognition; and if this were removed, all that would be indicated would be a pure direction, or a blank space.'[283]

When the instrument or medium employed by human reason is removed, *everything else is removed*. Cognition and reality cannot be detached from one another because they are inherently reciprocal or dialectical determinations: every act of cognition is by definition directed at a given reality, and every reality is by definition the correlate of an act of cognition. Each is what it is *through the other*, which means that a serious quest for truth and true knowledge must consider both of them *simultaneously*.

[282] *HW* III, 69; *PS*, 46f.
[283] *HW* III, 69; *PS*, 47

The *Phenomenology*'s starting point is therefore the dialectical correlation between consciousness and its most immediate object, i.e. *das erscheinende Wissen* — *phenomenal* knowledge, or the knowledge that first *appears* to consciousness. Instead of deploying a specific kind of critique, or aiming at a specific standard of truth, Hegel rejects all of the previous assumptions and takes up this correlation in its own terms. No method or aim is needed, he argues, because consciousness is already 'the notion of itself'[284]: it contains within itself the key to its own transformation, and must simply be led to recognize as much. Whereas natural or unconscious beings are confined to a fixed mode of identity — that is, whereas they *are what they are*, and cannot be anything else — consciousness is the source of its own identity — it is what it *determines* itself to be, and can rise therefore to new and more sophisticated modes of identity.

To understand what this means, let us look closer at Hegel's characterization of human cognition. Whenever consciousness cognizes a given reality, it 'simultaneously *distinguishes* itself from something and at the same time *relates* itself to it.'[285] More precisely, it distinguishes reality *in itself*, or *truth*, from reality *for consciousness*, or its *version* of truth. Now the assessment of the truthfulness of one's knowledge usually consists in a comparison between one's *version of truth* and *truth itself*. However, since both of these elements fall within consciousness, this assessment amounts to a comparison of consciousness with itself. Both what is examined and the standard of examination are *for consciousness*, which means that the exam is directed simultaneously at the truthfulness of one's thoughts and at truth itself.

This immanent movement is described as follows:

> Consciousness knows *something*; this object is the essence or the *in-itself*; but it is also for consciousness the *in-itself*. This is where the ambiguity of this truth enters. We see that consciousness now has two objects: one is the first *in-itself*, the second is the *being-for-consciousness of this in-itself*. The latter appears at first sight to be merely the reflection of consciousness into itself, i.e. what consciousness has in mind is not an object, but only its knowledge of that first object. But ... the first object, in being

[284] *HW* III, 74: 'Das Bewußtsein aber ist für sich selbst sein *Begriff*, dadurch unmittelbar das Hinausgehen über das Beschränkte und, da ihm dies Beschränkte angehört, über sich selbst.'

[285] *HW* III, 76; *PS*, 52

known, is altered for consciousness; it ceases to be the in-itself, and becomes something that is the in-itself only for consciousness.[286]

Consciousness is simultaneously the consciousness of its object (the *in-itself*) and the consciousness of its knowledge of the object (the *being-for-consciousness of the in-itself*). If these moments are shown to disagree with one another — if the object *in itself* is shown to be merely *for consciousness* — one's usual standpoint is radically transformed.

At first, this transformation appears to be merely *negative*. If one's version of reality does not coincide with reality itself, the problem seems to lie in the former, rather than the latter. 'It seems that consciousness cannot, as it were, get behind the object as it exists *for consciousness* so as to examine what the object is *in itself*.'[287] As a result, consciousness is faced with the nothingness of its usual knowledge, and shut out from reality and truth. But this negative attitude fails to appreciate the dialectical nature of this transformation. Since consciousness is simultaneously consciousness of reality and consciousness of itself, the acknowledgement of its ignorance entails a transformation of the very criterion whereby knowledge and ignorance are measured. In other words, if 'consciousness thus finds that its knowledge does not correspond to its object, the object itself does no stand the test'[288]: what seemed to be *in itself* is now only *for consciousness*, and this discovery *is itself the source of consciousness' new object*.

Hegel refers to this dialectical movement as the *experience* of consciousness (*Erfahrung des Bewußtseins*) and highlights its *positive* significance. In his eyes, the critical tradition focused exclusively on the nothingness of natural reason, and prescribed therefore a purely negative kind of deliverance. It failed to see that the untruth of one's usual ideas and concepts is not just *any* untruth. It is not an absolute, generic or *abstract* untruth, but a *determinate* untruth, viz. the untruth that emerges from the negation of a determinate content. In the *Phenomenology*, on the other hand, the negation of natural knowledge is a *determinate* negation (*eine bestimmte Negation*), whose positive content gives access to a higher cognitive stage. Once consciousness realizes that its knowledge is not real knowledge, but merely the appearance of knowledge, the latter is itself posited as

[286] *HW* III, 78f.; *PS*, 55
[287] *HW* III, 78; *PS*, 54
[288] Ibid.

the new object. 'This new object is the nothingness of the first, it is what experience has made of it.'[289] But since consciousness is simultaneously objective and subjective, knowledge *in itself* and knowledge *for itself*, this new object gives rise to a new criterion of truth, and to a new dialectical agreement. Truth is the result of a new self-comparison, which leads in turn to a new disagreement and to a new dialectical reversal.

Whenever this cycle is renewed, a new 'shape' or 'pattern' of consciousness (a new *Gestalt des Bewußtseins*) comes on the scene. A new world is revealed to consciousness, shaped by a new cognitive alphabet and a new set of aims and priorities. But although this world is entirely new, or entirely irreducible to consciousness' previous cognitive situation, it is nonetheless the necessary result of what went before, or a new solution to the same founding predicament. This dialectical mixture of difference and identity, or novelty and repetition, is the key to understanding Hegel's answer to the question raised at the beginning of this section: namely, how can each cognitive alphabet, or each phenomenological stage, be *at the same time* a whole and a part, self-sufficient and dependent on something else? According to Hegel, this paradox ceases to be one when the notions of identity and difference cease to be construed in rigid or dichotomical terms. Identity is never simple or immediate, but always the result of the agreement between a given version of reality and a given standard of truth. Likewise, difference is never simple or absolute, but always the inversion of this agreement, leading to the position of a new version of reality and a new standard of truth.

Each phenomenological transition is both a revelation and a repetition: on the one hand, each new stage introduces a different object and a different criterion of truth; on the other hand, each amounts to a more or less sophisticated translation of the same original notion. But although this guiding notion is present throughout the entire progression, it is not preserved in a simple or direct manner. What is maintained from stage to stage is not a specific object or a specific criterion of truth, but the very drive to overcome consciousness' internal contradictions. The negation of a given standpoint is never final, or abstract, because it preserves the need to accommodate its initial demands in some other way. In proper Hegelian language, this positive, or determinate negation is

[289] *HW* III, 79; *PS*, 55

called *Aufhebung*, or sublation: it is the movement whereby a given standpoint is simultaneously *negated*, *preserved* and *raised* to a higher form.[290] More specifically, it is the movement whereby a specific worldview is negated, while its underlying notion is preserved and raised to a new degree of realization.

The *Phenomenology*'s first chapter, devoted to the stage of 'sense-certainty' (*sinnliche Gewißheit*), starts out with a simple equation: true knowledge is the immediate correspondence between 'what simply is' and one's direct apprehension of it.[291] By comparing its object with itself, however, consciousness finds out that this agreement is only apparent. Since the knowledge it claims to possess is not able to account for reality, the sensory world is not real in itself, but only for consciousness. The simplicity or singularity ascribed to reality is not objectively true, but subjectively asserted, and the negation of singularity, i.e. universality, becomes consciousness' new object.

This first *Aufhebung* leads to the emergence of a new phenomenological stage, namely 'perception' (*Wahrnehmung*) and the realm of 'things'.[292] Once again, by comparing its object with itself, consciousness comes across a new contradiction. Since perception is unable to account for reality, the perceptual world is not real in itself, but only for consciousness. The 'conditioned' universality ascribed to reality is turned into its opposite, i.e. 'unconditioned' universality, and posited as consciousness' new object.

A new self-comparison and a new *Aufhebung* will lead, afterwards, to the stage of 'understanding' (*Verstand*), which will then pave the way to the stages of 'self-consciousness', and then 'reason', and then 'spirit', and finally science or 'absolute knowing'. Throughout this journey, consciousness is faced with new objects and with new criteria of truth, but all of them prove contradictory and provisional. According to the progression's own logic, the *experience* of consciousness is bound to continue until its

[290] As even the most basic introduction to Hegel's philosophy is sure to point out, the word *Aufheben* conveys both a *negative* meaning, which can be translated by the English verbs 'to suppress', 'to annul' or to 'subsume', and a *positive* meaning, related to idea 'maintaining' or 'preserving' something. Furthermore, the radical *heben* can be translated as 'to lift', 'to elevate' or 'to hoist', which reinforces the idea of a transition from a hierarchically inferior to a hierarchically superior stage or situation. Yet despite of the emphasis usually placed on the Hegelian meaning of this term, one should be aware that the meaning of *Aufheben* is not the same throughout the Hegelian corpus, and its usage is not always meant to refer to the technical procedure usually translated as 'sublation'.

[291] Cf. *HW* III, 80ff.

[292] Cf. *HW* III, 93ff.

knowledge and its object no longer disagree with one another. When reality *in itself* is no longer merely *for consciousness* — that is, when one's version of truth coincides with one's criterion of truth — subject and object will no contradict one another. In this final stage, 'knowledge no longer needs to go beyond itself', for the 'notion corresponds to the object and the object to the notion.'[293]

6.2. *Science and experience*

Generally speaking, Hegel's phenomenological critique is a double critique, directed simultaneously at two major cognitive attitudes. Its first and most immediate target is natural reason, or the kind of knowledge held by most of us in the course of our lives. The *Phenomenology* aims to denounce and overcome the *naïve* dogmatism of our usual account of reality, and to replace our daily standpoint with a true or scientific one. But Hegel's critique is also targeted at philosophy and philosophical reason. The *Phenomenology* also aims to denounce and overcome the *cultivated* dogmatism of critical philosophy, and to replace the 'critical' standpoint with a true or scientific one.

In the first case, Hegel's diagnosis coincides with the critical diagnosis. Both highlight natural consciousness' usual self-confinement, and both point out its incapacity or unwillingness to recognize this problem. Usually, whenever we consider our own ignorance, we tend to construe error and untruth as regional faults, due to a simple lack of correspondence: if our knowledge fails to capture a given object, or to account for a given event, it is replaced by a new one. Reality remains fixed, so to speak, and it is up to consciousness to find its true image. Returning to my previous analogy, this model rests on the assumption that the cognitive alphabet employed by natural consciousness is the same in which reality is written. Error and untruth amount to wrong words, or to wrong sentences, whose letters must be combined in a different way. But this assumption, according to Plato and to a vast philosophical lineage, is far from self-evident. The critical tradition shifts the focus from the words and sentences to the letters themselves, and

[293] *HW* III, 74; PS, 51

questions their ultimate meaning. If the ideas and concepts that support natural reason are defined by other ideas and concepts, and if the latter's meaning is bound, in turn, to other ideas and concepts, it might just be that reason's alphabet is entirely closed in on itself. In Hegel's language, it might just be that all of one's ideas are merely *for consciousness*, and not *in itself*, and thus that natural reason is constitutively unable to reveal reality *as it actually is*.

For Hegel, this discovery is necessary, and the only true starting point of a serious philosophical critique. Only a scepticism 'that is directed against the whole range of phenomenal consciousness', can render 'the Spirit for the first time competent to examine what truth is.'[294] But this discovery is by no means sufficient. If one's criticism stops here, philosophy is reduced to a simple and insoluble contradiction: on one side, although reason is purely *for itself*, it is bound to posit a real world beyond itself, if only to decree its inaccessibility; on the other side, although reality is purely *in itself*, it is bound to be conceived as the negative of consciousness, if only to reiterate its self-confinement. For Hegel, this sceptical impasse is the final result not only of Platonism, or Pyrrhonism, or Kantianism, but also of every critical project predicated on a rigid or non-dialectical conception of negativity. This kind of critique cannot go any further because it is bound to return, over and over again, to the dualism it seeks to overcome. Reduced to this pure negativity, consciousness 'must wait to see whether something new comes along and what it is, in order to throw it too into the same empty abyss.'[295]

This negative attitude, criticized in the introduction, is also described near the end of the preface, where Hegel reiterates the opposition between a purely negative kind of scepticism, whose result is a pure nothingness, and a positive or dialectical kind of scepticism, whose result is a determinate nothingness.[296] The first of these attitudes is linked to the procedure of 'argumentation' (*das räsonierende Verhalten*), and to what

[294] *HW* III, 73; *PS*, 50

[295] *HW* III, 73; *PS*, 51

[296] See *HW* III, 56ff. The preface was written after the rest of the *Phenomenology*, and reprises, albeit from a somewhat different perspective, the methodological discussion offered in the introduction.

Hegel calls 'representational thinking' (*vorstellendes Denken*). The second one is described as a form of 'notional thinking' (*begreifendes Denken*)[297].

The first attitude consists in a radical form of criticism, bent on refuting every idea or content it comes across. Once again, the critic's task is to refer everything back to the self, or to denounce consciousness' complete lack of objectivity. Whenever something is assumed to be real or true in itself, the critic is quick to discredit this claim: what is thought to be real is only so *for consciousness*, and hence unreal or untrue. But this attitude is once more dismissed as sterile and one-sided. This kind of reasoning 'imagines that by establishing the void it is always ahead of any insight rich in content'[298], but its argumentation is only 'a reflection into the empty "I", the vanity of its knowing.' [299] Its relationship with the content is still purely negative, and bound to return, again and again, to the same dead end. Since each new object is reduced to the same 'void', consciousness comes away empty-handed, and must look for a new object to negate.

To the 'notional' thinker, on the other hand, negativity is not abstract, but determinate. Every refutation is simultaneously a *positive* refutation because it *posits* something other than a complete nothingness. The negated content is not merely nullified, but also revealed as a new object, and as the source of a new cognitive relationship. This kind of scepticism is the one favoured throughout the *Phenomenology*, and the source of its various transformations. In order to emphasize the novelty of his method, Hegel goes on to spell out its actual stages, in a description that is in many ways evocative of the one offered throughout the introduction (and the previous section), but which is nonetheless worth considering.

The former contrast between notion and reality, or truth and appearance, is now reframed by the grammatical contrast between subject and predicate. The difference between 'representational' scepticism and 'notional' scepticism hinges on the opposition

[297] I chose this translation in order to maintain the link between *begreifen* and *Begriff*, previously translated as Notion. However, the word *begreifen* also evokes the more concrete movements of 'grasping' or 'seizing', which the adjective *notional* fails to convey. A. V. Miller (*PS*) translates *begreifendes Denken* as 'speculative thinking' – an ambiguous choice, which overlooks both his previous use of 'Notion' and the ideas of firmness and concreteness conveyed by the German expression. J. B. Baillie translates *begreifendes Denken* as 'ratiocinative thinking'.

[298] *HW* III, 57; *PS*, 36

[299] *HW* III, 56; *PS*, 36

between two different ways of understanding the subject-object relation, and hence on the logical structure of the predication process. In the first instance, consciousness sees itself as an inert subject, to which objects are appended in an accidental manner, as mere predicates. This mode of thinking can be illustrated by the following scheme:

Subject 1 ⟶ Predicate A
Subject 1 ⟶ Predicate B
Subject 1 ⟶ Predicate C
…

When something is predicated of a given subject without its essence being altered, that can only mean one of two things: either the predicate in question is indeed purely accidental and does not truly alter the subject it comes into contact with, or the predicate is already contained in the subject, and amounts to the mere isolation of an otherwise silent element, already comprised in the subject's definition. The proposition *A is B*, for instance, does not reveal anything new about A insofar as a) A is just as well C, D, or E, in which case its being B is not one of its defining features; or b) B is already part of A's definition, in which case the proposition is but a development of the identity principle, A = A. In both cases, the predication does nothing more than to confirm what was already known about the subject, without questioning the validity of its established identity. And that is why the procedure of argumentation appears so free and versatile. Its virtuosity seems boundless because its starting point and its finishing point are the same. There is no real predication, in the Hegelian sense, because the argumentative subject is really only affirming itself, or its superiority over the objective world.

Notional thinking, on the contrary, triggers a dialectical inversion of the grammatical elements in play. The example provided by Hegel — the statement *das Wirkliche ist das Allgemeine*, i.e. 'the actual is the universal'[300] — is the conceptual result of the *Phenomenology*'s first major *Aufhebung*, viz. the transition from the stage of sense-certainty to the stage of perception.[301] In this statement, the predicate 'universal' is not merely absorbed or discarded by the subject. The link between actuality and universality

[300] *HW* III, 60; *PS*, 39
[301] Cf. *HW* III, 93.

is neither tautological nor accidental, as in the previous model: universality is not a mere predicate of actuality, but the *substance*, or the *notion*, or the *essence* of the whole relationship.[302] Indeed, although actuality starts out by being actual (A = A), it loses its identity during the predication process (A ≠ A). What comes to light when actuality is equated with universality is not that the actual is *also* the universal, or that the actual *happens to be* universal, but that the actual is, *after all*, the universal — and has always been so. In other words, universality is not merely 'represented' (*vorgestellt*) by the subject, but grasped (*begriffen*) as its actual truth.

Hegel describes this movement in the following manner:

> [Representational thinking] suffers, as we might put it, a counterthrust. Starting from the Subject as though this were a permanent ground, it finds that, since the Predicate is really the Substance, the Subject passed over into the Predicate, and, by this very fact, has been sublated; and, since in this way what seems to be the Predicate has become the whole and the independent mass, thinking cannot roam at will, but is impeded by this weight.[303]

Unlike the procedure of argumentation, notional predication is not a reversible process. It destroys itself, so to speak, as it goes along.[304] Whereas initially the subject extended peacefully to the predicate, recognizing itself therein, it is now absorbed by the predicate and forced to rethink its own identity. The subject no longer sees itself with its own eyes, but with the eyes of the predicate, and the latter becomes thereby the new subject:

> Here thinking, instead of making progress in the transition from Subject to Predicate, in reality feels itself checked by the loss of the Subject. […] And now, having returned into itself in the Predicate, instead of being in a position where it has freedom for

[302] Most of the difficulties in this passage are due to the profuseness of nomenclature. In a seemingly interchangeable way, Hegel employs the terms *Begriff*, *Wesen* and *Substanz* to name that which, in the introduction, was mainly called *Begriff*, i.e. the excessive element recognized and preserved at each new phenomenological transition. The word *Substanz*, in particular, seems to have been chosen in view of its etymology: it is what *sustains* or *underlies* (*sub-stans*) the relationship between subject and object.

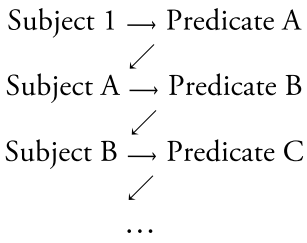
[303] *HW* III, 58; *PS*, 37

[304] Accordingly, although it can be stated that the 'actual is the universal', it cannot be stated that 'the universal is a predicate of the actual'. The truth of the term 'actual' has migrated to the predicate, and transformed thereby the entire relationship.

argument, it is still absorbed in the content, or at least is faced with the demand that it should be.[305]

The subject is no longer free because its notion is held hostage, as it were, by the predicate. But the latter also loses its original freedom, or the ability to be associated, generically, to other subjects. The predicate is now bound to the determinate negation of *this particular subject*: only as such is it in fact the new subject, as well as the source of a new proposition and a new dialectical inversion.

Once more, this movement can be illustrated by a simple scheme:



The *Phenomenology* is modelled on this dynamic form of predication. It consists in a global system of subjects and predicates, whose meaning is successively transferred to new predicates and to new subjects. As Hegel puts is, it stages the repeated acknowledgement that in thinking or saying something 'we meant something other than what we meant to mean; and this correction of our meaning compels our knowing to go back to the proposition, and understand it in some other way.'[306] This procedure takes the following form: consciousness starts out by assuming that all knowledge is immediate knowledge. It begins by claiming that the 'true is the actual' (Subject 1 ⟶ Predicate A). However, by

[305] *HW* III, 59; *PS*, 38

[306] *HW* III, 59; *PS*, 39. This self-correcting procedure is one of the main causes of the frequent complaint regarding Hegel's unintelligibility. In an essay devoted to the clarity of Hegel's writings, Adorno argued that this kind of complaint is due in great measure to the utterly un-Hegelian assumption that '[philosophical] terms must remain within the meaning once given them by definition' (op. cit., 112). This assumption is un-Hegelian because it confines philosophical reason to the static poles of self-identity and difference, and regards intelligibility itself as a fixed measure of the identity of a given idea with itself. However, 'once it is acknowledged that clarity and distinctness are not mere characteristics of what is given, and are not themselves given, one can no longer evaluate the worth of knowledge in terms of how clearly and unequivocally individual items of knowledge present themselves' (ibid., 100). Rather, 'only when the life of the thing expressed by the concept is compared with the meaning specified and when the old meaning is thereby dishonoured as invalid, is the other meaning constituted.' (ibid., 113)

pursuing this claim and fleshing out its consequences, it realizes that it must turn back and correct its original proposition. The 'actual' is not what it seems, but its opposite, namely the 'universal' (Subject A ⟶ Predicate B). This new predicate becomes the new subject, but its meaning is once again called into question: the 'universal' is not what it seems, but the 'infinite' (Subject B ⟶ Predicate C); and the latter, in turn, will also prove inconsistent and in need of replacement. This cycle can only be broken once consciousness says what it actually means to say, and no longer needs to correct its initial formulation. This can only happen, according to Hegel, in the stage of science, or absolute knowing, where '*Subject* is just as much *Substance*',[307] and predication finally comes to an end.

In the preface, as in the introduction, Hegel highlights the idea that truth cannot be known immediately, or directly, but only as a *result*. Immediate truths, precisely because they are immediate, are intrinsically unable to vouch for their truthfulness. They amount to bare assumptions, liable to be challenged and replaced by other bare assumptions. Asserting a preposition by 'adducing reasons for it, and in the same way refuting its opposite by reasons, is not the form in which truth can appear.'[308] If truth and science are ever to be reached, consciousness must rise above this arbitrariness, and it can only do so in an immanent way, through a 'scepticism directed at the whole range of phenomenal knowledge', or through 'the exposition of untrue consciousness in its untruth.'[309] But although the acknowledgement that one's assumptions are untrue is already a step further from believing them, it is still a long way from actual truth. What distinguishes Hegel's critical approach from conventional scepticism is the idea that untruth is not only negative, but also positive — that it can be set in motion, as it were, and lead the way to its own overcoming.

Untruth is not the end, but the beginning. However, the success of Hegel's undertaking is conditional on at least three major requirements, which follow from the *Phenomenology*'s own methodological programme, and failing which true knowledge will

[307] *HW* III, 587; *PS*, 489. Cf. *HW* III, 22f.: 'Es kommt nach meiner Einsicht, welche sich nur durch die Darstellung des Systems selbst rechtfertigen muß, alles darauf an, das Wahre nicht als *Substanz*, sondern ebensosehr als *Subjekt* auszufassen und auszudrücken.'

[308] *HW* III, 47; *PS*, 28

[309] *HW* III, 73; *PS*, 50

indeed remain out of reach. Firstly, since truth is incompatible with the immediacy or arbitrariness of human assumptions, it can only be secured when the latter is fully overcome. For this to happen, consciousness must expose the untruth not only of this or that assumption or set of assumptions, but of every single one of them. It must traverse the entire realm of appearance and make sure that no assumption is left untested or unrefuted. Otherwise, when it finally believes to have found true knowledge, it will have no way of knowing whether the latter is indeed so, or just a new assumption disguised as truth.

Returning to Hegel's grammatical model, the *Phenomenology* must cover the entire range of propositions liable to be endorsed by consciousness. Only when every subject and every object have been confronted and sublated, and there is nothing left to 'mean', is the realm of untruth truly exhausted and overcome. The 'recognition of the failure of all criteria is the negative side of an argument which leaves Hegel's dialectical holism as the positive side, the only alternative. Emergent holism is not one truth among others, but the only truth there can be.'[310]

Secondly, the success of Hegel's project is conditional on the ability to identify the correct outcome of each proposition and the correct sequence of sublations leading to the standpoint of science. For the realm of appearance to be truly overcome, consciousness must follow a specific route, or a specific guiding thread. If the predicative chain is broken, or any of its links overlooked, the whole progression will be compromised. Consciousness will once more be led to rely on mere assumptions, brought over from the realm of appearance, and the attainment of truth will no longer be possible. Against this attitude, Hegel insists on the immanent nature of the progression and on the uselessness of external criteria. The road to science *is itself already science*: its specific direction is not contingent or arbitrary, but determined by the notion itself. Accordingly, the *Phenomenology* does not lay out only the *experience* of consciousness, but also the *science*

[310] WILLIAMS, op. cit., 112. This issue evokes the classical notion of ἰσοσθένεια, or *equipollence*. In line with the sceptical tradition, Hegel argues that the truthfulness of a given claim can only be established once all claims to the contrary have been refuted. Until then, truth and untruth are merely dogmatic labels, arbitrarily pinned on a series of unverified beliefs. Unlike the Sceptics' equipollence, however, Hegel's equipollence is not a cul-de-sac, but a principle of motion. The *Phenomenology*'s strategy is to overcome scepticism by refuting all of consciousness' *positive* claims about reality, and reaching thereby the *negative* of untruth, i.e. truth or science.

of the experience of consciousness. 'The length of this path has to be endured, because … each moment is necessary' and 'has to be *lingered* over.'[311] The scrupulousness with which consciousness fulfils its mission — 'the seriousness, the suffering, the patience, and the labour of the negative'[312] — determines how far it is able to progress.

Thirdly, in line with the previous requirements, the success of this undertaking is also dependent on there being, in fact, an ending or a finishing line. Even if Hegel is right about the self-evolving nature of human reason, and even if the progression's different stages are indeed necessarily linked to one another, the original disagreement between subject and predicate might never be truly eliminated, and the progression might go on indefinitely. This scenario entails, once again, the complete unfulfilment of Hegel's project: if the progression has no ending, all of its stages, however complex or sophisticated, are liable to be refuted, and science remains out of reach.

Aware of this possibility, Hegel argues not only that the 'the *goal* is as necessarily fixed for consciousness as the serial progression'[313], but also that 'the interconnection of the forms of the unreal consciousness will by itself bring to pass the *completion* of the series.'[314] He claims, furthermore, that the progress towards this final conclusion, when 'knowledge no longer needs to go beyond itself', is unhalting 'and short of it no satisfaction is to be found at any of the stations on the way.'[315] These and similar claims are grounded, once more, in the idea of totality.

The *Phenomenology*'s stages are not simply untrue standpoints wrongly assumed to be true. They are also *fragments*, or *moments* of a wider, but usually invisible or inaccessible whole. The progression is not simply about truth and error, but also about a gradual movement of expansion, whereby consciousness is led not only to rephrase what it originally 'meant', but also to redimension the scope of its sentences. Once again, however, this movement must not be construed in a quantitative way, as the gradual enlargement of the same cognitive standpoint, or the simple accumulation of different cognitive fragments. Each fragment is perceived, at first, as the whole world, and its

[311] *HW* III, 33; *PS*, 17
[312] *HW* III, 24; *PS*, 10
[313] *HW* III, 74; *PS*, 51
[314] *HW* III, 73; *PS*, 50
[315] Ibid.

refutation amounts, therefore, to a global refutation, leading to the endorsement of a new cognitive totality. But although this new standpoint is just as global as the first one, and just as untrue, it offers nonetheless a wider (or less limited) glimpse into the progression's final standpoint.

This model can be further clarified with the aid of Leibniz and his own characterization of human cognition. In Leibniz's epistemology, human cognition is also grounded in a global account of reality. However, each individual consciousness, or monad, is confined to an individual cognitive perspective: it is a single *and yet* a global standpoint, a fragment *and yet* the whole universe. This paradox is illustrated by the image of an observer looking at the same city from different angles: depending on his or her position — on his or her *situs* — the observer sees the same elements combined in a specific way, according to a specific point of view.[316] This idea is reiterated in the *Phenomenology*, where each *Gestalt* is also a unique totality, or a 'situated' outlook on the whole of reality. For Hegel, however, this situation is not definitive. Whereas Leibniz's description is focused on the simultaneous diversity of different individual perspectives, Hegel's is focused on the self-evolving diversity of the same cognitive perspective. Instead of different observers placed in different positions, eternally confined to the same version of reality, the *Phenomenology* documents the movement of the same observer as he or she takes up different cognitive angles.

In Hegel's account, not all of these angles are equivalent. Although they are all equally one-sided, each builds on the information provided by the previous ones, like a cubist composition with increasingly faceted shapes. As each new angle is criticized and sublated, it 'ranges itself as a moment having its own place in the whole'[317], and is thereby integrated into the cognitive ground whereupon a new standpoint is made to rest. Due to the nature of this method, true knowledge can only be attained once all different angles have been viewed and criticized, which brings us back to the third requirement listed above, and to another major difference between the Hegelian and the Leibnizian models. Unlike Leibniz's spatial metaphor, where the same city is liable to be viewed from an

[316] See *Philosophische Schriften*, ed. cit. I, 10, 19, 20, 69, 150f.; II, 19, 57, 98, 438; III, 357, 623, 636; IV, 434, 438, 477, 484, 530, 553f.; V, 50, 65, 378; VI, 197f., 327, 599, 603, 616; VII, 120, 452, 542f., 556, 566f.

[317] *HW* III, 52; *PS*, 32

infinity of different perspectives, Hegel draws up a finite progression, comprising a limited number of stages. He argues that the whole underlying every cognitive perspective can be divided into a specific number of parts, or sets of parts, which sum up all of reason's cognitive attitudes. This does not mean, of course, that consciousness is reduced to a finite number of thoughts or judgements: to be sure, each phenomenological stage allows for an infinite number of instantiations, prompted by the infinite variety of one's actual experience.[318] But all of one's thoughts, Hegel argues, fall under a finite number of 'shapes' or 'patterns of consciousness', whose sublation implies the automatic sublation of all of their ramifications.

6.3. *Further hermeneutical issues*

The *Phenomenology of Spirit* documents the long transition from a natural, uneducated or untrue mode of cognition to a scientific, educated or true one. In the previous sections, I have outlined Hegel's methodological strategy and the main requisites for the successful completion of his project. In what follows I will return briefly to the beginning and examine the kind of ignorance or untruth Hegel places at the start of the progression.

At first sight, since natural or phenomenal consciousness (*das erscheinende Bewußtsein*) corresponds to one's usual cognitive perspective, it would seem natural to regard Hegel's progression as an actual history of the development of human reason. If this were the case, the different phenomenological stages would amount to actual cognitive strata, leading from the most primitive to the most sophisticated mode of cognition. Initially, in the stage of sense-certainty, knowledge would amount to a pure apprehension (*Auffassen*), devoid of comprehension (*Begreifen*), and consciousness would be exclusively focused on the sheer *being* of things (*das Sein der Sache*).[319] With the onset of perception, it would acquire the notions of selfsameness and difference, along with the power to distinguish between truth and falsehood. Afterwards, in the stage of

[318] Cf. *HW* VII, 370: 'Vollständigkeit heißt die vollendete Sammlung alles Einzelnen, was in eine Sphäre gehört, und in diesem Sinne kann keine Wissenschaft und Kenntnis vollständig sein.'

[319] *HW* III, 82; *PS*, 58

understanding, it would learn about laws and forces, establish causal relations, deduce logical inferences, etc. Its knowledge would be gradually transformed and perfected, until all of its contradictions were overcome.

The problem with this kind of reading, however, is all too apparent. Although one's usual cognitive perspective is far from perfect, and might even be fully challenged, it is certainly not as limited as the stage of sense certainty. In its primitive or natural form, human consciousness is already capable of much more than simply *apprehending* reality: it is able to recollect past events, to anticipate future ones, to confront different ideas and to form complex judgments. As seen above in connection with the *Philebus* and Plato's cognitive scale, human reason is intrinsically irreducible to the stage of αἴσθησις, or sense perception. The human mind is incapable of envisaging a purely sensuous account of reality and can only consider such a possibility in a negative way, as the inner contraction of a more complex cognitive standpoint. The same is valid, moreover, for the stage of μνήμη, or recollection, viz. for a purely mnemonic account of reality. The ground zero of human cognition is the stage of experience (ἐμπειρία, or δοξάζειν), already bound by normativity and predication.[320]

In the beginning of the *Phenomenology*'s first chapter, Hegel claims, nonetheless, that the kind of knowledge 'which is *at the start* or is *immediately* our object' is sense-certainty.[321] But these adverbial expressions must not be construed in a literal manner, as direct qualifications of reason's actual starting point. The progression does not start with the most immediate form of knowledge there is, but with *one's usual conception* of the most immediate form of knowledge there is. In other words, although natural consciousness is aware that its actual knowledge is much more 'advanced' than sense-certainty, it tends to recognize the αἴσθησις model as the simplest form of knowledge, and its own knowledge as a direct development thereof. What is as stake in the beginning is thus a *retrospective* diagnosis: starting from its actual cognitive standpoint, consciousness retraces its own cognitive history all the way to the very beginning. And the very beginning, for natural consciousness, is sense certainty.

[320] See above section 2.1.
[321] *HW* III, 79; *PS*, 58. Emphases added.

Hegel's progression does not consist in a series of actual cognitive stages, but in a series of cognitive models, or 'patterns of consciousness'. It sets out to offer a complete anatomical and morphological description of human cognition, starting with its most basic elements. Unlike other similar initiatives, however, this description is not univocal, nor offered by a neutral or disinterested observer: the different models taken up and criticized throughout the *Phenomenology* are forged by consciousness itself, to account for its own internal constitution. The elevation to the standpoint of science must start, according to Hegel, with consciousness' own view of itself, and particularly with its own view of the cognitive framework underlying its usual account of reality.

The acknowledgement of this aspect is paramount, in my view, to a proper understanding of Hegel's project. It can help explain not only the disparity between the progression's initial standpoint and consciousness' usual standpoint, but also the apparent incongruity of some of Hegel's thematic and historical transitions. Since the logic binding the different stages together is an immanent logic, determined by consciousness' own understanding of itself, the progression's guiding thread is not factical or historical: it does not reflect the sequence of cognitive attitudes and doctrines actually adopted throughout human life, or human history, but a necessary chain of discoveries, which leads consciousness to overcome its usual contradictions and to recognize its true self.[322]

If the phenomenological progression is a *retrospective* progression, in the aforementioned sense, its starting point is not an actual cognitive stage, or an actual historical stage, but a specific cognitive regime, or a specific cognitive alphabet. By placing sense-certainty at the basis of human cognition, consciousness declares its own knowledge to be a development of sense perception, that is, a highly complex web of simple sensory units, or αἰσθήματα, brought together to form diverse empirical constructs. This initial stage corresponds, therefore, to a simple form of empiricism, liable to produce highly developed forms of meaning while remaining faithful to the

[322] Cf. *HW* III, 591; *PS*, 493: The preservation of the progression's different stages, or 'Spirits' 'regarded from the side of their free existence appearing in the form of contingency, is History; but regarded from the side of their [philosophically] comprehended organization [*ihrer begriffenen Organisation*], it is the Science of knowing in the sphere of appearance [*die Wissenschaft des erscheinenden Wissens*, i.e. the Phenomenology of Spirit].'

simplest of cognitive formulas. Following this logic, one can even envisage different versions of the progression's subsequent stages based on sense-certainty: an empirically-based reason, an empirically-based self-consciousness, an empirically-based morality, an empirically-based, religion, and so on. Unlike the new cognitive models introduced in the stages of reason, self-consciousness, morality or religion, all of these versions presuppose the atomic understanding of reality that is typical of sense-certainty.[323]

Accordingly, when the stage of sense-certainty is sublated and replaced by a new one, consciousness finds out that this simple form of empiricism is not the real source of human reason. More precisely, it is led to recognize that the complexity of its actual mode of cognition cannot be reduced to a mosaic of sensuous units, and hence that the progression's real starting point must be looked for elsewhere, in a different and more sophisticated cognitive regime. As sense-certainty gives way to perception, consciousness endorses a new idea of what the simplest form of knowledge might be. Instead of a simple empirical atomism, it is now led to embrace a more sophisticated form of atomism, made up of 'things' and their attributes. Once again, however, consciousness realizes that its actual knowledge is irreducible to this new cognitive framework: it is already a mode of understanding — and then a variation of self-consciousness, and afterwards a modality of reason, and so on and so forth. With each new transition, the progression's real starting point is pushed upwards, again and again, until the very last stage. Since the entire progression is about finding a true cognitive foundation, unsusceptible of being debunked as yet another form of *Erscheinung*, and since this discovery does not take place in any of the progression's intermediate stages, Hegel's point seems to be that the true foundation of human reason is nothing less than absolute knowing, viz. the dialectical coincidence of notion and object, essence and appearance, for-itself and in-itself etc. In other words, his point seems to be that human beings' usual

[323] In the introduction to the *Encyclopaedia*'s third part, as Hegel discusses the immanent nature of the progression leading from 'subjective spirit' to 'absolute spirit', he points out that 'the determinations and grades of Spirit … are essentially only moments, states, determinations in the higher stages of development. As a consequence of this, a lower and more abstract determination of the mind [*Geist*] reveals the presence in it, even empirically, of a higher phase. In sensation, for example, we can find all the higher phases of the mind as its content or determinacy. And so sensation, which is just an abstract form, may to the superficial glance seem to be the essential seat and even the root of that higher content, the religious, the ethical, and so on; and it may seem necessary to consider the determinations of this content as particular species of sensation.' (*HW* X, 16f.; *PM*, 8f.)

perspective is a scientific perspective *in germine*, unaware of its true nature and forced to go through a long journey of self-discovery.

The *Phenomenology*'s global architecture remains invisible, however, throughout the progression. The acknowledgement that the first phenomenological model is untenable leads consciousness to look for a new one. And when the latter is refuted, a third model is sent over to replace it. Nevertheless, each new solution is construed as an entirely new alternative, disconnected from the previous ones. The immanent necessity binding the different stages together is not yet apparent and the progression is not yet seen or lived as a progression. In the eyes of consciousness, the whole *Phenomenology* amounts to a series of independent cognitive attitudes, brought in to fill the gap left open by the first phenomenological refutation. Consciousness only sees the negative side of its own experience, and not yet the *positive* or *determinate* value of negativity and untruth.

But although this uneducated standpoint is the *Phenomenology*'s main protagonist, it is not the progression's only standpoint. In the introduction, when discussing the work's methodology, Hegel distinguishes two different angles from which to look at the progression. The first one reveals a pathway *for consciousness* (*für das Bewußtsein*), corresponding to the negative or uneducated attitude just mentioned. The second one reveals a pathway *for us* (*für uns*), philosophers or 'phenomenologists'. Whereas *it*, i.e. consciousness, is unaware of the progression's immanent logic, *we*, i.e. philosophical or phenomenological observers, know each stage to be the necessary result of the previous one, and a mere piece within a wider dialectical puzzle. For *us*, each *Aufhebung* is both a suppression, a preservation, and a transition to a higher and more comprehensive cognitive stage.

This hermeneutical contrast is summed up in the following passage:

> [With each transition,] a new pattern of consciousness comes on the scene …, for which the essence is something different from what it was at the preceding stage. It is this fact that guides the entire series of the patterns of consciousness in their necessary sequence. But it is just this necessity itself, or the *origination* of the new object, that presents itself to consciousness without its understanding how this happens, which proceeds for us, as it were, behind the back of consciousness. […] *For it*, what has thus

> arisen exists only as an object; *for us*, it appears at the same time as movement and a process of becoming.[324]

Consciousness, absorbed in its own experience, is unaware of its underlying logic. All it sees is a succession of objects, or *Enstandene*, already formed and ready to use. Their origination, *Entstehen*, remains out of sight, or 'behind consciousness' back'. It can only be viewed from a wider hermeneutical angle, for which the movement itself takes precedence over the temporary shapes it comes to assume. Such a standpoint — *our* standpoint — is no longer dominated by a simple or tautological conception of identity. It knows each phenomenological transition to be both a negation and a position, a complete refutation and the entry point into a new cognitive world. Therefore, it regards each new object not merely as something that *happened to emerge*, but as something that was *bound to emerge*, and which is *bound to disappear* in a determinate way. This is what Hegel means when he claims, near the end of the introduction, that only *for us* is 'the way to science' 'already 'science' — or that only from *our* point of view is the 'experience of consciousness' simultaneously the 'science of the experience of consciousness.'[325]

In light of this contrast, the *Phenomenology* is often divided into two different textual layers. The first, *empirical* one, is the layer of action, so to speak, and usually occupies the central section of each chapter. It comprises 'the passages where consciousness is urged, as it were, to put itself in motion, and to *convert* itself, in the proper sense, to itself, in order to adapt its vision of things to the things that are actually brought into view.'[326] The second, *scientific* or *phenomenological* layer, usually limited to the moments of transition, provides an interpretation of what goes on in the first layer. It comprises 'the developments during which consciousness takes a break, so to speak, while the philosopher steps to the foreground, in order to account for the conceptual continuity that makes each experience the *necessary* product of the preceding one, as well as to add up all of these experiences and their unitary meaning.'[327] This second layer is purely exegetical: it does not interfere in any way with the progression's actual unfolding. Since

[324] *HW* III, 76; *PS*, 56
[325] Ibid.
[326] LABARRIÈRE, *La phénoménologie de l'esprit*, ed. cit., 37
[327] Ibid.

the critique pursued throughout the *Phenomenology* is immanent and self-moving, all that is left to the phenomenological observer is 'simply to look on' (*das reine Zusehen*).[328] But this passive role is just as essential as the active role played by consciousness: only through this added interpretation, 'contributed by *us*', is 'the succession of experiences through which consciousness passes [raised] into a scientific progression'.[329]

The contrast between these two layers is indicative of a wider structural distinction, which poses an additional and complex hermeneutical challenge. If the only 'way to Science' is a 'phenomenology of spirit', the 'scientific' sections of the text cannot be understood by natural consciousness. Only a standpoint that has already underwent the phenomenological progression and attained its final stage can grasp the true meaning of these passages. Following this conclusion, the progression is not a simple, but a double progression, which requires a double interpretation: after the elevation of natural consciousness to the stage of science, the resulting standpoint must go back to the beginning and acknowledge the immanent necessity underlying its original transformation. In this second, retrospective reading, 'Science sets forth this formative process in all its detail and necessity, exposing the mature configuration [*Gestaltung*] of everything which has already been reduced to a moment and property of Spirit.'[330]

But although this hermeneutical approach seems to agree with many of Hegel's methodological indications, it is contradicted by at least three main reasons. Firstly, although the progression's initial standpoint corresponds to consciousness' usual standpoint — or rather, to its usual conception of the most simple form of knowledge — the 'phenomenologist''s remarks are not entirely unintelligible to the reader that first takes up the *Phenomenology*. To be sure, the logical arguments and connections 'contributed by us' are not immediately clear. They fall outside natural consciousness' usual cognitive horizon, and require therefore the adaptation to a new language and to a

[328] *HW* III, 77; *PS*, 54.

[329] *HW* III, 79; *PS*, 55.

[330] *HW* III, 33; *PS*, 17. Returning to the parallel between the *Phenomenology* and the allegory of the cave, this retrospective standpoint can be compared to the liberated prisoner's standpoint as he returns to the cave and attempts to release the other prisoners. Having regained his initial position, his second ascent is very different from the first one: he already knows the shadows to be mere shadows, the cave to be a mere cave and the various forms of illusion found inside and outside the cave as the intermediate stages of a long dialectical progression, whose final stage is already known.

new set of concerns. Nevertheless, their meaning is taken into account from the very beginning, and becomes increasingly familiar as the progression unfolds.

Secondly, the alternation between the progression's empirical and scientific layers does not follow, on the whole, a fixed pattern. Whereas the first chapters are structured in a relatively similar way, directly evocative of the introduction's methodological description, the more advanced stages are not as uniform. The balance between both narratives layers is at times highly uneven, and the length, importance and meaning of their occurrences are very unequal.

Thirdly, the 'double reading theory' presupposes the (relative) independence of the empirical and the scientific layers. It presupposes the possibility of a first reading exclusively focused on the first layer, free from the retrospective element introduced by the second one. But in most cases, this separation is very difficult, if not downright impossible. The phenomenologist's considerations are usually integrated into the argumentative structure, and cannot be eliminated without sacrificing the intelligibility of the whole text. Moreover, the very frontier separating these two registers is often hard to determine. As the *Phenomenology* progresses, the clear-cut opposition between *being-for-us* (*Fürunssein*) and *being-for-consciousness* (*Fürsichsein*) grows increasingly blurred, and the two layers end up by coinciding with one another.[331]

For all of these reasons, the *Phenomenology* is best construed as a single or unified movement, made up of different, but simultaneous hermeneutical layers. The central distinction between phenomenal and scientific knowledge, reiterated and transformed at each new stage, is accompanied and complemented by the secondary, but no less important distinction between consciousness' and the philosophical narrator's understanding of the progression's immanent logic. With each new transition, the distance separating these two standpoints is increasingly reduced, and finally eliminated at the progression's final stage. The standpoint of science, or 'absolute knowing', signals therefore not only the coincidence between reality *in itself* and reality *for consciousness*, but also the fact that this coincidence is made evident both *for us* and *for consciousness*. In other words, the elevation to the standpoint of science entails the simultaneous

[331] See LABARRIÈRE, op. cit., 35-38.

acquisition, on the part of consciousness, of a retrospective understanding of the entire progression.

7.

SELFHOOD AND OTHERNESS

Broadly speaking, the *Phenomenology*'s first great section, entitled *Bewußtsein*, and dedicated to 'consciousness' in general, offers a series of variations on one central theme: the automatic and widespread belief in the intrinsic independence of the objective world. According to Hegel, the most distinctive trait of human consciousness, in its usual or natural form, is the assumption that reality exists in its own right, regardless of one's coming to know it. Things are what they are, and can either be known or ignored, perceived in a clearer or less clear way. The world is a show, as it were, in which the subject is a mere spectator, confronted with ready-made objects and events. This does not mean, of course, that reality cannot be changed or transformed. Although the subject is a spectator, he or she is also an actor, or an agent, and his or her actions have real effects and consequences. Nevertheless, all of these changes presuppose the acknowledgement of a fixed objective setting, on which subjectivity is usually grounded. The objective world is usually already *there*, or already *given*, and this assumption is the starting point of most human judgements and actions.

In the beginning of Hegel's progression, this assumption is taken up in its most primitive form. The object of human consciousness is one that simply or immediately *is*, 'regardless of whether it is known or not',[332] and one that remains what it is 'even if it is

[332] *HW* III, 84; *PS*, 59

not known'.[333] However, although natural consciousness is ready to buy into this version of things, a closer look at objectivity will bring to light its various contradictions. By attempting to capture the object in its independence or immediacy, consciousness is led to realize that the object is neither independent nor immediate: instead of simply *being there*, unaffected by one's cognitive access to it, objective reality is already implicated in a dialectical relationship with the knowing subject. Its identity is not simple, but mediated, and its objectivity is stained, so to speak, with subjectivity.

This discovery is the fuelling principle of the first three chapters. The realization that objectivity is not objective prompts the abandonment of the progression's first cognitive standpoint and the adoption of a new one, in which consciousness hopes to find a truly objective account of reality. This change, however, proves unsuccessful and leads to a new stage and a to a new disappointment. As consciousness moves from one standpoint to the other, the subjective contamination of objectivity is found to increase, rather than decrease. Finally, at the end of this first section, subjectivity takes over altogether. The world is revealed as the very opposite of what it was claimed to be, namely, a fixed objective setting, and consciousness finds out that its object is in fact its own self. This new transformation opens the *Phenomenology*'s second great section, dedicated to self-consciousness, and launches a new series of cognitive stages. The genesis of this new shape of consciousness, as well as its own phenomenological experience, will be examined throughout the following pages.

7.1. *The road to self-consciousness*

At the very beginning, cognition is nothing but 'sense-certainty' and its object is the purely immediate — sensuous reality as a *factum brutum*, prior to any intellectual or categorial distinction.[334] On the subjective side of the relationship, knowing is simply *apprehending* reality. Consciousness amounts to a purely transparent filter, which reveals

[333] Ibid.
[334] See *HW* III, 82ff.

the world as it really is. On the objective side of the relationship, reality is made up of simple concrete things, which reveal themselves immediately to consciousness. On closer inspection, however, the first contradictions start to emerge. Reality may well be simple and concrete, but this simplicity and this concreteness are not within the reach of sense-certainty. Consciousness cannot conceive of such a world because the very words or concepts it uses to refer to reality — 'this', 'here' and 'now' — are already universal and abstract.[335] Their meaning is not simple or self-contained, but dependent on something *other* than themselves, which guarantees their intelligibility. Although 'this', 'here' and 'now' are used to refer to simple things, they are in fact undifferentiated conceptual operators, which can be applied with equal justness to an infinity of objective contents.

The acknowledgement of this contradiction leads consciousness to alter its initial strategy and to endorse a new cognitive thesis: simplicity and concreteness do not reside in reality, but in consciousness; it is consciousness itself the carrier, so to speak, of the simplicity it mistakenly apportioned to he objective world. The initial model is therefore preserved, but transferred to the realm of subjectivity: it is *for consciousness* that a tree is 'this tree' and not another, and that the present moment is 'this now' and not another. However, like the 'here' and the 'now' of the previous stage, the 'I' of whom these claims are predicated — the 'I' which guarantees the singularity of each object — is itself revealed as an undifferentiated conceptual operator. Since the 'I' is always an 'I that x', or an 'I that y', or an 'I that z', etc., it is never a simple or self-sufficient determination, but a universal and abstract one, which can be applied with equal justness to an infinity of different contents.

The truth of the singularity envisaged by consciousness cannot be found either in the subject or in the object. In a last attempt to preserve the model, consciousness turns to

[335] As the verbs *abs-trahere* or ἀ-φαιρεῖν indicate, to *abstract* from something means to 'draw away' from something, or to 'remove' something from something else. For Hegel, something is *abstract* when its meaning, which can only be grasped as part of a wider dialectical whole, is separated from it and posited as a self-sufficient determination. Something is *concrete*, on the other hand, 'when it contains not only its *unitary*, immediate determination, but also what is other than itself.' (*HW* XX, 475). As we will see below in more detail, the phenomenological progression consists in the systematic discovery of concrete or mediated meanings falsely perceived as abstract or immediate ones. In the present case, the recognition that the knowledge provided by sense-certainty is not concrete, but abstract, or based on abstract conceptual operators, will lead consciousness to adopt a new cognitive model, and look therein for the concreteness in failed to find in the first one.

the relationship itself: if both the subject and the object, when isolated from one another, turn out to be universal operators, the loss of singularity can only arise from the very act of isolating them and considering them separately. In order to correct this, consciousness must focus on the two elements at the same time: viz. an individual 'I' — 'this I' — facing a 'here' and a 'now' — 'this here' and 'this now' — and preventing any form of otherness to undo the immediacy of this knowledge. But this new procedure also proves untenable. In its effort to isolate the concreteness of a single spatial point and a single temporal moment, the thinking 'I' merely 'indicates' or 'points out' (*zeigt auf*) the 'here' and the 'now' it has in mind, making them present for consciousness. In doing so, it attempts to remain within the realm of 'pure intuition' (*reine Anschauung*), immersed in the absolute simplicity of αἴσθησις and immune to all forms of speculation or self-reflection. However, in the very act of singling out a specific temporal moment, consciousness loses what it means to grasp. Every 'now', in being indicated, or pointed out, ceases to be what it was: it has already slipped into the past and been replaced by a new 'now', doomed in turn to the same disappearance. Likewise, in the very act of singling out a specific spatial point, consciousness loses sight of what it means. Every 'here', in being indicated, is revealed as part of a complex system of spatial relations: it is simultaneously 'above' something and 'below' something else, a 'right side' and a 'left side', an 'inside' and an 'outside', etc. Single spatial moments, like single temporal moments, are pure abstractions, artificially removed from a universal continuum where every element is intrinsically dependent on all others.

Just like the subject and the object of sense-certainty, the whole relationship is found to rest on universal and abstract elements, whose meaning lies elsewhere, in something other than themselves. This new acknowledgement leads consciousness to let go of this model. But although the concreteness posited at the start was refuted, it is nonetheless preserved as something 'envisaged' or 'meant' (*gemeint*) by consciousness. Although the progression's initial standpoint was overcome, this *Meinen* has not disappeared and still requires an appropriate mode of expression. In other words, consciousness realizes it 'meant something other than what [it] meant to mean', and this correction of its

meaning compels it to 'go back to the [initial] proposition, and understand it in some other way.'[336]

With this self-correction, the simplest form of cognition is transferred from the stage of sense-certainty to the stage of perception, or 'thinghood'. Whereas in the first model reality was a universal aggregate of simple sensuous units, in this new model it is a complex system of 'things', endowed with universal and mutually related properties. Just like before, the experience of this new stage follows a triadic scheme.[337] Firstly, the 'thing' is born out of the sublation of sensuous singularity. Unlike a single sensuous unit, each thing is both simple — it is *this* thing as opposed to *other* things — and universal — it partakes of a series of universal properties, shared by many other things. The first of these aspects is the thing's *negative* or *exclusive* unity. It is that in virtue of which a given thing is perceived as unique and self-identical: as an object whose colour, for instance, is *its particular colour* and not a generic colour, and whose shape is *its particular shape* and not a generic shape. The second of these aspects is the thing's *passive* or *inclusive* unity. It is that in virtue of which a given thing is perceived as an assemblage of universal properties: as an object whose colour is a mere instantiation of a generic colour, whose shape is a mere instantiation of a generic shape, etc. As consciousness explores these two aspects, it comes to recognize their incompatibility. If something is unique, the universality of its properties cannot be maintained. The thing's predicates must be subsumed under a truly exclusive and self-identical unity. Conversely, if something is multiple and varied, its uniqueness cannot be maintained. It ends up being lost, reduced to a more or less complex aggregate of universal properties. The 'thing' is thus an inherently contradictory object: its identity is caught between absolute selfsameness (nominalism) and absolute otherness (indetermination).[338]

As in the stage of sense-certainty, this first contradiction corresponds to the objective moment of experience. It is followed by a subjective moment, in which consciousness

[336] *HW* III, 59; *PS*, 39

[337] See *HW* III, 93ff.

[338] Hegel's argument leads back to Plato's notion of μέθεξις. Although the definition of a thing is necessarily supported by a series of εἴδη, or γένη (viz. redness, roundness, bitterness, etc.), each εἶδος or γένος can only be defined by providing examples of its individual occurrences (viz. a red apple, a round apple, a bitter apple, etc.). See notably *So* 251a-259c, *Re* 476a, *Phd* 100c, or *Phi* 15d.

takes responsibility for the thing's inconsistencies and attempts to save the perceptual model. However, in contrast to what happened in the first chapter, consciousness does not intervene as the essential moment of the relationship, but as the source of its untruth. More precisely, it is led to acknowledge that its apprehension of the thing is not a true or transparent apprehension (viz. not an actual *Wahrnehmen*, or *Nehmen des Wahren*), but a faulty or translucent one. Instead of perceiving the thing as it actually is, without adding anything to it, perceptual cognition adds a reflective element to the relationship, which distorts the thing's true image. The contradiction between universality and singularity is thus not inherent to the thing itself, but the result of an observational defect, which consciousness must now try to correct.

This correction is itself divided into two symmetrical moments, which correspond to the second and third parts of Hegel's overall argument. Firstly, consciousness accepts the thing's exclusive unity and regards its different properties as subjective additions. This means, for instance, that a grain of salt is not white in itself, but only in the eyes of the subject; that it is not cold in itself, but only because it is perceived as such; that it is not tart in itself, but only when the subject tastes it, etc. Things in themselves are only varied because the subject's corporeal senses break them down into different aspects and perceive them in different ways. This attempt to preserve the thing's simplicity is once again unsuccessful, as it fails to acknowledge that simplicity itself is always a *specific* or *determinate* simplicity, viz. that it is dialectically dependent on difference and otherness. There is no identity, however simple, without determination: for something to be what it is, it must be different from everything else, and this difference can only be found in its properties.

Consciousness then reverses its original position, accepts the thing's intrinsic multiplicity and recognizes itself as the source of the thing's unity. The grain of salt is indeed white *and* cold *and* tart — but *insofar* as it is white it is not cold, and *insofar* as it is cold it is not tart. It is the perceiving subject who, in order to escape the contradiction, isolates each of the properties and sets them up as independent categories. Once again, however, this hypothesis is untenable. If the link binding each thing to its properties is merely imaginary, the thing's identity is itself imaginary. Things are reduced to mere 'enclosing surfaces', which can hold an endless variety of contents.

By focusing alternately on the objective and the subjective poles of the relationship, consciousness attempted to isolate the contradiction between identity and difference, or singularity and universality, and prevent thereby the collapse of the perceptual model. But experience has shown that this contradiction is embedded in the heart of cognition, that it 'circulate[s] everywhere in every thought that is uttered'[339] and cannot be cast aside. Whereas the perceptual model is still guided by a tautological mode of cognition, in which self-identity is reduced to selfsameness, the *Phenomenology*'s third chapter introduces a different cognitive framework, in which selfsameness and otherness are intimately and irrevocably intertwined. The transition to the stage of understanding entails the discovery that there are no simple identities, or that every single identity implies its own negation.

Although the singularity pursued in the stage of sense-certainty passed over into universality, the latter is still, as Hegel puts is, a *conditioned* universality. With the endorsement of the perceptual model, the main contradiction of sense-certainty was not removed, but merely transferred to the more complex domain of thinghood. Perceptual universality is indeed more sophisticated than natural consciousness' naïve conception of singularity, but it is still at odds with the problem of determination: the thing is universal and *at the same time* determinate; it is one and *at the same time* varied. This universality is burdened with an antinomy that can only be overcome once the two competing terms cease to be posited alongside one another. The truth of perception is not disjunctive, but dialectical: it is not one of these two alternatives, nor the two of them together, but the very movement that separates and unites them. It is the continual dialectical transit whereby identity is turned into otherness, and otherness led back to identity.

With this new discovery, the simple universality of sense-certainty and the composite universality of perception are replaced by a genuine or *unconditioned* form of universality.[340] Whereas the two previous models attempted to maintain the subject's and/or the object's independence, proposing different forms of relationship, the understanding favours the relationship itself over its terms, or the dialectical flow over its

[339] *Phi* 15d4-5

[340] *Unbedingte Allgemeinheit*, or simply *das Unbedingte*. Hegel plays with the radical '*-ding*' and its negation — unconditioned universality is the negation of *Dingheit*, i.e. 'thinghood'.

limiting edges. This conceptual reversal introduces a new kind of object, i.e. 'force' (*Kraft*), which can no longer be thought of in simple or dichotomical terms. Unlike a thing, a force is a dynamic unity, whose identity is both self-identical and self-differentiating. [341]

The dialectical impasse which led to the implosion of thinghood is now recast as the internal duplicity of the same object: on the one hand, the movement whereby force loses its unity and is dissolved into the multiplicity of its properties is force's 'expression' or 'exteriorization'; on the other hand, the counter-movement whereby this plurality is reunified is 'force proper', or force 'driven back into itself'.[342] Unlike the previous model, however, the interaction between these two poles is not added to the poles themselves. It is the primary element of the relationship. Therefore, contrary to a thing and its properties, force and expression cannot be considered separately: a force exists because it is expressed, but its expression is itself a moment of force.

To fully understand Hegel's description, it is useful to mind the historical references scattered throughout this chapter. The transition from the dialectic of thinghood to the dialectic of force entails a departure from the world of common-sense and the embracement of a strictly metaphysical or 'scientific' explanation of reality. Whereas things and their properties, despite all of their contradictions, are nonetheless familiar determinations, forces belong to a different analytical level, no longer reducible to common experience. They are the operating units of a new approach to the natural world, whose historical emergence can be traced back to the 17th century and to the development of modern natural sciences. More specifically, the abandonment of the perceptual model and the discovery of the notion of force can be linked to Leibniz's critique of Descartes' mechanical physics. While the Cartesian model reduces objectivity to extension, and defines every object in terms of its shape, size and motion, Leibniz argues that the identity of objects cannot be grasped in this way. It requires a dynamic understanding of reality, grounded in the notion of force.[343] For Leibniz, shape and size

[341] See *HW* III, 107ff.

[342] *HW* III, 110

[343] See notably *Correspondance avec Arnauld*, ed. cit., 138: 'Le mouvement en tant qu'il n'est qu'une modification de l'étendue et changement de voisinage, enveloppe quelque chose d'imaginaire, en sorte qu'on ne saurait déterminer à quel sujet il appartient parmi ceux qui changent si l'on n'a recours à la force

are relative qualities, and hence an insufficient basis on which to ground the notions of identity and difference. Whereas a pure *res extensa* is nothing more than a multiplicity of physical or geometric atoms, joined together as a simple collection, the true nature of objectivity rests on a different kind of unity, which transcends the material realm. It is grounded in a series of self-determining principles — 'formal atoms', or monads — whose identity is not contingent or composite, but simple and necessary.[344]

Like Leibniz, Hegel replaces a world made up of things or atoms, whose qualities, movements and transformations are determined by external sources and circumstances, with a thoroughly dynamic world, in which all events and transformations are the 'expression' or 'exteriorization' of different forces. Moreover, like Leibniz, Hegel insists on the *global* or *systematic* nature of this model. Whereas natural consciousness tends to construe forces as regional or exceptional phenomena, added to an originally static or inert reality, the understanding takes over the entire realm of objectivity and envisages a world in which all traces of thinghood or atomism have been eliminated.

The transition from the consideration of a single force to the description of a global system of forces is not contingent or accidental, but a necessary result of the very notion of force. The starting point of Hegel's argument is the dialectical movement in which a given force expresses itself and retreats back into itself.[345] These two moments are part of the same force, but the passage from one to the other seems to imply the intervention of an external element. If the force were isolated, there would be no movement, for there would be no motor or motivation behind it. The unexpressed force expresses itself

qui est cause du mouvement et qui est dans la substance corporelle'; or *Discours de Métaphysique*, ed. cit., 229f.: '[La] force est quelque chose de différent de la grandeur, de la figure et du mouvement, et on peut juger par là que tout ce qui est conçu dans les corps ne consiste pas uniquement dans l'étendue et dans ses modifications, comme nos modernes se persuadent.'

[344] See, for ex., *Système Nouveau de la Nature*, ed. cit., 66f.: 'Je m'aperçus qu'il est impossible de trouver *les principes d'une véritable Unité* dans la matière seule ou dans ce qui n'est que passif, puisque tout n'y est que collection ou amas de parties à l'infini. Or, la multitude ne pouvant avoir sa réalité que des *unités véritables* qui viennent d'ailleurs et sont tout autre chose que les points mathématiques qui ne sont que les extrémités de l'étendu et des modifications dont il est constant, que le continuum ne saurait être composé. Donc pour trouver ces unités réelles, je fus contraint de recourir à un *point réel et animé* pour ainsi dire, ou à un Atome de substance qui doit envelopper quelque chose de formel ou d'actif, pour faire un Être complet. (...) Je trouvai donc que leur nature consiste dans la force et que de cela s'ensuit quelque chose d'analogue au sentiment et à l'appétit, et qu'ainsi il fallait les concevoir à l'imitation de la notion que nous avons des âmes.' See also BELAVAL, op. cit., 492-497.

[345] See *HW* III, 111f.

because it is *solicited* to do so by something other than itself; likewise, the expressed force retreats back into itself in response to a similar solicitation. However, when this external motor is examined more carefully, its action is found to coincide with the action of force. Like the latter, it also alternates between an unexpressed form (as it solicits the force to express itself) and an expressed form (as it solicits the force to retreat back into itself). The acknowledgment of this symmetry leads the first force to break up into two forces, which both express (and solicit), and are expressed by (and solicited by), the other.

But these two forces do not exist in isolation. They are part of a wider 'play of forces' whose scope is gradually increased, in the course of experience, until it coincides with the whole world. The identity of each force is expressed by another force, and the latter's identity is expressed in turn by another one. What emerges with this process is a truly systematic or organic mode of identity, in which every individual object is implicated in the definition of all others. In this global dialectical web, what everything is defined by its degree of interference with everything else.[346] There is no longer any fixed point or pole, no independent or self-sufficient source of meaning. Identity is henceforth absolute reciprocity — 'a pure being-posited by another' [*ein reines Gesetzsein durch ein Anderes*].[347]

As forces merge into each other, however, their reality is lost. Whereas force emerged as a determinate, albeit dynamic unit, endowed with a specific identity and a specific presence, its essence consists now 'simply and solely in this, that each *is* solely through the other, and what each thus is it immediately no longer is, since it *is* the other'.[348] The experience of force leads to a pure evanescence, or to a pure 'vanishing' (*Verschwinden*): force evaporates, as it were, into otherness, and what remains is merely the 'notion' or the 'thought' of it. This new acknowledgement is very important, and lays the ground for all of the chapter's subsequent developments. The truth of force is no longer an object, but

[346] In Leibniz's vocabulary, this systematic identity is determined by specific forms of *entr'expression* and *entr'empêchement*. See *Discours de Métaphysique*, ed. cit., 224: 'Ainsi une substance qui est d'une étendue infinie, en tant qu'elle exprime tout, devient limitée par la manière de son expression plus ou moins parfaite. C'est donc ainsi qu'on peut concevoir que les substances s'entr'empêchent ou se limitent, et par conséquent on peut dire dans ce sens qu'elles agissent l'une sur l'autre, et sont obligées, pour ainsi dire de s'accommoder entre elles. Car il peut arriver qu'un changement qui augmente l'expression de l'une, diminue celle de l'autre.'

[347] *HW* III, 114

[348] *HW* III, 115; *PS*, 86

an idea, devoid of objective content. Or, in other words, the truth of the understanding is no longer the unconditioned universal envisaged as an object (i.e. force), but the *idea*, or the *notion* of universality.

As in the previous phenomenological stages, the objective moment of experience gives way to a subjective moment, in which the object is no longer meaningful in itself, but only for consciousness. With the vanishing of force, the understanding loses its object, but this void cannot be filled in a simple or direct manner, by bringing in a new object or a new version of reality. To do so would be to go back to the progression's previous stages and to revive the world of things or the world of forces. This new transition no longer allows for any of that: having posited and tested different kinds of objects, consciousness was led from contradiction to contradiction, until the elimination of objectivity itself. What emerges, therefore, with the disappearance of force is the acknowledgement that objectivity — *all* objectivity — is unreal or illusory. The reality cognized by consciousness is a mere thought, or a mere notion, and *not* reality itself. It is the *appearance* of reality, devoid of objective truth.

Since the understanding is now reduced to this simple acknowledgement, its new object can only be envisaged in a negative way, as that which lies *beyond* appearance. This negative region, placed outside the reach of the understanding, is the 'inner essence of reality', or 'the true background of things'[349] Here, and nowhere else, is true objectivity to be found. Yet precisely because this new object is purely negative, it can only be conceived through what it is *not*, namely appearance. Therefore, the new object of the understanding is really a double object, or a dialectical relationship. It is 'the syllogism which has for its extremes the inner being of Things and the Understanding, and for its middle term, appearance.'[350]

This syllogism is the same one I have previously translated through the equation *R* (reality in itself) + *C* (cognition) = *R'* (cognized reality), and its different variations.[351] It is the basic formula of Kant's critical diagnosis, which Hegel now brings to the fore. As the understanding decrees the globally apparent nature of phenomenal reality, and transfers

[349] *HW* III, 116; *PS*, 86
[350] *HW* III, 117
[351] See above section 6.1.

truth to an 'inner essence', or an 'inner world', it replicates the main dualism underlying Kant's critique of natural reason. In both cases, consciousness is confined to a global cognitive perspective and bound to cognize things through the same transcendental filter. It is unable to step outside of itself, as it were, and peak at the world's 'inner essence'. The latter can only be conceived, therefore, as a pure *Jenseits*, or a pure *An-sich*.

But whereas Kant's diagnosis stops here, leaving consciousness at odds with pure negativity, Hegel aims to show that this standstill is not definitive, but provisional. 'Certainly', he claims, 'we have no knowledge of this inner world as it is here in its immediacy; but not because Reason is too short-sighted or is limited, or however else one likes to call it — on this point, we know nothing as yet because we have not yet gone deep enough — but because of the simple nature of the matter in hand, that is to say, because in the *void* nothing is known, or, expressed from the other side, because the inner world is determined as the *beyond* of consciousness.'[352] We find here an echo of the same criticism raised in the introduction: the critical tradition, in Hegel's eyes, is still grounded in an abstract or dichotomical conception of reason; by separating reason from reality, and building its whole argument upon this distinction, it takes for granted the very obstacle it seeks to overcome.

The remaining sections of the *Phenomenology*'s third chapter are devoted precisely to overcoming this obstacle. Initially, since the 'inner world' is defined as the opposite or the nothingness of appearance (*das Nichts der Erscheinung*), it is still an empty world. However, in the course of experience, it will 'fill itself out for the understanding'[353] and reunite with consciousness. Hegel's description of this process is extremely nuanced, and can only be presented here in outline.

Firstly, consciousness envisages the 'inner world' as the *law* that governs appearance.[354] Although the phenomenal world is dominated by difference and change, this difference and this change are not contingent or arbitrary, but necessary. They are the unstable image of a stable realm of laws. Nevertheless, although the understanding

[352] *HW* III, 117f.; *PS*, 88

[353] *HW* III, 119; *PS*, 89

[354] The inner world is the *law* [*Gesetz*] of phenomenal reality because the latter is *posited* [*gesetzt*] by it. Hegel plays with the similarity between *Gesetz* and different forms of the verb *setzen* throughout the entire chapter.

manages hereby to raise objectivity to truth, this first movement is not entirely successful, because this realm of laws does not yet extend to the entire realm of appearance. The phenomenal world retains a quantum of contingency which derives from the ambiguity of the very definition of law: on the one hand, the law is universal — it is the global law of appearance; on the other hand, it mirrors the variety of the phenomenal world — it is divided into the various laws that govern the physical world. Although only a universal law can ensure the unity of the phenomenal world, its universality is too broad to determine anything specific. Conversely, a law that seeks to translate the specific nature of phenomena splits up into a contingent set of laws, whose actuality is altered with every change of circumstance.

The contradiction arises from the opposition between amplitude and determinacy: the broader the law, the less determinate its content, and vice-versa. The law by which a stone falls and the law by which heavenly bodies move can be joined under the same general law, namely universal attraction. But since the latter applies to everything, its content is no longer determinate. It no longer expresses the two previous laws, but merely the *pure notion of law*, i.e. the generic declaration that 'all reality is conformable to law'. The discovery of this pure notion leads to a transformation of the dialectic's initial terms: instead of the contrast between a general law and a multiplicity of particular laws, consciousness focuses on the contrast between the different laws and their *notion*, or their *inner necessity*. In the case of universal attraction, the law's notion is the simple factor 'gravity', but the law of gravity says something more specific, namely that two bodies are attracted to one another by means of a force that is directly proportional to their mass and inversely proportional to the distance between them. Likewise, *electricity* corresponds to a simple notion, but the law of electricity postulates the attraction between a positive and a negative electric field. In both of these examples, the law does not follow necessarily from its notion. There is no real necessity because nothing in the notion anticipates the specific form assumed by the law: there is no logical reason why gravity should stipulate this specific relation and not another, or why electricity should be divided into positivity and negativity. If these things are necessary, their necessity is not logical, but factical and arbitrary.

The difference between law and notion is not necessary because it does not belong to the law itself. Law and notion are still indifferent to one another, and their unity must be

brought in from the outside. Consequently, the phenomenal world is not yet reducible to the inner or supersensible world, and the separation is maintained. The attempt to correct this discrepancy leads to the subjective moment of the dialectic of law. In order to free its object from contingency, the understanding declares the difference between law and notion to reside only in itself. Law and notion are not different in themselves, but only for consciousness, and their difference is hereby eliminated. This procedure amounts to what Hegel names *das Erklären*, i.e. the 'explanation model', and consists in equating the law's factical or external necessity with the notion's logical or intrinsic necessity: a law has this specific form because it is the immediate translation of its notion; and the notion is nothing other than the need for this kind of translation. Hence, for instance, the reason why electricity is divided into opposite poles lies in a notion defined precisely by such a division. In other words, electricity is divided into opposite poles because it is divided into opposite poles. Law and notion are equated, but in a merely tautological way, which reveals nothing concrete about either one of them. The ruse of this strategy consists in positing two beings that are not really different in order to decree, afterwards, their coincidence. It amounts to a disguised reiteration of the principle of identity.

The phenomenological critique of the 'explanation' model evokes a wider Hegelian motif, taken up in different part of the *Phenomenology* and elsewhere. In denouncing the tautological nature of this procedure, Hegel is calling attention to what he views as the general one-sidedness of conventional scientific thought. In the preface, this same theme comes up in connection with the contrast between mathematical and philosophical truths.[355] The static or 'lifeless' nature of mathematics and the so-called exact sciences, which Hegel opposes to the dialectical or self-differentiating nature of philosophical knowledge, derives from a simple understanding of identity and difference. Mathematical theorems amount to more or less complex developments of the same fundamental tautology, and can all be reduced, ideally, to the identity principle. They aspire to a kind of truth that is measured solely by its degree of self-coincidence, or by its power to exclude difference and otherness.

[355] Cf. *HW* III, 42ff.

The same is valid for the understanding, at this particular stage of the progression. Since the alternative between law and notion is still construed in simple or dichotomical terms, consciousness can only unite the two by reducing them to a simple self-coincidence. Law is the same as notion because notion is the same as law: this 'is an explanation that not only explains nothing, but … while it pretends to say something different from what has already been said, really says nothing at all, but only repeats the same thing.'[356] The contrast between law and notion can only be overcome through a radical transformation of the identity model employed by the understanding. Instead of construing difference as the *external difference* between two selfsame units, artificially cancelled or 'explained' away, consciousness must look for difference *within* identity — it must come to realize that difference is really *internal difference*, or that identity is really *self-moving*, or *self-differentiating*. If both terms are considered separately, the antithesis between law and notion, or appearance and truth, is bound to reappear.

This realization takes place throughout the dialectic's third and final stage. Initially, the equation between law and notion leads to a tautology. The latter, nevertheless, when subjected to a closer examination, turns out to be more than a simple or static self-coincidence. In Hegel's words, while the explaining procedure, as regards the object of the understanding, 'gives rise to nothing new', as regards the understanding itself it reveals 'the very thing that was missing in the law, viz. the absolute flux itself [*den absoluten Wechsel selbst*].'[357] Although the immediate unity between notion and law appears at first as a mere tautology, 'what is present here is not merely bare unity in which *no difference* would be *posited*, but rather a *movement* in which *a distinction is certainly made but*, because it is no distinction, *is again cancelled*.'[358] In order to demonstrate the necessity binding notion and law together, the understanding is forced to posit their difference and to negate it afterwards. As it does so, however, it fails to see the necessity of these two operations. Focused on its attempt to guarantee the unity of the supersensible world, it does not yet distinguish what is really necessary in the whole procedure: namely, the very movement whereby difference is *simultaneously* posited and

[356] *HW* III, 126; *PS*, 95
[357] *HW* III, 126; *PS*, 95
[358] Ibid.

cancelled. The recognition of this necessity is now revealed to the understanding: 'In the process … of explaining, the to and fro of change, which before was outside the inner world and present only in the appearance, has penetrated into the supersensible world itself. […] The Understanding thus *learns* that it is a *law* of *appearance itself*, that differences arise which are no differences, or that what is *selfsame repels* itself from itself.'[359]

Looking back on the whole chapter, the strategies devised by consciousness to deal with the collapse of thinghood — i.e. the notion of force, the difference between force's expressed and unexpressed form, the opposition between a soliciting and a solicited force, the contrast between the world of appearance and an inner or supersensible world, the antithesis between universal and particular laws, and finally the difference between law and notion — were all grounded in a binary model, in which identity and difference were kept on opposite sides. To be sure, inasmuch as all of these solutions were born out of the sublation of thinghood, they were no longer based on the rigid atomism of the first and second chapters, but on a fluid or dynamic model, in which identity and difference remained in constant contact with each other. Nevertheless, this reciprocity was still determined by two different poles, defined in increasingly sophisticated ways. They were still the primary source of movement, the grounding elements of the understanding. Now, on the contrary, movement itself emerges as the primary element. It is no longer an effect, or a result, but the *universal law* whereby anything whatsoever comes to be known.

As the movement between identity and difference is transferred to the inner world, the latter's definition is radically altered. Whereas the first law of appearance stood for a stable image of the phenomenal world, this new law decrees the suppression of all stability. What is selfsame, *by virtue of its very identity*, repels itself from itself, and what is different, *by virtue of its very difference*, draws itself to itself. Identity is immediately resolved into difference, and difference is immediately resolved into identity. This new law corresponds, therefore, to a second supersensible world, directly opposed to the first one. This 'inverted world' has absorbed in itself the difference which the first world had

[359] *HW* III, 126f.; *PS*, 95f.

left over to the realm of appearance, and in doing so reclaimed the quantum of contingency that the phenomenal world insisted on keeping to itself.

In this second supersensible world, everything is contrary to the first one: what in the first world was selfsame is now different; what was positive is now negative; what was white is now black, what was honourable is now dishonourable, etc. However, it is important to look closer at what is at stake in this contrast.[360] At first sight, considering Hegel's examples, it might seem that the inverted world is opposed to the first world as positive is opposed to negative, or black to white. If this were so, the two worlds would correspond to two opposite sets of contents, which would lead once again to the contrast between two independent things, or two independent poles, and hence to the previous stages of the progression. But the discovery of a universal law of difference no longer allows for this kind of reading. What is essential, then, about the notion of *verkehrte Welt* is not the opposition between two things, but between two cognitive regimes. The two worlds are not endowed with simple, but opposing identities — rather, they stand for different approaches to the very notion of identity: whereas the first world possessed a simple identity, the identity of the second one is at the same time the opposite of itself. Consequently, the inverted world is not simply a new world, but the dialectical union of the first world and itself:

> One must think pure change, or *think antithesis within the antithesis itself* [*die Entgegensetzung in sich selbst*], or *contradiction*. […] The supersensible world, which is the inverted world, has at the same time overarched the other world and has it within it; it is *for itself* the inverted world, i.e. the inversion of itself; it is itself and its opposite in one unity. Only thus is it difference as *inner difference*, or difference *in its own self*, or difference as an *infinity*.[361]

While in all of the previous stages difference was construed as the disagreement between selfsame units — sensations, things, forces and laws — such units are no longer conceivable. Wherever one looks, difference has already taken over. It is the fabric in which reality is sewn. This discovery provides an answer to the old debate concerning the

[360] For a detailed discussion, see notably GADAMER, op. cit., 31-47; FLAY, op. cit., 662-678; and MONETI, op. cit.

[361] *HW* III, 130f.; *PS*, 99. Translation altered.

origin of difference, viz. the issue of how difference can be said to emerge or *issue forth* (*herauskommen*) from identity. For Hegel, this question is already biased: difference is not a secondary or supervening element, but the ground on which all things originally rest. It is not external, but *embedded* or *inlaid* in the very notion of identity. Therefore, what appears simple or selfsame is already the result of a self-division. The Parmenidian ἕν is a mere abstraction, a divided moment whose otherness has been negated. Like all independent unities, it can only be thought of as such by being removed from the infinite flow of identity and difference, or from the 'absolute unrest of pure self-movement [*die absolute Unruhe des reinen Sichselbstbewegens*].'[362]

With the notion of infinity, or absolute difference, the understanding's journey comes to an end. 'Law completes itself into an immanent necessity, and all the moments of [the world of] appearance are taken up into the inner world.'[363] Whereas initially the sensible world was regarded as the real world, and the supersensible world as an empty and unknowable *Jenseits*, what now comes to light is the very opposite. The phenomenal realm is unmasked as the mere reflection of a law issued by consciousness, a law that posits the universal difference of natural phenomena. Difference is no longer something that occurs or is generated in a contingent or circumstantial way: it is itself the law that makes reality what it is. Accordingly, the phenomenal realm no longer exists by its own right, as something contingent or arbitrary, while the understanding struggles to translate it through mental rules and ideal patterns. The emptiness that was first placed on the 'inner' side of things is now transferred to the outside world, and eliminated.

At this stage, all that is left for consciousness is to return to its initial syllogism and draw out the final consequences. Since the phenomenal world (*R'*) has lost its independence and vanished into the inner world (*R*), the understanding (*C*) and the inner world are now directly facing each other (*C* = *R*). Consciousness realizes at last that the absolute interlacement of identity and difference is not an attribute or a feature of the objective world, but the founding truth of objectivity. It realizes, in other words, that its true object is nothing other than its own law, or rather *consciousness itself*. Therefore, 'behind the so-called curtain which is supposed to conceal the inner world, there is

[362] *HW* III, 133; *PS*, 101
[363] *HW* III, 130; *PS*, 99

nothing to be seen unless *we* go behind it ourselves, as much in order that we may see, as that there may be something behind there which can be seen'.[364] The dichotomy is finally lifted and a new shape of consciousness is born. The understanding now knows that its object, i.e. reality itself, is really only its own self: *the truth of consciousness is self-consciousness*.

7.2. *The dialectic of self-consciousness*

The emergence of self-consciousness puts an end to the naïve objectivism that has dominated, in one way or another, all of the previous phenomenological models. In the stages of sense-certainty, perception and understanding, truth was placed outside of consciousness. Reality was detached from cognition and confined to a series of increasingly sophisticated objects, corresponding to a series of increasingly sophisticated worldviews. Experience has shown, however, that truth cannot be conceived in this way. What was deemed true in itself turned out be so only for consciousness, and the realm of objectivity passed over into subjectivity. Consciousness found out that the truth of its previous objects is its own self, and that all knowledge is self-knowledge, or self-consciousness.

It might appear, at this point, that the whole progression has come to an end. If the divide between reality and cognition has indeed been surmounted, and if consciousness is now simply the coincidence of the self with itself, the truth searched for in the previous stages seems to have been found. The simple empiricism endorsed at the outset resolved into absolute idealism, and what lies ahead is only pure self-contemplation. If this were the case, the cognitive confinement denounced by the critical tradition would have been overcome. Consciousness would finally be free, albeit in a radically subjective way.

This is not the case, however, because objectivity was not simply eliminated, but sublated, or *aufgehoben*. The objective world was negated, but it was simultaneously preserved as a moment of self-consciousness, and raised to a new cognitive stage. To

[364] *HW* III, 135f.; *PS*, 103

grasp the meaning of this transition, we must return briefly to the understanding's final metamorphosis. As seen earlier, self-consciousness emerged out of the emptiness of the phenomenal world. When difference was raised to absolute difference, or to infinity, the inner world was finally filled. It was revealed, at last, as the universal law of the understanding, or as consciousness itself. Yet precisely because this law is the law of absolute difference, or infinity, the coincidence between consciousness and reality is not a simple or immediate coincidence, but an *infinite* one ($C \leftrightarrow R$). Consciousness is reunited with reality, but this reunion takes the form of a dynamic or dialectical exchange: on the one hand, it distinguishes itself from reality, and posits the latter as its object; on the other hand, it is immediately aware that this object is nothing other than itself, and cancels its initial distinction.[365] In this toing and froing, *self* and *other* are both equally essential moments and both completely dependent on each other. As if facing a mirror, consciousness is only conscious of itself by putting an 'other' ahead of itself and distinguishing itself from it; but this positing is only possible because consciousness is already conscious of itself to begin with — which in turn implies the sublation of a previously posited 'other', and so on indefinitely.

Unlike a thing, or a substance, self-consciousness is both identical to, and different from, itself. It is what it is through the simultaneous suppression of what it is not. In this endless circularity, the self-conscious 'I' is not a fixed or permanent ground, but merely the negation of otherness, and likewise the tautology 'I am I' really only means that 'I am not another'. The 'I' of which identity is predicated is nothing more than the 'return from *otherness*' (*die Rückkehr aus dem Anderssein*)[366], or *what is left* when all otherness was suppressed and subjectivity was restored. However, since the 'I' is defined negatively by

[365] The movement whereby consciousness distinguishes itself from itself and the counter-movement whereby it reunites itself with itself are not two different, chronologically successive movements. The interchange between self-identity and self-otherness is absolute: at any given moment, consciousness is already simultaneously in the process of dividing itself and returning to its unified form. It is this absolute simultaneity, reminiscent of Fichte's self-identity *Trieb*, that Hegel aims to highlight when insisting on the *immediate* character of this double movement: 'Ich unterscheide mich von mir selbst, und es ist darin *unmittelbar* für mich, daß dies Unterschiedene nicht unterschieden ist. Ich, das Gleichnamige, stoße mich von mir selbst ab; aber dies Unterschiedne, Ungleichgesetzte ist *unmittelbar*, indem es unterschieden ist, kein Unterschied für mich.' (*HW* III, 134f. Emphases added.)

[366] *HW* III, 138; *PS*, 105

what it is not, it is doomed to return to its opposite and to distinguish itself from it. With every 'I', a new 'other' is revealed, and objectivity is once again revived.

As it first comes on the scene, self-consciousness is torn between these two moments. While its truth is its own self, or pure subjectivity, the latter can only be achieved by returning to objectivity. In this centrifugal movement, the sensuous world is preserved for consciousness, but no longer as a positive or independent world. 'It is an enduring existence which, however, is only *appearance*; or a difference, which, *in itself* is no difference.' We have here the same syllogism discussed in the previous stage, save for one important difference. While in the previous case the syllogism had for its middle term appearance and for its extremes consciousness and reality, in this case the middle term is also appearance, but both extremes are occupied by consciousness. In other words, consciousness is now the unity of the self with itself mediated by appearance. In order to find its truth (i.e. self-consciousness), it must return to the sensuous world, exposes appearance *as* appearance and thereby eliminate it. This movement, at this stage of the progression is called *Begierde* (i.e. desire, or appetite).

The notion of desire is the starting point of a new dialectical cycle — the dialectic of subjectivity, or selfhood, in which the 'I' will be raised from a mere vanishing point to a more substantial form of identity. Throughout this new set of developments, Hegel aims to show that what consciousness knows in knowing itself is really much more than it initially supposes. Instead of a simple or selfsame being, lost in the pure contemplation of itself, self-consciousness is already a movement, or a *Trieb*, or a peculiar form of desire. The key to grasping its true identity lies, therefore, in the examination of this movement, which requires both the analysis of different definitions of desire and the analysis of the different objects required to satisfy them. As usually, Hegel's exposition follows two simultaneous threads, focusing on the subjective and on the objective implications of each phenomenological transition.[367]

To begin with, self-consciousness is 'desire in general' (*Begierde überhaupt*).[368] That is to say, it is the *desire to be itself*, or to raise itself out of otherness. On the subjective side,

[367] For a more detailed version of the following exposition, see my *A Compreensão Hegeliana do Desejo* (*Begierde*), ed. cit., esp. 182-195.

[368] *HW* III, 139; *PS*, 105

this movement is accomplished by suppressing objectivity and re-establishing the unity of the self with itself. However, as we have seen, this procedure is inherently self-defeating: if self-identity is still nothing more than the opposite otherness, the subject's self-coincidence is at the same time self-alienation. The satisfaction of desire can only be obtained by reviving, again and again, the original source of desire. On the objective side, the same circularity is repeated, but in the opposite direction: if otherness is still nothing more than the opposite of self-identity, the object is merely a 'not-I', reduced by consciousness to itself.

Unable to find satisfaction in this way, consciousness abandons this first solution and moves on to a new dialectical equation. This time around, however, the transformation is brought about through the relationship's objective side. Since the self-unification afforded by the negation of the first object led to a dialectical impasse, the object must be defined in a different way. For desire to be truly satisfied, objectivity must be construed as something more than a mere abstraction, or a fleeting moment. With this new acknowledgement, 'self-consciousness which is simply for itself, and directly characterizes its object as a negative element, or is primarily *desire*, will therefore, on the contrary, learn through experience that the object is independent.' More specifically, it will learn that the object has 'not merely the character of sense-certainty and perception, but [is] being that is reflected into itself', or 'a living thing' (*ein Lebendiges*).[369]

This new development evokes, once again, one of Hegelian philosophy's most important themes, already echoed in the previous phenomenological stage, and discussed at length in many other texts.[370] The dialectical contrast between life and death stands, in Hegel's writings, for the wider conceptual contrast between an abstract or tautological mode of identity and a dynamic or dialectical one. In the element of organic life, identity and difference are not rigid, but fluid determinations. They are the limiting edges of the ceaseless alternation between union and dissolution, concentration and dispersion, growth and decline. Accordingly, living organisms are not mere things, or collections of things, but self-moving or self-determining beings. Each organism is *at the same time* the

[369] *HW* III, 139f.; *PS*, 106

[370] Apart from the *Phenomenology*, see, for ex., *HW* VI, 436-461, 469-486; *HW* VIII: 359-365, 373-377; or *HW* IX, 337-539.

unity of its members and their self-differentiation, the universal fluidity of life and the division into new members and new particular beings.

It should be noted, however, that Hegel's description, in the passage we are now considering, does not apply simply to a new class of objects, juxtaposed to self-consciousness' first version of objectivity. What is here at stake is not a change of focus from a universe of 'dead' beings (viz. sticks, stones, etc.) to a universe of 'living' ones (viz. plants, animals, etc.). The term 'life' applies not to a part of objectivity, but to the whole of it, and the present phenomenological transformation is therefore, as always, a *global* transformation. The initial conception of objectivity has now been abandoned because *all* objectivity is, after all, *life*. Consciousness realizes that the distinction between identity and difference it had ascribed to the natural world is groundless, and that objectivity is really a dynamic, self-developing whole (*ein sich entwickelndes Ganze*).

On the subjective side of the relationship, consciousness has also been transformed, and has also acquired a new depth. The fleeting self of the initial model is now an organic unity, part of a wider organic whole. Its identity is no longer simple or abstract, but concrete or 'self-reflected', and otherness is for it no longer a purely negative limit. This new self contains its other within itself, and its life consists in the dialectical alternation between the two. On the one hand, the subject loses itself in the multiplicity of the objective world and is led to assume different shapes and functions. This expansion is mirrored by nature itself, which also displays a succession of independent shapes and operations. On the other hand, consciousness knows that this independence and this differentiation are nothing *in themselves*. The truth of the objective world is still self-consciousness, and the independence of its different shapes is a spurious independence, which must be negated and reabsorbed.

Hegel's somewhat cryptical description of this new form of desire may be illustrated by a variety of physiological examples: consider, for instance, the need for oxygen and water that are daily felt by every living being, and usually overcome through breathing and drinking; consider, moreover, the example of hunger, overcome by the consumption of plants and animals. In these cases, desire is not optional or arbitrary. The subject's survival (and the survival of self-consciousness) is actually dependent on the consumption and self-integration of external objects or elements (viz. air, water, plants, animals, etc.) Yet this concession to objectivity does not lead to the elimination of the subject. On the

contrary, consciousness is still the essential moment of the relationship, and otherness is merely the means or the instrument whereby the self reaffirms its sovereignty.

Hegel returns to this theme in the second part of the *Encyclopaedia*, dedicated to the Philosophy of Nature, and the analysis of this later text sheds a revealing light on the *Phenomenology*'s present stage. In the *Encyclopaedia*, consciousness' first contact with the natural world is also mediated by *Begierde*, and the suppression of this initial longing depends on the human ability to 'use nature to [one's] own profit, … that is, to destroy it.'[371] In a way, the history of human science and human technology is the history of desire and the different ways of satisfying it. But although consciousness develops increasingly complex ways of consuming and absorbing the natural world, its ultimate end is not the object of desire, but the satisfaction is affords. The truth of the object as object of desire does not reside in itself, but in the use that consciousness makes of it. Each natural thing and each living being exist on their own right, but only because they exist *for consciousness*: water, before it is a chemical compound, is 'that which allows human beings to quench their thirst'; and plants and animals, before they are independent living organisms, are 'that which allows human beings to quench their hunger.' Likewise, the end of desire is 'our satisfaction, or our self-feeling [*unser Selbstgefühl*], which was perturbed by a privation of some sort. The self-negation which occurs in me through hunger is at the same present as an other [*als ein Anderes, als ich selbst bin*], which must be consumed; my action is to sublate this opposition and make this other identical with me, or to restore, by sacrificing of the thing, the unity with myself.'[372]

In the case of hunger, this 'return to selfhood' is achieved in a literal way: animals and plants, in being consumed and digested, become 'flesh of my flesh' and 'blood of my blood'. Their identity is merged in my own identity and becomes the subject of a new desire, aimed at a new object. But this movement, repeated over and over again, rests on a circularity that is similar to the one found in the previous model. The end of desire is indeed the satisfaction obtained through the suppression of objectivity, but for this suppression to take place there must be a world to suppress. Just like before, the very

[371] *HW* IX, 13
[372] *HW* IX, 14

form of the relationship leads to the continual resurrection of the original source of desire. Self-consciousness is doomed, therefore, to an eternal state of dissatisfaction — a fate which Hans-Georg Gadamer, in his reading of this chapter, aptly compared to the perpetual oscillation in Goethe's *Faust*:

> So tauml' ich von Begierde zu Genuß,
> Und im Genuß verschmacht' ich nach Begierde.[373]

In order to overcome this circularity, self-consciousness must abandon this model and redefine, once more, the nature of desire and the meaning of its object. Just like before, this new transformation is brought about through the objective side of the relationship. Although consciousness' previous object was no longer a thing, or a pure abstraction, but a living or self-moving being, it proved unable to stop desire's vicious circle. The satisfaction of thirst and hunger are necessary conditions for the maintenance of self-consciousness, but they are not yet its essence or truth. If the self's return to itself is not to end, once again, in circularity, it must be countered by an equally strong symmetrical movement, capable of opposing subjectification. This kind of resistance cannot be offered by things or living beings, for the latter, as was shown in the course of experience, are ultimately but mere means or instruments for satisfaction. This counter-movement can only be ensured by an object that is no longer a means or an instrument in the aforementioned sense, but an end in itself — by an object that 'is in its own self negation, and in being so is at the same time independent.'[374] In other words, this counter-movement can only be ensured by *another subject*, or by *another self-consciousness*.

In this crucial turning point, the subjective and the objective moments are finally brought together and desire's unilateral logic is overcome. The 'I' is no longer faced with a 'not-I', destined to be consumed or subjectified. Its object is now 'another I', which is itself faced with a self-conscious subject, and the dialectic of desire gives way to the dialectic of recognition.

[373] 'Thus I reel from desire to satisfaction, / and in satisfaction I long for desire.' (GOETHE, *Faust. Erster Teil: Wald und Höhle*, vv. 3249f.) Cf. GADAMER, op. cit., 53f.

[374] *HW* III, 144; *PS*, 109f.

7.3. *Selfhood as recognition*

Each self-consciousness is what it is by recognizing, and being recognized by, another. Desire is no longer directed at an objective content placed outside the sphere of subjectivity, but at another subject's desire, or at another subject's demand for recognition. This new model re-enacts, at a higher level, the self-duplication of force, examined in the previous chapter. Just as force's expression was at first solicited by an external element, and found afterwards to depend on a second symmetrical force, the self now breaks up into two selves, whose identity is irrevocably intertwined. Each subject, in being conscious of itself, is simultaneously conscious of the other subject's self-consciousness. Once again, the most basic cognitive unit is already a double unit: to speak of *one* recognizing self is already to speak of *two* recognizing selves and of their mutual articulation.

This transformation is encapsulated in two short sentences, which mark the beginning of a wholly new dialectical cycle:

> A *self-consciousness exists for a self-consciousness*. Only so is it in fact self-consciousness; for only in this way does the unity of itself in its otherness become explicit for it.[375]

But this is not all. This first revelation is immediately followed by a second one, whose full implications will only become apparent much later. Just as the play of forces was not reduced to two forces, but extended to a universal web of forces, so too will the dialectic of self-consciousness eventually lead to a similar universalization. As the remaining stages of the progression will render more explicit, to speak of *two* mutually recognizing selves is already to speak of a *global system* of mutually recognizing selves. The discovery that singularity is rooted in universality, or that subjectivity is rooted in intersubjectivity, is the first defining feature of the Hegelian notion of Spirit, or *Geist*:

[375] *HW* III, 144f.; *PS*, 110: 'Es ist ein *Selbstbewußtsein für ein Selbstbewußtsein*. Erst hierdurch ist es in der Tat; denn erst hierin wird für es die Einheit seiner selbst in seinem Anderssein.'

> A self-consciousness, in being an object, is just as much 'I' as 'object'. And with this, the notion of *Spirit* is already available to us. What still lies ahead for consciousness is the experience of what Spirit is — this absolute substance which is the unity of the different independent self-consciousnesses which, in their opposition, enjoy perfect freedom and independence; 'I' that is 'We' and 'We' that is 'I'.[376]

This new thesis, undoubtedly one of Hegelian philosophy's most original and controversial motifs, amounts to nothing less than a complete reversal of the way in which we tend to conceive personal identity. Instead of an original self to which the consciousness of an other is added, Hegel argues that the latter is the very condition for the emergence of the former. Intersubjectivity is not a result, arising from the mutual interaction of simple autonomous selves, but the founding instance of personal identity. Self-consciousness is from the very start a social or collective phenomenon.

But this thesis is far from self-evident. Indeed, it seems to contradict the very experience of subjectivity: if my own self stands for the most intimate and secret dimension of my whole existence, something that only I know and only I can communicate, it is hard to accept that it might be in any way subsidiary of an external act of recognition. Moreover, if there is no individual self prior to otherness, a self which would oppose otherness and take note of that opposition, how can the acknowledgement of the existence of an 'other' ever come about? These and other perplexities bear witness to the persistence, within human consciousness, of the *Phenomenology*'s previous dichotomical framework, and tend to cloud one's understanding of this new cognitive stage. Any interpretation of this passage that assumes the priority of self-consciousness over the demand for recognition entails a return to the very misconception that is now being criticized. And yet this is precisely the kind of approach our ordinary standpoint tends to fall back on. The claim that consciousness is only self-conscious in the wake of a social exchange is liable to be read in light of the simple contrast between realism and solipsism. Since our usual account of reality is dependent (at least to some degree) on the belief in the actual existence of other people, on the assumption of a shared ground of natural qualities, on a collectively agreed system of linguistic conventions, etc., intersubjectivity is certainly an essential feature of personal identity. But this kind of

[376] *HW* III, 145; *PS*, 110

reading only shows that sociality is a necessary condition of self-expression or self-actualization. It still allows for the possibility of an incommunicable but no less real core of selfhood, immediately intuited, but absent from the sphere of social relations.

According to this version of the argument, Hegel would be merely pointing out that human consciousness cannot develop in isolation — that each human being is a social being, a ζῷον πολιτικόν, and that personal identity is just as much defined by one's sense of self as it is by one's circumstances and encounters, by one's social context and cultural constraints, etc. Now although all of this may be true, Hegel's point is much more radical. The notion of *Anerkennung* is not merely an adjective notion, but a substantive one. That is to say, recognition is not to be envisaged as something merely added to self-consciousness, as one of its predicates or qualifications. Rather, it entails the very dissolution of the antithesis between selfhood and self-expression. At this point, self-consciousness is no longer an independent participant in a social exchange (viz. a subject alongside other subjects), but the connecting movement itself (viz. the act of recognizing, and being recognized by, other subjects). It 'is *in* and *for itself* when, and by the fact that, it is in and for itself *for another*; that is, it is [what it is] only in *being recognized as such* (*als ein Anerkanntes*).'[377]

Hegel's unique conception of recognition is the basis for the *Phenomenology*'s further developments. In the course of the progression, the dialectical grammar of self-identity, with its different and increasingly complex declensions, will coincide with the self-evolving grammar of human relations. The same demand for recognition will give rise to a long succession of cognitive shapes — from the original encounter between two recognizing selves to more complex forms of social interaction, and to the shifting roles of master, slave, citizen, vassal, believer, actor, artist and philosopher, among many others. For Hegel, 'it is the nature of humanity to press onward to agreement with others; human nature only really exists in an achieved community of consciousnesses'[378], and the phenomenological progression is the long road towards this achievement.

Before moving on, however, let us return for a moment to our initial query and determine how this new transformation fits into the more general debate concerning the

[377] *HW* III, 145; *PS*, 111. Translation altered. Emphasis added.
[378] *HW* III, 65; *PS*, 43

relationship between freedom, power and knowledge. At the beginning of the progression, consciousness was led to acknowledge the contradictions embedded in its usual mode of cognition and driven to a 'state of despair.' 'Hampered' by 'the so-called natural ideas, thoughts and opinions'[379], its original condition can be compared to that of the prisoners held inside Plato's cave and its initial self-criticism to the breaking of their chains. Afterwards, throughout the first three chapters, consciousness begins its ascent. The dialectic of objectivity is fuelled by the disagreement between reality *in itself* and reality *for consciousness*, or truth and appearance: whenever consciousness believes to have found the truth about its object, and to know the world as it really is, the latter turns out to be something completely different, and what seemed like the end of the path is revealed as a new beginning. This experience, re-enacted at different cognitive levels, leads to the emergence of self-consciousness, that is, to the complete integration of objectivity within the sphere of subjectivity. This turning point offers the first, albeit provisional solution to Plato's predicament. By abolishing the difference between reality and appearance, or between an 'inner' and an 'outer' world, consciousness enters 'the native realm of truth'.[380] In other words, it claims to have left the realm of δοξάζειν and entered the realm of ἐπιστήμη.

The experience of self-consciousness will then show that this victory is only apparent, since the 'I' is still irrevocably bound up in objectivity, stranded to the circular logic of desire. What follows, however, is not simply the collapse of another mode of cognition, but the introduction of a whole new philosophical theme, which seems to distance Hegel from Plato and the critical tradition. Indeed, the dialectic of desire leads to what Axel Honneth dubbed an 'intersubjectivist innovation'[381], devised to transcend the idealist framework previously set up by Hegel and to pave the way for what is arguably the *Phenomenology*'s most original philosophical contribution.[382] According to

[379] Ibid.

[380] *HW* III, 138; *PS*, 104

[381] *Die Kampf um Anerkennung*, ed. cit., 31

[382] Hegel's argument, however, is not altogether without precedent. The idea that individual self-consciousness is necessarily mediated by a social exchange is already found in Fichte's writings. Ever since the earliest formulation of his *Wissenschaftslehre*, dating from 1794, Fichte defines self-consciousness as an inherently dialectical process, fuelled by the continual interplay between a primordial core of subjectivity (*Ich*) and an opposing but equally primordial core of objectivity (*nicht-Ich*). In this model, it is the subject's

Alexandre Kojève, one of the first Hegelian exegetes to acknowledge the revolutionary nature of this transition, 'in having discovered the notion of recognition, Hegel found himself in possession of the key idea of his whole philosophy.'[383]

Whether or not one endorses Kojève's sweeping claim, the conceptual change introduced by the dialectic of *Anerkennung* is hardly innocuous. Whereas previously the *Phenomenology*'s fuelling principle was the disagreement between subjectivity and objectivity — the alternative between a purely subjective and a purely objective standpoint, or the search for some kind of compromise between the two — consciousness is now at odds with the disagreement between subjectivity and intersubjectivity. Its new task is to show that self-identity is not an individual phenomenon, but a collective one; and, moreover, that it encompasses not only two individuals, but the whole world. The fate of self-consciousness is the systematic process of 'giving oneself over and finding oneself in other'[384], and its final outcome is a universally systematic mode of identity, in which the 'I' is at last 'We', and vice-versa.

constitutional need to posit an 'other' ahead of itself that provides the 'impulsion' (*Anstoß*) necessary for the emergence of self-consciousness, as well as the basis for the recognition of an outside world. The 'I', torn between the positing of otherness, or *in itself*, and the acknowledgement that the latter is something merely posited, or *for itself*, is trapped within a 'circle' which can be developed and expanded, but cannot be escaped. (*Werke* I, ed. cit., 281) But although this account anticipates Hegel's characterization of self-consciousness, it can still be construed in a purely immanent manner. A more explicit approach to sociality and recognition can be found in Fichte's *Grundlage des Naturrechts*, where the notion of *Anerkennung* acquires a new and more decisive meaning. Individuality is defined as a 'reciprocal' or 'communal' concept (*ein Wechselbegriff* or *gemeinschaftlicher Begriff*) which 'can only be conceived in relation to another', and personal freedom is likewise grounded in a reciprocal or communal interaction (*Werke* III, ed. cit., 47f.)

[383] *Hegel, Marx et le christianisme*, ed. cit., 352. The change introduced by the stage of self-consciousness is an important element in the debate regarding the continuity of Hegel's text. The advent of self-consciousness brings about not only an inversion of the conceptual relationship between the subject and the object of human knowledge, but it also seems to transform the thematic scope of the *Phenomenology*: instead of the dialectical evolution of epistemologically-oriented 'shapes of consciousness', Hegel turns his attention to concrete self-conscious individuals, to their mutual interaction and to their physical and social surroundings. It was largely due to this change of context that the self-consciousness chapter became the *Phenomenology*'s most popular chapter and a privileged vehicle for Hegelian philosophy's long-lasting cultural impact. Influential authors such as Jean Wahl or Alexandre Kojève centred their entire readings of Hegel's progression around this chapter, treating the first three stages as a sort of critical prologue. Moreover, the notions of desire, recognition and labour introduced by Hegel were the basis for important developments of his philosophical project, from Marx's or Adorno's reassessment of its economic, political and cultural implications to Habermas' or Honneth's focus on the social and political significance of *Anerkennung*.

For a brief overview of the exegetical history of the *Phenomenology*'s self-consciousness chapter, see RAUCH, op. cit., 125-222.

[384] FULDA, *Der Begriff des Geistes bei Hegel und seine Wirkungsgeschichte,* ed. cit., 192

This new task does not imply, however, the elimination of the first one. Henceforth, the phenomenological progression will deal simultaneously will both of these themes: as we will see below in more detail, the transition from the stage of self-consciousness to the stage of Reason, and from the latter to the stage of Spirit, will lead to the acknowledgement both that the subject is not independent from the object and that the subject is not independent from other subjects. Spirit is consciousness that is 'conscious of itself as its own world, and of the world as itself'[385]; furthermore, it is 'the unity of the different independent self-consciousnesses which, in their opposition enjoy perfect freedom and independence.'[386] Once both of these truths have been acknowledged and harmonized, the experience of consciousness is replaced by the experience of Spirit, which will then lead the way to the progression's final standpoint.

[385] *HW* III, 324; *PS*, 263
[386] *HW* III, 145; *PS*, 110

8.

FREEDOM AND SERVITUDE

In the previous phenomenological stages, the contrast between subject and object, cognition and reality, for-itself and in-itself, etc., was spelled out in different and increasingly sophisticated ways. Throughout the *Phenomenology*'s first three chapters, in an effort to preserve the independence of the objective world, consciousness held on to the distinction between appearance and reality, or a phenomenal and a noumenal realm. With the emergence of self-consciousness, this distinction was transferred to the heart of subjectivity. Consciousness became its own object, or its own truth, and its knowledge became self-knowledge, or the unity of the self with itself. Finally, this new standpoint was also shown to be untenable. The unity of self-consciousness was compromised by the dialectic of desire and its truth was placed once again in the outside of world. This time around, however, experience ended up showing that consciousness' true object is not really an object, but another consciousness, which is itself both subject and object, for-itself and in-itself. The dialectic of self-consciousness gave way to the dialectic of recognition, which will now lead to a new string of transformations.

As always, consciousness' progress is slow and surgical. When the notion of recognition comes on the scene, it is not yet fully understood. Each self-consciousness sees itself as the essential moment of the relationship, or as a self that is only *recognized* faced with a self that only *recognizes.* Still guided by the logic of desire, each regards the other as a mere object, or a mere individual, confined to the contingency of natural life. Yet since the truth of self-consciousness has been revealed as a reciprocal exchange, this asymmetry must now be eliminated. In order to become truly self-conscious, each subject

must appear to the other as the opposite of an object or an individual. Each must be recognized as a pure act of recognition, which can no longer be seized and objectified. For this to happen, each must undertake 'the pure negation of its objective mode' and show 'that it is not attached to any specific *existence*, not to the individuality common to existence as such, that it is not attached to life.'[387] Only through this negative movement can the logic of desire be overcome.

As long as one's interest in life is maintained, one's existence is still a natural existence, determined by a natural drive for enjoyment and self-preservation. Although the self has risen to the stage of self-consciousness, it is still perceived as a natural being, and bound to be recognized as such. Once this self-attachment is gone, however, what emerges is pure self-determination, or a pure demand for recognition. Each individual shows the other he is willing to risk his own life in order to be recognized. But since this demonstration occurs simultaneously on both sides of the relationship, what takes place is a struggle for recognition, where both parties seek the death of the other and the confirmation of their own independence.[388]

This mortal struggle is the most basic form of social interaction. It is the first solution devised by self-consciousness to attain recognition. On closer inspection, though, it turns out to be self-defeating. For if either self-consciousness succeeds in

[387] Ibid. Hegel's idea is remarkably modern. By contrasting one's automatic or spontaneous attachment to natural existence and the deliberate effort required for the attainment of self-consciousness, or self-awareness, Hegel anticipates one of post-idealism's and post-modernism's central philosophical themes. Much like Kierkegaard or Heidegger would later do, Hegel is drawing attention, among other things, to the fact that self-consciousness is not given and cannot be taken for granted. It must be pursued or fought for, and can only be conquered by breaking free from the immediate logic of naturalism, viz. by questioning existence as a whole and facing death, for the first time, as a real possibility. This moment of absolute dread, which entails 'the absolute melting-away of everything stable' (*HW* III, 153; *PS*, 117), anticipates the role later played by anguish and despair in the existential tradition.

[388] Given the obvious disparity between the theoretical nature of the previous stages and the concreteness of this new dialectical interchange, it is tempting to interpret this mortal struggle as an actual depiction of humankind's remotest infancy. This temptation is heightened by the close similarities between Hegel's text and Hobbes' account of a primordial 'state of nature', famously defined as the 'war of all against all' (cf. *De Cive* I, 2; or *Leviathan* XIII). What is here at stake, though, is not a historical or anthropological description, but a specific cognitive model. By raising itself to the stage of recognition, consciousness has come to adopt a new perspective on its own understanding of reality, and more precisely on what the most immediate form of knowledge amounts to (see above section 6.3.). It has come to realize that the most primitive form of cognition is not sense-certainty, nor perception, nor the understanding, nor self-consciousness in its initial form. The true foundation of human cognition is a mortal struggle between two mutually recognizing selves. Although consciousness is not reduced to this primitive struggle, or to this primitive standpoint, all of its thoughts and judgements are now perceived as immanent developments thereof.

eliminating its rival, recognition itself comes to an end. Although the winning self does risk its life, it is no longer able to recognize, or be recognized by, the other, and its own selfhood sinks back to immediacy. Its victory is an empty victory, whose result leads back to the starting point.

The dialectical link between both contenders cannot be severed because self-identity and otherness are equally essential to the relationship. This initial strategy must therefore be revised. Although each self-consciousness must assert its independence by negating the independence of the other, each can only become independent by being recognized as such. Consequently, each must attempt to reduce the other to a state of nothingness which nonetheless preserves its ability for recognition. In other words, the struggle must not end in death, but in domination, and the vanquished party must submit to the winner and recognize his or her independence.

8.1. *Mastery and slavery*

With this new transition, the contrast between life and death gives way to the contrast between sovereignty and servitude. The mortal struggle envisaged in the previous stage is replaced by the confrontation between a master (*Herr*), or an 'independent consciousness whose essential nature is to be for itself'[389], and a slave (*Knecht*), or a 'dependent consciousness whose essential nature is simply to live or to be for another.'[390] Initially, what binds the master to the slave is the fear of death. Both have risked their lives in the struggle, but their willingness to sacrifice themselves was unequal. Whereas the master, in the face of death, refused to give up his independence, the slave chose to submit and save himself. His attachment to life is what holds him in bondage, the 'chain from which he could not break free in the struggle.'[391]

The simple dichotomy of the previous model was abandoned in favour of a more nuanced solution. Self-consciousness is now construed as a synthesis between dependence

[389] *HW* III, 150; *PS*, 115
[390] Ibid.
[391] *HW* III, 151; *PS*, 115

and independence, divided between a sovereign and a slavish standpoint. Accordingly, the initial opposition between selfhood and objectivity is no longer mutually exclusive and the natural world is readmitted once more into the relationship. The model becomes, therefore, a triangular model: apart from fear, the relationship between master and slave is also mediated by 'thinghood', that is, by the objective element which the master managed to overcome and which the slave remained attached to during the struggle. This 'thing', viz. the material world, is now the point of contact between both extremes, viewed differently from each side of the equation.

The master, on his side, engages in a mediate exchange with both remaining edges of the triangle: his relationship with the slave is mediated by the thing, inasmuch as he is now the thing's proprietor and forces the slave to work on it; likewise, his relationship with the thing is mediated by the slave, inasmuch the latter transforms the thing and makes it ready to be consumed and enjoyed by the master. Conversely, on the slave's side, the relational scheme is also twofold: the slave can only secure the master's mercy (and thus his own survival) through his work on the thing; however, his access to it is granted by the master, who own both the thing and the fruits of the master's work.

As usual, Hegel undertakes a systematic survey of the model's internal contradictions, in order to unveil its inherent instability and precipitate its downfall. To begin with, although the master did succeed in subjecting his rival and obtaining recognition, his victory is not as complete as he might have expected. To be sure, he managed to rise above objectivity and to achieve the satisfaction denied in the dialectic of desire. The slave's work, demanded by the master in exchange for his survival, transformed the latter's material existence into a life of sheer consumption and enjoyment. However, the master's independence is not grounded in his enjoyment, but in the recognition afforded by the slave, of which consumption and enjoyment are the indirect results. Recall that the main condition for the achievement of independence, at the beginning of the dialectic, was the recognition of each self on the part of another recognizing self, equally eager to prove his independence. Only two rivals willing to put their freedom above their particular existence were liable to grant each other, by means of a mortal struggle, the independence they both craved. This boldness, however, is precisely what the master no longer finds in the slave. After the struggle, the slave relinquished his independence and chose a life of subjection. Therefore, the recognition he is now able to offer is no longer

the recognition of a free or independent self-consciousness, but that of an unfree or dependent one. He can no longer provide the mirror, as it were, on which the master sees his own truth, and the latter's self-consciousness is once again compromised.

From the point of view of the slave, the situation is no less ambiguous. Although his subjection seems at first absolute, his defeat is nonetheless the result of a thwarted claim to independence whose essence is very similar to the master's. After the struggle, the slave was consigned to a position of inferiority, but he has also staked his life and has also faced the possibility of death. As Hegel puts it, he was fearful 'not of this or that particular thing, or just at odd moments, but [his] whole being has been seized with dread'. In this experience, he 'has been quite unmanned, has trembled in every fibre of [his] being, and everything solid has been shaken to its foundations'[392]. Now it should be reminded, once again, that this radical self-detachment, this 'rooting-out of all immediate being', is the very movement out of which self-consciousness is born. To rise above particularity, self-consciousness must abandon the realm of natural existence and divest itself from every trace of contingency. Only as absolute negativity, or as 'pure being-for-self', can it find its truth.

By coming face to face with death, the slave is left, but for one split second, with nothing left to lose. Yet this foretaste of freedom is soon cut short and his defeat forces him to submit once more to the contingency of natural life. Unlike the master's standpoint, however, which allows for no further self-development, the slave's standpoint is full of promise. He is now free *in posse*, and his work will eventually lead to the replacement of his current subjected self with a truly independent mode of identity. Servile labour will bring about the synthesis, within the same self-consciousness, of the master's and the slave's opposite stances.

Before looking closer at Hegel's argument, two brief remarks are in order. Firstly, to grasp the meaning of this new dialectical transition, it is useful to bear in mind Hegel's preliminary claims regarding the phenomenological progression's immanent necessity. Just like before, Hegel is not simply hazarding a set of possible scenarios, which may or may not ensue from the situation at hand. He is deducing the necessary outcome of the

[392] *HW* III, 147; *PS*, 117

previous stage's inner contradictions. And although it is certainly possible that servile labour leads to no liberation whatsoever, and that the inequality between master and slave is perpetually maintained, this possibility should be put down, following Hegel's logic, to self-consciousness' inability to grasp the implications of its own nature. The slave's labour harbours a dialectical tension whose logical outcome is freedom and independence. And likewise the master's sovereignty harbours a dialectical tension whose logical outcome is servitude and submission. But whether or not the latter takes place is a matter for speculation. The *Phenomenology* follows the slave's future metamorphoses but says nothing about the master's ultimate fate.[393] It is up to the slave, rather than the master, to carry the phenomenological torch.

Secondly, since the master's authority is dependent on the slave's activity, and since his enjoyment is in fact produced and conditioned by the slave, one might expect the master-slave dialectic to evolve in a symmetrical way and to culminate in an inversion of the initial dialectical situation: the slave would recognize his true power and rise to sovereignty, whereas the master would succumb to his impotence and become enslaved.[394] To be sure, Hegel is aware of this possibility and does argue that 'the *truth* of the independent consciousness is … the servile consciousness of the slave.'[395] However, he is also aware that this kind of strategy is ultimately unable to bring about an actual phenomenological transformation. Indeed, the acknowledgement that the master is no master and that the slave is no slave can only lead as far as a simple role reversal. And although this transformation is undoubtedly relevant — Marx and Engels will later argue that it is the motor of human history — it is still an immanent development of the same phenomenological model. It may be re-enacted again and again without ever questioning the basic premises upon which the model is built: new masters will be subjected to new slaves, the latter will give way to new masters, and so on indefinitely. Hegel avoids this impasse by reversing this basic strategy: the sublation of the master-slave scenario is not achieved through the slave's rebellion, but through his submission. By abandoning his

[393] Although in the *Encyclopaedia*'s master-slave dialectic, which amounts in many ways to a condensed version of the *Phenomenology*'s dialectic, Hegel suggests that the slave's liberation entails the liberation of the master. Cf. *HW* X, 226f.

[394] Cf. GADAMER, op. cit., 59f.

[395] *HW* III, 152; *PS*, 117. Translation altered.

initial demand for recognition and by exploring the true meaning of his new condition, the slave will undergo a more profound metamorphosis and discover a more sophisticated kind of freedom.

At first, the slave's independence is only the seed of independence, planted inside him by his brief confrontation with death. Although his attachment to natural existence proved stronger, during the struggle, than his drive for independence, the former will now be eradicated and the latter will regain the upper hand. This process is once again twofold, divided into an objective and a subjective side. Objectively, through manual labour, viz. through the concrete action of handling and reshaping external things, the slave opposes their independence. He draws them out of their immediate existence and grants them a specific form and a specific meaning. Which brings us to the subjective side of this process: through work, the slave comes to acknowledge this transformation as his own doing, and to find therein a new source of independence. His 'formative activity [*das formierende Tun*] is simultaneously the individuality or the pure being-for-self of consciousness, which now, in the work outside of it, acquires an element of permanence; it is in this way, therefore, that the working consciousness comes to see in the independent being [of the object] *its own independence.*'[396]

What Hegel has in mind is a peculiar form of transfer, in which the independence of the object is gradually recognized as the slave's own independence. More than manual labour in general, which can nonetheless be carried out in an automatic and thoughtless way, this procedure is perhaps best illustrated by the work of an artist. Initially, the object is nothing more than pure materiality. It is what it is: a brute presence, hostile to any kind of transformation. In this early stage, the work has to adapt to the object's mode of being, yield to the contingent peculiarity of its nature. But as the worker's efforts increase, matter takes on a definite shape and its initial hostility is overcome. The object's materiality retreats to the background and the author's design is brought to the fore. The object is no longer a mere thing, but an *artefact* or a *token* of the author's identity.

[396] *HW* III, 153f.; *PS*, 118. The use of *Form*, *Formieren* and other related expressions suggests, in the English language, an ambiguity that speaks to the ambiguity of this whole process: the slave *forms* the object in the sense of providing the object with a definite shape, but it is also *formed* by it in the sense of learning the truth about himself. Work *forms* the object but also the character, as it were; it is both a physical and an educational enterprise.

Like the master before him, the slave is now able to restore his independence vis-à-vis the object. But while the master's independence rested on a preformed object, worked on by the slave, the latter is the author of his newly shaped world. The difference between these two attitudes reflects, once again, the *Phenomenology*'s recurrent distinction between a simple or abstract conception of negativity and a concrete or dialectical one. Whereas the master's enjoyment amounts to a simple negation of objectivity, devoid of movement or mediation, the slave's labour brings about a determinate negation thereof, or an *Aufhebung*. The slave alienated his own identity, but only to re-integrate it within a wider dynamic whole. Through obedience and labour, he 'gave himself over' to another, but only to 'find himself' as its ultimate truth.

Looking back on this entire movement, it is easy to see why the *Phenomenology* would play such a decisive role in the future development of social and political philosophy. Hegel's dialectic of self-consciousness paves the way for the Marxist tradition not only because of its focus on the reversibility of power, or the master-slave dialectic, but also because of the specific emphasis it places on the notion of work, or labour. For Marx, as well as for Hegel, labour represents a force far superior to the master's sterile domination. In both cases, it is up to the worker to take possession of his work and, in doing so, to break free from alien domination. Unlike traditional Marxism, however, Hegel is not interested in alienation as such, but in self-alienation. More precisely, he is interested in the worker's voluntary subjection to an external form of authority. At this stage of the progression, work is necessarily servile, for only as such can it lead to a dynamic or dialectical form of liberation.

If the slave is to win back his freedom, he must first give it up altogether. The object of his work must appear to him as in a mirror, endowed with the same independence he himself longed for before the moment of subjection. If this initial contrast is not radical enough, the transfer of power will not be complete. The slave will hang on to a residue of self-identity and his negation will remain a simple one. In this case, his work may lead to an immediate or material liberation, but not to a true form of independence. Negativity will be directed at this or at that particular object, but not at objectivity as such. The

slave will be free, but his freedom will still be 'enmeshed in servitude' (*eine Freiheit, welche noch innerhalb der Knechtschaft stehenbleibt*).[397]

Absolute freedom requires, on the contrary, absolute fear. Only through such visceral, all-encompassing dread can work become a negation of negation and bring about the transformation envisaged by Hegel:

> If consciousness *forms* the thing without the initial absolute fear, it is only a vain, self-centred attitude; for its form or negativity is not negativity *in itself*; and therefore its formative activity cannot give it a consciousness of itself as essential being. If it has not experienced absolute fear but only some lesser dread, the negative essence has remained for it something external, its substance has not been infected by it through and through.[398]

But fear in itself is not enough. If the slave succumbs to a state of paralysing dread, his liberation is also compromised. Only servile labour can turn his fear into a positive and cathartic force. In order to stress this point, Hegel speaks of the 'discipline of service and obedience'[399], foreshadowing more recent and less ambiguous views on the virtues of servitude.

The unsavoury connotations of Hegel's argument were not lost on many of his critics. But although his emphasis on servitude is indeed controversial, one must not lose track of the argument's scope, nor of the role it is supposed to play in the *Phenomenology*'s overall progression. Firstly, it should be reminded that Hegel is not referring to an existing situation, nor is he advocating servitude as such. He is arguing, rather, that servile labour is a necessary stage in consciousness' immanent path of self-recognition. His argument is not strictly political, which is why all references to Marxism, contractarianism or authoritarianism must be handled with care. As always, in order to unmask each stage's inner inconsistencies, consciousness has to *experience* them and acknowledge their effects. Just as the life and death struggle was the logical follow-up of desire, so too servitude follows from consciousness' view of itself as a blind demand for recognition. In like manner, the next phenomenological stage will reveal that the present

[397] *HW* III, 155; *PS*, 119
[398] *HW* III, 154f.; *PS*, 119
[399] *HW* III, 154; *PS*, 119

model is itself inconsistent, and that true independence is to be found elsewhere. Therefore, the freedom provided by servile labour is but a transient, self-sublating construct, whose truth is a different and more sophisticated form of freedom.

Furthermore, it should also be noted that the kind of servitude Hegel refers to in the last three paragraphs of this section is no longer directly caused by the domination of the master over the slave. While in the previous paragraphs servitude was defined 'only in relation to mastery', it is now taken for what it is 'in and for itself.'[400] Even though the master emerged initially as the winning party, the slave is eventually revealed as the true protagonist of the relationship. For the slave exerts on himself, voluntarily, the action that the master had forced upon him, and by doing so drives the master into obsolescence. It is the slave who, in order to save his own life, renounces his independence and embraces a life of servitude. His 'absolute master', as Hegel puts it, is not the master, but life itself.[401] The master-slave dialectic is therefore the dialectical struggle of the slave with himself, or with his natural existence. In this immanent movement, the master acts merely as a sort of trigger: his true role is to exert the pressure necessary for the slave to retreat into himself and to re-emerge as a truly independent self-consciousness.

8.2. *Slavery and philosophy*

Through work, the slave fights off the resistance put up by the natural world. As work increases, and as the worker becomes more skilled, the object is divested from its objective nature. Its initial appearance becomes an increasingly faithful image of the worker's design, or an increasingly clear expression of subjectivity. Once this process is completed, the subject takes over the objective element and recognizes the world's independence as its own independence. The master-slave dialectic gives way to a new phenomenological stage, in which self-consciousness is raised to a new level of freedom.

[400] *HW* III, 152; *PS*, 117. Translation altered.
[401] *HW* III, 154

At first sight, this new development appears to replicate the discovery made at the beginning of the chapter, namely that self-consciousness is the truth of consciousness, or that subjectivity is the truth of objectivity. But whereas in the beginning the 'I' was purely evanescent, or the mere 'return from otherness', self-consciousness is now aware of itself as the truth of otherness. The dissolution of objectivity is no longer something that merely happens, as a result of the self's inner contradictions, but something that is actively brought about by consciousness. The 'I' is now 'its own pure movement', or the awareness 'of itself as essential being.'[402] In other words, the self is now a *thinking* self, to whom the outer world is the content of a free, manipulable set of thoughts. While in the previous stages the relationship with otherness was still one of subordination, grounded in different forms of desire and recognition, consciousness has now succeeded in breaking this cycle and reversing the elements of the equation. Instead of yielding to objectivity, or adapting to a preformed version of things, the self is now the primary source of reality's identity, the grounding instance of its meaning and value.

By subordinating the outer world to the world of thought, self-consciousness rises for the first time to a truly self-reflexive or philosophical stance. Its object is no longer objectivity as such, but its own perception of it, and its thoughts are directed, therefore, at thinking itself. 'In thinking', Hegel argues, 'I *am free*, because I am not in an other, but remain simply and solely in communion with myself'. Hence, 'my activity in grasping [something] is a movement within myself' (*meine Bewegung in Begriffen ist eine Bewegung in mir selbst*).' [403] In this passage, Hegel calls upon the distinction between 'representational thinking' (*vorstellendes Denken*) and 'notional thinking' (*begreifendes Denken*) already discussed in the preface.[404] As we have seen earlier, these two cognitive attitudes correspond to two different ways of understanding the relationship between subjectivity and objectivity. In the first case, the subject is conceived as a simple or selfsame 'I', confronted with a series of simple or selfsame objects. Since the identity of both elements is already presupposed, their interaction is purely superficial. Subject and object remain the same before and after the relationship, i.e. mere abstractions. In

[402] *HW* III, 152; *PS*, 120
[403] *HW* III, 156; *PS*, 120. Translation altered.
[404] See *HW* III, 56ff. and above section 6.2.

Begreifen, or notional thinking, on the other hand, the relationship is concrete or dynamic. The object is no longer perceived as a mere object, but as the truth about the subject, and the link binding them together is no longer contingent, but necessary. Subjectivity is just as much subjective as it is objective: or rather, the subject is a *self-moving* or *self-differentiating* subject, whose 'other' is already itself.

In Hegel's words, whereas 'what is [merely] represented [*das Vorgestellte*] … has, as such, the form of being something other than consciousness', and is thus set up as the essential moment of the relationship, a notion (*ein Begriff*) is 'a content grasped *in thought*', which renders consciousness '*immediately* aware of its unity with [it].'[405] Throughout the progression's previous stages, consciousness' natural inclination for the first of these attitudes was continually challenged, but not altogether eliminated. With this new transition, however, the notional attitude achieves a new and important victory. Following the slave's withdrawal into himself, self-consciousness is no longer attached to otherness as such, or to a world that merely *is*. From its new point of view, everything is the correlate of a specific thought, or a specific notion. Everything that is *is so* insofar as it is *thought to be so*.

Stoicism is the first phenomenological incarnation of this attitude. To the Stoic thinker, reality is either true or false, good or bad, pleasant or painful — but only insofar as he or she thinks it so. These attributes do not belong to reality in a substantial or intrinsic manner. They are the results of a complex web of cognitive claims, or ὑπολήψεις, whose content and value are determined by consciousness. The Stoic's freedom consists, therefore, in the power to review and redefine the scope and configuration of this web. If reality's 'imprint' on consciousness is not absolute, but agreed upon, what usually appears natural or self-evident is merely the correlate of a specific thought, or set of thoughts, whose content can be changed. Once this is acknowledged, 'what is essential is no longer the difference as a *specific thing*, or as *consciousness of a specific natural existence*, as a feeling, or as desire and its object, but only

[405] See *HW* III 156; *PS*, 120. Translation altered.

the difference posited by *thought*, or the difference which from the very first [*unmittelbar*] is not distinguished from myself.'[406]

If difference, or meaning, is grounded in thought, it is not fixed or invariable, but open to change. And this change is all the more necessary, according to Stoicism, since consciousness is usually led to accept as true, or good, or relevant, what lies beyond one's control. Personal freedom is compromised by a series of assumptions that are not perceived as assumptions, but as facts, and Stoicism's aim is precisely to denounce this misconception and to restore human thought's original plasticity. It highlights the need to roll back our usual expectations about reality, to reduce life's share of contingency and to limit our existential horizon to what is really up to us (ἐφ'ἡμῖν). This transformation entails a radical revision of the importance usually ascribed to external goods, to physical pleasure, to material wealth, and to all things whose attainment cannot be guaranteed, and the pursuit of which corresponds, therefore, to a form of imprisonment.

The freedom professed by Stoicism can only be found in the element of thought, away from the commotion of human passions and mundane achievements. This negative ideal amounts to a radical form of indifference, or resignation (ἀπάθεια), which transcends both the master's material comfort and the servant's physical hardship:

> Whether in the throne or in chains, in the utter dependence of its individual existence, [the Stoic's] aim is to be free, and to maintain that lifeless indifference which steadfastly withdraws from the bustle of existence, alike from being active as passive, *into the simple essentiality of thought.*[407]

Apart from Hegel's explicit association between this new standpoint and historical Stoicism, the argumentative path followed throughout this chapter also evokes the one we have encountered earlier in the *Gorgias*. In Plato's dialogue, the dispute between Socrates and Callicles hinged on the dialectical interplay between freedom, knowledge

[406] *HW* III 157; *PS*, 121. Translation altered.

[407] Ibid. Hegel is alluding to Marcus Aurelius, who sat 'in the throne' of the Roman empire from 161 to 180 A.D., and to Epictetus, who was born 'in chains' and released from slavery in 68 A.D. In order to grasp the full implications of Hegel's argument, it is useful to compare the *Phenomenology*'s brief characterization of Stoicism with the lengthier discussion offered in the *Lectures on the History of Philosophy*, where similar arguments are used to describe and refute Stoicism's conception of freedom (cf. *HW* XIX, 249-296).

and ignorance. More precisely, it revolved around the issue of whether human freedom requires a special kind of knowledge — and, if so, what kind. Callicles embodied a domineering attitude that is in many ways similar to that of Hegel's 'master': his idea of freedom was linked to pure enjoyment or gratification, secured through the domination of a weaker or baser sort of men on the part of a stronger or nobler sort. Socrates' attitude, on the other hand, anticipates the servile standpoint portrayed in the *Phenomenology*, whose immanent development leads to Stoicism and to the 'freedom of self-consciousness'. For Socrates, inasmuch as the alleged benefits of physical strength, material wealth or political power, praised by Callicles, rest on immediate assumptions, whose truthfulness has not been proven, they are neither empowering nor liberating. They amount to mere δόξαι, or to misplaced ὑπολήψεις, which leave human beings at the mercy of what they cannot control. Because of this, Socrates rejects Callicles' power and the definition of freedom on which it is grounded. Like Hegel's 'slave', he aspires to a different kind of deliverance, no longer dependent on concrete pleasures or concerns. His aim, like that of the Stoics, is not to dominate others, but to achieve self-mastery (ἐγκράτεια ἑαυτοῦ), or the 'simple essentiality of thought.'[408]

In the *Phenomenology*, this new standpoint is the logical outcome of the master-slave dialectic. Unlike the master and the slave, the Stoic is aware that enjoyment and domination are not immediate or essential determinations, but only *thought* to be such. Yet although this discovery represents a decisive turning point in the progression, and the first step towards a truly dialectical conception of freedom, it is not yet the truth of self-consciousness. The freedom unveiled by the slave's intellectual emancipation is at first but an immediate result, or the freedom of thinking *in general*. Consciousness has risen above the realm of contingency and proclaimed its independence vis-à-vis the objective world, but its standpoint is reduced, in the stage of Stoicism, to this very proclamation.

Whenever the Stoic comes across a determinate content, it immediately points out its lack of essentiality. Whatever is deemed true or good is promptly unmasked as the reflection of a subjective judgement and led back to the element of thought. The Stoic

[408] I cannot elaborate here on the parallels between the *Gorgias* and Hegel's master-slave dialectic, but I have done so in FERRO, *The dialectics of domination*, ed. cit. For an analysis of the relationship between the *Gorgias* and Stoic philosophy, see, for ex., ALESSE, op. cit., LONG, op. cit., 70ff., and STRIKER, op. cit.

refuses to abide by any particular judgement because his freedom is as wide as thought itself. It cannot be reduced to this or to that specific truth without relapsing into the dogmatism of the previous stage. Nevertheless, the continual dissolution of particularity is only a negative movement, in which nothing positive gets determined. Although Stoicism knows what the true and the good are *not*, it is incapable of determining what they actually are. Its ideal is thought itself, that is, a thought whose only content is the fact that it has none.

The Stoic's position, as Hegel conceives it, is untenable. Thought *as such*, devoid of particularity, is also deprived of practical or existential depth. It is constitutively unable to provide a criterion with which to distinguish what is true from what is false, what is right from what is wrong, what is good from what is evil, etc. But since such criteria are nonetheless needed, the Stoic must find them where it cannot look for them, namely in the realm of objectivity. Since 'the notion as *abstraction* cuts itself off from the multiplicity of things, and thus has *no content in its own self*, one has to be '*given* to it.'[409] And this predicament replicates, at a higher level, the one faced earlier by the subject of desire: even though the Stoic rejects particularity and refuses to acknowledge the essentiality of the objective world, he or she is nonetheless forced to return to it again and again.

Stoicism is not successful because it is still only an incomplete negation of objectivity. The freedom it affords is not yet true freedom, but only a reaction to the servitude endured in the previous stage. Although for the Stoic the self's particular existence only counts as thought, the latter is merely the dissolution of something determinate. It is not yet self-moving, or infinite thought, but a form of negation that is still fuelled by an external source. Only a cognitive model capable of overcoming this dependence can turn Stoicism's conception of freedom into the actual experience of freedom. And that model, the next in line in the phenomenological progression, is Scepticism.

Unlike the stoic self-consciousness, the sceptical self-consciousness calls into question not only the determinate contents that make up the objective world, but also the very

[409] *HW* III, 158; *PS*, 122

relationship between cognition and objectivity. Whereas the Stoic preserves reality's referential status, but looks for its truth in the realm of thought, Scepticism no longer allows for this dualism. True freedom requires a self-referential mode of cognition, no longer fuelled by the negation of otherness. For the first time, the objective world to which Stoicism, with each new negation, is forced to go back to is taken over by thought and nullified.

Hegel's focus on scepticism, however, is not confined to a specific phenomenological stage. The aim of the entire progression, as we have seen earlier, is to overcome the dogmatism of natural cognition through a 'scepticism directed against the whole range of phenomenal consciousness.'[410] Generally speaking, every phenomenological stage leads to scepticism: each cognitive model set forth by consciousness ends up being refuted and replaced by a new admission of ignorance. But this ignorance is never final or irrevocable. It does not lead to a pure or simple nothingness, but to a determinate nothingness, which is the starting point of a new cognitive experience. On the one hand, the entire *Phenomenology* is grounded in the conviction, inherited from Plato and the critical tradition, that the attainment of truth can only be envisaged once all cognitive presuppositions have been identified and suspended. Philosophy must bring about 'a state of despair about all the so-called natural ideas, thoughts, and opinions, regardless of whether they are called one's own or someone else's.'[411] On the other hand, if this procedure is to yield a positive result, it cannot be reduced to a simple form of scepticism. The attainment of truth requires a dynamic or self-evolving scepticism, whose results are both negative and positive.

The 'exposition of the untrue consciousness in its untruth is not a merely negative procedure.' To think otherwise is to give in to the 'scepticism which only ever sees pure nothingness', and 'cannot get any further from there.'[412] In the introduction, Hegel points out that this one-sided attitude is not the *Phenomenology*'s final destination, but only one of the many stops lying ahead. He is referring to the stage of Scepticism and to the radical stance we are now considering.

[410] *HW* III, 73; *PS*, 50
[411] Ibid.
[412] *HW* III, 73f.; *PS*, 50f.

Unlike the Stoic, the Sceptic is no longer content with intellectual freedom. Having uncovered the contradictions underlying the stoic model, his conception of freedom is now more demanding. In particular, the Sceptic rejects the idea that consciousness contains the truth of the objective world. Whereas the Stoic looked for truth in the element of thought, but failed to come up with a concrete criterion of truth, the Sceptic denies the existence of such a criterion and decrees the impossibility of true knowledge. While the Stoic is still attached to the possibility of a rational redemption of irrationality, the Sceptic regards this hope as naïve. In his eyes, each thought is just as good as the next one, for none is able to convey anything certain or definite about reality.

The seriousness of this new charge is also reflected in the *Phenomenology*'s global architecture. Heretofore, the phenomenological progression remained faithful to one fundamental rule. Consciousness started out from what it believed to be the simplest or most immediate cognitive regime, namely sense-certainty. Afterwards, realising that its actual cognitive powers were irreducible to the kind of atomism that is typical of sense-certainty, it devised a new form of self-recognition, namely perception. When this proved illusory, it moved on to the stage of understanding — and then to the stage of self-consciousness, and then to the dialectic of desire and recognition, etc. In all of these transitions, the refutation of each model left behind a residue of meaning — an unfulfilled *Meinen*, or an unrealized *Begriff* requiring a new form of expression. This residue was the fuelling principle of a long succession of negations and re-integrations, whereby consciousness was led to reform and to widen its usual standpoint, to revise many of its convictions and to accommodate new elements found along the way.

But the emergence of Scepticism is liable to put an end to this pattern: in its most radical form, Scepticism represents, if not the exhaustion of the phenomenological mechanism, at least the exhaustion of the method that was hitherto followed by consciousness. For the sceptical outlook entails a voluntary sabotage of the very movement of adhering to a cognitive model, whichever it may be. Instead of looking for new ways of correcting the shortcomings of a given cognitive regime, the Sceptic turns his back on the whole undertaking and refuses to assent to any truth other than that there is none, or that all standpoints are ultimately the same.

In the *Phenomenology*'s first chapter, consciousness assumed the existence of a reality independent from thought and stated naïvely that things 'are what they are'. Later, with

the emergence of Stoicism, it abandoned this naïve realism and came to acknowledge that things *are not what they are*, but only *thought of* as such. Now although these two standpoints are radically different, the Sceptic will argue that both conceive knowledge as a reaction to an independent or pre-established 'other', which must be either embraced or rejected. For these two standpoints, to be or not to be is still the question. But the sceptical self-consciousness rises above this dichotomy and regards being itself as a misconception. To claim, categorically, that something *is* or *is not* is to construe identity and difference as fixed or dogmatic categories, falling outside the infinity of thought. For the Sceptic, to be absolutely free one must abolish difference itself, and hence the very idea of predication — 'because the 'different' is just this, not to be in possession of [oneself], but to have [one's] essential being only in another.'[413]

What is achieved through Scepticism, Hegel argues, is the recognition that 'reality contains no permanent element, and must vanish before thought.'[414] But this vanishing, unlike the one brought about by Stoicism, is no longer merely replaced consciousness' version of reality: the self does not merely *acknowledge* its own primacy, nor does it merely *detect* a general lack of correspondence between thought and reality. Scepticism is a mode of cognition 'to which it does not *happen* that its truth and reality vanish without its knowing how, but which, in the certainty of its freedom, *makes* its other which it claims to be real, vanish.'[415] In other words, Scepticism amounts to the self-conscious denial of both the realist claim that truth is to be found in objectivity and the idealist claim that truth is to be found in thought. By decreeing the absolute untruth of objective reality, the Sceptic frees thought from the very alternative between being and not-being, selfhood and otherness, objectivity and subjectivity.[416]

[413] *HW* III, 161; *PS*, 124. Translation altered.

[414] Ibid.

[415] *HW* III, 160; *PS*, 124

[416] Of course not all forms of scepticism need to be this radical. One might accept that true knowledge is unattainable and still admit, on different grounds and with different degrees of conviction, the notions of being, identity, or difference. Here, however, Hegel is not interested in this kind of compromise. As always, his aim is to test the limits of this cognitive model and to bring out its inner contradictions. Moreover, his critique is conducted at a stage where there is no longer room for partial or incomplete forms of negativity: the freedom pursued by self-consciousness is a radical kind of freedom, opposed to objectivity as such, and scepticism is therefore a radical kind of scepticism, equally opposed to objectivity as such. Once again, in order to grasp the full implications of Hegel's argument, it is useful to compare the *Phenomenology*'s brief characterization of Scepticism with the lengthier discussions offered in

Opposed to difference and determinacy, consciousness becomes a self-enclosed maze. Sceptical freedom is absolute because it is completely unbounded: since nothing is different from anything else, there is no good reason to claim this or that, to hold on to this truth instead of another. In this 'absolute dialectical unrest', every difference is dismissed as spurious and eliminated. But inasmuch as this elimination is itself grounded in a determinate judgement, it is once again dismissed, and coincidence is thrust back into difference. Thought is as much free and malleable as it is gratuitous and arbitrary, and self-consciousness consists in the perpetual transition from each thing to its opposite.

Scepticism is unable to provide consciousness with a stable cognitive perspective. Nay, by refusing the possibility of a determinate version of reality — grasped in a determinate way, endowed with a determinate set of properties, governed by a determinate set of principles, etc. — what the Sceptic is refusing is the very possibility of human cognition. His cognitive horizon is reduced to a single, uninhabitable punctum, of which meaning, identity and determination have all been expelled. Therefore, he can only maintain this negative attitude at the price of contradicting himself: the Sceptic abolishes determination, but does so in a determinate way; he rejects being, but must *be* in order to do so; he dismisses everything stable and objective, but cannot help relapsing into stability and objectivity.

The Sceptic's inability to uphold his own standpoint, reiterated with each new act of cognition, leads to the phenomenological collapse of Scepticism and to the endorsement of a new shape of consciousness. But this new stage is not yet the deliverance craved by Scepticism. The 'unhappy consciousness', as Hegel calls it, amounts rather to the full acknowledgement of Scepticism's contradictions, and to a new and decisive turning point in the debate concerning the possibility of freedom.

the 1802 essay *Verhältnis des Skeptizismus zur Philosophie* (cf. *HW* II, 213-272) and in the *Lectures on the History of Philosophy* (cf. *HW* XIX, 249-296).

8.3. *The unhappy consciousness*

Sceptical freedom entails a performative contradiction: on the one hand, the Sceptic claims to have thrown off the shackles of meaning and determination; on the other, this proclamation is itself something meaningful and determinate. The paradox is unavoidable because determinacy lies at the very heart of human cognition and conditions all aspects of human life. The sheer fact of possessing a body, or reacting to sensuous stimuli, or taking up a specific space during a specific period of time — all of this forces the Sceptic to return to objectivity and assent to the world he repudiates. The mirage of a disembodied self, devoid of context and particularity, cannot be extended in time, let alone mentioned or discussed. Every qualification is a disqualification and every concrete reasoning leads back to identity, difference and predication.

This ceaseless interchange is Scepticism's first reaction to its own contradictions. Caught between freedom and servitude, self-consciousness jumps from content to content, refusing to linger in any of them. Whereas at one time it 'recognizes that its freedom lies in rising above all the confusion and contingency of existence, … at another time [it] equally admits to a relapse into occupying itself with what is unessential.' As a result of this admission, however, it immediately 'lets the unessential content in its thinking vanish', and yet in doing so it is once again 'the consciousness of something unessential.'[417] In this endless fluctuation, the Sceptic's life becomes a 'confused medley, the dizziness of a perpetually self-engendered disorder.'[418] Like a child who rejoices in contradiction, he negates everything for the sake of negation and turns everything into its opposite.

This attitude brings about a new phenomenological deadlock, which leads to a new change of perspective. Although the Sceptic is aware of his inner dividedness, and

[417] *HW* III, 162; *PS*, 125

[418] *HW* III, 161; *PS*, 125. This sentence contains a double meaning which both Miller's and Baillie's translations fail to acknowledge. The term *Schwindel*, rendered as 'dizziness', or 'giddy whirl', also means 'swindle' or 'fraud' — the *swindle* of scepticism, as it seeks to separate the intellectual and existential dimensions of human cognition. In the previous paragraph, Hegel speaks of *Sophisterei* with the same idea in mind.

although his whole existence consists in reaffirming, again and again, his disagreement with himself, his conception of freedom is still one-sided. At this point, freedom still stands on one side, as something to be accomplished, while difference and determinacy still stand on the other, as the obstacles preventing this accomplishment. The former is still the central moment of the relationship, while the latter are still the vehicles whereby thought is reunited with itself. But the experience of Scepticism will put an end to this asymmetry. According to Hegel, the logical outcome of Scepticism is the recognition that these two moments cannot be envisaged separately: thought's absolute freedom is only conceivable as a reaction to the determinacy of the objective world; and the latter, in turn, only exists as a product of thought. Both moments are essential because they are both part of the same thinking self. 'It is in fact *one* consciousness which contains within itself these two modes'[419], and for which the two are both necessary and incompatible. This acknowledgement leads to the abandonment of Scepticism's dichotomical worldview and to the emergence of a new shape of consciousness, 'which *knows* that it is the dual consciousness of itself, as self-liberating, unchangeable, and self-identical, and as self-bewildering and self-perverting.' [420]

This new cognitive standpoint, born out of Scepticism's self-unification, is appropriately named *das unglückliche Bewußtsein*, that is, the 'unhappy' or 'unfortunate' consciousness. Unlike the dilemma previously faced by Stoicism and Scepticism, this 'unhappiness' is now twofold: it is due, on the one hand, to self-consciousness' continual urge to reconcile thought and reality, and on the other hand to the acknowledgement that such reconciliation is impossible. Consciousness is now fully aware that its remedy is also the cause of its evil, but cannot escape either one of them: it attempts to free itself from the objective world, but only to reiterate its captivity and to renew the source of its unhappiness.[421]

Like Scepticism, the unhappy consciousness is initially torn between the pure universality of thought and the contingency of individual existence. One side is now

[419] *HW* III, 163; *PS*, 126

[420] Ibid.

[421] The unhappy consciousness *knows* that its pursuit of freedom is also the source of its unhappiness, but is unwilling or unable to escape the contradiction. Its impossible situation brings to mind the closing lines of Beckett's *L'Innommable*: 'Je ne peux pas continuer, je vais continuer.'

referred to, in Parmenidean vein, as 'the simple Unchangeable' (*das einfache Unwandelbare*), whilst the other is now defined in Heraclitean terms as 'the protean Changeable' (*das vielfache Wandelbare*).[422] However, these two aspects are not to be construed as independent or juxtaposed aspects. Although the self is divided, it is itself directly both sides of the division. Yet 'since for it both are equally essential and contradictory, it is merely the contradictory movement in which one opposite does not come to rest in *its* opposite, but in it only produces itself afresh as an opposite.'[423]

The unhappy consciousness *is* and *is not* both sides of the division, and the aim of the ensuing dialectic is precisely to clarify this paradox. At the beginning of a long and complex exposition, Hegel sums up the dialectic's three main steps, corresponding to the different ways in which individuality relates to universality:

1. Self-consciousness sees itself as a mere individual opposed to an abstract and unattainable beyond. The Unchangeable is the essential moment of the relationship;

2. The Unchangeable is brought closer to the changeable moment by assuming an individual form. The essentiality is now shared by both sides;

3. The Changeable finds itself within the Unchangeable and both terms are reconciled.

At first, the meaning of these stages remains somewhat obscure. As the exposition unfolds, however, Hegel's use of words such as 'elevation', 'hope', 'devotion' or 'incarnation' brings to mind the specific context of religious experience. Through this change of vocabulary, the initial opposition between intellectual freedom and concrete individuality gives way to the opposition between a universal or divine mode of existence and a concrete or individual mode of existence. But although this line of interpretation can prove very useful in decoding some of Hegel's most cryptic formulations, it must not overshadow the more general implications of his argument.

In the previous sections, Hegel had already pointed out the parallel between the phenomenological stages of Stoicism and Scepticism and the actual historical incarnations of these doctrines. In the present section, this historical dimension is

[422] *HW* III, 164; *PS*, 127
[423] Ibid.

rendered even more explicit. As practically every commentator has recognized, the historical backdrop of the unhappy consciousness is the evolution of early Judaism, the rise of Christianity and its main developments throughout the Middle Ages. This historicized reading, encouraged by numerous textual hints, is further supported by the fact that this whole section amounts in many ways to a direct follow-up of Hegel's early theological investigations, 'transposed to a properly philosophical plane.'[424] In those early texts, the dilemma faced by the unhappy consciousness is already discussed in all its complexity, although the discussion is more firmly rooted in its concrete historical context. Hegel's arguments revolve around the mythical figures of Moses and Abraham, the religious practices of the first Christian communities, the life of Christ and the historical evolution of the Christian Church.[425]

But how far should one rely on history to interpret Hegel's text? On the one hand, given all of its historical echoes, as well as its closeness to Hegel's early writings, it is tempting to suppose that the dialectic of the unhappy consciousness was originally conceived *ex historia*, and that the obscurity of Hegel's text is partly due to an active effort to conceal the argument's historical skeleton, so to speak, and preserve thereby its universal validity. Moreover, it can hardly be denied that in the absence of previous knowledge about the historical and theological themes Hegel is alluding to, his exposition would be less easily and/or less thoroughly understood. On the other hand, however, it must be borne in mind that the *Phenomenology* is not a history of human consciousness, nor does it aim to retrace human reason's actual development. It consists rather in a succession of cognitive models, devised by consciousness to account for its own cognitive abilities. Accordingly, the unhappy consciousness is first and foremost a cognitive model, in the peculiar sense I have previously indicated,[426] which means that Hegel's description must be able to stand on its own, as it were, regardless of its historical allusions. Its meaning does not derive primarily from external or contingent associations, but from the progression's immanent logic.

[424] HYPPOLITE, op. cit., 187

[425] See notably *Die Positivität der christlichen Religion* (1795-96) and *Der Geist des Christentums und sein Schicksal* (1798-1800) (*HW* I, 104-217 and 274-418).

[426] See above section 6.3.

In the following pages, for the sake of clarity and expediency, I will focus simultaneously on the argument's logical and historical dimensions. Historically, the transition from Scepticism to the unhappy consciousness is echoed by the decline of Greek culture and the rising influence of Jewish and Christian monotheism.[427] For Hegel, broadly speaking, the history of Greek philosophy is the history of subjectivity and its conquests. From Thales to Sextus Empiricus, the Greeks discovered and explored the infinite power of thought, redefining the scope and boundaries of human self-consciousness. During this long *Ausbildung*, thought strived to 'tear itself away from what is sensuous' and to attain 'a totality beyond the sensuous and the imaginary' (*jenseits des Sinnlichen und der Phantasie*)[428]. But this one-sided movement was bound to be overcome: Scepticism was both its climax and its dying hour, paving the way to a new philosophical age. The idea that thought could rise above objectivity and overcome the contingency of the natural world was definitively abandoned and the search for intellectual freedom gave way to a new conception of servitude. The Greek world was replaced by the Jewish world and the focus on freedom and independence was replaced by a new fixation on the sensuous and concrete aspects of human life.

In the *Phenomenology*, this transformation corresponds to the sublation of Scepticism and to the unhappy consciousness' first dialectical equation. Initially, self-consciousness regards the Unchangeable as the essential side of the relationship and places itself on the opposite side. It is now a purely changeable or unessential being, facing a divine and unattainable *beyond*. Like the sceptical self-consciousness, the unhappy self-consciousness repudiates the contingency of its own existence, but unlike the Sceptic it accepts this contingency as its own fate, or as the nothingness to which life is naturally doomed. This first attitude brings to mind the opening lines of the *Ecclesiastes*, or the wretchedness endured by Job. Moreover, it replicates the servitude described in the master-slave dialectic, but on a different scale: slavery is now the self's natural condition, and the master is now a divine and all-powerful master.

[427] For Hegel's account of this historical movement, see notably *HW* XIX, 487-514. In the *Lectures on the History of Philosophy*, the transition from Scepticism to Medieval theology is mediated by Neoplatonism, which is not explicitly mentioned in the *Phenomenology*, but may also be contained in the dialectic of the unhappy consciousness.

[428] *HW* XIX, 487

Yet although self-consciousness views itself as the unessential moment of the relationship, it is not opposed to the other moment as a substance is opposed to another substance. The unhappy consciousness is by definition both sides of the equation, which means that neither the Changeable nor the Unchangeable can be eliminated or cast aside. Although self-consciousness is at first a merely contingent or accidental moment, it is immediately conscious that its essence lies elsewhere, namely in the Unchangeable. Therefore, it is bound to reach out and claim possession of it. But since the latter is by definition transcendent and unattainable, the self's yearning cannot be fulfilled. It can only reach the divine element through worship and adoration, that is, through the continual admission of its own servitude. Its life is therefore an eternal blend of hope and wretchedness, of uplifting devotion and degrading humiliation.

Despite the sinuosity of Hegel's exposition, the problem is really a simple one: the unhappy consciousness is bound to acknowledge the simultaneous validity of two self-excluding standpoints. Its unhappiness stems, therefore, from the need to reconcile what appears irreconcilable: on the one hand, if the immutable side is placed at a finite distance, its immutability disappears, and self-consciousness returns to the naïve realism of ordinary consciousness; on the other hand, if the mutable side passes over into the immutable one, self-consciousness returns to Scepticism and to the contradiction that led to its present condition. The only way out of this paradox is to devise a form of relationship that nonetheless preserves the independence of the two extremes. The solution, if it does exist, must allow for the subsistence of a contingent *here* and a necessary *beyond*, but in such a way that both can be recognized as different sides of the same self-consciousness.

The search for this solution leads to the unhappy consciousness' second variant, which Hegel associates with the Christian tradition. More specifically, the Christian concept of incarnation is offered as the answer to the Jewish dilemma and the figure of Christ is brought in to unite both sides of the antithesis. Through Christ, the 'embodied Unchangeable', the immutable side approaches the mutable one without relinquishing its immutability — for Christ is *simultaneously* Father and Son, divine and human, eternal and temporal. With this transition, the first Unchangeable, which was merely 'an alien

Being who passes judgement on the particular individual', is replaced by a second Unchangeable, 'a form of individuality into which the entire mode of existence passes.'[429]

Moreover, since the unhappy consciousness is simultaneously both sides of the equation, this transition also entails a transformation of the human or changeable element. By conceiving God simultaneously as a universal ideal and an individual presence, human beings are brought closer to the Unchangeable without foregoing their changeable nature. Their essence is no longer to be found in a pure beyond, posited at an infinite distance, but in a synthetic unity (i.e. Christ), which is both concrete and ideal, finite and infinite.

At first glance, this second solution seems to accommodate the contradictions of the previous solution: the concept of divine incarnation manages to overcome both the pure contradiction of Scepticism and the simple self-coincidence of ordinary consciousness; it offers a solution in which both terms come into contact with one another without sacrificing their independence. Experience will show, however, that this compromise is once again illusory. Christianity's embodied divinity is just as inaccessible as Judaism's abstract one, for although the Christian is granted the possibility of experiencing the Unchangeable in its concrete or physical form (the doubt of Saint Thomas), the object of this experience is a sensuous object, and hence a changeable one. In Hegel's own words, when the Unchangeable 'is sought as a particular individual, it is not a universal individuality in the form of thought, not a Notion, but an individual in the form of an object, or an actual individual'. Therefore, the body of Christ is really only 'an object of immediate sense-certainty, and for that very reason only something that has already vanished.'[430]

Like all physical bodies, the body of Christ only displays its bodily nature. It confronts the observer with the silent obstinacy of sensuous objects, bound to a specific form and a specific appearance. The idea that such a body is at the same time a living image of eternity is something merely added to it — an abstract label, as it were, irreducible to its physical presence. And although self-consciousness is attached to this label, it is forced to acknowledge its arbitrariness: it can neither explain why this

[429] *HW* III, 165; *PS*, 128
[430] *HW* III, 169; *PS*, 131f.

particular body is an immutable essence, nor why such an essence should be compelled to assume a bodily form. Both of these ideas are brought over from outside of the sensuous realm. They are the products of a new beyond, which is just as unattainable as the first one.

This new dialectic leads self-consciousness back to its previous dilemma: if incarnation is simply an idea, the Unchangeable remains distant and unattainable; but if divinity is successfully identified with individuality, its immutability is reduced to the contingency of human life and the dialectical bond is once again severed. In the latter case, the death of the historical Christ leads to the simple elimination of transcendence, and Christianity is reduced to a particular historical event, occurred during a particular historical period, in a particular geographic location.

In order to escape this contradiction, the unhappy consciousness moves on to a third and final solution, which is again evocative of the master-slave dialectic, and particularly of the slave's final emancipation. Since the previous attempts to set up the Unchangeable as the essential element have failed, the unhappy consciousness returns to the changeable element and looks therein for the essence of the relationship. In a way, this new attitude is an acknowledgement of defeat: unable to ground its inner restlessness in a stable or convincing form of transcendence, consciousness returns to the stronghold of sensuous existence. From now on, its life is to consist in labour and the enjoyment of earthly goods. The spiritual heights of divinity are replaced by the concrete depths of sensation: for here, at least, the fruits of one's labour are truly one's own, and the objects of enjoyment are within one's reach.

Unlike the previous modes of self-consciousness, however, this return to the sensuous world is not immediate or one-sided. Objectivity is no longer merely the negative of self-consciousness, destined to be consumed or nullified, but 'an *actuality broken in two*, which is only from one aspect intrinsically null, but from another aspect is also a sanctified *world*.'[431] Once, again, since the unhappy consciousness is simultaneously both sides of the relationship, its new focus on particularity entails a redefinition of the universal or unchangeable moment of the relationship. The incarnate Unchangeable,

[431] *HW* III, 170f.; *PS*, 133

which was formerly confined to a single individual, is now extended to the whole of objectivity, and Christianity assumes hereby a pantheistic dimension: the entire world is now the 'form of the Unchangeable', and its previous individuality 'has in general the significance of all actuality.'[432]

What slowly comes to light, with this new development, is the possibility of a truly reciprocal dialectical exchange. At first, only the individual moment seems to matter: the activity of labour and enjoyment contrasts with the passivity of the natural world. But since the latter is now a 'sanctified world', labour and enjoyment are only made possible because the Unchangeable has offered itself to self-consciousness and provided the basis for its activity. Although the unhappy consciousness sets about transforming the natural world and reaping the benefits of this transformation, it is simultaneously aware that nature is a divine gift. Therefore, instead of proclaiming its independence vis-à-vis the objective world, as the Stoic or the Sceptic did, it gives thanks for its personal achievements and regards them as a sign of God's favour.

According to Hegel, this acknowledgement brings consciousness closer to the solution of its long-lasting dilemma. Looking back on the previous developments, the path towards happiness follows the *Phenomenology*'s usual triadic scheme: position, negation and negation of negation (or *Ansichsein*, *Fürsichsein* and *An-und-fürsichsein*). Firstly, consciousness posited the Unchangeable as the essential moment — either at an infinite distance (step 1 of my initial list) or as an incarnate individual (step 2). Afterwards, aware of the instability of both of these solutions, it turned away from the transcendent element and came to regard its concrete existence, actualized in labour and enjoyment, as the essential moment of the relationship. But this change of focus led to the simultaneous recognition that the concrete world is also a divine world, whose 'sanctification' is the underlying condition of labour and enjoyment. This recognition will now lead to a new contradiction, whose resolution amounts to the third and final step of the dialectic.

The unhappy consciousness has come to acknowledge that its world is a holy world, and that its life is a holy gift, but this new solution is still dominated by the particular

[432] *HW* III, 171; *PS*, 133

end of the relationship. Although consciousness knows itself to be blessed, and gives thanks for what it receives, its humility remains superficial. Personal identity is still grounded in an individual core of self-certainty, which the self refuses to hand over to the Unchangeable. In Hegel's words, consciousness still views itself 'as this particular individual, and does not let itself be deceived by its own seeming renunciation, for the truth of the matter is that it has *not* renounced itself.'[433] The fruits of its labour may well be provided by an alien source, but it can still use them up and enjoy them. Their divine provenance does not compromise their actual value, nor the pleasure and the autonomy they afford.

> Though consciousness renounces the *show* [*Schein*] of satisfying its feeling of self [*Selbtsgefühl*], it obtains the *actual* satisfaction of it; for it *has been* desire, work and enjoyment; As consciousness it has *willed*, *acted* and *enjoyed*. Similarly, its *giving of thanks*, in which it acknowledges the other extreme as the essential Being and counts itself nothing, is its *own* act which counterbalances the action of the other extreme, and meets the self-sacrificing beneficence with a *like* action.[434]

In other words, consciousness is willing to acknowledge the superiority of the universal element, but it refuses to give up its own individuality. This self-attachment revives the self-division maintained in the previous stages and constitutes a new source of unhappiness, which calls for a new dialectical compromise. In the remainder of the chapter, self-consciousness will be led to negate its previous negation, to alienate its own individuality and to acknowledge once more the absolute dominion of universality. But inasmuch as this negation is a *determinate* negation, this new alienation is different from the first one. Instead of reducing itself to a simple nothingness, immediately taken over by otherness, the unhappy consciousness will attempt to *find itself in its other*, and turn its self-division into a truly reciprocal exchange.

We have seen this same movement play out in the beginning of the self-consciousness chapter, in the dialectic of recognition, and again in the master-slave dialectic. In all of these cases, Hegel insists on the fundamental difference between an

[433] *HW* III, 173; *PS*, 134
[434] *HW* III, 172; *PS*, 134

abstract or immediate form of self-consciousness (the simple tautology 'I am I') and a concrete or mediated form of self-consciousness (the 'identity of identity and difference', or the unity of selfhood and otherness). As he argues in different ways throughout the *Phenomenology*, self-identity is not something given, but the result of a dialectical exchange. In its simplest or most immediate form, selfhood is still an abstraction. In order to become concrete, it must give up its initial self-certainty, lose itself in otherness and return from otherness to itself — no longer as a simple 'I', but as a dynamic or self-moving 'I', made up of identity and difference.

The same is now valid for the unhappy consciousness, but on a higher level. The 'other' in which it must lose itself is no longer another individual, or another self-consciousness, but a divine Being, or universality as such. In order to free itself from contradiction, self-consciousness must step out of itself and give itself over to the Unchangeable. Like the slave in the previous stage, it must sacrifice its independence in order to find it anew, in a higher cognitive plane.

Let us follow Hegel's description of this final renunciation. Firstly, personal freedom is counted as vain and relinquished. Consciousness gives up its pretence of autonomy, forfeits the authorship and the responsibility of its actions and places itself at the mercy of an alien will. Secondly, life itself is divested of its former meaning and reduced to a purely physiological process. The 'petty actions' and 'animal functions' which used to be performed naturally, in view of something else, are now the self's primary focus, and the source of embarrassment and shame. Thirdly, personal property and the fruits of labour are likewise rejected. Enjoyment and pleasure are denounced as sinful and replaced by fasting and abstinence. In a word, self-consciousness embraces the 'ascetic ideal' that Nietzsche will later revile.

But this is not all. This radical renunciation also entails a profound cognitive transformation. Apart from forfeiting the right to choose and to act on its choices, consciousness also gives up its quest for a meaningful existence, and sets about 'practising what it does not understand' (*das Treiben eines unverstandenen Geschäftes*).[435] This aspect is extremely interesting, and would certainly profit from a more detailed exposition.

[435] *HW* III, 175; *PS*, 137

What seems to be here at stake is not simply the observance of sacred rites or traditions, but also the adherence to religious dogmas, that is, the acceptance of truths whose validity no longer depends on one's rational assent. And this new demand, even more than the previous ones, represents a complete break with the previous forms of self-consciousness. Whereas earlier the subjective moment had tried to obliterate, determine, influence, or at least accommodate itself to, the objective moment, the notion of dogma entails an inversion of this pattern. Henceforth there is nothing left for the subject to choose or decide, no criterion with which to ground its perspective. Self-renunciation culminates in a pure leap of faith, which demands unconditional acceptance.

This point must be clearly spelled out: by 'practising what it does not understand', consciousness is not simply acknowledging its own ignorance, or surrendering to a higher epistemic authority. If that were the case, the *incomprehensible* would amount to *what has not yet been comprehended*, but might come to be so. Consciousness' self-alienation would remain incomplete, spoiled by the hope of acquiring a more thorough cognitive standpoint, which would reinstate the subject as the essential moment of the relationship. But Hegel's point is more radical. At this stage of the progression, the dialectical resolution pursued by self-consciousness requires something like a *credo quia absurdum*: an unconditional acceptance of the incomprehensible *qua incomprehensible*, viz. a form of adhesion whose force rises as the hope of a rational explanation grows weaker.[436]

This complete cognitive surrender is self-consciousness' final response to the unhappiness endured in all of its previous forms. Unable to find a rational principle in which to ground its identity, it was led to adopt a wholly different strategy. Whereas Stoicism, Scepticism, and the unhappy consciousness were all bent on 'explaining away' the self's inner contradictions, their efforts have only given rise to more contradictions. Their intrinsic faith in rationality has now been identified as the primary source of self-division, and dismissed as vain and presumptuous. Consciousness is now aware that its unhappiness cannot be overcome through reason or self-certainty. In order to rise to a

[436] The Christian dogma of transubstantiation, for instance, conveys a kind of truth that can be accepted or denied, but which is not open to refutation. And it cannot be refuted precisely because it does not rest upon a logical argument or an empirical observation. It is a statement directly issued by God, and its truth lies therefore in the authority of its source.

new cognitive standpoint, it must give in to complete irrationality and embrace the mystery of faith.[437]

Before this transformation is accomplished, however, one last element must be considered. As the self renounces its will and understanding, it is led to revive once more the incarnation model and to bridge the gap between particularity and universality with the aid of a minister or a priest (*ein vermittelnder Diener*). Having sacrificed the power to choose and to interpret the meaning of its actions, consciousness entrusts its freedom to a divine spokesman and relies on his authority to distinguish right from wrong, true from false, good from evil, etc. We have here a new syllogism, which has for its extremes the individual consciousness and the divine element, and for its middle term the priest (or the Holy Church in general). The latter is now the official representative of the Unchangeable, and the vehicle whereby the individual is allowed to join the universal.

But this new solution is once again short-lived. It amounts to one last subterfuge, that only succeeds in delaying the dialectical resolution which by now appears inevitable. Although Hegel's description, at this point, is extremely concise, the argumentative itinerary leading to the collapse of the unhappy consciousness may be summed up as follows.

Initially, the individual's self-negation is construed as a simple negation. 'In the sacrifice actually carried out, consciousness, having nullified the *action* as its own doing,

[437] This attitude is echoed by the Christian condemnation of intellectual or philosophical hubris, voiced in different ways throughout the Holy Scriptures. Recall, for example, Christ's intent, in the First Epistle to the Corinthians, to 'destroy the wisdom of the wise' and 'bring to nothing the understanding of the prudent' (1 Cor 1:19); or his recommendation to Timothy to 'avoid profane and vain babblings, and oppositions of science falsely so' (1 Tim 6:20); Or even Isaiah's famous injunction: 'if you do not believe, you will not understand' (Isa 7:9), which proclaims more clearly than any other the absolute subordination of reason to the authority of faith.

To be sure, one might conceive the relationship between faith and reason in a more nuanced way, and envisage different forms of reconciling these two cognitive modes, namely a) by construing philosophical or scientific knowledge as a 'natural' or 'secular' translation of the immediate knowledge granted by faith (Augustine's *fides quaerens intellectum*); b) by regarding science and faith as different kinds of knowledge, directed at different kinds of objects (see, for ex. Aquinas' *De Veritate* 14, 9); or c) by reconciling both terms through the pursuit of a philosophically or scientifically informed faith (see, for ex., Leibniz's *Nouveaux essais*, ed. cit., 391). In the present phenomenological stage, however, philosophical insight and religious faith *must* appear as self-excluding attitudes, for they have both been driven to their most radical form. Since consciousness, throughout the previous stages, was led to embrace increasingly radical forms of subjectivism, its unhappiness can only be overcome by an equally radical form of self-alienation. Or rather, since consciousness' attempts to 'understand' itself have led to restlessness and contradiction, it can only find peace by relinquishing this attitude and 'practising what it does not understand'.

has … obtained relief from its *misery*.'[438] But this relief appears at first to have been granted by the Unchangeable. Consciousness is still convinced that 'its own misery is only *in itself* the reverse, i.e. that its own deeds bring it only *in itself* self-satisfaction or blessed enjoyment; that its own pitiable action too is only *in itself* the reverse, viz. an absolute action.'[439] These insights are not yet *for consciousness* because it persists in duplicating itself and ascribing its accomplishments to the Unchangeable.

Nevertheless, insofar as relief and enjoyment were brought about by consciousness' self-sacrifice, they are at the same time recognized as its own doing. Consequently, self-negation is finally revealed as a determinate negation, and the self places itself on both sides of the relationship. Just as self-consciousness, in its first appearance, has recognized itself as the author of the 'inner world'[440], and just as the slave, in the master-slave dialectic, has recognized itself as the 'sculptor' of the natural world[441], so too the unhappy consciousness has now returned into itself. It is aware, for the first time, that its essence is also its object, and that its object is also its essence. And this awareness coincides with the phenomenological concept of Reason, 'the certainty that consciousness, in its particular individuality, has being absolutely *in itself*, or is all reality.'[442]

In the stage of Reason, the relationship between the self and the world no longer amounts to an exclusive disjunction. Whereas previously consciousness had to settle for one of two opposite moments, both promising an independence they failed to deliver, the advent of Reason marks the end of this dualism. In this new cognitive stage, consciousness and self-consciousness are finally brought together. Thought is no longer divorced from actuality and the world is both an idea and a lasting presence.

[438] *HW* III, 176; *PS*, 138
[439] Ibid. Translation altered.
[440] See above section 7.1.
[441] See above section 8.1.
[442] *HW* III, 177; *PS*, 138

9.

TOWARDS FREEDOM

From the stage of self-consciousness onwards, Hegel's progression is burdened by three fundamental antitheses. Firstly, the phenomenological subject is faced with the general opposition between subjectivity and objectivity, or between reality *in itself* and reality *for consciousness*. This was the contrast that launched the phenomenological progression, and the reason why sense-certainty was abandoned in the first place. Unable to reconcile the spheres of selfhood and otherness, natural consciousness was led to abandon its initial standpoint and to travel from stage to stage, in search of its lost unity. Along the way, the pure objectivity of sense-certainty was replaced by the pure subjectivity of self-consciousness, which later degenerated into scepticism.

Secondly, the phenomenological subject is also faced with the more specific, but no less decisive opposition between subjectivity and intersubjectivity, or between individual selfhood and a socially mediated identity. As we have seen earlier, whereas the contrast between subjective certainty and objective truth is present throughout the entire progression, the notion of recognition was first made explicit in the self-consciousness chapter, as the subject of desire came to recognize its constitutive dependence on another desiring subject. This dependence evolved into a mortal struggle, which was then replaced by the dialectical exchange between a master and a slave. After this dispute, however, the theme of sociality was set aside, in favour of a new focus on individuality and transcendence.

This focus on transcendence brings us to the third main antithesis at play throughout the progression. With the experience of scepticism, the cognitive approach

that had dominated the preceding stages was exhausted and a new world opened up before consciousness. Whereas previously the dialectical progression had been confined to a horizontal plane, inhabited by human subjects and natural objects, the unhappy consciousness transferred the debate to a vertical plane and unveiled for the first time the idea of God. What thus came to light was a transcendent or absolute yearning, buried in the heart of self-consciousness and irreducible to the finite element of human life. The regional sphere of existence was articulated with the universal sphere of divinity, leading to a series of complex, but equally unstable dialectical equations, and this mounting tension was finally resolved in the stage of Reason, but not altogether eliminated.

From now on, all of these elements will be revived, developed and revealed as the subplots of a wider overarching narrative. The stage of Reason, brought about through the immediate synthesis of consciousness and self-consciousness, is the meeting point of the first two contrasts. Its task is therefore a double task: on the one hand, it must be able to harmonize the conflicting demands of subjectivity and objectivity, and to reveal human cognition as a truly dialectical phenomenon; on the other hand, it must move beyond the individual or dualistic stances of the previous models, and honour Hegel's definition of human consciousness as an inherently social phenomenon — the 'I' that is 'We' and 'We' that is 'I'. The articulation of these two movements will generate a series of increasingly sophisticated shapes of consciousness, which will eventually lead to the stage of Spirit (*Geist*), where Reason is finally raised to truth. This new phenomenological conquest will then reawaken the divine element left dormant in the previous stages, raise the progression to a new degree of complexity and set out its final itinerary.

9.1. *Reason and Spirit*

Throughout the previous stages, self-consciousness was at war with the world, forced either to abolish objectivity in order to assert its own independence, or to give up its independence and surrender to objectivity. In an effort to break this cycle, the unhappy consciousness relinquished its will and understanding, placed itself at the mercy of a transcendent Being and divested itself of its usual self-certainty. As a result, its relation to reality turned round into a positive relation. Instead of retreating once more into

selfhood, consciousness has now come to find itself in the world it had previously repudiated, and to feel at home in it. It has come to realize that 'what *is*, or the in-itself, only *is* in so far as it is *for* consciousness, and what is *for* consciousness is also *in itself*.'[443]

The acknowledgement of this reciprocity is the gateway to a new phenomenological stage, and the basis for a radically different understanding of reality. But the full extent of this transformation is not immediately apparent. As usually, the effects of this discovery are broken down into a succession of more modest phenomenological leaps, which together make up the *experience* of Reason, or the elevation from its initial certainty (*die Gewißheit der Vernunft*) to its final truth (*die Wahrheit der Vernunft*).

As Reason comes on the scene, it is still an immediate result. Consciousness has overcome its former self-division and is now 'certain of being all reality', but this certainty is at first nothing more than a simple or subjective form of idealism. For consciousness, reality is now *its own* reality, that is, a thoroughly *phenomenal* realm whose meaning is provided by Reason. Hegel is alluding, once again, to Kantianism and reiterating the criticism already raised in the introduction and in the transition from the dialectic of consciousness to the dialectic of self-consciousness. If the truth of knowing lies in the unity of apperception, the object of knowledge must be either thoroughly phenomenal, and hence unreal, or placed outside of the phenomenal realm, and hence unknowable. This idealism is 'empty', Hegel argues, because it is stuck in this contradiction and forced at the same time to assert and deny reality's essential nature. While Reason provides the *form* of reality, its categories must be 'filled' with, or 'activated' by, an objective *content*. But since the latter is brought in from the outside, from the noumenal world, it is by definition unknowable. 'The pure reason of this idealism, in order to reach this *other* which is *essential* to it, and thus is the *in-itself*, but which it does not have within it, is therefore thrown back by its own self on to that knowing which is *not* a knowing of what is true.'[444] It is forced, therefore, to move back and forth between its own subjective truth, or the unity of apperception, and an independent objective beyond, be it a Thing-in-itself (Kant), or an extraneous *Anstoß* (Fichte).

[443] *HW* III, 180; *PS*, 140f.
[444] *HW* III, 184f.; *PS*, 145.

To overcome this impasse, consciousness must recognize that the 'I' of idealism is still an abstraction. If the 'certainty of being all reality' is to be raised to truth, subjectivity must abandon this pure formalism, immerse itself in objectivity and come out of this relationship as a truly concrete or mediated self. Instead of claiming *in abstracto* that 'everything is its own', it must set out to find itself in the objective world and to fill out its initial emptiness with the richness of natural existence. This movement is expounded at length in the dialectic of 'observing Reason', which amounts to a more sophisticated version of the dialectic of consciousness. In this new phenomenological stage, to know the world is to know oneself, and to know oneself is to know the world. Contrary to what happened in the stages of sense-certainty and perception, objectivity is no longer an 'other' opposed to consciousness and existing in its own right. Whereas 'the perception and *experience* of various aspects of the Thing were something that merely *happened* to consciousness, … here, consciousness *makes its own* observations and experiments.'[445] Its aim is to establish and bring forth the world's intrinsic rationality, and to find itself therein as *actual* or *embodied* Reason. This project entails a long and painstaking journey, during which consciousness comes to recognize itself as an extraordinarily complex and differentiated being, whose rationality is embodied and articulated in a myriad of different ways. This journey is divided into three main stages, devoted to the observation of nature, the observation of self-consciousness and the observation of their mutual connection.

The first of these stages is in many ways the embryo of the *Encyclopaedia*'s Philosophy of Nature. This segment of the progression is fuelled by the continual tension between Reason's theoretical constructs and the contingency of natural phenomena. While consciousness attempts to subsume reality under universal patterns and laws, the natural world keeps eluding these attempts, proving to be more complex and unpredictable than originally supposed. The limitations of the observational approach are initially felt in connection with the study of physics and geology, and aggravated with the study of biology and zoology. In its quest for the universal principles underlying the morphology and behaviour of living beings, Reason sets out to test a series of explanatory

[445] *HW* III, 185; *PS*, 145

models, leading to complex and convoluted debates about the influence of environmental conditions on living organisms, the nature and scope of internal and external teleology, the joint action of sensibility, irritability and reproduction. In the end, however, none of these approaches is capable of bridging the gap between the universality of thought and the particularity of the natural world. Consciousness is left with a series of 'interesting connections' and 'clever remarks', whose interconnection is loose and contingent, and whose conflicting insights cannot be raised to a coherent rational system.

This failure leads consciousness to turn inward and to search for rationality in the laws that govern the human mind. The second and third stages of 'observing Reason' deal with the observation of self-consciousness as regards both a) the nature and purpose of its thoughts and actions (observational and social psychology) and b) the meaning of its bodily expression (physiognomy and phrenology). In the first case, the debate is centred on the issue of how and to what extent human thoughts and actions are influenced by one's natural and social environment. Once again, Hegel argues that 'influence' (*Einfluß*) is a vague and ultimately self-contradictory notion, incapable of grounding a truly scientific account of human behaviour. He also claims that human beings' power to determine the content of their thoughts and actions is ultimately irreducible to a constellation of natural or social circumstances. To be sure, 'if these circumstances, way of thinking, customs, in general the state of the world, had not been, then of course the individual would not have become what he is'[446]; but since these circumstances are only relevant insofar as they are recognized and incorporated into the individual's outlook on reality, he can choose either to allow 'free play to the stream of the actual world flowing in upon [him]', or break it off and transform it.[447] This ambiguity undermines the alleged necessity of psychological laws and leads to a new divorce between the reasoning self and the objective world.

In the second case, which amounts to the third and last stage of the dialectic, consciousness turns once more to objectivity, in a final attempt to harmonize its psychological insights with the natural laws uncovered in the first stage. Hegel explicitly addresses two pseudosciences which became popular during the 18th century and

[446] *HW* III, 231; *PS*, 184
[447] *HW* III, 232; *PS*, 184

proposed to draw psychological conclusions from the observation of human beings' physical features — in the case of Physiognomy, advocated by Johann Kaspar Lavater, by observing people's anatomy (particularly their faces); in the case of Phrenology, championed in the writings of Franz Joseph Gall, by observing the shape and size of their skulls. As in the previous stages, Hegel's refutation consists in exposing the absence of a necessary link between physical features and mental dispositions. Although Reason is indeed as much 'mental' as it is 'real' or 'actual', it cannot be placed inside a muscle, or a bone or a brain fibre. This possibility corresponds to the last stage of Reason's objective alienation and calls for a reversal of its initial strategy.

The failure to unify thought and being leads consciousness to realize that human beings are not 'things' to be 'observed'. Their essence is irreducible both to a combination of circumstances and to the complexity of human anatomy and physiology. Human thoughts and actions are dependent on a principle of self-determination that eludes observation, and this fundamental insight alters the entire focus of the enquiry: henceforth the rational subject will no longer be regarded as an *observer*, but as an *agent*. Its aim is no longer to uncover the rationality of the natural world, but to find satisfaction in it.

As regards the *Phenomenology*'s global architecture, this new stage can be interpreted as a more sophisticated version of the dialectic of self-consciousness. In Hegel's words, 'the category which, in the course of the observation, has run through the form of *being*, is now posited as being-for-self [*in der Form des Fürsichseins*]: consciousness no longer aims to *find* itself *immediately*, but to produce itself by its own activity.'[448] What is now at stake, therefore, is no longer the subject's intrinsic nature, but its self-actualization. And since the latter is to take place in a world inhabited by other self-actualizing subjects, the theme of social recognition is reintroduced and re-emphasized, as an increasingly loud echo whose full resonance will coincide with the sublation of Reason.

Rational self-consciousness starts out as a purely egoistic standpoint, whose self-actualization is fuelled by an individual urge for satisfaction. This initial characterization evokes the dialectic of desire, but Hegel is careful to distinguish between his previous

[448] *HW* III, 261; *PS*, 209

definition of *Begierde* and his current insistence on *Lust*, or 'pleasure': unlike the desiring self-consciousness, intent on annihilating the objective world and reaffirming its absolute sovereignty, rational self-consciousness is aware that the objective world is its own world, and that its essence is to be found therein. Yet despite this reciprocity, the self's outlook on reality is just as selfish and one-sided as before: the world is still conceived in a purely instrumental manner, as a vehicle for the attainment of enjoyment and satisfaction. Hegel's description is also reminiscent of the master-slave dialectic in that it focuses not on the pleasure of eating or drinking, but on the pleasure of social and sexual domination. Rational self-consciousness, he argues, 'attains to the enjoyment of *pleasure*, to the consciousness of its actualization, in a consciousness which appears as independent, or to the vision of the unity of the two independent self-consciousnesses.'[449] This unity, however, is not yet the mediated unity of *Anerkennung*, but the immediate unity of lust and gratification. In this context, the other amounts to a 'negative essence', 'devoid of reality and content.'[450]

But this shape of self-consciousness is short-lived, for the self-actualization it affords is shown to be illusory, and ultimately intolerable. Since the object of pleasure is devoid of reality, each act of gratification is the source of a new urge for gratification, and so on indefinitely. This inescapable cycle is not so much gratifying as it is self-restricting and stifling. Its true result is not a free or fulfilling life, but a life of obedience, whose permanent restlessness is eventually cut short by death.

Faced with this gloomy prospect, self-consciousness gives up its initial egoism, widens the scope of its outlook and develops a philanthropic interest in the welfare of mankind. In this new transition, the idea of law is retained, but its meaning is transformed: the blind law of hedonism is replaced by a self-made law, the 'law of the heart', which asserts human beings' universal right to happiness and fulfilment. Whereas personal pleasure was essentially base, the law of the heart is noble and high-minded. It represents the first phenomenological attempt to formulate a universally valid moral imperative.

[449] *HW* III, 272; *PS*, 218
[450] *HW* III, 272; *PS*, 219

For this new rational agent, suffering and iniquity are not intrinsic or unavoidable, but the products of an unjust world, ran by tyrannical leaders and corrupt institutions. Reason's new mission is therefore to *reform* the world and to bring out humanity's original goodness. But this noble project is once again discredited, and for two main sets of reasons: firstly, inasmuch as the law of the heart is still the law of *one* heart, abstractly proclaimed as the law of *all* hearts, its reforming task is doomed to failure. The philanthropist is either bound to recognize the naïvety of his efforts, and to give up his mission, or to lapse into fanaticism, and to try and force his or her own righteousness on the whole world. Secondly, since the law of the heart is a pure moral imperative, or the negative image of the world's moral degeneration, its actualization amounts to a self-betrayal. Hegel makes this point in the following manner: 'the law of the heart, through its very realization, ceases to be a law of the *heart*. For in its realization it receives the form of an [affirmative] *being*, and is now a universal *power* for which this particular is a matter of indifference, so that the individual, by setting up his own ordinance [*Ordnung*], no longer finds it to be his own.'[451] In other words, once the reformer's abstract righteousness is actualized, the heart is deprived of its moral authority and becomes entangled in the very dynamic it was supposed to eliminate. As a result, self-consciousness' reforming efforts lead to social estrangement — inasmuch as other individuals reject the reformer's dogmatism, and choose to obey the laws of their own hearts — and self-estrangement — inasmuch as the reformer himself (or herself), torn between moral 'activism' and moral 'puritanism', proves unable to bring together the model's conflicting claims. This failure leads to a new dialectical reversal and to the emergence of the last shape of 'active Reason', entitled 'virtue and the way of the world' (*Die Tugend und der Weltlauf*).

The previous models collapsed as a result of Reason's inability to turn its individual aspirations into universally valid goals: whereas the individualism of pleasure was crushed by the universal law of necessity, the law of the heart failed to harmonize its universal ideals with the actual diversity of other people's interests and aspirations. Yet although these models have proven untenable, they represent a real, albeit timid departure from

[451] *HW* III, 277; *PS*, 223

Reason's initial egoism. In the stage we are about to consider, this movement away from the individual and towards the universal becomes consciousness' central concern. In contrast both to the individualism of pleasure and to the individual universalism of philanthropy, it embraces a truly *universal universalism*, in which individuals *qua individuals* no longer play any role.

The standpoint of Virtue can be defined as the standpoint of pleasure *in reverse*: while the latter was selfish and self-centred, the former is entirely altruistic; and while the latter sought to achieve satisfaction *at the expense of universality*, the former's universal design must be accomplished *at the expense of individuality*. Virtuous self-actualization consists in a selfless pursuit of the common good, culminating in a universal concord reminiscent of Rousseau's *volonté générale*.[452] The very idea of particularization is rejected and every personal aim is seen as illegitimate and perverse.

It is hardly surprising that this new standpoint should once again be abandoned. Virtue's emphasis on universality is just as abstract as hedonism's emphasis on individuality, and therefore just as ineffective. The contradiction in question is also similar to the one faced previously by the stoic self-consciousness: since virtue's only mandate is to abolish particularity, its ideal is an empty one, namely universality *as such*. But human beings cannot live or act on such a mandate. Their thoughts and actions require a specific content, and are thus bound to relapse into particularity. In Charles Taylor's words 'man cannot suppress his particularity, and act just as a vehicle of the universal; for he cannot simply act on the motive of conforming his action to universal maxims, setting aside all other motives. To refuse all other forms of action is to do nothing.'[453]

[452] See notably ROUSSEAU, *Du Contrat Social*, ed. cit., 277.

[453] TAYLOR, *Hegel*, ed. cit., 166. As most commentators have recognized, the whole dialectic of Active Reason evokes the period of Enlightenment, and particularly Rousseau's moral and political philosophy. While the 'law of the heart' mirrors Rousseau's belief in the fundamental goodness of human nature, as well as his denouncement of the corrupting effects of culture and civilization (as discussed, for instance, in the *Discours sur les sciences et les arts*), the dialectic of 'virtue' echoes his views on justice, democracy and popular sovereignty (famously expounded in the *Social Contract*). In order to grasp the relevance of this historical parallel, it is useful to confront this segment of the *Phenomenology* with the subsequent chapters on Enlightenment, Absolute Freedom and Terror. In many ways, Hegel regards the French Revolution as the historical materialization of 'Active Reason'. But whereas in the *Phenomenology*'s fifth chapter Reason's obsession with virtue is described as naïve and 'boring', in the *Phenomenology*'s sixth chapter Spirit's obsession with virtue leads to terror and fanaticism (see *HW* III, 431-441): although Rousseau and

Virtue's immobilism is opposed to what Hegel calls the 'way of the world', that is, the acknowledgement that human motivations are in fact intrinsically selfish, and that the common good is really only the global sum of countless individual efforts. This atomistic view, characteristic of modern political liberalism, is brought in to contradict virtue's sterile utopianism. But Hegel's critique is just as much directed at the virtuous self as it is at the 'way of the world'. Its aim is to overcome the antithesis expressed by their opposition and to reveal reality as the undivided unity of particularity and universality. On one side of the antithesis, the 'fatuous rhetoric' of virtue is bound to be 'dropped like a discarded cloak'[454]; on the other side, the individuality of the 'way of the world' is also bound to be abandoned: when it 'acts in its own interest, it simply does not know what it is doing; and when it avers that everyone acts in his own interest, it is merely asserting that no one knows what action is.'[455]

The fate of Active Reason is to move beyond the empty idealism of virtue and the resigned selfishness of liberalism. But this can only come about, Hegel argues, through an 'interfusion of individuality and universality' — that is, through the sublation of the long-standing antithesis between self-determination and self-actualization, or between being-for-itself and being-in-itself. In order to overcome this contradiction, self-consciousness must cease to distinguish between individual beliefs and the 'real world' over against them. Each of these terms can only be conceived *through the other*, and their opposition is *simultaneously* no opposition at all.[456]

Accordingly, in Reason's third and last main section (appropriately named 'self-consciousness which takes itself to be real in and for itself'), the subject's task is no longer to determine whether his or her actions agree with an underlying essence or criterion —

Robespierre have deemed it 'rational' to free mankind from the oppression of an unjust political order, their democratic zeal gave way to an equally oppressive pursuit of virtue. Just like before, the common good was immediately opposed to particularity, and raised to an abstract and ultimately untenable ideal.

[454] *HW* III, 290; *PS*, 234

[455] *HW* III, 291; *PS*, 235

[456] In a sense, although Hegel still speaks of Reason (and will continue to do so throughout the three remaining subsections), the *principle* of Spirit is already before us. What is now left to determine is how the 'interfusion of individuality and universality' is to take shape in a world inhabited by other 'universal individuals'. As Charles Taylor puts it, we have yet to shift the 'centre of gravity' and 'to see the individual no longer as the centre with his social relation as a peripheral fact about him' — an individual that is *also* universal — 'but rather see the social whole as central, and the individual as a manifestation of it' — a universal that is *also* individual. See TAYLOR, *Hegel*, ed. cit., 168.

be it the selflessness of 'virtue' or the selfishness of the 'way of the world'. Individuality has become 'intrinsically real', which means that the end of one's actions must now be looked for in the actions themselves. There is no longer an independent self presiding over the world, nor an independent world presiding over selfhood. From now on, the role of self-consciousness is to display through its own actions the unity of reason and reality. Self-identity is equated with *self-expression* and the subject becomes a creator or a producer, whose truth lies in his or her work.

The ensuing dialectic is once again reminiscent of the master-slave scenario. Just as the slave had 'put himself in his work', and 'shaped the Thing in his own image', so too will the rational self express its identity through its activity. Yet just like before, the coincidence of selfhood and self-expression is not as straightforward as it may appear. For one thing, since the product of one's work is usually finite and contingent (a text, a painting, a house, etc.), it is by definition incapable of matching the infinite complexity of self-consciousness. Furthermore, since this product usually has a fixed shape, it can only capture the nature of self-consciousness at the time of its creation — or not even that, for the creative impulse may be shorter than the actual creation, and the work may already be outdated in the very moment of completion. These and other contradictions open up a new gap between self-identity and self-expression. They suggest that what one *is* is irreducible to what one *does* (or has done, or will do), and frustrate thereby Reason's quest for full self-expression.

In addition, Hegel points out that the truth or value of one's work is not determined solely by the nature or intensity of one's creative efforts. In a social environment, one's creations are made visible to others. Their meaning is influenced by other people's appreciation, and the latter influences the creator in return. Every act of self-expression entails an intersubjective exchange which alters the very identity one is supposed to be expressing. This co-implication is described in terms of the dialectical opposition between one's commitment to 'the heart of the matter' (*die Sache selbst*) — viz. a selfless immersion into the specific challenges and requirements of one's work — and an undeclared concern for success and recognition — which turns one's professional

commitment into 'deceit' (*Betrug*).[457] What one expresses through one's work is not simply what one is, or what one is actually capable of, but how one wants to appear to others.

Hegel's argument anticipates some of his later insights on culture (*Bildung*) and self-alienation. Here, however, his chief concern is to show that true self-identity cannot be conveyed in this manner. After establishing that the product of one's work is not an adequate vehicle for self-expression, and that it is bound to offer an incorrect or deceitful image of one's true intentions, Hegel goes on to argue that this inadequacy is ultimately due to the limited scope of the whole dialectic. The individual is incapable of overcoming the contradiction between seriousness and deceit precisely because he still thinks in individual terms. Granted, the dialectic has already shown that self-expression is not a private process, but one that is influenced or affected by public opinion. However, this is not enough. Reason must move beyond this superficial insight and reveal the whole dynamic of self-expression as an intrinsically *social* process.

Each individual, in seeking to express himself through his work, is confronted by other individuals who seek to do the same. What we have, then, is not one rational agent opposed to others *in general*, but a highly complex web of self-expressing agents, all of which aim simultaneously at the 'heart of the matter' and at social recognition. But Hegel's point is not simply that one's work is influenced (or perverted) by social exposure. He also claims that the work of each individual is a *condition*, in a more or less direct manner, of other people's work. And this is made clear, he argues, by observing how one's interest in a specific subject is often the starting point of other people's interest on the same subject, or how one's findings often rest on, or derive from, the findings of others. These observations alter the nature of the 'heart of the matter' and reveal each individual's activity as the contribution to a *universally shared activity*. If, on the one hand, what is essential in work is the expression of each agent's particularity, on the other hand 'it is learned by all parties that they *all* are affected, and regard themselves as invited

[457] Cf. *HW* III, 294ff. In order to illustrate this paradox, R. Solomon gives the example of 'a scholar laboring for years on a study of a philosopher who has been dead for 150 years; on the one hand, no task could be more disinterested and free from the suggestion of personal gain; on the other hand, no enterprise would be more personally cautious since, for the scholar — as Hegel indelicately puts it — "the most important thing is his own vanity."' (*In the Spirit of Hegel*, ed. cit., 517)

[to participate], and instead of a *mere* 'doing', or a separate action, peculiar to the individual [who opened up the subject matter], something has been opened up that is *for others* as well.'[458]

This last sentence sets the stage for Reason's final metamorphosis. It has now been established that individual activity is not an independent or self-sufficient process, but a mere moment within a wider social web. Yet this conclusion does not exclude the possibility of a return to the 'way of the world' — that is, the dependence of each on all may simply result in a global mosaic of self-interested standpoints, led out of necessity to collaborate with each other. This is not, however, the kind of interdependence Hegel has in mind. He aims at a much stronger universal bond, and regards this new transition as the first true break from the liberal logic of the previous standpoints. The 'heart of the matter' 'is such 'only as the action of *each* and *everyone*: the essence which is the essence of all beings, viz. *spiritual essence* [*geistliches Wesen*] […] [It] no longer has the meaning of a predicate, and loses the characteristic of lifeless abstract universality. It is rather substance permeated by individuality, *subject* in which there is individuality just as much as *qua* individual, or *qua this particular* individual, as *qua all* individuals.'[459]

This decisive passage, where the notion of *Geist* is already fully present, brings to mind Hegel's well-known claim that the True should be expressed both as Substance and Subject. The methodological description offered in the preface,[460] which I have discussed earlier when considering the *Phenomenology*'s global structure,[461] can now be repeated almost word for word: 'starting from the Subject, as though it were a permanent ground, [consciousness] (*i.e.* the rational agent) finds that, since the Predicate (i.e. 'heart of the matter') is really the Substance (i.e. the ethical substance, *Spirit*), the Subject passed over into the predicate, and, by this very fact, has been sublated.'

This new sublation paves the way for the emergence of Spirit, and for the *Phenomenology*'s second half. But before closing the net on individuality, Hegel dwells a little longer on the inner structure of this newly found standpoint. Since individual activity was revealed as the 'action of each and everyone', human actions can no longer

[458] *HW* III, 310; *PS*, 251. Translation altered.
[459] *HW* III, 310f.; *PS*, 252
[460] See *HW* III, 56ff.
[461] See above section 6.2.

answer to the individual demands of self-expression. They must now be guided by universally valid laws, which can be acknowledged and shared by everyone. These laws are *moral* laws, and their content is the 'absolute heart of the matter' (*die absolute Sache*).[462]

In the remainder of the Reason chapter, Hegel focuses on the dichotomy between an individual conception of morality — namely *Moralität*, based on the Socratic-Kantian imperative of rational self-determination — and a social conception of morality — namely *Sittlichkeit*, or 'ethicality', based on a wider system of communally shared values. This key Hegelian theme, already discussed in earlier writings (notably in the *Naturrecht* essay (1802-3) and in the *System der Sittlichkeit* (1803)), is further developed in the *Encyclopaedia* and the *Philosophy of Right*, and its importance cannot be stressed enough. The idea that self-determination is from the very beginning a social or communal matter, grounded in a dialectical process of recognition, is the main reason behind Hegel's transition from a philosophy of Reason to a philosophy of Spirit, and the key to understanding the articulation between the *Phenomenology*'s first and second halves.

Yet despite the importance of this transition, Hegel's exposition is brief and somewhat obscure. It starts with Reason's acknowledgement of the need for a morally-grounded theory of action: if human actions are to rise above mere arbitrariness, human agents must be in possession of a universal criterion with which to distinguish what is right from what is wrong, what should be done from what should be avoided. Hegel's primary interlocutor here is Kant, and particularly his *Metaphysik der Sitten*. The issue can thus be translated in Kantian terms as the need to guide oneself by a 'universal law of action', which 'must serve human will as its only principle.'[463]

Initially, this law appears to be self-evident: 'sound Reason knows immediately what is right and good', and the individual's only task is to act accordingly. However, for this universal law to provide actual moral guidance, it must be rendered determinate, that is, it must be translated into a series of concrete moral precepts. Hegel undertakes a brief discussion of two of these precepts, namely 'everyone ought to speak the truth' and 'love thy neighbour as thyself.' As usual, his aim is to show that these mandates are not as

[462] *HW* III, 311
[463] *AA* IV, 402

straightforward as they appear. On closer inspection, their content turns out to be ambiguous and contingent, and this discovery undermines their allegedly universal authority. Furthermore, since these precepts are supposed to be 'immediately given', their origin remains a mystery, as does their mutual articulation: it is at best unclear why these specific precepts were deemed universally binding, as opposed to a myriad of other possible precepts, and not a word is said about whether or how they are supposed to condition one another.

These difficulties lead consciousness to abandon its initial standpoint and to devise a new and less ambitious solution. In the chapter's third and last subsection, Reason is no longer said to *provide* the content of moral laws, but rather the *criterion* with which this content can be tested. However, the criterion in question is not just any criterion, but a *moral* one, which means that the test carried out by Reason must focus exclusively on each precept's logical consistency, and disregard the actual consequences of its endorsement. This procedure is directly evocative of Kant's so-called 'universalization' test — that is, his contention that the validity of a moral precept must be assessed not on empirical or contingent grounds, but by determining whether it can be successfully transformed into a universal maxim.[464] For Hegel, this criterion is inconclusive and incapable of grounding a truly operative moral system. To illustrate his point, he turns to a brief discussion of the notion of property, aiming to show that the law that there should be private property, as well as the law that there should *not* be private property, can both be said to fulfil and to violate Kant's formal criterion: when considered in isolation, they are both logically consistent (and hence 'universalizable'), but in the concrete, when it comes to specific rulings and regulations, they are easily shown to be contradictory (and hence 'non-universalizable'). Hegel argues that this ambiguity extends to every other moral law and proceeds to announce (perhaps too swiftly)[465] the *Aufhebung* of Reason's last phenomenological resort.

[464] See notably *AA* IV, 402, 421, 437ff.

[465] Many commentators have denounced the brief and ultimately superficial way in which Hegel, at this stage, addresses Kant's moral philosophy. Indeed, it is not at all clear that Hegel's passing discussion of the right to own property (and its opposite) should result in a complete or convincing refutation of Kant's universalization principle. For readings that are sympathetic to Hegel's objection, in the *Phenomenology* and elsewhere, see, for ex., SMITH, *Hegel's Critique of Liberalism,* ed. cit., 73ff.; SOLOMON, op. cit., 528-533; or TAYLOR, *Hegel*, ed. cit., 168ff., 371; or WALSH, op. cit., chapter 4. For further reading on Hegel's

In the two previous solutions, as Reason sets out to provide and then to test the content of moral laws, it is still confined to the sphere of *Moralität*. The principles that guide human actions are still construed as a series of commandments prescribing what 'ought' to be, which can be accepted or rejected according to individual rational judgements. The moral examiner places himself above actuality, as though he did not belong there, and sets out to determine what is good and right by passing judgement on human life. His mandates, however, are still 'only a *willing* and *knowing* by this particular individual'[466], unilaterally raised to universality. They are not yet the expression of life's true or actual universality, which can only be grasped as an ethical, or *spiritual* phenomenon.

Unlike morality, *Sittlichkeit* is not abstract or arbitrary, but concrete and actual. It is not the result of a decree, but the native element of human life. As Hegel puts it, ethical laws simply *are* — they are *already there*, *already at work*, woven into one's very identity, embodied from times immemorial in the institutions and traditions that shape human life. Their content is not something to be posited or tested, but something to be acknowledged. Accordingly, if these laws 'are supposed to be validated by *my* insight, then I have already denied their unshakeable, intrinsic being and regard them as something which, for me, is perhaps true, but also is perhaps not true.'[467] This, for Hegel, is contradictory, since the 'ethical disposition [*die sittliche Gesinnung*] consists just in sticking steadfastly to what is right, and abstaining from all attempts to move or shake it, or derive it.'[468] Now these and similar statements might be taken to mean that 'ethicality' is really about obedience, or about submitting unreservedly to the *statu quo*. If that were the case, this new stage would surely amount to a step back rather than a step forward in consciousness' phenomenological journey. But Hegel's point, at this stage, is simply that *Sittlichkeit* is the necessary starting point for any serious consideration of human life. In order to be examined and judged, ethical values must first be acknowledged, which can

engagement with Kantian morals, see, for ex., AMERIKS, op. cit.; PATTEN, op. cit., 82-103; PIPPIN, *Idealism as Modernism*, ed. cit., 92-128; SIEP, *The Aufhebung of Morality*, ed. cit; or WOOD, op. cit., esp. 154ff.

[466] *HW* III, 320; *PS*, 260

[467] *HW* III, 322; *PS*, 261f.

[468] *HW* III, 322; *PS*, 262

only truly happen once Reason ceases to be regarded as an individual phenomenon, detached from its social and historical environment. As Robert Solomon points out, 'this is not to say that *Sittlichkeit* has not rules, or that the rules cannot be criticized on even rejected on rational reflection; but it is to insist that first there must be such rules and the practices in which they are embodied. Society is not, as in Hobbes and Rousseau, first of all a conference for the purpose of establishing the rules; society is that set of rules and practices. Their justification and legitimacy, the search for reasons, comes later, if at all.'[469]

With the dawn of the 'ethical world', the atomism that presided over the previous phenomenological stages is finally sublated and the subject sees itself in a new light. Consciousness, Self-Consciousness and Reason, falsely posited as concrete cognitive principles, are now unmasked as the abstract semblances of Spirit — that is, as the abstract, pre-social and pre-historical semblances of a *concrete*, *socially mediated* and *historically situated* self:

> Spirit is thus self-supporting, absolute, real being. All previous shapes of consciousness are abstracts forms of it. They result from Spirit analysing itself, distinguishing its moments, and dwelling for a while with each. This isolating of those moments presupposes Spirit itself and subsists therein; in other words, the isolation exists only in Spirit which is a concrete existence.[470]

Every human self is a 'spiritual' self — this decisive acknowledgement marks the accomplishment of the *Phenomenology*'s first main task. All of its previous stages, criticized and abandoned in favour of more comprehensive shapes of consciousness, are now brought together within Spirit's global perspective. What was previously abstract and isolated is now the element of a concrete and dynamic whole, in which every particular moment determines, and is determined by, all the others.

This global reunification is the result of a double movement, whose advances and setbacks were documented in the previous chapters. Consciousness has struggled to bring together, on the one hand, the spheres of subjectivity and objectivity (viz. selfhood and

[469] *In the Spirit of Hegel*, ed. cit., 534f.
[470] *HW* III, 325; *PS*, 264

otherness), and on the other hand the spheres of subjectivity and intersubjectivity (viz. individuality and universality). Accordingly, Spirit is not a simple, but a double synthesis, or the simultaneous sublation of these two dialectical contrasts.

As regards the first one, recall that the notion of Reason, as it first came on the scene, was described as consciousness' 'certainty of being all reality'. The antithesis between subject and object, maintained throughout the dialectics of consciousness and self-consciousness, was replaced by the dialectical unification of selfhood and otherness, but this new coincidence was not immediately understood. Confined at first to an empty form of idealism, Reason was led to move out into the world and to search for truth in the concrete element of natural existence. The transition to the stage Spirit is the result of this long and eventful journey. In Hegel's words, 'Reason is Spirit when its certainty of being all reality has been raised to truth, and it is conscious of itself as its own world, and of the world as itself.'[471] Whereas initially Spirit was consciousness that only *has* Reason, it has now been grasped as 'Reason that *exists*, or as Reason that is *actual* in Spirit and is its world.' Only now, then, is the notion of Spirit fully realized: 'it *is* Spirit, the *ethical* essence that has an *actual* existence.'[472]

Moreover, Reason's journey has also been centred on the contrast between subjectivity and intersubjectivity. Throughout its various stages, consciousness has attempted to articulate its own identity and interests with the identity and interests of others, by means of different conceptions of personal and common good. But these efforts took the form of arbitrary principles, issued unilaterally by Reason and disconnected from the actual laws and traditions that shape the existing world. The acknowledgement of this disagreement led to the replacement of the *moral* framework adopted by Reason with the *ethical* framework introduced by Spirit, recognized at last as the underlying principle of human life. Whereas the dialectic of Reason was still about individual model self-determination, Spirit is 'the *ethical* life of a people [*Volk*], in so far as it is the *immediate truth* — the individual that is a world.'[473]

[471] *HW* III, 324; *PS*, 263
[472] *HW* III, 326; *PS*, 265
[473] Ibid.

Henceforth, the progression will no longer deal with individual shapes of consciousness, but with collective ones. This does not mean, though, that the *Phenomenology* will now focus on crowds of people instead of individuals, or that it will now offer a general exposition of world history. Hegel's point is rather that the essence of human thoughts and actions cannot be grasped through an examination of their personal motivations. It calls for a wider analytical grid, which is able to capture both the specific nature of each individual self-consciousness and the universal element on which it is grounded. Since each individual is now a whole world, Spirit is likewise the dialectical unity of individuality and universality: on the one hand, it is 'the unmoved solid *ground* and *starting-point* for the action of all, and it is their purpose and goal, the in-itself of every self-consciousness expressed in thought'; on the other hand, 'it is a fragmented being, self-sacrificing and benevolent, in which each accomplishes his own work, rends asunder the universal being, and takes from it his own share.'[474] The immanent cohabitation of these two moments will now fuel the experience of Spirit, whose main steps will be briefly outlined in the following pages.

9.2. *The trials of Spirit*

Spirit is at first 'true spirit' and its original seat is the ethical world. In this new phenomenological setting, individuality and universality are immediately joined together. The private and the common good are not at variance with one another because each finds in the other its natural expression. The dialectic of *Sittlichkeit* draws on a theme that is dear to Hegel and to a whole generation of German thinkers. The harmony of the ethical world is linked with the harmony of the ancient Greek world, and with the moral innocence of ancient Greek life.[475] For Hegel, the Greek world represents the first

[474] *HW* III, 325; *PS*, 264

[475] Among the various sources that might have influenced Hegel's characterization of ancient Greek *Sittlichkeit*, it is worth highlighting Herder's essay *Vom Erkennen und Empfinden der menschlichen Seele* (1778), which deals with the opposition between the primitive unity of human consciousness, embodied by Greek culture, and the pathological dividedness of modern consciousness; Schiller's philosophical writings — namely the series of letters *Über die ästhetische Erziehung des Menschen* (1795) — and his

incarnation of the 'ethical substance'. It is an 'immaculate world' 'unsullied by any internal dissension'[476], where individual desires are naturally aligned with communal aspirations.

But the ethical world is nonetheless a self-differentiated world. Spirit's universal and individual moments are present within it in the form of two general laws: a divine law, associated with the private sphere of family life, and a human law, associated with the public sphere of political life. The first of these laws proceeds directly from the gods and regulates the most basic aspects of human existence. Its content is passed on from generation to generation, honoured by ancestral rites and ceremonies. This law is primarily linked to the female sex, inasmuch as women were traditionally in charge of the household and the family, destined to watch over the most intimate dimension of human life. The human law, on the other hand, is upheld by the state, and regulates the life of the *polis*. Here, individuals are not regarded as family members, but as citizens, called upon to fulfil their duties towards the community. The human law is primarily linked to the male sex, inasmuch as men were traditionally in charge of politics, trade, warfare, and all other public affairs.

In the first stage of the dialectic, these two laws complement each other. Whereas the family's main role is to educate the individual and to prepare him for the duties of citizenship, the state's main role is to include him in society and to ensure that his actions serve the interests of the community. At the end of the citizen's life, however, this movement is reversed: his body is given back to the family, in order to be buried, and the divine law claims back its initial sovereignty. This cycle is harmonious because both laws have jurisdiction over different aspects of human life, and guarantee thereby the maintenance of the ethical order. But how stable is this alternation? What happens if both laws are called upon to judge the same human deed? This is the scenario envisaged in the next stage of the dialectic, as Hegel moves from the stability of the 'ethical world'

'classical' poetry — notably the poem *Die Götter Griechenlandes* (1788-1800), among many others; Hölderlin's views on the opposition between the lost innocence of Greek culture and the abstract disenchantment of the modern world, poignantly depicted in the epistolary novel *Hyperion, oder der Eremit aus Griechenland* (1797-99). For a useful discussion of Hegel's interpretation of ancient Greek culture, see JANICAUD, op. cit. For an overview of German idealism's engagement with ancient Greece, from Kant to Hölderlin and Hegel, see TAMINIAUX, op. cit.

[476] *HW* III, 341; *PS*, 278

(*die sittliche Welt*) to the instability of the 'ethical action' (*die sittliche Handlung*) and starts unpacking the model's inner contradictions.

The main reference for this new dialectical stage, already anticipated by Hegel's insistence on the importance of burial rights, on the social dichotomy between man and woman and on the conflict between kinship and citizenship, is of course Sophocles' *Antigone*.[477] Following the death of Polynices, pretender to the throne of Thebes, King Creon decrees that his body shall remain unburied. Polynices is deemed an enemy of the city and sentenced to perish as such.[478] His sister Antigone defies this order, buries his body and is subsequently incarcerated. The conflict expounded in the play stems from the fact that both contenders — Antigone and Creon — are legitimized by a different law and unwilling to abandon their position. Antigone, the woman, represents the divine law, and views her brother as a relative to be honoured. Creon, the man, represents the human law, and views Polynices as a citizen to be punished. This conflict is insoluble because both sides are completely immersed in their natural roles and unable to consider each other's point of view. They are confined to an immediate ethical mandate, and have no choice but to reaffirm what they know to be right.

The ethical world has reached a deadlock, which it is unable to overcome. Since law is still construed in an immediate or natural way, social roles are still merely given. They amount to one-sided stances thrust against each other and bound for destruction. The acknowledgement of this limitation reveals 'the germ of deterioration inherent in the beautiful harmony … of the ethical Spirit'[479] and precipitates the collapse of the Greek world and the transition to a new social order. As the bond that unites the citizens is severed, the ethical community breaks down into a simple aggregate. Citizens are replaced by free individuals, or free-standing 'persons', whose life is devoted to the pursuit of self-interest. This new solution, which Hegel equates with the juridical

[477] Another subject dear to Hegel. Cf. *HW* XV, 544, 549f.; and *HW* XVII, 133.

[478] Polynices and Eteocles, sons of Oedipus, had both been offered the throne of Thebes, but Eteocles managed to expel his brother from the city with the support of the Theban people. Polynices travelled to Argos, where he rounded up a small army, and launched a campaign to reconquer Thebes. In the ensuing battle, which is the subject of Aeschylus' *Seven Against Thebes*, the two brothers ended up killing each other in the battlefield and Creon was crowned king of Thebes. In his eyes, Polynices is a traitor because he turned against his native city, and must not be given any burial rites.

[479] *HW* III, 354; *PS*, 289. Translation altered.

freedom of the Roman world, echoes the previous transition from the stage of 'virtue' to the 'way of the world'. Like before, society is now reduced to the total sum of its members and the common good is reduced to a mosaic of particular aims and aspirations.

While Greek citizenship rested on a positive and robust conception of self-identity, personhood is conceived in a negative and contingent way. A person is no longer defined by its convictions or by its place within the community, but by the sheer fact of being a person. And since this qualification is still an empty qualification, it must be filled up with the enjoyment and acquisition of external things. Social recognition becomes dependent on the acquisition of property, which leads in turn to a redefinition of personal freedom. Whereas previously the welfare of others was a condition for personal welfare, the latter is now the primary element of the relationship. Freedom is reduced to a negative concept, that is, to the freedom of pursuing one's interests and enjoyment without the interference of others. And since this new prerogative is no longer grounded in a natural social equilibrium, it must be upheld in an artificial manner, through the legal regulation of civil rights.

Each person is now a self-interested, legally protected, property-owning individual. This definition is not Greek, but thoroughly modern. Historically, it leads as far as the advent of civil society, the rise of political liberalism and the development of modern capitalism. But Hegel's aim here is not to paint this wider historical picture. As always, he is interested in the model's immediate contradictions, and finds them at hand within the Roman imperial world. His argument consists in showing that a society exclusively predicated on self-interest is bound to degenerate into inequality and despotism. The contradiction arises from the fact that the legal right to accumulate property, and hence guarantee personal freedom, is compromised by the very right to accumulate property, inasmuch as the latter leads to a disproportionate concentration of wealth. In a society ruled by the rich and the privileged, political power is disconnected from the common good. Civil rights are compromised and the legal order gives ways to a thoroughly arbitrary world, where law and morality are outweighed by greed and caprice.

This experience opens the door to a whole new series of contradictions, which Hegel associates with the moral and political history of the modern world. Led from the 'beautiful harmony' of the Greek world to the complete disharmony of the Roman

world, Spirit must now attempt to restore its self-integrity. It cannot do so, however, by returning to the purity of the ethical world and ignoring its contradictions. The solution must be born out of the ruins of the legal world, and requires a new set of concepts and articulations.

The dialectic of 'true Spirit' is replaced by the dialectic of 'self-alienated Spirit' (*der sich entfremdete Geist*), whose first seat is the realm of modern culture (*Bildung*). Hegel's argument here consists once more in a series of variations on the relationship between the individual and the state, with consciousness alternating between the praise of state power over private wealth and the praise of the latter over the former. This first impasse is overcome when both terms are joined together and deemed either wholly beneficial or wholly prejudicial. The first of these stances is held by a 'noble consciousness' (*das edelmütige Bewußtsein*), while the second is held by an 'ignoble consciousness' (*das niederträchtige Bewußtsein*). For the latter, the individual and the common good can be effectively reconciled, provided that each contributes to the preservation and development of the other. The state is envisaged as a just institution, whose primary goal is to increase the wealth of its subjects, whereas wealth is viewed as a vehicle for individual happiness, which is in turn a condition for the welfare of the state. In order to help turn this vision into reality, the noble individual puts his energy and wealth at the service of others. He renounces his immediate self-interest and embraces a life of heroism, which ennobles and strengthens the authority of the state. In doing so, however, he unwillingly sets in motion a new version of the dialectical contradiction unveiled in the previous model. As the state grows stronger, its power increases and the distance separating it from individuality is widened. A more powerful state demands a more complete submission on the part of its subjects, and the individual's noble heroism is thus no longer enough. He must now recognize the authority of an absolute monarch, and replace his noble attitude with a subservient and flattering one, so as to adequately express his complete subjection.

The transition from nobility to flattery, accompanied by suggestive allusions to French absolutism and the court of Versailles, prompts the re-emergence of the 'ignoble consciousness'. Through flattery, the noble character is corroded and demeaned. At the same time, however, this self-abasement gives rise to a feeling of revolt, directed both at the power of the state and at the power of wealth. Spirit now realizes that these two

moments cannot be reconciled: the state is perceived as a despotic and arbitrary power, opposed to private welfare, and wealth is no longer a source of happiness, but an instrument of subjection. The nobleman, caught up in this impossible cycle, regards his own situation as ignoble and humiliating. The vanity of his previous sacrifices is gradually impressed upon him, but his material subjection prevents him from breaking free. Instead, his moral integrity degenerates into a scornful cynicism, aimed at everything supposed to count as serious, important or noble. This new attitude is the exact opposite of the idealism expressed at the beginning of the dialectic. Spirit is no longer noble or base, but the very consciousness of this opposition. Whenever something is deemed noble or admirable, it points out its intrinsic baseness; whenever something is deemed base or shameless, it immediately praises its nobility. By reducing everything to its contrary, it celebrates the 'absolute perversion' of modern social life. Its only truth is a 'destructive judgement' aimed at the whole range of accepted norms and values, 'a nihilistic game … [which] overpowers everything'.[480]

This 'disrupted consciousness' (*das zerissene Bewußtsein*) is not merely faced with a contradictory world. 'As self-consciousness in revolt', 'it is aware of its disrupted state, and in thus knowing it has immediately risen above it.' Through the conscious dissolution of everything serious and stable, 'all content is turned into something negative, which can no longer be grasped as having a positive significance.' Consequently, 'the positive object is merely the *pure "I" itself*, and the disrupted consciousness *in itself* this pure self-identity of self-consciousness that has returned to itself.'[481] The next dialectical stage documents, therefore, a new return to self-consciousness, but one that can take two different forms. On the one hand, Spirit proposes to find a refuge from the vanity of the outside world within itself, and decrees the truth of objectivity to be the pure essentiality of thought. On the other hand, it transfers this same truth to a divine supersensbile world, placed beyond subjectivity and objectivity. In the first case, Spirit

[480] *HW* III, 386; *PS*, 317. Hegel's primary source here is Diderot's *Le neveu de Rameau*, which documents the subversive dialogue between an 'honest', philosophically-minded narrator (closely identified with Diderot himself), and the extravagant nephew of the French composer Jean-Baptiste Lully, whose incendiary views on morality, art, human relations and social conventions encapsulate the various contradictions of modern culture.

[481] *HW* III, 390; *PS*, 321

becomes 'pure insight' and endorses a rationalist worldview; in the first case, Spirit becomes 'faith' and endorses a religious worldview. What follows is the dialectical confrontation between Christian faith and the revolutionary ideals of the Enlightenment (*Aufklärung*), in a direct evocation of the intellectual debate that dominated Europe from the late 17th century up until Hegel's lifetime.

As always, the dispute starts out as a simple antithesis, with each side rejecting the other's professed standpoint. For the Christian, the Enlightenment's unconditional faith in rationality is naïve, for human reason, just like the objective world, is ultimately a divine gift, whose contradictions can only be solved by appealing to a higher cognitive plane. For the enlightened thinker, on the other hand, religious faith comes down to a series of mistakes and superstitions, imposed on the ignorant masses by 'the will of the deceiving priesthood and of the oppressive despot.'[482] The Enlightenment's aim is to undo this domination, to reawaken men's rational instincts and to win them over to its side of the barricade.

Although Hegel maintains, throughout this section, that the opposition between faith and reason is still an abstract opposition, and that both sides have much more in common that they are willing to acknowledge, this is not yet the place for a reconciliation. The dialectic will focus, rather, on what Hegel views as the overall superficiality of the Enlightenment's understanding both of religion and of its own conceptual project. Among the various accusations levelled against Christianity, Hegel singles out the denouncement of faith's supposed unrealism (the idea that God is either a mental fabrication and/or a political deception), of its contradictory relapse into empiricism (the worship of physical objects, the reliance on the Bible's historical text) and of its 'foolish' asceticism (the irrational renouncement to natural enjoyments and private property). All of these ideas reveal, in Hegel's eyes, the Enlightenment's constitutive inability to grasp the true essence of religious faith.

Afterwards, the exercise is reversed and the Enlightenment is called upon to offer its own version of truth. Averse to all forms of transcendence, the enlightened thinker

[482] *HW* III, 401; *PS*, 330. Hegel plays with the opposition, lost in the English translation, between *Glaube* and *Aberglaube*. For the Enlightenment, Christian faith is not a real faith, but a *pseudo*-faith, or a *deception*.

reduces God to a mere 'vacuum' and turns his focus on the natural world. He confines reason to a thoroughly secularized standpoint, whose main ideals are concrete objectivity and utility. *Das Nützliche*, the 'useful', becomes Spirit's true essence, and with it the Enlightenment claims to have put to an end to faith's vacuous musings and to have 'transplanted heaven to earth down below.'[483] This presumption will be tested in the third stage of the dialectic of self-alienated Spirit, where this 'earthly heaven' is transferred to the social realm and transformed into a concrete political doctrine.

The Enlightenment's political worldview is brought in to correct the contradictions of the previous social models, where the individual and the state have remained at odds with one another. In the ethical order, private and communal interests were equated, but there was yet no room for self-determination. Social roles were immediately given and social conflicts ended in strife and death. In the legal order, on the other hand, the emphasis was placed on individual freedom, or free personhood, but the common good was thereby sacrificed. Free particularity led to the unchecked accumulation of wealth and power, which generated tyranny and self-alienation. Faced with this dilemma, the new model must retain the positive aspects of the two previous two models and at the same time correct their inner contradictions. It must be able to accommodate both the Greek concern with society's global harmony and the modern concern with the freedom to pursue individual aims and aspirations. As Bruce Gilbert has recently put it, 'the singular members of this society must choose, but choose in such a way as to ensure that particularity "makes no difference," and thus choose for the sake of the universal — the common good. Any particularity that does make a difference is a faction — a particularity that defeats the rationality of society.'[484]

This new synthesis of personal and universal self-fulfilment is notoriously associated, in the *Phenomenology*, with the French Revolution and the so-called Reign of Terror, which lasted from September 1793 to July 1794. Throughout this relatively short section, Hegel's description echoes Rousseau's conception of a purely participatory democracy, grounded in the abolishment of all forms of political representation, on the elimination of the state's intermediate administrative structures and on an equal

[483] *HW* III, 431; *PS*, 355. Translation altered.
[484] *The Vitality of Contradiction*, ed. cit., 123

distribution of property. In Hegel's terms, Rousseau's ambitious project, as well as the political efforts of key revolutionary figures such as Robespierre, Danton or Saint-Just, were aimed at the attainment of 'absolute freedom'. 'All social groups or classes, which are the spiritual spheres into which the whole is articulated, are abolished', and the purpose of each individual consciousness 'is the general purpose, its language universal law, its work the universal work.'[485] This very formulation, however, is already indicative of the project's inner inconsistency. For if personal ideas and deeds are nothing else than general ideas and deeds, their content is simply the formal demand that they become universal, and hence no actual content. The absolute freedom of Enlightenment is really only the movement whereby determinacy and particularity, which are bound to re-emerge again and again in the course of social life, are cancelled and raised to universality. Since everything concrete is born out of particularity, and perceived therefore as a subversion of the 'general purpose', all that is left is pure negativity. Society degenerates into a police state, obsessed with detecting and denouncing all remaining traces of individuality. The universal harmony promised at the outset leads once again to disharmony, tyranny and death.

This experience has driven Spirit to a new peak of self-alienation, and will now prompt a new return to itself. Just as the 'disrupted consciousness' sought refuge in faith and pure insight, so too will absolute freedom 'leave its self-destroying reality and pass over into another land of self-conscious Spirit, where … freedom has the value of truth.'[486] This new 'land' is the stage of Morality, the first in a whole new series of phenomenological transformations. Up until now, Spirit was unable to find the harmony it seeks in the concrete realm of social relations. Henceforth, it will attempt to set up a new system of moral values, on which to ground its thoughts and actions. The richness and complexity of this new string of chapters, which would merit an entire book of its own, must be reduced here to a series of general indications.

In this new dialectic, Hegel returns to Kantian morality and expands the criticism sketched in the final stage of the dialectic of Reason. His overall strategy is taken over from the dialectic of absolute freedom: in the Rousseaunian model, particular and general

[485] *HW* III, 433; *PS*, 357. Translation altered.
[486] *HW* III, 441; *PS*, 363

aims were immediately joined together, but not truly synthesized, which led to their final divorce; in Kant's account of practical reason, according to Hegel, moral virtue is likewise juxtaposed to the concrete element of human life. Kant's moral doctrine is founded on the antithesis between the actual content of human choices and what that content *ought to be*. Since moral goodness is defined as that which all actions must aspire to, but can never really attain, it is bound to remain a pure *beyond*, or a mere *Sollen*. This general line of criticism is illustrated by a critique of Kant's three postulates of practical reason, namely the practical need to believe in personal freedom, the practical need to believe in immortality and the practical need to believe in the existence God.[487] For Kant, although these beliefs cannot be theoretically demonstrated, they are the necessary foundations of every moral judgement. If human life were confined to a purely eudaemonistic framework, unpleasant choices and decisions, however moral, would never be made. For morality to prevail over happiness, its motivation must lie outside the eudaemonistic sphere, in the ideas of freedom, immortality and God.

Firstly, Hegel argues that the Kantian postulate of freedom arises from the spurious presupposition that the natural world, governed by natural causality, is separated from the moral world, governed by moral mandates. This separation is spurious because reality's agreement with morality is grounded in human actions themselves, and not in some abstract postulate. Although moral consciousness 'starts from the idea that, *for it*, morality and reality do not harmonize, … it is not in earnest about this, for in the deed the presence of this harmony becomes *explicit for it*.'[488] Consciousness is not 'in earnest' about morality because what it deems morally necessary is ultimately also what it wants to see happen. 'What it really holds to be most desirable, to be the Absolute, is that the highest good be accomplished, and that moral action be superfluous.'[489]

This 'duplicity' (*Verstellung*)[490] is also brought out in Hegel's critique of the postulate of immortality, which locates moral goodness at an infinite distance, beyond the confines

[487] Cf. *AA* V, 132. These the three main declensions of Kant's notion of *metaphysica naturalis*, mentioned at the outset of this work (see above fn. 1).

[488] *HW* III, 456; *PS*, 377

[489] Ibid.

[490] New play on words: the *Verstellung* that characterizes the 'moral worldview' is both a *dissimulation* — insofar as morality is not what it seems, but a dissimulation, a pretence, a ruse, etc. — and a *displacement*

of human existence. Once again, this attitude is grounded in a separation which is not taken seriously. Since the truth of morality is shifted (*verstellt*) to a 'nebulous remoteness, where nothing can any more be distinguished or comprehended, […] the truth of morality is really only the imperfection of actual existence, 'supposed to be a *progress towards* perfection.'[491] Ultimately, then, the moralist 'is concerned not a bout morality, but solely about happiness as such.'[492]

Finally, as regards the postulation of the existence of God, Hegel moves away from Kant's actual argument and bases his critique on the distinction between a 'pure duty' (*die reine Pflicht*), which corresponds to a universal moral imperative, and a variety of 'determinate duties' (*bestimmte Pflichten*), which convey what is morally right in each individual's particular situation. Faced with this distinction, the moralist must find a way of harmonizing the contingency of his specific duties (for ex., his obligation to help his friends or relatives) with the absolute moral necessity of pure or universal duties (for ex., his obligation to help every human being, or to end human suffering). This opposition is bridged with the aid of a divine consciousness, or a divine lawgiver, to whom the moralist entrusts the role of sanctifying particular duties and rendering them pure and necessary. For Hegel, however, this new expedient amount to a new form of *Verstellung*, which opens the door, once again, to an eudaemonistic contamination of the whole model. By legitimating, through God, the adaptation of morality to circumstances, the moralist evades the responsibility of discerning the true content of duty and sticking fast to it.

All of these contradictions lead the moral Spirit to abandon this model and to reject the various dichotomies on which it is based. The truth of morality is brought back, once again, to the objective world, and all duties are regarded as determinate duties, grounded in each individual's determinate reality. This new shape of Spirit is called 'conscience' (*das Gewissen*) and its conception of duty is radically different from the previous one. Whereas morality had based its worldview on the necessary divorce between a *Sein* and a *Sollen*, conscience looks for the content of duty in the concrete element of life, certainty and action. In Hegel's words, 'I act morally when I am *conscious* of performing only pure

— insofar as this dissimulation consists in 'displacing' the unity between reality and morality into an abstract *beyond*, unilaterally postulated by the moralist.

[491] *HW* III, 458; *PS*, 378

[492] *HW* III, 459; *PS*, 379

duty and nothing else but that; this means, in fact, when I do *not* act. But when I really act, I am conscious of an *other* [*etwas anderes*], of a reality which is already in existence, and of a reality I wish to produce; I have a *specific* purpose and fulfil a *specific* duty in which there is something else than the pure duty which alone should be intended.'[493] We find here a new and decisive shift from the sphere of *Moralität* to the sphere *Sittlichkeit*, where duty is just as much a subjective claim as it is a reflection of existing values, conditions and circumstances. In contrast to the moral Spirit, conscience is now immersed in the real world. It observes reality, acts upon it, acknowledges the consequences of its actions and draws its own conclusions. Its behaviour is no longer guided by abstract mandates, but by specific *convictions*. The distinction between right and wrong judgements is no longer determined by moral postulates or deductions, superimposed on the concrete challenges of human life, but by the 'voice of conscience'.

Yet although this new model represents, according to Hegel, an important step forward from the abstract world of morality, it is not yet free of contradiction. For one, insofar as conscience's determination of what is right and wrong is exclusively based on its own personal judgement, it is naturally fallible. A single conscience cannot possibly know all that it needs to know about reality in order to determine whether a given judgement is really good or commendable, much less the best one available. Moreover, insofar as conscience's convictions are personal and arbitrary, they are bound to clash with the convictions held by other consciences.

The remaining part of the dialectic will concentrate on the second of these objections. On one side, conscience is immediately sure (*gewiss*) of its insights. Its self-confidence evolves into a special regard for its own judgements and intuitions — a 'moral genius' which it longs to see recognized.[494] However, precisely because these insights are mere intuitions, their truth cannot be fully shared. When made public, they are met with distrust and incomprehension. In its efforts to overcome this difficulty, conscience is led to refrain from acting, so as to avoid falling into contradiction or betraying its inner convictions. Retreating into itself, it now 'lives in dread of besmirching the splendour of its inner being by action and existence; and, in order to preserve the pureness of its heart,

[493] *HW* III, 468; *PS*, 386
[494] *HW* III, 481; *PS*, 397

it flees from contact with the actual world, and […] vanishes like a vapour into thin air.'[495] In doing so, however, it is gradually becomes the very opposite of what it first appeared to be: instead of a concrete ethical Spirit, immersed in the element of life, conscience has become as vague and elusive as the moralism it first criticized.

On the whole, we are now faced with the opposition between this solitary conscience — or this 'beautiful soul', whose only aim is to remain true to its beliefs — and the anonymous mass of society — a 'universal other' bound to the established moral order.[496] While the former refuses to abide to social conventions, which it sees as falling below its inner standards, the latter condemns the former's self-isolation and resents its sanctimonious attitude. For the conscientious individual, the other is a mere critic, who fails to grasp its pure integrity. For the majority, conscience is *sinful*, or *evil*, and must be brought back from isolation. The two sides are pitted against each other and can only come together through a new and more sophisticated form of *Anerkennung*, which leads to the conclusion of the dialectic of Spirit. In this final and somewhat obscure section, Hegel speaks of reconciliation in terms of *confession* and *forgiveness*. In the course of the argument, both parties are shown to be essential to the definition of Spirit, but incapable, on their own, to provide its truth. Whereas the individual 'I' of conscience stands for pure self-certainty, devoid of actuality, the universal 'I' of society stands for pure actuality, devoid of self-certainty. In other words, whereas the first moment is Spirit *for itself,* the second moment is Spirit *in itself*. Their final reunification comes about when conscience *confesses* its obstinacy to the other, and becomes part of it, and when the latter *forgives* this obstinacy, and finds itself therein. The result of this process is describes as follows:

> The reconciling *Yes*, in which the two 'I's let go their antithetical *existence*, is the *existence* of the 'I' which has expanded into a duality, and therein remains identical with

[495] *HW* III, 483f.; *PS*, 400

[496] The term *schöne Seele* is borrowed from Goethe's *Wilhelm Meisters Lehrjahre* (1795), which together with Rousseau's *Confessions* (1782) seems to be on the back of Hegel' mind throughout this entire section. On the context and meaning of this section, see NORTON, op. cit.; PAHA, op. cit.; or SICKENDIECK-MEYER, op. cit., 221-238.

itself, and in its complete alienation and opposite possesses the certainty of itself: it is God manifesting in the midst of those who know themselves as pure knowledge.[497]

The sinuosity of this final section, liable to be decomposed and reconstructed in many different ways, should not make us lose sight of the dialectic's global meaning. Throughout the chapter devoted to Morality, Spirit claimed to be 'certain of itself' in different, but equally inconsistent ways: it looked for its essence in a series of moral postulates, in the immediacy of concrete existence and in the purity of its own convictions. In all of these cases, however, the self remained self-divided. Whether cut off from the objective world or cut off from other subjects, it was finally led to realize that self-certainty is but one moment of Spirit. To be 'certain of itself', Spirit must also be certain of its world. Or rather, it must be certain of being the unity of itself and its world. Or again, it must find itself in its world as both subject and object, individual and universal, certainty and truth. Only the acknowledgement of this reciprocity can reveal Spirit's true identity: 'the human subject who creates his own history, who no longer finds the universal outside of himself, who no longer opposes the universe, but carries it within and absorbs it into himself.'[498]

With the sublation of conscience, this unity appears to have been reached. Spirit's individual and universal moments have said 'yes' to one another and were finally merged together. But Hegel now suggests that this unity, albeit complete, was accomplished at a horizontal level, and not yet at a vertical one. The unity of the self with its world has assumed the form of a living social organism, but its life is still a worldly or secular life. It has yet to work out its relationship with divinity, left unresolved in the previous stages of the progression. This is why Spirit's 'pure knowledge' finds 'God manifesting itself in its midst', and why this new standpoint still calls for a higher form of synthesis.

[497] *HW* III, 494; *PS*, 409. Translation altered.
[498] HYPPOLITE, op. cit., 477

9.3. *Tragedy and comedy*

The relationship between reason and 'faith' was introduced in the stage of unhappy consciousness, discussed in connection with the ethical world and redefined with the Enlightenment's critique of religion. Since all of these stages led to contradiction, Spirit must now revive this debate and look for a higher form of reconciliation. This is the task undertaken in the stage of Religion, the last one in the progression. Like always, the dialectic begins with the most basic form of religious experience, directly associated with the natural world. The contact between the human and the divine realms is initially mediated by sunlight, expressed afterwards through the adoration of plants and animals, and finally replaced by the fabrication and worship of man-made artefacts, whose initial geometric shape gives way to the depiction of plants, animals and human figures. These objects, however, still offer a limited image of divinity, unable to convey the complexity of the self's religious feelings. They represent god as a mere 'other', opposed to the worshipper and confined to a chunk of wood, or a block of stone. For a higher communion between Man (Spirit *for itself*) and God (Spirit *in itself*), the craftsman must put himself explicitly into his work and transform the object into a living expression of self-consciousness. Only by being 'purged of the immediate shapes of Nature' and 'dissolved into spiritual shape'[499] can the object become a true meeting point of the self's earthly existence and its divine calling: the concrete interface 'where Spirit meets Spirit', or 'the unity of self-consciousness Spirit with itself.'[500]

In response to this new phenomenological demand, the craftsman becomes an *artist* and 'natural religion' is replaced by 'religion in the form of art' (*Kunstreligion*). Spirit will now assess different forms of artistic expression and the extent to which they succeed in bridging the gap between the human and the divine. Starting with the 'abstract' experience afforded by the adoration of sculpted deities and the intonation of religious hymns, Spirit moves on to the 'living' synthesis of image and speech provided by

[499] *HW* III, 512; *PS*, 424
[500] *HW* III, 511f.; *PS*, 424

religious cults and festivals. Finally, it embraces the properly 'spiritual' experience afforded by the literary work of art, where gods and men are put in direct relation with each other. This dialectical crescendo also reflects the development of an increasingly social religious model: whereas natural religion still allows for an individual exchange between Man and God, religious cults and festivals are necessarily shared by a whole nation (*Volk*). Literary works, on the other hand, are neither personal nor national, but *universal*: their raw material is the 'single pantheon' of human language and their main subject humanity *as such*.[501]

The three literary genres considered by Hegel — epic, tragic and comic poetry — illustrate different conceptions of the relationship between self-consciousness and a divine or transcendent *An-sich*. More than mere descriptions of specific art forms, these three models represent three different answers to the question initially raised in the stage of unhappy consciousness and left open ever since: namely, how free or independent is self-consciousness? Or rather, what is the real extent of human self-determination? Can the self really rise above all forms of alienation and become purely *for itself*? Is it bound to relinquish its independence and submit to an absolute and all-powerful *in itself*? Or is it doomed to remain trapped in this contradiction, thrust back and forth from self-unity to self-division, with no 'happiness' in sight?

In epic poetry, this dilemma takes the form of a syllogism 'in which the extreme of universality, the world of the gods, is linked with individuality, with the bard, through the middle term of particularity.'[502] This middle term is filled by the deeds of epic heroes, whose concrete individuality, raised to an ideal form, embodies the values of an entire *Volk*. In the course of the narrative, however, this heroism is shown to be entirely dependent on the will of the gods. The epic hero may see himself as a free, self-determining individual, but the epic tale will show that he is *not*. The gods are the real source of determination, the 'substantial powers' that rule human life. But this is only one side of the issue. Since the gods themselves are conceived as individuals, animated by the same kind of passions and desires that animate human beings, their power is not as substantial after all. In a 'comical self-forgetfulness of their eternal nature', they also feel

[501] *HW* III, 529; PS, 439
[502] *HW* III, 531; PS, 441. Translation altered.

love, rage, envy and powerlessness. In short, they are also *human*, and the hero's confrontation with divinity becomes, therefore, 'an empty and arbitrary showing-off', which 'transforms the apparent earnestness of the action into a harmless, self-confident play, without result or outcome.'[503]

Following this transformation, the syllogism's middle term is eliminated and its extremes are redefined. As regards universality, the authority of the gods is transferred to a higher sphere, to which the gods themselves are subordinated, namely the law of fate or necessity. This is now the divine moment of the relationship, the 'substantial power' facing the individual. As regards individuality, the self-confidence of the epic hero is transferred to the bard himself, who is led to abandon his passive role and to take part in the action. Unlike the epic hero, this new hero 'is himself the speaker, and the performance displays to the audience … *self-conscious* human beings who *know* their rights and purposes, the power and the will of their specific nature, and know how to *assert* them.'[504] The epic dispute is now a *tragic* dispute, which revives the opposition between the human and the divine laws discussed previously in the beginning of the Spirit chapter. This time around, however, the stakes are much higher. Whereas the antagonistic stances illustrated earlier by Creon and Antigone stood for immediate ethical mandates, the tragic standpoint that will now be considered is a concrete spiritual standpoint, which has tested, negated and reabsorbed all of the *Phenomenology*'s previous developments. This point is extremely important and must be clearly emphasized. The kind of knowledge that will now be put to the test is not simply the knowledge held by ordinary consciousness, or the knowledge concerning a specific situation or course of events. It is the necessary culmination of the long and complex journey we have hitherto considered. Consciousness, Self-consciousness, Reason, Spirit, Natural Religion, Art: what is now at stake is *all of this* — and nothing less. The confrontation between the tragic hero and his fate is a test to human knowledge as a whole, in all of its aspects, spheres and dimensions.

In the beginning of this work, when discussing the relationship between freedom and knowledge, I have suggested that Hegel's understanding of tragedy can be compared

[503] Ibid.
[504] *HW* III, 531; PS, 444

to Aristotle's, insofar as both rely on the notion of περιπέτεια, that is, on the idea of a complete *reversibility* of human knowledge.[505] This idea is made clear in the present stage of the *Phenomenology*, where the essence of tragedy is defined precisely as the 'inversion of the known into its opposite' (*die Verkehrung des Gewußten in sein Gegenteil*), or the 'changing round of the rightness of character and of knowing into the rightness of the opposite' (*das Umschlagen des Rechts des Charakters und des Wissens in das Recht des Entgegengesetzten*)'[506] For Hegel, what defines the tragic experience is the movement whereby Spirit, convinced it knows the truth about itself and the world, is disabused of this conviction. The disillusionment experienced countless times before is repeated once again, but this time at the very summit of the progression. As the tragic situation unfolds, Spirit is caught up once more in the antithesis, which it claimed to have left behind, of self-certainty and truth. What it took to be true is not true, and what it took to be untrue is true after all.

However, this περιπέτεια is not a regional or contingent phenomenon, which can be corrected through a mere swap of beliefs. This impossibility is immediately made clear by Hegel's description of the other side of the equation. If, on the one hand, the hero's knowledge is contradicted by Fate, the latter is also affected by this movement. For the truth revealed by the 'Furies' is just as one-sided as the truth proclaimed by self-consciousness, and could just as easily have been different. In other words, if the hero could have anticipated the actual result of his actions, and altered their content, Fate would once again have eluded him, for it is *by definition* opposed to self-consciousness. What is learned through tragedy is not simply that the knowledge held by Spirit is *wrong*, but that it is *intrinsically open to the possibility of error*, and *intrinsically unable to eliminate this possibility*. In tragedy, '[consciousness] pays the penalty of trusting a knowledge whose ambiguity, *for such is its nature*, also becomes explicit *for* consciousness and a *warning* to it.'[507] The hero learns that *all knowledge*, be it individual or universal, subjective or objective, earthly or divine, is equally one-sided, and liable to change round to its opposite.

[505] See above section 1.1.
[506] *HW* III, 538; *PS*, 447. Emphasis added.
[507] *HW* III, 539; *PS*, 448

> The truth … of the opposing powers of the content and of consciousness is the result that both are equally right, and therefore, in their antithesis, which is brought about by action, both are equally wrong. The course of the action demonstrates their unity in the natural downfall of both powers. […] Both are *oblivion*, the disappearance of actuality and the action of the powers of substance, […] for none is for itself the essence.[508]

The experience of tragedy is hereby consummated. The stage is cleared, as it were, and self-consciousness is left alone with itself. It has seen everything solid or substantial dissolve before its eyes, every notion and every law lose their value. Truth, divinity and fate have all been exposed as the props of a fictional world, imagined and sustained by self-consciousness, and all that is left for the hero is to 'step forth from his mask', acknowledge the deception *as deception* and exhibit *himself* as 'the fate of the gods … and of the absolute powers themselves.'[509] This global ἀναγνώρισις marks the transition to the *Phenomenology*'s last and most radical form of self-alienation, the final abyss into which all of the previous stages are thrown. Whereas tragedy is the *negative* acknowledgement that nothing is what it seems, comedy is the *positive* embracement of this new spiritual order, whose content is precisely the fact that it has none. Whereas tragedy is the emptying out of what was once full and stable, the *perplexity* of seeing one's world disappear, comedy is the experience of emptiness *as such*, the *certainty* of this absolute loss. More than that, it is the certainty that there never was anything to be found, and that the progression's previous tribulations were really 'much ado about nothing'.

In the stage of comedy, the seriousness of Spirit's earlier attempts to find itself is turned to ridicule. Self-consciousness has left the tragic stage: it is no longer engaged in the plot, no longer willing to suspend disbelief. Its standpoint is no longer that of the tragic hero, but that of the actor who has given up his role, or that of the spectator who knows that the play is just a play, and that nothing is *really* worth fretting about. What it now sees is a series of *unsubstantial* powers, values and notions propped up as something substantial, and herein lies the source of their comic effect. Whatever is deemed sacred,

[508] *HW* III, 539f.; *PS*, 448. Translation altered.
[509] *HW* III, 541; *PS*, 450

solemn or serious is revealed as a mere *impersonation* of solemnity or seriousness, whose ridicule is directly proportional to its (alleged) degree of importance. Therefore, although comedy brings out the underlying emptiness of every aspect of human cognition, its full power is directed at the 'final' or 'sacred' truths of natural and social life. In the first case, it turns against the purported independence of the empirical world, of its laws and principles. In the second case, it targets the customs and values that shape human society. This double movement is 'the consciousness of the dialectic contained in these maxims and laws themselves, and, consequently, the consciousness of the vanishing of the absolute validity previously attached to them.' As a result, the notion of truth, beauty and goodness 'thus display a comic spectacle: through their liberation form the opinion which contains both their specific determinateness as content and also their absolute determinateness, … they become empty, and just for that reason the sport of mere opinion and the caprice of any chance individuality' (*das Spiel der Meinung und der Willkür der zufälligen Individualität*).[510]

Through Hegel's phenomenological conception of tragedy and comedy, we are brought back, once again, to Plato, to the critical tradition and to the relationship between freedom and philosophy. As we have seen earlier, Plato's dialogues are also built around the notion of περιπέτεια, or *Verkehrung*, or *Umschlagen*, extended to every aspect of human life. As Socrates reveals the 'emptiness' of human opinions (*Meinungen*, δόξαι), he is describing human existence as a *tragic* existence, in the Hegelian sense of the word. Moreover, since what characterizes human opinions, according to Plato, is not simply the fact they are 'empty', but also the fact that they are usually taken to be 'full' — i.e. serious, solemn, substantial, etc. — the experience staged in the dialogues is also a *comic* experience, in the Hegelian sense of the word.

In the phenomenological stage of comedy, as in Socrates' enquiries, human ideas are stripped of their usual 'mask' and exposed as the *impersonation* of a series of contents they are unable to convey. But although in Plato's dialogues this experience is just as much tragic as it is comic, for there is ultimately no way of knowing whether reason's masks do conceal a true face — and, if so, what that face looks like — in the phenomenological

[510] *HW* III, 543f.; *PS*, 451f.

stage of comedy this experience is not as ambiguous. 'Unlike what happens with tragedy, comedy's unmasked masks are not the masks of an αὐτό whose density is equal to, or higher than, the density of the masks themselves; they are not the masks of an *enigmatic self-essentiality, still left to discover*, which holds everything in its power. No. […] The mask contains only the mere pretence to be more than it is, and conceals nothing other than the *inanity of this pretence*. … [It] only conceals its own emptiness, and is totally in the hands of the subject, who understands it as such and can dispose of it completely.'[511] Therefore, the phenomenological experience of comedy is no longer comparable to the perplexity felt by Socrates' interlocutors throughout his interrogations, nor to the apparent hope displayed by Socrates himself in his ceaseless search for a true conceptual content. The comic self-consciousness is an utterly cynical or disenchanted mode of cognition, which no longer expects to find, in itself or in others, anything other than emptiness and deceit. Although it is explicitly associated with 'irony', its 'wantonness' (*Mutwille*) and 'levity' (*Leichtsinn*) are decidedly more Aristophanic than Socratic, while its 'contempt' (*Spott*) and 'bitterness' (*Bitterkeit*) vis-à-vis the natural and the social realms anticipate Nietzsche's attitude towards science and morality. What is more, instead of fleeing the acknowledgement of its own condition, the comic self revels in it, as the anarchist revels in disorder. Its negative attitude, verging on nihilism, resembles the absolute 'vertigo' later described by Baudelaire:

> Aussitôt le vertige est entré, le vertige circule dans l'air; on respire le vertige; c'est le vertige qui remplit les poumons et renouvelle le sang dans le ventricule.
>
> Qu'est-ce que ce vertige? *C'est le comique absolu*.[512]

The stage of comedy, more than all others, poses a problem of continuity. Consciousness is now a purely negative power, whose essence is precisely that it has none, and allows for none to be posited. The self has been driven to the edge of itself, as it were, and condemned to absolute derision. Since its negative power extends to every dimension of human life, any attempt to rescue the self from itself appears doomed to

[511] CARVALHO, *Verse actuar a sí mismo*, ed. cit., 213
[512] *Oeuvres Complètes* II, ed. cit., 383. Emphasis added.

fail. Any 'serious' stride forward, aimed at a new and less sterile cognitive attitude, is bound to be greeted with laughter and contempt.

If the progression were to stop here, the philosophical release promised at the outset would remain out of reach. Freedom would be reduced to the negative freedom of comedy and Hegel's phenomenological project would be reduced to a critical project, or to a peculiar form of scepticism. This, however, is not the case: consciousness' dissolution of everything stable or substantial is only at first perceived as a dead end. As so many times before, the self's complete self-alienation will be revealed as the moment of a wider dialectical synthesis, in which self-certainty and truth will once again be reunited. This final reconciliation will take place in the stage of 'Revealed Religion', leading to the *Phenomenology*'s grand finale.

9.4. *A new beginning*

The dialectical movement unrolled in the Religion chapter can be read, once again, as a re-enactment of the itinerary followed by the unhappy consciousness. Here, however, we have already moved past the individual relationship between Man and God, as well as the abstract forms of worship, asceticism and self-sacrifice documented in the Self-consciousness chapter. What is described in the dialectic of Religion is the inner division of the concrete Spirit that has emerged from the stage of Morality — a shape of Spirit in which individuality and universality have been reunited, but which is still confined to a 'mundane existence', and still burdened with a divine calling.[513] The religious conceptions discussed throughout this stage do not stand merely for individual standpoints or attitudes, but for global developmental stages, where the phenomenological subject is simultaneously an independent, self-determining agent, and the product of a specific cultural and historical context.

Initially, the divine element is located beyond the natural world. God is the essential moment of the relationship, worshipped and summoned by the faithful. With tragedy

[513] *HW* III, 498f.; *PS*, 412f.

and comedy, however, Heaven is gradually 'depopulated' and the self emerges as the essential moment of the relationship.[514] The comic consciousness dissolves every external form of authority and becomes itself the Absolute Being, or the new 'God'. But Hegel will now argue that this new independence is also illusory and destined to be overcome. Since comedy's power is pervasive, its poison is bound to contaminate its own source and to dissolve the very will that keeps it alive. Indeed, although comedy is 'the consciousness of the loss of all essential being in this *certainty of itself*', it is also 'the loss … of this knowledge about itself — the loss of substance as well as the loss of Self.'[515] In other words, comedy's derision leads to self-derision, which leads to self-destruction. What was at first perceived as a vast joke gives way to a spiritual void, and the self's previous arrogance is replaced by misery and destitution.

Comedy leads to a new version of the unhappy consciousness, or to 'the grief which expresses itself in the hard saying that "God is dead."'[516] The outcome of this experience is echoed, in Hegel's description, by the decline of the ancient Greek world, by a relapse into Stoicism and Scepticism, and by a general loss, on the part of religion, of the ability to fulfil humanity's divine calling. Since consciousness has turned its back on God, and failed to find in itself a worthy replacement, the human and the divine elements are once again separated. In order to overcome this new conflict, consciousness must cease to regard these two elements as independent from each other and devise a new form of synthesis. The solution lies in a return to the notion of *incarnation*, whereby the two previous conceptions of God — namely God as *absolute Being*, or the abstract deity of Natural Religion, and God as *absolute Subject*, or the negative power of Comedy — are finally brought together. In contrast to the two previous forms of divinity, an *incarnate* or *embodied* God is simultaneously negative and positive, self-conscious and actual, for-itself and in-itself. Its dialectical nature mirrors the dialectical nature of the human self and replicates its own self-division.

With the notion of incarnation, the divine Being is finally *revealed*. The phenomenology of religion, which hitherto attempted to accommodate consciousness'

[514] *HW* III, 540ff.; *PS*, 451ff.
[515] *HW* III, 547; *PS*, 455
[516] Ibid.

self-certainty with the certainty of an absolute essence, placed either outside of the natural world or inside the thinking self, has now overcome this dualism. Henceforth consciousness 'does not start from *its* inner life, from thought, and unite *within itself* the thought of God with existence; on the contrary, it starts from an existence that is immediately present and recognizes God therein.'[517] That is to say, once God is 'made Man', the unity between humanity and divinity is no longer *strived for* (as in Natural Religion) or *asserted* (as in Comedy), but *directly beheld* (*angeschaut*): in the figure of Christ, consciousness is immediately aware that the divine element is *not* different from the human element — that it was only *posited* as such by the previous forms of experience, and has now been raised to truth.

This new development is completely misunderstood if the embodied divinity is regarded merely as an objective presence *endowed* or *burdened* with a divine meaning. If that were the case, this divine individual would only differ from the icons and statues worshipped in the previous stages by its human appearance and its ability to move, speak, see, etc. It would be a living image of divinity, but an image all the same, whose essence would still be located elsewhere, beyond the natural world. Hegel's point, however, is that God is now revealed as a *self-conscious* being — a being that is both conscious of itself as divinity and conscious of its actual existence. In other words, God is now revealed as *Spirit*, for Spirit is precisely 'the knowledge of oneself in the externalization of oneself; the being that is the movement of retaining its self-identity in its otherness'[518] This declaration is the turning point of the whole dialectic, and must be clearly emphasized: Hegel is now pointing out that consciousness, in having arrived at a truly dialectical conception of divinity, *has simultaneously arrived at a true definition of itself*. The contradiction between self-certainty and transcendence it was unable to resolve has now been sublated by a new religious shape, which emerges before consciousness as in a mirror and reveals at last its true identity. '[God] is immediately *present* as Spirit'[519], and through God's presence consciousness will also be revealed as Spirit — no longer the

[517] *HW* III, 551; *PS*, 458
[518] *HW* III, 552; *PS*, 458
[519] *HW* III, 554; *PS*, 461

incomplete Spirit that had come out of the Morality chapter, but a Spirit that has fulfilled its transcendent calling and joined together the human and the divine realms.

Once the religious consciousness realizes that the incarnate divinity is identical to itself, Spirit will at last be known as Spirit, and the progression will reach its final stage. For now, however, we still have the human self opposed to the divine self, not yet aware of their coincidence. For the former to see itself in the latter, or for the latter to come out of the mirror, as it were, and join the former, two last steps need to be taken. Firstly, the image in the mirror is not yet fully faithful to the original because the embodied divinity is still a particular individual. As we have seen, however, the religious consciousness is no longer simply an individual standpoint, but the individual reflection of a universal standpoint. The embodied divinity, on the other hand, is still '*this individual* self-consciousness', still 'in an antithesis to the universal self-consciousness.' It is already Spirit, but 'Spirit as an individual Self', and 'not yet equally the universal Self, the Self of everyone.' [520] In order to correct this discrepancy, God must be divested of His individuality and raised to a universal form. He must die and be resurrected, abandon His human body and re-emerge as a Holy Spirit. In this new form, God's presence is not eliminated, but expanded: the divine Being reappears before consciousness as a truly universal Self, whose presence extends to the whole world. His death 'becomes transfigured from its immediate meaning, viz. the non-being of this *particular* individual, into the *universality* of the Spirit who dwells in His community, dies in it every day, and is daily resurrected.'[521]

The image in the mirror is now complete, but Spirit's return to itself is still dependent on one final task. According to Hegel, the language employed by religion is still anchored in representational thinking (*vorstellendes Denken*), and not yet in notional thinking (*begreifendes Denken*). This distinction brings us back, once more, to the methodological discussion expounded in the preface: although religion has unveiled the true nature of Spirit, it has done so by means of different representational predicates (*Vorstellungen*), corresponding to different spiritual events. The content of religion is 'the true content, but all its moments … have the character of being uncomprehended [*den*

[520] *HW* III, 555; *PS*, 461f.
[521] *HW* III, 571; *PS*, 475

Charakter, *nicht begriffen zu sein*], of appearing as completely independent sides which are externally connected to each other.'[522] In short, religion still views Spirit in a *narrative* way, as a succession of transformations predicated of the same divine Being. This kind of perspective, however, is not yet Spirit's *native* perspective. It is still a *borrowed* perspective, which has yet to grasp the *immanent* and *simultaneous* nature of these transformations. Although religion has helped Spirit rise to its true form, its phenomenological role has come to an end. Like all of the progression's previous stages, its specific language has outgrown its use, and must now be replaced by a new mode of expression.

Since Spirit was brought face to face with its own image, all that it must do now is step forward and reclaim its authorship. With this final movement, what in religion was still '*content*, or a form for representing an other', is revealed as 'the *Self*'s own *act*.'[523] Humanity and divinity, individuality and universality, Subject and Substance — all of these contrasts are divested of their narrative form and disclosed as the interchanging moments of the same overarching Notion. They are no longer merely placed before consciousness, but absorbed by it and dissolved in its inner movement. As a result, Spirit is no longer estranged from itself. The image in the mirror is finally brought to life, so to speak, and Spirit is finally raised to truth.

In the progression's final standpoint, boldly entitled 'Absolute Knowing', the Notion pursued by consciousness is finally *realized* (*realisiert*). Whereas in the previous stages knowledge and truth kept falling outside of each other, leading to new and increasingly complex cognitive models, their separation has finally been overcome. This coincidence, however, is not a simple or static coincidence, but a *dynamic* or *infinite* one. The truth revealed at the end of the progression is no longer a true *content*, or a true *object*, singled out from a wider universe of contents or objects. This kind of interpretation leads back to the dualistic attitude which has prevailed throughout the progression, and has now been left behind. The truth unveiled in this final stage consists rather in the emergence of a truly systematic mode of cognition, in which all previous dichotomies have been eliminated, and where everything is joined together with everything else. On the other

[522] *HW* III, 556; *PS*, 463
[523] *HW* III, 582; *PS*, 485

hand, though, this does not mean either that difference and variety have been eliminated, or that experience was reduced to a simple reiteration of the same universal content. Knowledge and truth have been united, but this unity is a *self-moving* unit, in which identity is achieved through a continual return to otherness. Spirit is what it is by being what it is not, or by finding itself in what is other than self.

In Hegel's own terms, 'Absolute Knowing' is reached when every remaining *Vorstellung* is dissolved into the absolute flow of *Begriff*. And this movement encompasses all of the main dichotomies we have considered throughout the previous chapters: with regard to the contrast between subjectivity and objectivity, Spirit has become its own object, or 'the movement of retaining its self-identity in its otherness'[524]; with regard to the contrast between subjectivity and intersubjectivity, Spirit has become both 'an individual Self' and the 'universal Self, or the Self of everyone'[525]; with regard to the contrast between humanity and divinity, Spirit has emerged as the truth of Religion, or 'the Divine's intuition of itself.'[526] In its final form, Spirit is the reconciliation of these three main syntheses, joined together in the element of Science.

If the *Phenomenology*'s final standpoint is indeed a scientific standpoint, the education of consciousness appears to be complete. But this final success cannot be granted without at least a brief consideration of the various conditions required for its actualization. According to the methodology set out by Hegel at the beginning of the progression, the standpoint of science can only be reached once every main form of ignorance has been identified and overcome. Science can only emerge if the 'exposition of the untrue consciousness in its untruth' has been exhaustive, and there is no longer any form of untruth left to examine. Or again, Spirit in its fully fledged form can only be revealed once the 'natural ideas, thoughts and opinions' with which consciousness is initially 'hampered' have all been removed, and there is no longer any illusion at work in human life. This general requirement can be broken down into a series of more specific requirements, all of which must be met in order for Hegel's project to be successful.

[524] *HW* III, 552; *PS*, 459
[525] *HW* III, 555; *PS*, 461f.
[526] *HW* III, 580; *PS*, 483

Firstly, the elevation to science is conditional on the correct identification of every form of untruth liable to be adopted, whether implicitly or explicitly, by human consciousness. The vast kaleidoscope revealed throughout the *Phenomenology* must cover not only every cognitive angle or attitude available throughout human life, but also every combination of angles or attitudes available to us. If this inventory is not complete, there might still remain within us a hidden presuppositive core, incompatible with Hegel's final aim. Secondly, the elevation to science is also conditional on the correct identification of the dialectical sequence whereby all of these different angles and attitudes reflect and replace one another. If this global chain is incorrectly assembled, the necessary succession of abstract and determinate negations envisaged by Hegel is bound to be compromised, along with the progression's final goal. Thirdly, the elevation to science also requires that all of the *Phenomenology*'s stages have been thoroughly sublated and overcome, so that no presupposition is allowed to exceed its proper phenomenological life-span. If each new sublation does not correspond to a whole new beginning, or to a whole new worldview, its critical role is irrevocably compromised. Fourthly, in line with the previous requirements, for consciousness to reach the stage of science it must be able to make sure that this final stage is indeed the last one, and not the product of a hidden set of assumptions, liable to extend the progression into new and unknown territories. Fifthly, although all of the aforementioned conditions refer to the progression itself, and to how it is crafted, they must also be applied to the progression's reader or interpreter. For even if the *Phenomenology* itself fulfils all of the previous requirements, we readers may not be able to understand or assimilate all of its elements and arguments. In this case, even though we may be provided with all of the means necessary to arrive, with Hegel, at the standpoint of science, we might be incapable of doing so — whether out of ignorance, inattention or some other personal limitation.

Furthermore, all of this must be related to the problem of freedom. If the *Phenomenology*'s final standpoint is indeed a scientific standpoint, the liberation of consciousness appears to be complete. If Spirit, in its fully fledged form, is free from ignorance and self-deceit, it is also endowed with a new form of power. Whereas previously its plans and intentions, grounded in a series of cognitive presuppositions, were liable to be thwarted by an error of judgement, or by a misguided interpretation of reality, this danger appears to have been removed. The elevation to the standpoint of

science entails a complete cognitive release and a new degree of self-determination. Once again, however, this transformation is entirely dependent on the fulfilment of all of the aforementioned requirements. The failure to observe any one of them compromises the release promised by Hegel, as well as the power it is apt to provide.

Finally, apart from the previous list of conditions, the possibility of an actual deliverance of human consciousness is also conditioned by two further ambiguities, concerning the *Phenomenology*'s final standpoint. First, near the end of the last chapter, Hegel argues that Absolute Knowing, in its accomplished form, entails a return to the beginning of the progression. 'Science', he claims, 'contains within itself the necessity of externalizing the form of the Notion, … for the self-knowing Spirit … is the *certainty of immediacy*, or *sense-consciousness* — the beginning from which we started.'[527] This new outlook on the sensuous world is of course radically different from the outlook available in the stage of sense-certainty. For a scientific standpoint, the sensuous world is not merely sensuous, but the correlate of science, that is, the reflection of a truly concrete or dialectical mode of cognition. Accordingly, this scientific return to sense-certainty is no longer a source of doubt and contradiction, as it was in the beginning, but a triumphant affirmation of Spirit's 'supreme freedom and assurance.'[528] Immediately afterwards, however, Hegel points out that this externalization is still incomplete, and that the freedom it affords is likewise incomplete. What is more, he claims that this externalization entails a 'sacrifice', which forces Spirit to 'display the process of its becoming in the form of *free contingent happening*.'[529] Although the exact meaning of these last paragraphs cannot be discussed here, Hegel's point seems to be that the outlook of Absolute Knowing, insofar as it is tied to the contingent element of History, and continually confronted with new empirical material, is bound to reinvent itself *ad eternum*. If this interpretation is correct, what emerges in these final paragraphs is the idea of an open-ended progression, which must be revised and updated in the course of

[527] *HW* III, 589f.; *PS*, 491
[528] *HW* III, 589f.; *PS*, 491
[529] *HW* III, 590; *PS*, 492. Translation altered.

human history. And if this is so, the extent of consciousness' final release is also limited and continually open to revision.[530]

Secondly, consciousness' release must also be considered in light of the role played by the *Phenomenology* in Hegel's philosophical system. And here, once again, there is ample room for speculation. In the introduction, Hegel describes his work as the complete journey from the standpoint of natural consciousness to the standpoint of 'true existence', or 'the authentic Science of Spirit.'[531] In the preface, though, written at a later date, he makes clear that the progression constitutes only the first part of science, to be followed by more advanced forms of enquiry. Its role is to prepare 'the element of [true] knowing' and to raise the 'moments of Spirit' to their true form, whereas 'their movement, which organizes itself in this element into a whole' belongs to 'Logic or Speculative philosophy.'[532] The *Phenomenology* is therefore an introductory work, whose final result is not yet the final stage of reason's self-development. Its task is the negative task of freeing consciousness from its 'natural ideas, thoughts and opinions' and elevating it to a truly unhypothetical standpoint, in the Platonic sense of the word. Accordingly, the freedom attained at the end of the phenomenological progression is the freedom from one's pretence of knowledge and the possibility of a pure or unspoiled apprehension of reality. On the other hand, insofar as the 'element of true knowing', according to Hegel, is liable to be further developed and detailed, this initial freedom is also liable to assume new and more robust forms.[533] Their exact meaning, however, can only be known at the end of the *Phenomenology*'s itinerary, if its main purpose is indeed fully accomplished.

[530] The *Phenomenology*'s final paragraphs are notoriously ambiguous. If Hegel is indeed introducing the idea of an open-ended progression, subordinated to the contingent element of human history, the problem arises as to how this model is to be articulated with his previous claims about the progression's immanent necessity. For a useful discussion of this section, see HYPPOLITE, op. cit., 554-583.

[531] *HW* III, 81; *PS*, 57

[532] *HW* III, 21f.; *PS*, 491

[533] On the articulation between the *Phenomenology* and Hegelian Science, see FORSTER, op. cit., 259-288.

9.5. *Hegel and Nietzsche*

The *Phenomenology of Spirit*'s unique appeal, amply felt throughout the 19th and 20th centuries, is due in great part to Hegel's tremendous faith in the power of human reason. Indeed, few other modern thinkers have placed such high hopes in philosophy's ability to reform human consciousness and to free human beings' usual standpoint from its constitutive limitations. This attitude is all the more remarkable since, as was shown in the previous chapters, it can hardly be put down to a superficial engagement with human reason, or to an overly indulgent diagnosis of consciousness' intrinsic limitations. If Hegel does stand out for his philosophical optimism, his work is no less notorious for its radical denouncement of human ignorance.

In the *Phenomenology*, these two aspects are mutually related. On the one hand, following a long lineage of critical philosophers, Hegel argued that ignorance cannot be simply corrected or replaced by true knowledge because it is not usually perceived as ignorance, but precisely as true knowledge. The attainment of truth is conditional, therefore, on the exposure of the groundlessness and unintelligibility of all accepted knowledge, as well as on the possibility of guaranteeing that this task was indeed extended to all of the main principles and ideas that guide human life. On the other hand, departing from the critical tradition, Hegel argued that this negative procedure is liable to have a positive nature, and a positive outcome. He sought to show that the untruth of natural knowledge is not simply untrue, as it were, but an undeveloped or unravelled form of truth. Irrationality is itself bound to a rational principle and the critic's task is to identify this principle, follow where it leads and abandon the realm of untruth.

Accordingly, Hegel's phenomenological scepticism is a dynamic or self-moving scepticism, in which every particular criticism is the direct consequence of a previous criticism and the necessary antecedent of a new refutation. Human ideas and judgements are no longer taken up at random, as in Plato's dialogues, or contingently associated, but placed within a global cognitive map, whose main itinerary allows for no detours. This ambitious strategy is grounded in two revolutionary convictions, which set the *Phenomenology* apart from all previous philosophical works. Firstly, Hegel attempts to

show that all aspects, dimensions or areas of human knowledge amount to different views or to different perspectives on the same cognitive whole. Whereas most of us tend to regard the fields of experimental science, morality, psychology, politics, religion or art as different cognitive worlds, whose aims, laws and methods are unique and irreducible to each other, Hegel regards all of these worlds as different semblances of the same overarching Notion. The *Phenomenology*'s aim is to dissolve these distinctions and to bring about a truly global cognitive perspective, as rich and as complex as life itself. Secondly, Hegel also argues that this final perspective, due to the very form of the progression, is necessarily set for consciousness. If the different phenomenological stages were merely added together, in the form of an aggregate or a compound, the progression might go on indefinitely, with no guarantee of success. But this, according to Hegel, is not the case. Human reason is inherently systematic, which means that the standpoint of Science is not merely possible, but actually necessary. Although consciousness may never leave the realm of untruth, out of weakness or idleness, true knowledge is no less real and within human reach.

With these two convictions, tested throughout the *Phenomenology*, Hegel raised philosophy's old faith in reason and rationality to a new and unprecedented height. At the same time, however, his systematic solution can be seen as the final expression of this faith. Hegel was one of the last, if not the last great Western thinker to believe in the possibility of a truly sovereign mode of cognition, free from the chains of illusion and self-deceit. He was one of the last to hold that the contradictions that plague human life can be cured *through reason*, and not *against it*, or *beyond it*, or *in spite of it*. After Hegel and idealism, the philosophical canon shifted in an entirely new direction. Hegelian optimism was replaced by a new antagonism towards reason and science. From Kierkegaard's existentialism to Husserl's and Heidegger's 'pre-theoretical' turn; from Schopenhauer's pessimism to Stirner's nihilism, Hegel's insistence on harmony and reconciliation were inherited by a historical age at odds, perhaps more than any other, with the ideas of fragmentation, disharmony and conflict.[534]

[534] The philosophical shift away from Hegel and systematic reason was mediated, of course, by Marxism, whose position in this global itinerary is ambiguous. On the one hand, broadly speaking, Marx's dialectical materialism is also a systematic doctrine, envisaged as a global solution to the contradictions of

To borrow Karl Löwith's formula, the revolutionary *Bruch* in 19th century German philosophy coincided, to a great extent, with the disintegration of Hegel's dialectical model and with the emergence of new and varied alternatives to systematic rationalism. In this context, Nietzsche is often pointed out as Hegel's most obvious countertype. This contrast is based on a number of well-known differences, which refer both to the content and to the form of Hegel's and Nietzsche's writings: whereas Hegel is ultimately an optimist, Nietzsche is famous for his 'Dionysian pessimism';[535] whereas Hegel is the systematic philosopher κατ'ἐξοχήν, concerned with order and symmetry, Nietzsche is renowned for his dispersed aphorisms; whereas Hegel is a difficult and often abstruse writer, Nietzsche's prose is elegant and mordacious.

But Hegel's and Nietzsche's works also display important structural similarities. Firstly, both of their philosophies are built around the idea of science, or true knowledge, and both can be interpreted as an attempt to answer the main challenge posed by the critical tradition. Nietzsche's views on truth and true knowledge, just like Hegel's, are deeply influenced by the critical diagnosis set forth in Plato's dialogues. In his eyes, human beings' usual cognitive perspective is a flawed and self-deluded perspective, burdened with a series of unwarranted assumptions about itself and the surrounding world. These assumptions are for the most part automatic, embedded in the heart of human cognition and hardly ever confronted or called into question. Furthermore, these assumptions are not restricted to a specific brand of ideas or judgements, nor to a specific domain of experience. They contaminate the whole edifice of human reason and condition every aspect of human life.

Secondly, like Hegel, Nietzsche also regards reason's inner contradictions as a form of imprisonment and servitude. In line with the Platonic tradition, he also describes human life as a downgraded or diminished life, and human beings as downgraded or diminished beings, reduced to a lesser version of their original selves, confined to a mode of cognition which compromises their power to choose and act freely. Moreover, he also construes philosophy's critical role as a liberating task, destined to reverse this state of

modern reason. On the other hand, Marx's faith in reason and rationality is very different from Hegel's. Instead of endorsing the latter's idealist framework, he famously sought to 'turn Hegel on his head' and introduce a new brand of philosophical criticism.

[535] *KSA* III, 622

affairs and to re-establish mankind's original independence.

Thirdly, as noted earlier, both Hegel and Nietzsche envisage this liberating movement as a historical and cultural process, and both have turned to the history of human reason in order to illustrate its main implications. What is more, the Hegelian and Nietzschean 'historiographies' of freedom share many important references. For one thing, both denote a clear fascination with the primordial innocence of the ancient Greek world, contrasted with the cultural degeneration of the modern Western world. In both accounts, the birth of Greek tragedy and the emergence of Socratic and Platonic dialectics constitute important conceptual innovations, with profound implications on the meaning and scope of human freedom. Furthermore, Hegel and Nietzsche also highlight the cultural and conceptual role played by religion in the definition of freedom, and particularly by the rise and evolution of the Jewish and Christian traditions. Finally, both authors view Kant's critical project as a defining moment in the history of modern reason, with important repercussions on the definition of freedom.

However, these generic coincidences can only tell us so much. Hegel's and Nietzsche's 'historiographical' appraisals are not only very different but, in many respects, opposed to one another. Hegel regards the phenomenological history of human reason as the necessary path towards the actualization of freedom. On his view, each step along the way is necessary and liberating, provided it is correctly understood and criticized: whereas the harmony of Greek *Sittlichkeit* was irredeemably lost, its sacrifice was needed for the attainment of a higher cultural synthesis, grounded in the modern principle of individual freedom; whereas the standpoint of primitive Christianity was a source of unhappiness and alienation, it played a decisive role in the dialectical transition to the stage of Reason; whereas Kant's critical standpoint led human consciousness to a higher stage of self-awareness, its inner contradictions still called for a new *Aufhebung* and for a more complete form of deliverance.

For Nietzsche, on the contrary, the idea that freedom is dependent on a systematic rational enquiry is seriously mistaken. Unlike Hegel, he rejects the basic premise of Plato's scientific project and regards the pursuit of scientific knowledge not as the remedy, but as the primary cause of human servitude. Whereas Plato and Hegel anchored their criticism of human cognition in the notion of truth, or science, and set about denouncing the untrue or unscientific nature of human concepts and beliefs, Nietzsche

argues that this grounding notion must itself be criticized. Despite its usual philosophical protagonism, the notion of truth is just a mere belief, or a mere assumption, just as groundless and unintelligible as all other beliefs and assumptions. Therefore, only a critical philosophy that is capable of acknowledging and overcoming this prejudice can be actually liberating. Whereas Plato and Hegel advocated a release *from untruth*, or a release *through science*, Nietzsche reverses this general plan. He advocates a release *from truth*, or a release *from science*, and the embracement of a wider and freer standpoint, no longer confined to its usual cognitive cage.

Consequently, as we will see in the following chapters, Nietzsche's engagement with the history of human reason is very different from Hegel's. Whereas Hegel viewed Greek dialectics, Christian morality and modern Reason as necessary stages in the road to human freedom, Nietzsche sees them as different and increasingly dogmatic incarnations of the scientific ideal, and hence as different and increasingly violent forms of imprisonment. In his eyes, the long odyssey of human culture reflects, for the most part, the long evolution of the same grounding prejudice — a philosophical and cultural *parti pris* whose genealogy must be retraced, exposed and rejected.

Part III

Nietzsche's Metacognitive Project

10.

NIETZSCHE'S IDEA OF TRUTH

Achtet mir, meine Brüder, auf jede Stunde,
wo euer Geist in Gleichnissen reden will:
da ist der Ursprung eurer Tugend.[536]

I have focused previously on the *Republic*'s seventh book and on Plato's most famous allegory. Since Nietzsche is also fond of allegories, it seems appropriate to start this new string of chapters with one of his most famous images. The setting is no longer a subterranean cave, but a rocky mountain-top, and the lighting is no longer dim, but overflowing. It is the moment when Zarathustra, the prophet and 'free spirit', abandons his own cave, steps into daylight and addresses himself to the sun. After ten years of isolation, his heart has changed. He has grown weary of his wisdom, 'like a bee that has gathered too much honey'[537], and wants to pass it along. He has decided to return to the 'underworld', the world of human beings, and his future task is summed up in a formula that encapsulates Nietzsche's entire philosophical programme: to distribute his wisdom 'until the wise among human beings have once again enjoyed their folly, and the poor once again their health.'[538]

As so often in Nietzsche's writings, this description is filled with Platonic allusions.

[536] *KSA* IV, 99
[537] *KSA* IV, 11; *TSZ*, 3
[538] Ibid.

Apart from the cave and the sunlight, two of Plato's chief allegorical elements, the whole picture is evocative of the *Republic*'s allegorical situation. Zarathustra's descent into darkness, with the aim of freeing human beings from their natural ignorance, echoes the return of Plato's liberated caveman to his original dwelling place, equally determined to share his discovery with his former companions. But just like Plato's caveman, Zarathustra's teachings are not taken seriously: his first speeches, directed at a crowd gathered in a marketplace, are greeted with incomprehension, laughter and hostility.[539]

In Plato's case, the allegory of the cave is meant to widen human beings' usual cognitive perspective and to unveil a possibility which is usually hidden and out of reach. Since all of us tend to regard the limits of our vision as the limits of the world, we are usually confined to our natural mode of cognition and bound to conceive difference as an immanent variation of the same cognitive alphabet. This peculiar form of imprisonment is the starting point of Plato's philosophical project. In the allegory of the cave, and in other texts, Socrates resorts to what I have named earlier, following Fränkel, a Heraclitean thought pattern[540]: starting from the comparison between human life and that of the prisoners inside the cave, Socrates goes on to claim that these prisoners stand for *us*, that their life is *our* life. This declaration is the crucial moment of his narrative, the pivotal point in which description gives way to allegory. *We* are now the lower terms of a new comparison, whose higher term can only be envisaged negatively, as an utterly different kind of life.

Nietzsche's strategy is identical, and re-enacted in different ways throughout *Zarathustra*'s prologue. Apart from the aforementioned description, directly evocative of Plato's allegory, Nietzsche resorts to two other comparisons, both built according to the same allegorical scheme. Firstly, he defines Zarathustra as 'an awakened one' (*ein Erwachter*) and other people as 'sleepers' (*Schlafenden*).[541] This image is also Platonic[542] and aims both to expose human beings' current awakenness as a *false* or *dream-like* awakenness and to suggest the possibility of a *genuine* awakenness, located somewhere beyond the confines of ordinary consciousness. Secondly, in one of many references to

[539] See *KSA* IV, 14ff. Cf. *Re* 516e-517a
[540] See above section 1.2. and below
[541] *KSA* IV, 12. Cf. *KSA* V, 12.
[542] See above section 1.3. and below section 11.1.

Darwinism, Nietzsche compares human beings to *apes* and introduces thereby a new allegorical triad: just as apes are inferior to humans, the latter are inferior to a higher form of being, namely the *Übermensch*:

> What is the ape to a human? A laughing stock or a painful embarrassment. And that is precisely what the human shall be to the overman: a laughing stock or a painful embarrassment.[543]

The term 'overman' is an allegorical term, whose meaning is not immediately intelligible. It has to be *deduced* or *derived* from the allegory's double comparison. Contrary to what is often suggested — and despite the ambiguous connotations of the term 'superman' — Nietzsche's *Übermensch* is not a stronger or aggrandized version of our regular selves. As in Plato's allegory, what is here at stake is not a *quantitative*, but a *qualitative* upgrade: just as human cognition is qualitatively different from animal cognition, overhuman 'cognition' lies beyond the sphere of human cognition, and can only be attained by means of a radical reform of ordinary consciousness.[544]

Nietzsche's Darwinian language is also meant to highlight the dynamic nature of this evolutionary scale. Ape, man and 'overman' are not fixed or independent stages, separated by clearly defined boundaries. They are the moments of a living process, continually open to revision. Here, once again, Nietzsche's vision is entirely Platonic: just as Plato defined human consciousness as the meeting point of two opposing forces (σῶμα and ψυχή, δόξα and ἐπιστήμη), Nietzsche compares mankind to 'a rope fastened between animal and overman' — and what is more, 'a rope over an abyss.'[545] Following Plato, he highlights humanity's *hybrid* or *transitive* nature: human beings are permanently torn between 'under-humanity' and 'over-humanity', and ultimately unable to escape this self-division. They are not stable or self-sufficient beings, but incomplete and self-contradictory ones. Their life is 'a dangerous crossing, a dangerous on-the-way, a

[543] *KSA* IV, 14; *TSZ*, 6. Apart from this passage's Darwinian overtones, Nietzsche is also referencing Heraclitus's fragments DK 82 and 83, where human beings are equally compared to apes. This is precisely one of the sources on which Fränkel based his notion of Heraclitean thought-pattern.

[544] In the 'overhuman' sphere envisaged by Nietzsche, the word 'cognition' no longer applies. As we will see below in more detail, what this transformation is meant to abolish is precisely human life's *cognitive* or *gnoseological* dimension, that is, its constitutive dependence on truth and true knowledge.

[545] *KSA* IV, 16; *TSZ*, 7

dangerous looking back, a dangerous shuddering and standing still.'[546] It is 'a bridge and not a purpose'[547], 'a *crossing over* and a *going under*' (*ein* Übergang *und ein* Untergang).'[548]

In short, human beings are *no longer* animals and *not yet* overmen — they are something *in between*, *auf-dem-Wege*, μεταξύ. But this intermediate position is not stable or static. And this point cannot be stressed enough: in Nietzsche's evolutionary scheme, human life is continually disrupted by a drive for self-transcendence; it is a permanent battle ground, as it were, where one can either stand or fall, rise to overhumanity or relapse into animality.

But these are still formal indications. Though it is now clearer *how* Nietzsche construes the relationship between *Mensch* and *Übermensch*, it has yet to be determined *why* he draws such a distinction in the first place, *what Übermenschheit* consists in and *why* he regards it as a form of liberation. In the following chapters, I will discuss the need, the causes and the consequences of this transformation: I will start by considering its epistemological implications, linked to Nietzsche's redefinition of the notions of truth and untruth, or wisdom and ignorance; afterwards, I will consider its moral implications, linked to Nietzsche's redefinition of the notions of good and evil, mastery and slavery, power and servitude; finally, I will discuss Nietzsche's redefinition of freedom, and his solution to Plato's original predicament.

10.1. *Truth and untruth*

In *Zarathustra*'s prologue, the transition from humanity to 'over-humanity' is said to depend on the attainment of *wisdom* — the *new* wisdom acquired during Zarathustra's retreat, which is contrasted with the *old* wisdom held by human beings. The dissemination of this newly found wisdom is supposed to make 'the wise among human beings enjoy once again their folly', and 'the poor enjoy once again their health.'

[546] Ibid.: 'Ein gefährliches Hinüber, ein gefährliches Auf-dem-Wege, ein gefährliches Zurückblicken, ein gefährliches Schaudern und Stehenbleiben.'

[547] *KSA* IV, 16f.; *TSZ*, 7

[548] *KSA* IV, 17; *TSZ*, 7

At first glance, this initial declaration could also be interpreted in Platonic terms: 'to lead the wise to enjoy once again their folly' is a close enough definition of Socrates' usual strategy, whereby human 'wisdom' is revealed as fundamentally illusory and human beings are led to acknowledge their natural 'folly'; in like manner, 'to lead the poor to enjoy once again their health' could be read, in Platonic vein, as the effort to 'cure' human beings' diseased standpoint through the acknowledgement of its own poverty. In either case, Zarathustra's revelation would coincide with Socrates' main philosophical insight: human beings are 'of no account in respect to wisdom'[549]; their usual conviction regarding the truthfulness of their thoughts is dogmatic and unwarranted; in other words, their idea of truth falls short of *truth itself*.

There are, indeed, important parallels between the Platonic and the Nietzschean characterizations of human cognition, which is worth considering in order to grasp the meaning and scope of Nietzsche's 'Zarathustrian' revolution. To begin with, following Plato and a long philosophical tradition, Nietzsche highlighted the importance of the role played by the notion of truth in human life. In particular, he called attention to the fact that human thoughts, desires and actions are never neutral or indifferent with regard to their truth-value. On the contrary, they are supported by a global version of reality, usually assumed to be true. Just like the 'vital atlas' mentioned earlier, in connection with Plato's *Philebus*,[550] or the 'cognitive alphabet' mentioned in connection with Hegel's *Phenomenology*,[551] this version of reality offers a global diagnosis of one's practical or existential situation: the diagnosis of what is at stake in one's life, the assessment of one's circumstances, the establishment of one's goals, etc. It provides the setting in which life usually unfolds and the insight needed for the success of one's decisions and actions.

Nietzsche refers to this phenomenon as the 'will to truth' and places it at the centre of his philosophical enquiries.[552] Human beings 'will' truth because they are dependent

[549] *Ap* 23a

[550] See above section 2.2.

[551] See above section 6.1.

[552] Nietzsche uses many different terms for this phenomenon: *Wille zur Wahrheit*, *Trieb zur Wahrheit*, *Zucht zur Wahrheit*, *Wille zum Wissen*, *Wahrheitstrieb*, *Wissenstrieb*, *Erkenntnistrieb*, *Wahrheitsstreben*, *Liebe zur Wahrheit*, *Leidenschaft der Erkenntnis*, *Pathos der Wahrheit*, *Hang zum Wahren*, etc. On these notions, see notably *KSA* I, 755ff., 875ff.; *KSA* III, 264f., 351f., 372, 575ff., 581f.; *KSA* V, 15ff., 23, 41f., 144f., 167ff., 398ff., 409ff.; *KSA* VII, 424ff., 433, 476, 486ff., 491f., 496f.; *KSA* IX, 471, 467f.; *KSA* X, 61; *KSA*

on a true diagnosis of what things are, of how they stand, of what they mean, etc. Yet Nietzsche also points out that this 'will' is not absolute or uniform, but liable to assume different degrees of intensity. Even though all human beings are concerned with the truthfulness of their usual account of reality, this concern is not the only, nor indeed the strongest motivation behind their thoughts and actions. To 'will truth' is not necessarily to will it *above all things*, or to will it *at any price*. While the ideal of complete or absolute truth might motivate philosophers and scientists — and even in their case only for a short while, if at all — it is certainly not what most people look for. In everyday life, the interest in truth is indeed pervasive, but not very demanding. Although human beings rely on the notion of truth, they are usually content with illusion; and although their adopted version of reality is usually assumed to be true, it is only as true as it needs to be in order to be believed. In most situations, most of us are either too busy or too lazy to get to the bottom of things. Unlike philosophers or scientists, we usually settle for a given definition or course of action without asking too many questions.

Most human 'truths' conceal the acceptance of a great deal of ignorance, falsehood and self-deception. Nietzsche refers to this phenomenon as the 'will to illusion' and highlights its practical or existential utility.[553] If man were merely 'a knowing animal' (*ein erkennendes Thier*), 'truth would drive him to despair and destruction.'[554] In daily life, however, epistemic claims are usually subordinated to practical demands and the accuracy of one's ideas is conditional on how far they will help one pursue one's practical interests. '"Truths" are demonstrated through their effects, not through logical proofs', and the falsest judgements are sometimes the most indispensable to human life. 'What is *true* and *what works* are regarded as identical.'[555]

In its customary form, then, human life is moved by two complementary drives.

XI, 109. 146f., 164f., 520f., 624, 648f., 699; *KSA* XII, 121, 352, 364ff., 384f.; *KSA* XIII, 440f., 446f., 500.

[553] Again, Nietzsche uses many different terms for this phenomenon: *Wille zur Illusion, Wille zur Täuschung, Wille zum Schein, Wille zur Dummheit*, *Trieb nach Glauben, Bedürfnis nach Glauben, Bedürfnis nach Illusion, Verlangen nach Glauben, Verlangen nach Illusion*, etc. On these notions, see notably *KSA* I, 760, 876f., 888ff.; *KSA* II, 51, 54f.; *KSA* III, 260f.; *KSA* V, 15ff., 41f., 56f., 113f.; *KSA* VII, 433, 435, 476, 498f., 526, 633; *KSA* IX, 466; *KSA* XI, 244, 648f., 699; *KSA* XII, 323; *KSA* XIII, 23, 193f., 226, 229, 522, 531.

[554] *KSA* I, 760

[555] *KSA* VII, 433. Emphasis added.

Usually, the 'will to illusion' has the upper hand: the practical or existential ease felt in everyday life is a clear evidence that the problem of truth is never seriously raised. But even though one's concern with truth is not very demanding, the 'will to illusion' is always grounded in a 'will to truth'. If one's ideas and judgements were not taken to be true, they would not be illusory, because an illusion is precisely a misplaced ascription of truth. Nietzsche's contrast between *Wille zur Wahrheit* and *Wille zur Täuschung* does not refer, therefore, to a simple dualism, but to a cognitive gradation, ranging from a *pure* or *strong* concern with truth — i.e. a will to truth proper, free of illusion — to an *impure* or *weak* concern with truth — i.e. a 'will to illusion', with varying degrees of intensity.

The 'will to illusion' is a specific mode of the 'will to truth'. Moreover, it is the mode that tends to dominate in everyday life. This general description is in many ways a recapitulation of Plato's doxastic scheme. But it is also the underlying scheme of every philosophical project grounded, in one way or another, in the ideal of truth. We still find it in Kant, where the main issue at stake is the possibility of true knowledge, and we also find it in Hegel, whose phenomenological project aims to reveal 'the actual knowledge of what truly is.'[556] Generally speaking, all of these projects acknowledge the disagreement between the alleged truthfulness of human beliefs and the standard of truth they are supposed to fulfil. According to Nietzsche, they all amount to different accounts of the cognitive scale ranging from truth to illusion.

Nietzsche argues that all previous philosophers, in their different efforts to expose the doxastic nature of human knowledge, were themselves subjected to the influence of a thoroughly doxastic presupposition — namely, the presupposition of truth's referential status. His critique is often articulated by means of visual metaphors, and particularly through the notion of *perspective*. Although the Platonic tradition and its subsequent developments aim to denounce ordinary consciousness' usual *lack of perspective* — viz. its usual tendency to regard the limits of its vision as the limits of the world — they do so from an equally provincial or limited standpoint. By positing the ideal of truth ahead of their examinations, most philosophers ask us to 'think an eye which cannot be thought of at all, an eye turned in no direction at all, an eye where the active and interpretative

[556] *HW* III, 68; *PS*, 46

powers are to be suppressed, absent, but through which seeing still becomes a seeing-something.'[557] This eye is the 'eye of truth', to which nothing is hidden or invisible. And yet, Nietzsche argues, 'there is *only* a perspectival seeing, *only* a perspectival "knowing"; the *more* affects we are able to put into words about a thing, the *more* eyes, various eyes we are able to use for the same thing, the more complete will be our "concept" of the "thing", our "objectivity".'[558]

Habitually, the notions of perspective and perspectivism tend to be understood in one of two generic ways. To have a perspective on something might mean, firstly, that one views only an aspect or facet of a wider or more complex reality. One is not in possession of a global picture, or a global understanding of things, and this limitation may entail some form of self-deception. This, however, is not Nietzsche's meaning, for such an interpretration still assumes that there is a specific *truth* to be known, which one cannot fully grasp. Secondly, in a stronger sense, the notions of perspective and perspectivism might be used to mean that there is no fixed or final version of things, and that a given object or event can only be known in partial and subjective ways. This interpretation, albeit closer to Nietzsche's meaning, is still not what he appears to have in mind. Instead of *one* truth, there are now several possible truths, but all of them are still grounded in the idea of cognitive adequacy. These different perspectives, despite their profusion, are still supposed to provide specific insights into what something *really is*, into its *true identity*, into its *actual meaning*, etc.

In its authentic and most radical sense, Nietzsche's perspectivism proposes to overcome this general attitude. It is not directed at specific objects or realities, whose true identity is more or less apparent, but at the *very notion of truth*. His point is that the tendency to think in terms of truth and untruth, or accuracy and inaccuracy, is itself a form of perspectival confinement. To adopt such an attitude is to reduce one's contact with reality to a cognitive perspective, whose *regional* or *limited* nature is not recognized as such. As Nietzsche puts it, to 'see' reality with the 'eyes' of truth and cognition is to forfeit the richness of one's original vision, and the various possibilities it affords.

[557] *KSA* V, 365; *GM*, 87

[558] Ibid. Cf., for ex., *KSA* II, 20f.; *KSA* III, 592f.; *KSA* IX, 466; *KSA* XI, 699ff.; *KSA* XII, 315; *KSA* XIII, 373f.

In order to understand the full implications of Nietzsche's criticism, it is important to return once more to the Platonic and the Hegelian projects. In the latter cases, the problem of truth came down to the correspondence between subjective certainty and objective truth, *Begriff* and *Gegenstand*. However, Plato and Hegel did not simply confront human beliefs with an independent ideal of truth, brought in from the outside. They both argued that a truly epistemic standpoint can only be attained through an immanent critique of natural consciousness. In Plato's case, this critique amounts to the *negative* procedure whereby Socrates attempts to 'purify' the minds of his interlocutors; in Hegel's case, it amounts to the *negative and positive* procedure whereby each phenomenological stage is sublated and replaced by a higher and more comprehensive one. But although Plato, Hegel and countless others questioned the alleged truthfulness of human beliefs, they failed to question the value of truth itself. According to Nietzsche, they failed to recognize that truthfulness is not the only standard — nor, indeed the most important one — with which to assess the value of human beliefs.

> This famous truthfulness that all philosophers so far have talked about with veneration: what questions this will to truth has already laid before us! […] We paused for a long time before the question of the cause of this will — until we finally came to a complete standstill in front of an even more fundamental question. We asked about the *value* of this will. Granted, we will truth: *why not untruth instead?* And uncertainty? Even ignorance?[559]

The *idea of truth* — the *Ansichsein* which for Plato and Hegel represents the necessary backdrop of every act of cognition — is for Nietzsche the result of a prejudice. It is not necessary, but contingent; not the standard of measurement, but something more to be measured. In other words, the philosophical demand that every knowledge be *true* knowledge (whether in its stronger, epistemic variant, or in its weaker, doxastic one) does not follow from the definition of knowledge, but from a specific and ultimately unjustified *attitude* towards knowledge. It is the product of a peculiar kind of 'will', which tends to conceal its volitional origin and to pose as something essential and self-evident.

[559] *KSA* V, 15; *BGE*, 5.

To look beyond the familiar contrast between truth and illusion is the task of a new breed of philosophers, of which Zarathustra is the forerunner. Their mission is to replace the opposition between truth and untruth with the more general opposition between 'validity' (*Wert*, *Gültigkeit*) and 'invalidity' (*Wertlosigkeit*, *Ungültigkeit*). More precisely, their task consists in showing that false judgements may be *valid* judgements and true judgements *invalid* ones. Once the 'will to truth' is overcome, the truth-value of a given judgement is no longer a valid argument for or against it. By then, human beings will have moved beyond this alternative and (presumably) learned to rely on new forms of valuation.

In this shift from *truth* to *value*, Nietzsche saw the key to a new and radically different outlook on reality, no longer bound to the narrow confines of cognition. But what exactly does this transformation entail? And what exactly is its end result? A brief survey of Nietzsche's writings is enough to reveal the multifarious and often ambiguous nature of his answers to these questions. In line with his chosen methodology, he addresses this issue in a truly 'perspectivistic' way, changing the tone, emphasis or angle of his description in order to make different points or to fit different contexts. Accordingly, his views on the 'transvaluation' of truth can be interpreted in at least three different ways, all of which are featured in his published and unpublished writings:

1. Firstly, to overcome the 'will to truth' may amount to embracing a 'will to untruth', that is, a conscious or voluntary 'will to illusion'. While most of us have no qualms in accepting various forms of illusion, we do not accept them *as illusions*, but precisely *as truths*. This does not mean, of course, that we do not deceive other people, nor that we are always against being deceived: dishonesty is an important part of human life, as is the aesthetic pleasure derived from fiction and fantasy. However, the kinds of deception we are ready to accept are usually very restricted. They refer to superficial elements of a much wider and complex account of reality, whose truthfulness is usually taken for granted. When we go to the cinema, for instance, we may agree to being transported to a world of fiction or make-believe, but this deception is very limited. Throughout the film, our idea of who we are, of where we are, of the identity of the

people around us, of the physical structure of the surrounding space, etc., is not usually affected.

In order to 'consciously remain in untruth'[560], we would have to relinquish our implicit faith in all of these judgements. The transformation envisaged by Nietzsche cannot be compared, therefore, to anything we know or have experienced. A life in which illusion — *all* illusion — were accepted and recognized *as illusion* would be very different from ours. It would resemble, perhaps, a fantastic dream, or a fantastic theatre play, where every incongruity is accepted as part of the plot. The falsity of an idea or judgement would no longer be an argument against it, nor would truthfulness be synonymous with value or legitimacy. In this new world, contradiction would be as welcome as coherence, and encouraged whenever it proved advantageous.

But admitting that this transformation is possible, or even conceivable, its end result still presupposes the contrast between truth and untruth. To will illusion *as illusion* is still only to will *the opposite of truth*, i.e. untruth. Although truth has been denied its referential status, it is still employed, if only negatively, as a standard of valuation.

2. A second way of construing the argument, closely related to the previous one, amounts to recognizing the subsidiary nature of the contrast between truth and untruth vis-à-vis other, more essential forms of valuation. This theme is specifically evoked whenever Nietzsche speaks of the need to replace the 'philosophy of truth' with a 'philosophy of life', or 'self-enhancement'. In *Beyond Good and Evil*, for example, the value of human judgements is said to depend not on their truthfulness, but on 'how far [they] promote and preserve life, how well [they] preserve, and perhaps even cultivate, the [human] type.'[561] *In the Gay Science*, Nietzsche extends this critique to philosophical judgements, and calls for the replacement of conventional philosophers with philosophical *physicians*, capable of exposing philosophical 'truths' and 'the unconscious disguise of physiological needs.'[562]

As we will see below in more detail, this general argument is both *descriptive* and *prescriptive*. On the one hand, Nietzsche is claiming that human life is ultimately not

[560] *KSA* II, 53f; *HAH*, 29

[561] *KSA* V, 18; *BGE*, 7: 'Die Frage ist, wie weit es [sc. ein Urtheil] lebenfördernd, lebenerhaltend, Art-erhaltend, vielleicht gar Art-züchtend ist.'

[562] *KSA* III, 348; *GS*, 5

about truth, but about health, future, growth, power, etc. The 'will to truth' is merely the disguise of a more essential or more primary existential drive — a 'will to life', or a 'will to power'. On the other hand, Nietzsche is prescribing the acknowledgement and the correction of this mistake. He is advocating a return to life's original or authentic from, beyond the artificial sway of truth and cognition.

Once again, the intended result of this transformation is the divorce between truth and value. Whenever false or illusory judgements prove empowering or life-promoting, they are accepted as valid judgements. Conversely, whenever an increase in clarity or accuracy entails a decrease in strength or vitality, the corresponding judgement or set of judgements is rejected. A life lived on these terms is indeed very different from our usual life, but the limitation found in the previous reading of Nietzsche's argument is still present here, albeit in a subtler form. For even if this new solution is possible, or even conceivable, it still presupposes the notion of truth.

Admitting that one can become entirely indifferent to the truthfulness of one's beliefs and care only about power, vitality or self-enhancement, this concern is still grounded in a global diagnosis of who one is, where one stands, what is at stake in one's life, what matters most in it, and so on. The pursuit of power and self-enhancement still presupposes a vital atlas, or a vital script, in which power and self-enhancement play the central role. It is supported by the general conviction that 'life is about power, or self-enhancement' — a conviction deemed *truer* than, say, the conviction that 'life is about truth', or about 'pleasure', or about 'money', or about 'fame', or about anything else. In this general sense, the 'will to power' is still but a specific instantiation of the 'will to truth'.

3. Finally, a third and more radical way of construing the argument is to take Nietzsche at his word and to envisage the overcoming of the 'will to truth' in *all of its forms*. For human beings to break free from the idea of truth, they must overcome their usual need of a vital atlas, or a vital script. A will that is no longer regulated by the idea of truth can no longer ground its claims in an assessment of how things *are* and how they *are not*, of who one *is* and who one *is not*, of what is *at stake* in a given situation and what is *not at stake* in it, etc. This extreme possibility raises a very difficult issue, concerning the criterion employed by this new form of consciousness. If the notions of truth and untruth cease to guide human judgements, which notions, if any, are liable to take their

place? Assuming that human cognition requires nonetheless *some kind of normativity*, what will the latter be based on? These and similar challenges call into question the very possibility of Nietzsche's 'Zarasthustrian' revolution, as it is now being interpreted: is the notion of truth detachable from the notions of value, will, choice, power, and so on? And if so, to what extent? Is it possible to value something over something else without automatically deeming it *truer*, or *less illusory*?

All of these interrogations point to a new and completely uncharted territory. They seem to require an insight into a kind of experience that our usual standpoint is simply not equipped to conceive. Accordingly, Nietzsche often places the goal of his journey beyond the limits of humanity: perhaps the overcoming of truth, as he understands it, is indeed an 'overhuman' task, whose result cannot be appreciated from this side of existence.

10.2. *Truth and metaphor*

The 'problem of truth' is arguably Nietzsche's central philosophical concern, revisited again and again throughout his writings. His solution to this problem, however, besides allowing for very different interpretations, is also characterized by a further and important ambiguity. In his *Zarathustra* and elsewhere, Nietzsche insists on the radical contrast between human beings' natural cognitive condition and a new, 'overhuman' condition. This contrast is meant to highlight the absolute incompatibility between human cognition's current form and the form it would have to assume if it were to leave behind its attachment to truth. There is no meeting point, no reconciliation, no common ground between these two existential attitudes: since human reason is by definition incapable of grasping the kind of transformation Zarathustra prescribes, the only way to do so is to acquire an altogether new identity or mode of being. To borrow Pyrrho's well-known formula, the only way to achieve this transformation is 'to strip oneself completely of being human'[563], and to put on a new existential garment.

[563] LONG; SEDLEY, op. cit. I, 14 C

In other texts, however, this evolutionary theme is replaced by a different approach to the problem of truth, namely a historical or genealogical approach. In these cases, Nietzsche's strategy is to put the 'will to truth' into *historical* perspective, by arguing that human beings' usual attachment to truth is not a fixed cognitive trait, but an acquired one, whose nature and intensity have changed considerably over time. According to this line of argument, humanity's current cognitive condition is a degraded or decadent condition, deviated from its original form and in need of rehabilitation.

This genealogical strategy is first developed in a brief and incomplete text entitled *On truth and lying in a non-moral sense*, written in 1873 and left unpublished during Nietzsche's lifetime. The text's initial focus is once again the origin of truth's normative value: *when* did human beings' interest in truth first emerge? *Where* does it come from? *How* did it become such a generalized phenomenon? According to Nietzsche, the answers to these questions require a journey to the infancy of human history. Before the advent of communal life, when human beings were still solitary creatures, truth and untruth were devoid of normative value. Primitive men and women were exclusively focused on their own survival and had no qualms in deceiving both other creatures and themselves. Their whole existence, Nietzsche argues, was based on a series of spontaneous, provisional and often self-contradictory beliefs, whose truth-value hinged solely on whether they were momentarily useful or useless, beneficial or harmful, self-promoting or self-diminishing. Human intellect was fully subordinated to human life and reason counted solely as a life-enhancing tool, just like physical strength or dexterity.[564]

At some point in time, due to 'necessity and boredom', this primeval solitude gave way to the creation of small communities, grounded in collective pacts of non-aggression.[565] Nietzsche's description is thoroughly Hobbesian, and he even speaks of a

[564] Nietzsche's criticism of the 'will the truth' is liable to be interpreted in more radical and less radical ways. Considering my previous inventory, the present line of argument can be linked to the first and second hermeneutical possibilities, but not yet to the third: the primitive life-experience Nietzsche is now documenting — which he will also associate with primitive Greek life — is already supported by a 'will to truth' in the more general sense, viz. by a need for what I have termed, in connection with Plato's and Hegel's philosophies, a vital atlas or a vital script. However, this primitive concern for truth is less intense or less automatic than a modern and more robust 'will to truth'.

[565] A more detailed account of the reasons behind the transition from solitary to communal life are offered, for ex., in *KSA* II, 95: 'Die Lustempfindung auf Grund menschlicher Beziehungen macht im Allgemeinen den Menschen besser; die gemeinsame Freude, die Lust mitsammen genossen, erhöht

bellum omnium contra omnes, replaced by a 'peace treaty' recognized and upheld by every individual. However, his primary concern is not the birth of political authority, nor the origin of social laws and institutions, but the historical evolution of the notions of truth and untruth.[566]

In a pre-social environment, truth is a personal matter. Since there is no real communication, language is not yet needed and reason is still made up of inarticulate beliefs, of spontaneous and changeable impressions. This does not mean, of course, that every human belief is equally valid, or that every human action is equally desirable. As Nietzsche will later write, 'humans first placed value into things in order to preserve themselves.'[567] Before truth was elevated to a logical and cultural ideal, the value of human beliefs was entirely determined by the will to preserve and cultivate life. Therefore, although the primitive men imagined by Nietzsche are not yet social or scientific beings, their cognitive activity is already guided by a complex system of practical assumptions. They do not live in a state of permanent wonder, as though the world were created anew at each given moment. They are continually engaged in the process of *interpreting* reality and adapting their behaviour to the results of their interpretations. Their intellectual activity consists in reducing the world to a series of concrete meanings, in breaking it down to a web of similarities and differences, in sorting what is important from what is irrelevant, what is beneficial from what is harmful, what is pleasant from what is painful, etc.[568] Primitive individuals know, for instance, that a cloudy sky is the sign of an approaching storm and realize they must take shelter; they associate the shapes and colours of different animals with different degrees of danger and different hunting techniques; they know, moreover, that some seasons are warmer than

dieselbe, sie giebt dem Einzelnen Sicherheit, macht ihn gutmüthiger, löst das Misstrauen, den Neid: denn man fühlt sich selber wohl und sieht den Andern in gleicher Weise sich wohl fühlen. Die *gleichartigen Aeusserungen der Lust* erwecken die Phantasie der Mitempfindung, das Gefühl etwas Gleiches zu sein: das Selbe thun auch die gemeinsamen Leiden, die selben Unwetter, Gefahren, Feinde. Darauf baut sich dann wohl das älteste Bündnis auf: dessen Sinn die gemeinsame Beseitigung und Abwehr einer drohenden Unlust zum Nutzen jedes Einzelnen ist.'

[566] Nietzsche's views on social contract theories are not univocal, and seem to range from a complete endorsement to a complete rejection (confront, for ex., *KSA* II, 89f., 95, 99f., 114 or 555ff. with *KSA* IV, 88; or *KSA* V, 324f.).

[567] *KSA* IV, 75; *TSZ*, 43

[568] In other words, they are already driven by a 'will to truth' in the general sense.

others, that the air is usually colder at higher altitudes, that the sight of smoke is usually the sign of a fire, and so on.

In his eagerness to highlight the contrast between primitive and modern reason, Nietzsche goes so far as to claim that for primitive men, 'any conception of *natural* causality is altogether lacking.'[569] Although their understanding of reality is based on a series of practical assumptions, their ideas lack the force of actual laws. 'In regard to everything external … no conclusion can be drawn that something *will* be thus or thus, *must* happen thus or thus; it is [*they*, namely these individuals] who are the more or less secure or calculable; man is the *rule*, nature is *irregularity*.'[570] In this primitive stage, human beliefs are inherently provisional and their articulation is loose and variable. There is no explicit demand for logical coherence and no actual premium placed on intelligibility or verisimilitude. Although reality is always interpreted to be thus or thus, and although most things are expected to happen in a certain way rather than another, these assumptions are liable to change just as easily as they came to be adopted.[571]

In a social environment, however, things are different. Personal assumptions must be externalized and shared with others. Their truth-value is no longer a personal matter, but a social one, and this change introduces the need for a coherent system of linguistic signs. Different things are accorded different names, and the latter must be recognized and adopted by an entire community. As a result, truth and untruth lose their former plasticity and become attached to specific linguistic combinations. They are henceforth

[569] *KSA* II, 112; *HAH*, 63

[570] *KSA* II, 112f.; *HAH*, 63

[571] It might be objected, of course, that this 'cognitive plasticity' is never as wide Nietzsche seems to think. Although some of the empirical connections that guide one's practical life are relatively weak, and thus more liable to be questioned — such as, for instance, the assumption that a cloudy sky is the prelude of a storm, or that smoke is the sign of fire — there are other, much stronger connections, whose validity is harder to challenge — such as, for instance, the assumption that heavy objects have a tendency to fall to the ground, or that the sun rises anew every morning. Surely even a primitive human being is bound to realize that these things happen *every single time*, and is therefore apt to feel their inevitability just as strongly as modern men. Nietzsche's point, however, is not that primitive human beings are not guided by empirical assumptions, but that they do not think of reality in terms of truth and untruth. Like modern men, they would be surprised if a heavy object, when dropped, would float in space instead of falling to the ground. But whereas for modern men this discovery would amount to the collapse of a whole system of beliefs, for the primitive man it would simply prompt the endorsement a new and equally provisional constellation of empirical connections. In the latter case, the belief in 'gravity' would be revised simply because it would have lost its *instrumental* value, that is, because it would have proven useless or unreliable when applied to the fulfilment of a specific task. Nevertheless, this general explanation is still open to several objections, to which I will return in the following pages.

the measure of the correspondence between a given thing or action and its appropriate designation.

This normative turn introduces for the first time the possibility of error, which can arise out of two different situations. A person can err involuntarily, by not knowing or forgetting how to use one or more linguistic signs, or voluntarily, by replacing a correct designation with an incorrect one.[572] In the latter case, error gives way to *deceit*: the linguistic contrast between truth and falsehood leads to the more personal contrast between guilt and innocence, and the practical concern for truth will eventually give way to a *moral* concern for truth. But this new development still lies ahead: when language is created, the notions of truth and untruth are still independent from both epistemological and moral concerns. Truth and honesty are not yet intrinsically valuable, nor desired for their own sake. Whereas human beings 'desire the pleasant, life-preserving consequences of truth, they are indifferent to pure knowledge if it has no consequences', and 'actually hostile towards truths which may be harmful and destructive.'[573] Likewise ignorance and dishonesty are not yet bad or reproachable in themselves, but only insofar as they harm other people or contradict the common good. 'Human beings do not so much flee from being tricked as from being harmed by being tricked.'[574] Hence, only when the failure or the unwillingness to conform to the accepted linguistic standards proves dangerous or harmful is the ignorant scolded and the liar banned.

In this primitive scenario, no one wills truth *as such*. The notions of truth and untruth are contrived as a means of breaking the usual soliloquy of human experience and finding a common ground for sociality and communication.[575] They are tools devised to assess the degree of correspondence between human experience and an arbitrary system of signs, established by decree and strengthened by habit. Consequently, to show that something is true or untrue is not to reveal what it is in itself, but simply to determine whether it conforms to the linguistic sign it was originally accorded. As

[572] Nietzsche only considers the second of these situations, but the first one is implied in his argument.

[573] *KSA* I, 878; *BT*, 143

[574] Ibid.

[575] Note the difference between Nietzsche's and Hegel's model, according to which sociality is not a consequence, but the very basis of self-interest and self-identity.

Nietzsche puts it, 'if someone hides something behind a bush, looks for it in the same place and then finds it there, his seeking and finding is nothing much to boast about.'[576] In the same manner, if truth is merely the confirmation of a pre-established convention, its significance is purely tautological. It is the product of a self-referential system whose different contents are bound to the same immanent rule.

However, Nietzsche's argument is not simply about words, but about the actual things they stand for. The notion of truth and all of its contradictions arise from the fact that words are never merely words, but the vehicles for specific theses about reality. Indeed, it is easy to see that the link between a given name and the thing it refers to is merely contingent: few people would endorse the claim that the English word 'tree', for instance, is in a better position to convey what a tree *really is* than, say, the German word *Baum*, or the Portuguese word *árvore*; these words amount to cultural and historical conventions, whose origin is ultimately just as arbitrary as that of any other word. The point, however, is that all of these words are usually uttered in the conviction that they refer to something real and definite, i.e. a tree καθ'αὐτό, placed somewhere beyond the linguistic and conceptual realms. *Tree*, *Baum* and *árvore* are usually regarded as approximations to an actual empirical content, whose concrete truth is irreducible to this or that specific term.[577]

[576] *KSA* I, 883; *BT*, 147

[577] Nietzsche's text is ambiguous as regards the specific role played by words and concepts in the genesis of the contrast between truth and untruth. Firstly, as we have just seen, he locates the origin of this contrast in a social compact, embraced out of 'necessity and boredom.' With the transition from the individual to the social sphere, 'that which is to count as 'truth' … becomes fixed, i.e. a way of designating things is invented which has the same validity and force everywhere, and the legislation of language also produces the first laws of truth.' (*KSA* I, 877; *BT*, 143) This initial claim is then used to back the second part of the argument, concerning the origin of the human concern for truth. As individuals grow accustomed to ascribing specific words to specific things, they tend to forget that the former are mere conventions and come to believe in an essential link binding each word to the thing it refers to. Now if one accepts the idea that truth's normative value is the result of a social compromise, one must also acknowledge that the latter could never have arisen if *some kind of normativity* was not already in place. Inasmuch as one's pre-social life is already filled with all kinds of assumptions and expectations, the pre-social world is already interpreted and hierarchized in countless ways, divided into specific shapes and meanings, adapted to various needs and routines. The solitary man may not need words or an alphabet, but he is just as capable of predicting specific events and ascribing (relatively) fixed meanings to things.

The question one should ask, then, is whether the will to truth is necessarily dependent on a social compact, or whether it can arise simply out of habit. In other words, one should ask whether the solitary man is not just as liable, with time and habit, to forget the subjective origin of his assumptions and to mistake them for actual truths. This ambiguity might help explain why Nietzsche decided to complement his Hobbesian approach with an alternative account of concept-formation, closer to a classical empiricist

Nietzsche's aim is to show that this assumption is unfounded. And what is more, that it amounts to a complete subversion of the way human concepts are actually engendered. The notion 'tree' is born as a reaction to a nervous stimulation, which is then associated with an image, which is then translated by a specific sound. Nowhere in this process is a tree *in itself* to be found: the whole movement starts within the sphere of subjectivity and only afterwards does it spill over, so to speak, into the sphere of objectivity. As we grow accustomed to linking a specific stimulation with a specific image or sound, we tend to assume that the former is caused by something outside us, namely by an actual tree. What started out as a subjective impression ends up as a reified presence: like Pygmalion, we are surpassed by our own creations and led to believe their intrinsic truth.

As he retraces the genesis of the contrast between truth and untruth, Nietzsche arrives at one of the central themes of his earlier writings, namely the self-enclosedness of human knowledge and the impossibility of reaching beyond the sphere of subjectivity. Since all human concepts are self-referential, words do not grant access to reality *as such* — nor are they supposed to. What they do reveal is the relation of reality to the human mind. They document the different ways in which we are affected by this or by that nervous stimulus, the inner hierarchy of our perceptions, the properties of things we choose to highlight and the ones we fail to notice, etc. In a way, human concepts are like mirrors: the image they display is our own image as seen through the different filters or lenses of everyday life.[578]

approach (cf. 387ff.) As the essay progresses, his emphasis on *sociality* (and *words*) is gradually shifted to an emphasis on *habit* (and *concepts*), and although the former line of enquiry is still necessary to explain the emergence of truth as a *moral* concept (an idea which is left for the most part undiscussed), the latter line of enquiry seems to acquire a life of its own and to provide an account of the origin of the 'will to truth' which is ultimately independent from the social compact mentioned at the outset.

[578] Bernard Williams seems to be referring, at least partly, to *Ueber Wahrheit und Lüge im aussermoralischen Sinne* when he claims that 'in his earliest writings about truth and error, Nietzsche sometimes spoke as though he could compare the entire structure of our thought to the real nature of things and find our thought defective'. This picture, which Nietzsche 'rightly rejected', would imply that 'we can somehow look round the edge of our concepts at the world to which we are applying them and grasp it as entirely unaffected by any descriptions.' (*GS*, xxf.) In my view, however, this is precisely the kind of dualism Nietzsche strives to overcome. If he were simply positing a world *in itself* on one side and the realm of subjectivity on the other, and if he were simply claiming that the two can never truly coincide, then his whole argument would indeed come down to a fairly conventional form of empiricism. But his point is rather that the whole business of concept-formation takes place *within the subject*: even when human concepts are said to be triggered by 'sensuous impressions' or 'nervous stimulations', Nietzsche is

To put it more plainly, the original source of the world's meaning is not the world itself, but our apprehension of it. *We* are the creators of the truths we put into words — the *artists* in whose image reality is forged. But this process is usually accompanied by a peculiar form of amnesia, favoured by time and habit, which leads us to forget our creative role and to regard truth and meaning as already given.

Nietzsche will later return to this same idea:

> Because we have for millennia made moral, aesthetic, religious demands on the world, looked upon it with blind desire, passion or fear, and abandoned ourselves to the bad habits of illogical thinking, this world has gradually *become* so marvellously variegated, frightful, meaningful, soulful, it has acquired colour — *but we have been the colourists*: it is the human intellect that has made appearance appear and transported its erroneous basic conceptions into things.[579]

To cognize is to *create*, and human reason is above all a *creative* process. Nevertheless, as words and concepts become more familiar, this acknowledgement tends to lose its force. As names grow attached to the things they name, and as they are passed on from generation to generation, their initial arbitrariness falls out of view and their truth acquires a new and illegitimate status. Trees, animals or cars are no longer regarded

careful to point out the immanent nature of the whole procedure. He claims that 'to infer from the fact of the nervous stimulation that there exists a cause outside us is already the result of applying the principle of sufficient reason wrongly.' (*KSA* I, 878; *BT*, 144)

Following Kant (and Plato), Nietzsche is pointing out the contradiction between the *transitivity* and the *self-enclosedness* of human cognition. Whereas on the one hand our knowledge of reality is necessarily mediated by words and concepts — we do not think of a tree as an abstract coloured mass, but as a particular instantiation of the εἶδος 'tree' — on the other hand the meaning of this concept can only be provided by pointing to specific trees — that is, to empirical events which are defined precisely by partaking of the εἶδος 'tree'. Although human concepts are usually *transitive* — that is to say, although they are usually burdened with the idea of an objective, trans-linguistic or trans-conceptual *beyond* — their meaning can only be determined *within* the conceptual sphere, by resorting to concepts which lead in turn to other concepts, and so on indefinitely. 'Zwischen zwei absolut verschiedenen Sphären wie zwischen Subjekt un Objekt giebt es keine Causalität, keine Richtigkeit, keinen Ausdruck, sondern höchstens ein *ästhetisches* Verhalten, ich meine eine andeutende Uebertragung, eine nachstammelnde Uebersetzung in eine ganz fremde Sprache.' (*KSA* I, 884).

[579] *KSA* II, 36ff.; *HAH*, 20. Emphasis added. Cf., for ex., *KSA* XII, 352f.: '"*Die wahre und die scheinbare Welt*" — diese Gegensatz wird von mir zurückgefürht auf *Werthverhältnisse* / wie haben unsere Erhaltungs-Bedingungen projicirt als *Prädikate des Seins* überhaupt / daß wir in unserem Glauben stabil sein müssen, um zu gedeihen, daraus haben wir gemacht, daß die "wahre" Welt keine wandelbare und werdende, sondern eine *seiende* ist'; or 385: 'Der Mensch projicirt seinen Trieb zur Wahrheit, sein "Ziel" in einem gewissen Sinn außer sich als *seiende* Welt, als metaphysische Welt, als Ding an sich, als bereits vorhandene Welt.'

as the arbitrary translations of subjective stimuli, but as *things in themselves* whose original or objective essence was translated into words and concepts. As a result of this change, human beings become convinced that their knowledge is wider and more efficient than it actually is. We come to believe that when we speak of trees, animals or cars 'we have knowledge of the things themselves, and yet we possess only *metaphors* of things which in no way correspond to the original entities.'[580]

This use of the word 'metaphor' leads back to Hegel, and to his own concept of *Metapher*, or *Vorstellung*.[581] Much like Nietzsche, Hegel resorts to these terms in order to highlight the fundamental contrast between natural consciousness' supposed knowledge and its actual ignorance. As seen earlier, the Hegelian critique of natural cognition, as developed throughout the *Phenomenology* and elsewhere, is set in motion by two complementary acknowledgements: first, the fact that human knowledge, in its natural or customary form, is usually taken to be clear and adequate, and to provide direct access to *things in themselves*; second, the fact that this assumption, when subjected to a closer examination, turns out to be groundless. In Hegelian terms, natural ideas are not clear or adequate, but mere *metaphors* of concepts whose actual content is out of reach. Yet Hegel's point is not simply that natural knowledge falls short of true knowledge, but that it is bound to *start out* as such. In other words, it is precisely because human ideas are intrinsically *metaphorical*, and act as the signposts of a meaning which is to be found *elsewhere*, that knowledge as such is possible. This centrifugal movement is not only the underlying basis of every natural idea, but also the starting point of the dialectical pathway towards true knowledge.

Moreover, although Nietzsche's formula is evocative of Hegel's, both authors' descriptions are primarily grounded in Plato's critique of human cognition — and particularly in the Platonic notion of δοξάζειν, and the related phenomena of μέθεξις and ἀνάμνησις. Nietzsche's (and Hegel's) main argument is thoroughly Platonic in that it highlights both the fundamental gap and the fundamental bond that separates and

[580] *KSA* I, 18; *BT*, 144. Emphasis added.
[581] Cf. notably *HW* VIII, 40ff.

unites human concepts and the meanings they are supposed to convey.[582] On the one hand, since each concept is part of an immanent or self-enclosed cognitive alphabet, which extends to every domain of human experience, things *in themselves* are just as conceptual as any other thought-content. To think otherwise is to fall prey to a conceptual fiction and to mistake the sphere of subjectivity for the sphere of objectivity. On the other hand, this fiction is itself a necessary condition for the intelligibility of most, if not all human concepts. If a tree were merely a nervous stimulation, and not also an image, a sound and a physical presence, it would be utterly unintelligible, and hence of no practical use. For one cannot lean against a mere concept, nor climb it, nor stand on its branches, nor chop them up into wooden logs. It is only because the concept 'tree' is not usually experienced as a mere concept, but as something real and definite, that a tree is a meaningful part of our existential horizon. The verisimilitude of the concept 'tree', as that of any other concept we happen to make use of, is a basic requirement for the practical viability of our daily existence.

Furthermore, like Plato (and Hegel), Nietzsche confronts these two clauses and concludes that the latter leads most of us to forget about the former. The practical need to assume the objective truthfulness of our concepts amounts to a peculiar form of amnesia, whose seriousness is aggravated with every new conceptualization. In other words, although human knowledge is entirely *metaphorical*, it is not lived or experienced as such. With time and practice, the metaphorical nature of our thoughts and ideas is gradually forgotten. Each of our concepts comes to be regarded as the *real thing* — or, at any rate, as an efficient or direct pathway into the *real world*. And if this habit is maintained for a long period of time, it will easily outgrow any form of resistance.

The constant need to translate personal stimulations into objective truths has led to the development of whole languages and systems of concepts, which have rendered

[582] In order to appreciate the various historical echoes that are here at stake, it is important to bear in mind that Nietzsche's use of the word *Begriff* — and my use of it, following Nietzsche — is different from Hegel's. For the latter, *Begriffe* are usually opposed to *Vorstellungen* and refer to the *real* or *true knowledge* which natural consciousness falls short of. Accordingly, Hegel often claims that natural knowledge is not yet *notional* or *conceptual* knowledge (*Begreifen*, or *begreifendes Denken*), but the mere *appearance* thereof. For Nietzsche, on the other hand, concepts are the equivalent of Hegel's *Vorstellungen*, or Plato's δόξαι. Hence, from Hegel's point of view, the parallel between his main argument and the Platonic and Nietzschean standpoints will only emerge if one replaces the word *concept* with the words *belief*, *representation*, *opinion*, etc.

humanity stronger, wealthier and more cunning. But the heights we have reached have blocked the view of the ground down below. Human reason has grown so vast and complex that human beings have lost sight of its moving foundations, and ended up mistaking it for a solid and natural construction. According to Nietzsche, this short-sightedness is the source of the epistemic hybris that has shaped human culture for such a long time.

> What, then, is truth? A mobile army of metaphors, metonymies, anthropomorphisms, in short a sum of human relations which have been subjected to poetical and rhetorical intensification, translation and decoration, and which, after they have been in use for a long time, strike people as firmly established, canonical, and binding; truths are illusions of which we have forgotten they are illusions, metaphors which have become worn by frequent use and have lost all sensuous vigour, coins which, having lost their stamp, are now regarded as metal and no longer as coins.[583]

Time and habit have led human beings to forget about the instrumental, metaphorical, or 'coin-like' nature of human concepts. As a result of this widespread amnesia, men and women have come to regard trees, cars, their own bodies and all other experiential contents as things in themselves and to assume they have a clear enough knowledge of what those things *truly* are. Nietzsche's whole argument consists in exposing the contingent nature of this attitude and to counter its usual popularity. His aim is to unearth the possibility of a standpoint to whom things are so or so for merely *instrumental* or *aesthetic* reasons: a standpoint whose beliefs are no longer supported by conceptual or epistemic claims; a standpoint to whom it is ultimately irrelevant whether things are real in themselves or not, true or fictional, subjective or objective; in short, a standpoint whose metaphors are acknowledged and lived *as metaphors*.

[583] *KSA* I, 880f.; *BT*, 146

10.3. *Science and Wisdom*

According to Nietzsche, every truth is the result of a metaphor, and human cognition is the creative process whereby each individual life is translated into images, words, and concepts. Although this definition is still somewhat vague, two important aspects are already clear: first, the 'drive to form metaphors' is not regional or optional, but constitutive, or transcendental. It 'cannot be left out of consideration for even a second without also leaving out human beings themselves.'[584] Secondly, it is as free and boundless as one's imagination. As the very word suggests, metaphors are much closer to artistic creations than they are to objective findings or scientific observations. Their origin is purely subjective and their ability to convey the meaning of a given reality or experience is ultimately an individual matter.

However, this last consideration only applies to *original* or *first-hand* metaphors, that is, to metaphors originally created by each individual to convey the meaning of his or her own experiences. Whereas Nietzsche's primitive men are indeed the creators of their own metaphors, and the sole judges of their value, modern men are usually not as creative or independent. They are born into a cultural world, grounded in several centuries of concept-formation. And though their experiences are in a way as individual or as personal as those of primitive men, they are no longer mediated by original metaphors. Their world is already labelled and documented, explained and hierarchized. Since cognition is now chiefly a matter of correspondence, objects and events are no longer as open to different meanings or interpretations: there is a *true* concept (or set of concepts) for every empirical content (or event), and every other combination is perceived as *less true*, or altogether *untrue*. Every failure to conform to this system is usually put down to ignorance, irrationality or artistic license: if one applies the term 'chair' to a table, for instance, or if one claims that human beings are 'four-legged animals', one's words will have a poetic meaning, or none at all.

As shown above, Nietzsche regards this rigidity as the result of a collective amnesia,

[584] *KSA* I, 887; *BT*, 151

which led to the elimination of human reason's original malleability. In his own words,

> Only by forgetting this primitive world of metaphor, only by virtue of the fact that a mass of images, which originally flowed in a hot, liquid stream from the primal power of the human imagination, has become hard and rigid, only because of the invincible faith that *this* sun, *this* window, this table is a truth in itself — in short, only because man forgets himself as a subject, and indeed as *an artistically creative* subject, does he live with some degree of peace, security and consistency; if he could escape for just a moment from the prison walls of this faith, it would mean the end of his 'consciousness of self' [*Selbstbewußtsein*].[585]

The idea that life's practical viability is conditional on a stable and relatively dogmatic system of truths is fairly intuitive. As Nietzsche himself concedes, if human reason would relinquish its normative dimension, and if all human meanings were suddenly open to revision, there would be no peace, security or consistency. If human imagination were to regain its fluidity, the notions of law and causality would be called into question, and experience itself would be radically transformed.

What is more, one does not have to imagine such an extreme scenario in order to appreciate the practical utility of conceptual dogmatism. In fact, if only half of human concepts were to lose their normative value — nay, if only one hundredth, or even one thousandth of all human concepts were to suffer a similar fate, our usual outlook on reality would already be significantly altered. If we ceased to rely on the implicit truthfulness of tables, windows, trees or cars, the specific role they play in our lives would be transformed. The practical ease with which we sit at a table, look out of a window, climb a tree or jump into a car is usually linked to our in-grained belief in the objective truth of these concepts. If tables, windows, trees and cars would appear to us as subjective creations, and no longer as *things in themselves*, the vast array of uses, operations, habits, qualities, and expectations we have come to associate with them would also be transformed. They would no longer blend into our existential landscape as they normally do, but stick out as new, unfamiliar and potentially unrecognizable elements.

[585] *KSA* I, 883; *BT*, 147. See also *KSA* II, 30f. As noted in Ronald Speirs' translation, the word *Selbstbewußtsein* can mean both 'self-consciousness' and 'self-confidence', or 'self-conceit'.

But this is not all. If the objective truthfulness of these and other concepts were radically questioned, the resulting transformation would not be limited to specific elements or regions of our existential horizon. It would affect, in one way or another, our entire understanding of reality. For unlike what we are usually led to assume, concepts are never atomic or elementary determinations: they cannot be reduced to simple cognitive events, nor are they liable to be understood or considered in isolation. Their meaning can only be grasped in context, as part of a complex semantic web whose ramifications extend to the whole conceptual sphere.[586] Returning to my previous example, a tree is always much more than a *simple* or *actual* tree — and this is so for at least three main sets of reasons: first, because the concept 'tree', in order to be understood, requires the simultaneous consideration of a endless series of related concepts, without which the concept 'tree' would not be intelligible (namely 'solidity', 'unity', 'identity', 'plant', 'life', 'trunk', 'leaf', etc.); second, because a tree is never simply an individual or isolated tree, but a universal genus, of which each individual tree is but a concrete instantiation; third, because the concept 'tree' is itself necessarily implicated in the meaning of many other concepts (namely 'trunk', 'forest', 'orchard', 'willow', 'tamarind', etc.).

This and similar examples demonstrate what Nietzsche argues by means of architectural images: namely, that human concepts are never loose or independent elements, but the constituents of a vast and intricate cognitive edifice, where every brick or beam has its proper place, and where all contribute in a specific way to the harmony of the ensemble. By 'dissolving images into concepts', or by 'sublimating sensuous metaphors into a schema'[587], mankind managed to escape the fleetingness of sense perception and the solitude of subjectivity. The normative turn from arbitrariness to abstraction enabled 'the construction of a pyramidal order based on castes and degrees, the creation of a new world of laws, privileges, subordinations, definitions of borders,

[586] One of the locus classicus of this idea is *So* 251ff., where Plato argues that human concepts are, by definition, *relational* determinations, naturally bound to *participate* in (κοινωνεῖν, μετέχειν), or to *comingle* with (μιγῆναι), other determinations.

[587] *KSA* I, 881; *BT*, 146

which … confronts the other, sensuously perceived world as something firmer, more familiar, more human, and hence as something regulatory and imperative.'[588]

This vast and admirable construction has been the source of much more than 'peace, security and consistency'. It was and still is the source of marvellous feats, the guiding force behind all sorts of practical and theoretical exploits, of technical and spiritual inventions. Throughout millennia, this conceptual dogmatism has helped human beings counter their natural limitations, heighten their physical and intellectual powers, extend and improve their lives. It has lifted human culture and civilization to unimaginable heights, endowing each human being with 'resourceful skill, a subtlety beyond expectation.'[589]

In light of all this, Nietzsche's claim that 'conceptual faith' amounts to a form of imprisonment is not as intuitive. For even if something has been lost, and human imagination has indeed been killed or stifled by the dogmatism of conceptual reason, a lot more seems to have been gained. And surely the primordial fluidity of human imagination, albeit free from the burdens of consequence and coherence, was itself a wider and much more damning form of imprisonment. In the absence of stable definitions, of fixed habits and laws, human beings were doomed to be 'swept away' by sudden impressions, and their lives were eternally unstable and nomadic. Like explorers travelling through an ever-changing landscape, they were unable to rest or to settle down, and their freedom — if one can call it so — was too wide to be liberating.

What, then, is the exact meaning, in this context, of Nietzsche's contrast between freedom and imprisonment? On the one hand, he does not seem to be advocating a simple return to ignorance, or a complete departure from the sphere of conceptual reason. Such a project would not only be self-contradictory — its very formulation is, after all, intrinsically conceptual — but also, in all probability, downright unfeasible. As we have seen earlier, with the aid of Plato and the *Philebus*, human cognition is irreducible to the spheres of sense perception (αἴσθησις) and memory (μνήμη).[590] Every act of cognition, however simple or primitive, already entails a specific diagnosis of

[588] *KSA* I, 881f.; *BT*, 146
[589] *Antigone*, line 365 (ed. cit., 37)
[590] See above section 2.1.

reality, that is, a silent claim (or constellation of claims) about how things are, what they mean, how they relate to one another, etc. And this diagnosis is by definition a *normative* diagnosis: it presupposes the distinction between adequate and inadequate portrayals of things, reliable and unreliable representations, accurate and misleading versions of events, etc. All of this is of course irreducible to the mere storing and combining of sensations. It is the work of an empirical standpoint, whose cognizing is already a form of distinguishing, sorting out, comparing, judging and putting up boundaries.

In addition, we have also seen that an empirical standpoint such as ours is never stuck in the present moment. Our diagnosis of reality is always a proleptic diagnosis, whose aim is just as much to determine what things are (and were) as it is to anticipate what they will be. To cognize is thus not simply to describe or to judge, but to plan and to anticipate; to harbour expectations, based on one's current and past experiences, about how things are supposed to evolve. And this tendency, as we have also seen, is the direct cause of another one of our standpoint's defining features: namely, the propensity, heightened by time and habit, to transform our empirical observations into fixed laws, endowed with intrinsic or objective truth. In Platonic terms, this propensity speaks to the *doxastic* nature of human cognition; in Nietzschean terms, it reflects the historical tendency to forget the *metaphorical* foundation of human cognition.

If human beings' native standpoint is not sensuous or mnemonic, but empirical (or metaphorical), Nietzsche's critique of conceptual reason must also be understood and interpreted within the bounds of ἐμπειρία. More specifically, if Nietzsche's liberated standpoint is to be at all conceivable, it must be located somewhere within the narrow margin that separates ἐμπειρία and δοξάζειν — that is to say, it must be able to take its own metaphors *seriously enough* to ensure the world's practical viability, and yet *not as seriously* as natural consciousness usually does. This is indeed a very delicate cognitive compromise, which is once again dependent on how literally or how radically one construes Nietzsche's understanding of basic notions such as truth, untruth, concept, experience, normativity, etc. On the one hand, if human consciousness ceases to regard its own account of reality, if only momentarily and for purely practical purposes, *as though it were real*, the world will become an utterly uninhabitable place. Assuming this extreme existential transformation is even possible, it can only be conceived in a negative way, as Plato's *post-doxastic* beyond, or as Nietzsche's *overhuman* beyond. On the other

hand, if human consciousness is led to believe its own creations as though they were real *in themselves*, it is bound to relapse into the doxastic imprisonment Nietzsche is criticizing.

Neither of these alternatives is compatible with Nietzsche's genealogical strategy. If his claims regarding ancient reason's increased freedom or malleability are to be taken seriously, there has to be a middle term between normative rationality and utter irrationality. His whole argument hinges therefore on the possibility of a standpoint that is aware of itself as a metaphorical standpoint, but still able to lead a practical life.

This is the compromise Nietzsche has in mind when contrasting the so-called 'man of reason' (*der vernünftige Mensch*) and 'man of intuition' (*der intuitive Mensch*).[591] Whereas the latter answers to the scientific imperatives of coherence and objectivity, and tends to forget himself as a creative subject, the latter is fully aware of his role as creator of truth and meaning. He also relies on empirical concepts and observations, but his attachment to objectivity is of a different nature. His idea of value is not confined to the alternative between truth and error, or to the contrast between a world of appearance and a world in itself. Concepts are for him mere tools, which can be used at certain occasions and tossed away in others. They bring out certain aspects of reality, just as certain colours or sounds are used by artists to produce certain visual or acoustic effects. Inasmuch as their value is irreducible to a logical or objective criterion, they are the opposite of science's cool abstractions: their meaning is as varied and incoherent as life itself, and although their infinite palette also includes scientific reason, the latter is but one amid a multitude of different cognitive hues.

This freedom is unknown to the man of reason, for whom metaphors are no longer lived or recognized as such. Nevertheless, the aesthetic foundations of human experience are still available to him in the form of art and artistic creation, where 'the intellect, that master of pretence, is free and absolved of its usual slavery.'[592] Through art, scientific men are offered a glimpse of their original identity, and led to abandon, if only momentarily, their usual dogmatism.

[591] Cf. *KSA* I, 886ff.
[592] *KSA* I, 888; *BT*, 151

> The intellect has now [through artistic creation] cast off the mark of servitude; whereas it normally labours, with dull-spirited industry, to show to some poor individual who lusts after life the road and the tools he needs, and rides out in search of spoils and booty for its master, here the intellect has become the master itself and is permitted to wipe the expression of neediness from its face. Whatever the intellect now does, all of it, compared with what it did before, bears the mark of pretence [*Verstellung*], just as what it did before bore the mark of distortion [*Verzerrung*].[593]

The key to understanding Nietzsche's point lies precisely in the difference between the voluntary *pretence* of artistic creation and the involuntary *distortion* usually involved in empirical conceptualization. While both cognitive attitudes are grounded in a thoroughly subjective procedure, only the first one is able to acknowledge and take advantage of the true power of human imagination.

> That vast assembly of beams and boards to which needy man clings, thereby saving himself on his journey through life, is used by the liberated intellect as a mere climbing frame and plaything on which to perform its most reckless tricks; and when it smashes this framework, jumbles it up and ironically re-assembles it, pairing the most unlike things and dividing those things which are closest to one another, it reveals the fact that it does not require those makeshift aids of neediness, and that it is now guided, not by concepts but by intuitions.[594]

But what does a life guided by metaphor and intuition look like? How does it differ, in practical terms, from the lives of modern, 'needy' men? The answer, according to Nietzsche, can be found in human history. The 'intuitive man' — viz. the 'free man', the 'artistic man', the 'liberated intellect', etc. — is not just a conceptual or philosophical anti-type. He is the product of a lost cultural era, whose cognitive freedom is no longer intelligible or within reach. This lost world is the world of Zarathustra, where reason and imagination, earnestness and playfulness, lucidity and madness went hand in hand. But it is also the world of ancient Greek mythology, where art and imagination used to rule over objective truth, and where each word and gesture were accompanied 'by pretence, by the denial of neediness, by the radiance of metaphorical visions, and indeed generally

[593] *KSA* I, 888; *BT*, 152
[594] Ibid.

by the immediacy of deception.'[595] To quote another of Nietzsche's well-known formulas, the 'Greeks were superficial — *out of profundity!*'[596] They drifted from logic to myth with the ease of artists, and this ingenuity was also visible in their unique cosmology: unlike the world of modern men, entirely made up of 'things in themselves', the Greek world was peopled with gods and other fantastic beings, subject to magical influences and hidden forces. It was a not a mere depository or objective truths, but the living mirror of human passions, fears and aspirations.

The contrast between the ancient Greek world and the modern world is a recurring Nietzschean motif, explored in different ways throughout many different works. It plays a particularly important role in Nietzsche's early writings concerning Greek tragedy, where the origin of classical theatre is famously located in the historical synthesis of the so-called Dyonisian and Apollonian principles.[597] In his 1886 preface to *The Birth of Tragedy*, Nietzsche points out that his analysis of Greek tragedy is already an attempt to solve 'the *problem of science* itself' (*das Problem der Wissenschaft selbst*). In it, he claims, science is 'grasped for the first time as something problematic and questionable.'[598] And the main novelty of his approach lies in his adoption of an artistic, rather than a scientific or epistemological approach to the problem of science. He has realized, better than anyone before him, that 'the problem of science cannot be recognized within the territory of science.'[599]

[595] *KSA* I, 889; *BT*, 152f. In these and similar passages, Nietzsche claims a) that ancient Greek men were naturally artistic beings, who favoured the *aesthetic* over the *concrete* aspects of life; and b) that the kind of ideas and values cherished by ancient Greek society can no longer be understood or shared by modern men. Both of these ideas are familiar motifs of German Romanticism's revival of ancient Greek culture, poignantly conveyed by Schiller in his 1788 poem *Die Götter Griechenlandes* (*Sämtliche Werke* II, 673-676):

Wie ganz anders, anders war es da!
[…]
Da der Dichtkunst mahlerische Hülle
Sich noch lieblich um die Wahrheit wand! —
Durch die Schöpfung floß da Lebensfülle,
Und, was nie empfinden wird, empfand.

[596] *KSA* III, 352; *GS*, 9

[597] This group of works includes *Zwei öffentliche Vorträge über die griechische Tragödie* (1870), *Die dionysische Weltanschauung* (1870), *Die Geburt des tragischen Gedankens* (1870), *Sokrates und die griechische Tragödie* (1871) and the well-known *Die Geburt der Tragödie* (1872, republished in 1886), which incorporates material from many of the previously mentioned texts.

[598] *KSA* I, 13; *BT*, 4f.

[599] *KSA* I, 13; *BT*, 5

Nietzsche proposes to carry out two main tasks, namely 'looking at science through the prism of the artist' and 'looking at art through the prism of life.'[600] The first of these tasks is also attempted, as we have just seen, in his 1873 opuscule concerning 'truth and lying in a non-moral sense'. Indeed, both texts display important similarities: in the *Birth of Tragedy*, Nietzsche also draws a clear distinction between the aesthetic freedom of ancient Greek culture and the scientific objectivism of modern culture; he also opposes a primeval 'kingdom of wisdom', where fables and metaphors are still cherished and taken seriously, to the 'logical despotism' of conceptual reason. In addition, he also contrasts the freedom and the sovereignty of the Greek standpoint to the slavery and the servitude of the scientific attitude.

The work's second task, however, adds a new element to Nietzsche's narrative, which is only briefly addressed in the second section of the 1873 opuscule, left unfinished.[601] In addition to the *epistemological* issue of 'where the will to truth comes from', discussed throughout the first section, Nietzsche also focuses on the *existential* issue of 'whether life is worth living', and on the role played by truth and untruth in this debate. The importance of this new theme is also reiterated in the 1886 preface, where the whole text is said to be about the 'value of existence', and in the new subtitle which accompanies the work's second edition: 'Hellenism and Pessimism'.

But how does the contrast between optimism and pessimism — or, as Nietzsche also puts is, between the *affirmation* and the *negation* of life — relates to our previous considerations regarding truth, freedom, and metaphor? We have seen that the starting point of human cognition is, according to Nietzsche, a transcendental 'drive to form metaphors'. Each human being, faced with the burden of giving shape and meaning to his or her own subjectivity, is led to translate his inner life into a palpable existential setting. The human intellect 'makes appearance appear' and endows it with a specific shape, a specific texture, a specific structure, a specific depth, etc. Moreover, we have also seen that this original drive can result, according to Nietzsche, in different degrees of freedom, depending on the extent to which one is aware of oneself as an artistically

[600] *KSA* I, 14; *BT*, 5

[601] See *KSA* I, 886ff. The 1873 opuscule breaks off when Nietzsche, after having focused on the *epistemological* implications of the contrast between metaphorical and conceptual reason, starts detailing its properly *existential* implications — and reiterating one of the *Birth of Tragedy*'s central arguments.

creative subject. Whereas the 'man of intuition' avails himself of all kinds of cognitive resources, and is thus able to express his subjectivity more freely and more completely, the 'man of reason' is locked within the 'prison walls' of objectivity.

Near the end of the *Truth and Lying* essay, Nietzsche goes on to argue that these two cognitive approaches are grounded in two different existential attitudes. Both the man of reason and the man of intuition desire to 'rule over life', but they attempt to do so in different ways. The former is a diligent, conscientious and thoroughly obedient creature. Like the ant in Aesop's fable, the rational man's only concern is to provide for the future and to guarantee a safe and regular life. He puts his faith in empirical concepts and abstractions in order to avoid pain and misfortune, and his efforts usually pay off. However, his obedience is unable to yield him happiness. It is only a means of defence, or a way of minimizing the uncertainty of an otherwise unbearable life. In direct contrast to this attitude, the intuitive man is careless and exuberant. He is not an intellectual creature, but an aesthetic one, and his whole life in lived in pursuit of joyous visions and charming effects. He 'only acknowledges life as real when it is disguised as beauty and appearance', and 'reaps directly from his intuitions not just protection from harm but also a constant stream of brightness, a lightening of the spirit, redemption and release.'[602] Accordingly, he suffers more and more severely than the rational man, as he is less able or less willing to learn from his past misfortunes. But his suffering is as just as unreasonable as his pleasure, and just as free from the constraints of objective truth.

The *Birth of Tragedy* is built around a similar motif, but Nietzsche's main argument is more clearly grounded in a pessimistic diagnosis of human life. Inspired by Schopenhauer, he assumes from the beginning that life is *not* worth living, but that it can be 'salvaged', 'redeemed' or 'justified' as an *aesthetic* phenomenon. In his eyes, this discovery is the greatest achievement of ancient Greek culture, and indeed of any other human culture. Because of their uniquely artistic nature, ancient Greek men were also uniquely sensitive creatures, and hence uniquely aware of the horrors of human existence. Their entire worldview, Nietzsche argues, can be interpreted as a reaction to the so-called

[602] *KSA* I, 886ff; *BT*, 152f.

'wisdom of Silenus', that is, to the acknowledgement that being born is not a blessing, but a burden, and that being alive is ultimately worse that *not to be* at all.[603]

In order to cope with this truth, the Greeks took refuge in a realm of disguise and appearance: they inoculated the world's primitive hostility with a *Mittelwelt* of beauty and divinity, filled with sacred beings and sublime shapes. With time, this *Apollonian* instinct for beauty succeeded in casting its 'pictorial veil' over every aspect of human existence, rendering life pleasant and worthwhile. But this love of disguise is utterly irreducible to the idea of *deception* or *distortion*, as it is understood by modern reason. And this point cannot be stressed enough. This veil of beauty is not to be regarded as a curtain drawn over a world in itself, whose objects were somehow *truer* or *more real* than the objects of Greek imagination. For this kind of reading is already contaminated by modern reason's objectivist framework. As Nietzsche continually points out, Greek gods and myths are not false because they are not true, in the objective or scientific sense of the word. They are metaphors *lived as such*, and indeed cherished *because of it*. Unlike modern men, ancient Greek men revelled in the act of pretending. They did not divide the world into real and unreal objects, but into tasteful and distasteful images, into charming and cumbersome metaphors. Their only concern was the extent to which each new image was able to impress and delight, to infuse pleasure and joy.

In like manner, Nietzsche's central contrast between the Apollonian and Dionysian principles must not be construed as an intellectual or scientific contrast. Neither of these stances claims to be truer than the other one, nor to offer an image of the world as it actually is. What is at stake is once again an *aesthetic* opposition, which must be understood from within the metaphorical sphere. The Apollonian element is the element of creativity and pretence, responsible for 'making appearance appear'. The Dionysian principle, on the other hand, stands for the destruction of appearance, and the elimination of the metaphorical plane. Dionysos is the god of intoxication and madness: under his spell, the principle of individuation is threatened and subjectivity itself is liable to disappear.

[603] Cf. *KSA* I, 34ff.

It must be noted, however, that the *Birth of Tragedy*'s initial argument, as summed up in the previous paragraphs, does not correspond to Nietzsche's final word on the relationship between art, truth and life. In his 1886 preface, already mentioned earlier, Nietzsche distances himself from Schopenhauer and his pessimistic worldview. During the more than ten years that separate his initial text from his self-critical preface, he altered his views on optimism and pessimism, and came to regard the latter as a specific symptom of modern reason. In most of his later writings, the grandeur of ancient Greek culture in located precisely in the readiness to accept and embrace life, whereas pessimism and nihilism are construed as the preeminent signs of modern Europe's cultural decadence. This change of strategy requires a reformulation of the *Birth of Tragedy*'s initial diagnosis and the abandonment of the original emphasis on the 'wisdom of Silenus'. And sure enough, in an attempt to harmonise his past and present views, Nietzsche opens his preface with a somewhat factitious distinction between different kinds of pessimism: a 'pessimism of *strength*', indicative of physical health and vitality, and a *weaker* sort of pessimism, associated with Schopenhauer and modern culture as a whole.[604]

The importance of this inconsistency will become more apparent in the next chapters, in the sections dedicated to Nietzsche's views on the relationship between life, health and power. For now, my present purpose is to highlight the continuity between Nietzsche's views on truth and metaphor, as expounded in the *Truth and Lying* essay, and his contrast between the Dyonisian and Apollonian principles. In both cases, the decline of Greek creativity is put to down to a peculiar kind of amnesia, whereby the freedom and fluidity of human imagination is replaced by the imprisonment and rigidity of conceptual reason. But whereas in the first case this movement is put down to time and habit, in the second case it is documented in a different and more detailed manner.

It all comes down, once again, to the peculiar balance between Apollo and Dionysos. Although these two principles represent opposite forces, they are intrinsically dependent on one another. Just as the violent excesses of the Dionysian *Trieb* must be tempered by

[604] Nietzsche returns to this same distinction in *The Gay Science*, where he contrasts the *romantic pessimism* of Schopenhauer and Wagner, born from an 'impoverishment of life', and the *Dionysian pessimism* of ancient Greece, born from a 'superabundance of life'. Cf. *KSA* III, 619ff.

the serenity of the Apollonian *Trieb*, so too the latter is only truly meaningful as a reaction to the former. Dionysos acts a counter-weight to the world of metaphor, reminding the subject-artist of his or her own condition. Its role is similar to that of a thunderstorm, which by repeatedly striking the same house reminds its owners of its structural frailty.

In Nietzsche's own words, Dionysos reveals 'the playful construction and demolition of the world of individuality as an outpouring of primal pleasure and delight.'[605] This central idea is further illustrated by the evocation of the Hercalitean fragment where the power that shapes the world is compared to the action of a child playing on the beach, building up piles of sand only to knock them down again.[606] What is important in these images is the fact that *both* moments, creation and destruction, are equally essential, and intrinsically related to each other. If the creative moment were ignored, subjectivity would be eliminated and Dionysos would reign supreme; conversely, if the destructive moment would be neglected, the metaphorical world would grow rigid and out of proportion. Whereas the first of these alternatives might correspond to some embryonal stage of human reason, wherein subjectivity was not yet fully formed, the second one is precisely what took place, according to Nietzsche, in the years that predated the demise of ancient Greek culture.

The course of the argument is well known: as Greek art became more intellectualised, the Apollonian element grew increasingly prominent. In the case of tragedy, it was finally chased by Euripides chased from the theatre stage.[607] But what happened to tragedy was merely a symptom of a much wider transformation, which struck the very heart of Greek culture. Free from the Dyonisian menace, human metaphors grew higher and sturdier. Piles of sand were replaced by wooden towers, by large stone bulwarks — and soon the world of metaphor was no longer recognizable. Apollonian visions crystallized into cold abstractions and aesthetic freedom gave way to epistemic truth. The artistic drive to impress and delight was replaced by the scientific wish to pierce the veil of illusion and to reveal the world *as it actually is*.

[605] *KSA* I, 30; *BT*, 18
[606] Ibid. The fragment in question is number 52 in the Diels-Kranz edition. Cf. also *KSA* V, 323f.
[607] See esp. *KSA* I, 74ff.

Nietzsche's historical description is now more precise. The death of metaphor is linked to a specific cultural event, brought about at a specific juncture of human history. Although the demise of tragedy, poetry and Greek art in general was a gradual process, it was not due simply to time and habit. In this new account, the birth of scientific reason coincides with the rise of one man — Socrates — and the endorsement of a new form of culture — Socratic culture.

11.

PHILOSOPHY IN CHAINS

And then Socrates came along — and his influence 'spread out across all posterity to this very day, and indeed into the whole future, like a shadow growing longer in the evening sun.'[608] Anyone who has but a superficial acquaintance with Nietzsche's works will know how important a role Socrates plays in them. Whether merged with Plato or standing on his own, he is spoken of with admiration and scorn, with awe and resentment, with fascination and disgust. Moreover, he is usually the object of a variety of different and sometimes contradictory reproaches: he is called a demon and a slave, a genius and a charlatan, a naïve optimist and a conceited despot. But in neither of these roles is the importance of his ideas any less remarkable.

For Nietzsche, Socrates and his followers are responsible for the demise of ancient Greek culture, grounded in a purely aesthetic conception of human knowledge, and for the birth of the modern European world, driven by a new and unprecedented faith in coherence, verisimilitude and accuracy. Whereas ancient Greek wisdom was both constructive and destructive, grave and gratuitous, Apollonian and Dionysian, Socratic wisdom is no longer fuelled by this antagonism. It is the result of Apollo's triumph over Dionysos and the beginning of truth's historical and philosophical empire. In this new cultural environment, spontaneity and arbitrariness are no longer encouraged or admired. The logical element takes hold of human imagination and metaphors are turned into

[608] *KSA* I, 97; *BT*, 71

concepts, precepts and laws. Reason is no longer a vehicle for happiness and delight, but a measuring instrument placed above human life.

Socrates is the first and most extreme exponent of this new cognitive attitude. He regards the attainment of true knowledge as the highest of human conquests and spends his entire life pursuing it. In his eyes, truth and value are one and the same thing: whatever is to count as beautiful, pleasant, just or noble must be submitted to a logical examination, and proved to be rational. What is more, life itself must be subjected to this critical exam: Socrates' main concern is no longer whether life is pleasant or captivating, but whether it makes sense. His philosophical ideal is a life free from contradiction and self-deception; a life where every choice is the result of a truly ration judgement.

11.1. *The Socratic turn*

In Nietzsche's writings, the revolution promoted by Socrates is often illustrated by the analogy between living and dreaming, or the idea that human existence amounts to an *oneiric* or *dream-like* existence.[609] This image is also a well-known Platonic motif, and the coincidence is certainly not accidental: its philosophical implications are closely related, once again, to the 'problem of science', which unites and separates the Socratic and the Nietzschean conceptions of truth.

Whereas ancient Greek life, under the aegis of Apollo, amounts to a long and blissful dream, where pain and suffering are redeemed by the transfiguring power of art, the arrival of Socrates is a moment of awakening. Apollonian men revel in their own inventions, like dreamers who know they are dreaming and wish to go on dreaming, but Socrates refuses this self-indulgence. Turning his 'great Cyclopian eye' on this fantastic dreamworld — an eye 'in which the lovely madness of artistic enthusiasm never glowed'[610] — he subjects its fabrications to a new and cruder kind of examination. With a few stern words, images turn into concepts, their truthfulness is challenged and their spell

[609] See notably *KSA* I, 38ff., 760, 886ff.
[610] *KSA* I, 92; *BT*, 67

is finally broken: magic and myth lose their power, gods and nymphs disappear into thin air; the whole edifice of human invention comes under threat. Its charm is lost on the man of reason, for whom clarity and utility are the only concerns.

At first sight, this destruction seems justified. Beauty was sacrificed, yes, but in the interest of truth and knowledge. A world of illusion has been replaced by a more real world, where ideas and judgements are closer to the actual being of things. But as personal experience usually shows, what is peculiar about dreams is precisely the difficulty in distinguishing what is real from what is not, what belongs to the oneiric sphere and what does not. In most dreams, the dreamer is convinced he is awake, and only by actually waking up does he recognize he his mistaken. Dreams are unsettling precisely because they mimic the feeling of awakeness that is typical of one's real, non-oneiric existence. And therefore, even when the dreamer feels he has awoken, this feeling may amount simply to a new and more sophisticated form of dreaming.[611]

In the initial sections of the present work, when discussing Plato's scientific project, we have come across a very similar phenomenon. We have seen that Socrates' denouncement of the doxastic nature of human knowledge does not amount to replacing a false account of reality with a true one. His discovery is a negative discovery, just like that of a dreamer who suddenly realizes he is dreaming. Socrates highlights the fact that none of the ideas and words we resort to in our daily life are as reliable or as intelligible as we take them to be. Their usual familiarity is an illusion whose spell, once it is recognized as such, loses its power. The experience felt throughout Plato's dialogues is therefore one of *awakening*. However — and herein lies the main point — it does not lead to an *actual* awakening, but merely to an awakening *to the idea that one is dreaming*. As soon as one realizes that one's account of reality is not real, but illusory, one is faced with a different, radically transfigured version of the same oneiric world. Gods and nymphs, but also everyday objects and ideas are suddenly rendered opaque and awkward, like the props of a theatre play that have lost their dramatic purpose. In Plato's words, this unusual

[611] On the ambiguity between dreaming and being awake, see, for ex. *KSA* I, 887; *KSA* V, 97, or 114f.

experience exposes the 'long day' of human existence as a 'nocturnal day', caught between sunlight and darkness.[612]

The world unveiled by Socrates is at best a 'middle world', located somewhere between waking and sleeping. It is the old familiar world devoid of its familiar meaning — a strange and unwelcoming world, where practical freedom is no longer automatic. In this new cognitive regime, our usual confidence in the ability to interpret reality is suddenly compromised, along with the power to choose and to anticipate the consequences of our choices. We are left as strangers in a strange land, or as fish out of water, in need of the vital oxygen provided by dreams.[613] And our main concern is thus to escape this confrontation and to go back to sleep.

Despite the obvious parallel between Plato's and Nietzsche's images, there are at least two major differences between them. And to grasp their meaning is to grasp the main divergence between what I have named Plato's scientific project and Nietzsche's metacognitive project.

Firstly, although Socrates' undertaking is a negative undertaking, he seems to think that it is nonetheless a *necessary* one. Or rather, although his critique is not a sufficient condition for the actual liberation from the dreamworld of human reason, it is nonetheless a necessary condition for a *positive* or *post-critical* transformation ever to come about. This belief can be expressed by one of Albert Camus' best-known statements, offered in a very different context, but with the same idea in mind: 'to accept the absurdity of everything around us is one stage, a necessary experience: it should not become an impasse. It arouses a revolt that may become fruitful.'[614] Translated into Platonic language, this means that the only way out of the doxastic realm is the 'cleansing' or 'purification' (καθαρμός) prescribed in the *Sophist* and the *Republic*.[615] To be sure, exposing the spuriousness of human beings' cognitive self-confidence will not automatically provide them with true knowledge. However, it will somehow bring them closer to it — or, in any case, less distant from it. As Socrates points out, in a passage already quoted above, the awakening proposed by philosophy can lead to one of two

[612] Cf. *Re* 521c
[613] Cf. *Phd* 109a-110a
[614] *Essais*, ed. cit., 1425
[615] Cf. *So* 230b-c and *Re* 533c-d. See above 5.2.

results: 'either we shall find what we are after [namely, true knowledge], or we shall be less inclined to think we know what we do not know at all; and surely even that would be a recompense not to be despised.'[616] The man who is awoken by philosophy is better off in either case, for he is *less imprisoned* or *less deluded* than the ordinary dreamer.

For Nietzsche, however, there is nothing to gain from this transformation. On the contrary. Since every human thought is grounded in a metaphor, human reason as such is intrinsically oneiric, and Socrates' critique only succeeds in replacing one kind of dream with another. The formal parallel between the two positions is clear: just as Socrates had sought to expose human beings' usual standpoint as a dream-like form of awakenness, Nietzsche aims to expose Socrates' shrewdness as a dream-like form of awakenness. But his point is not simply that all of us are dreamers, *no matter what*. Although human reason is indeed universally oneiric, not all dreams are alike. And while the first, pre-conceptual or pre-Socratic kind of dreaming is free and life-promoting, the second, conceptual or Socratic one, is slavish and destructive. Whereas Socrates' oneiric awakenness leads to a fruitless standstill, whereby life itself is called into question, Apollonian dreaming is smooth and fertile. It is not a rational phenomenon, but an artistic one, and its fruits are not conditional on a painful or laborious quest. Unlike the philosophical dreamer, the artistic dreamer hangs on to a ceaseless and ever-changing illusion, with no need or interest to let go.

> 'Leave him hanging', cries *art*. 'Wake him up', cries the philosopher, in the pathos of truth. But he himself, convinced that he is shaking the sleeper, falls into an even deeper magical sleep — and dreams perhaps of 'ideas' or immortality. Art is more powerful than knowledge, for the former wills life, while the latter achieves as its final goal only — destruction. —[617]

This antithesis leads to a further point of contention between the Platonic and the Nietzschean projects, related not only with what Nietzsche termed the 'will to truth', but also with what he termed the 'will to illusion', or the 'will to deception'. As both Plato and Nietzsche recognize, Socrates' standpoint is an exceptional standpoint. Its quest for

[616] *Th* 187c1-3
[617] *KSA* I, 760

truth is too radical, too exacting, and too contrary to life's practical demands. Consequently, human beings are hardly ever *awake*, in the Socratic sense of the world. Their usual state is an *oneiric* state, and their adherence to Socrates' ideas is always brief and reticent, like that of 'people awakened from a nap.'[618]

Yet although both ancient and modern men are natural born dreamers, the latter's dreams are very different from the former's. Whereas Apollonian men regard life itself as a dream, and human cognition as a globally metaphorical process, modern dreams are already spoiled by the earnestness of Socrates' quest. Ancient men do not long to be awoken because for them illusion is *itself* the truth. There is nothing beyond the world of dreams because the world itself is a dream: its truth is the beauty offered to the dreamer and its rules are the magic and incoherent rules of dreams.[619]

Ancient men 'knew how to dream and didn't even need to fall asleep first.'[620] Modern men no longer possess this ability. After being awoken by Socrates, and thereby losing touch with life's metaphorical or oneiric roots, they went back to sleep. But their sleep is heavier than that of ancient men: their dreams are no longer self-conscious, or *creative* dreams, but involuntary, or *conceptual* ones. Modern men are no longer aware of their dreaming condition. They mistake the world of dreams for a world in itself, whose meaning is a matter of truth, order and intelligibility. Their artistic powers are no longer free or spontaneous and illusion is no longer lived as illusion, but as the mere opposite of truth.[621]

In this new, self-deluded existence, dreaming itself is only tolerated as a state of exception, as a kind of insanity opposed to the seriousness of concrete everyday life. According to Nietzsche, however, it is one of the few bridges that still unites modern men with their former or original selves. Along with sudden outbursts of passion, bouts of madness, drunkenness or the ecstasy of artistic creation and enjoyment, dreams rank

[618] *Ap* 31a4

[619] See, for ex., *KSA* I, 26ff., 38, 44, 64, 553ff.

[620] *KSA* III, 423; *GS*, 70

[621] On the contrast between ancient and modern dreaming, see, for ex., *KSA* IV, 154; *KSA* III, 416f., or 495: '*Die grösste Veränderung.* — Die Beleuchtung und die Farben aller Dinge haben sich verändert! Wir verstehen nicht mehr ganz, wie die alten Menschen das Nächste und Häufigste empfanden, — zum Beispiel den Tag und das Wachen: dadurch, dass die Alten an Träume glaubten, hatte das wache Leben andere Lichter.'

among the few experiences that are still capable of unweaving reason's conceptual web and tearing down the 'prison walls' of objectivity.[622]

The dialectic of dream and reality is another important element of Nietzsche's overall characterization of the relationship between life, truth and freedom. In particular, Nietzsche's reinterpretation of the Platonic dialectic of dream and reality is another means of challenging Socrates' scientific revolution and his own views on the relationship between life, truth and freedom. As the previous discussion has shown, the different degrees of awakenness distinguished by Nietzsche correspond to different degrees of cognitive freedom, that is, to different degrees of practical and intellectual autonomy. Accordingly, the *oneiric* imagery employed throughout the previous discussion is akin to the *carceral* imagery employed, for instance, in Plato's allegory of the cave. In both cases, what is at stake is the assessment of human freedom 'as regards both education and the lack of it.' And in both cases the varying quality of human beings' outlook on reality translate into very different existential scenarios.

In Socrates' version of the argument, human beings' usual standpoint is an imprisoned standpoint. However, it is also a sleeping standpoint, whose understanding of reality resembles that of a dreamer. Philosophy's role is to awake the dreamer and to confront him with his imprisoned condition. This radical form of ἀναγνώρισις is precisely what takes place inside Plato's cave. Once the prisoners are allowed to turn their heads, they realize that their outlook on reality is seriously limited: what they used to regard as the whole world is but a small portion of it, and what they took to be things in themselves are mere shadows of unknown objects. With this transformation, the prisoners are disabused of their former convictions and left more awake than before and more capable of assessing their practical situation. But this awakening is not yet the end of human dreaming, nor the elimination of human captivity. It is merely the first, *negative* step of a wider philosophical revolution, whose intended result is an actual awakening, or an actual departure from Plato's cave.

In Nietzsche's version of the argument, human beings' original standpoint is a free standpoint. It is not free to choose or to decide, to interpret or to react to the

[622] On the idea that dreams provide an insight into humankind's aesthetic or oneiric infancy, see, for ex., *KSA* II, 31ff.; *KSA* III, 226, 423.

surrounding world, but it is not yet doomed to think in terms of true and false concepts, of accurate and inaccurate versions of reality. It is also a dreaming standpoint, but its dreams are not opposed to reality as appearance is opposed to truth. With the rise of 'Socratism', this freedom is gradually compromised and primitive consciousness is forced into the epistemic framework of modern reason. For Nietzsche, the awakening to the fact that one is unable to grasp reality *as it really is* does not represent a step towards the liberation of human consciousness, but the first stage of humankind's self-imprisonment. In his eyes, men and women are not originally imprisoned: they have *become* so due to an exaggerated and unhealthy fascination with truth. Socrates, as the main promoter of this new cognitive trend, is not the liberator, but the jailer of mankind: he has helped tie up human consciousness, only to lament afterwards its inability to break free.

Since this transformation is about freedom and autonomy, its effects are not limited to the domains of logic or epistemology. Nietzsche regards the rise of Socratism as the starting point of a philosophical movement with aesthetic, moral and political implications; a movement whose underlying principles have gained strength throughout the centuries and whose far-reaching influence is ultimately responsible for the cultural crisis of modern Europe.

11.2. *Athens and Jerusalem*

As is well-known, Nietzsche paid special attention to what he considered to be the moral, ethical, or religious implications of Socrates' and Plato's scientific revolution. In this new intellectual attitude he saw not only the origin of the modern distinction between a phenomenal and a noumenal world, but also the seed of a metaphysical tendency to endow truth with an eternal or transcendent value. Once reason was subjected to Socrates' impossible demands, its natural limitations came to the fore: human ideas were reduced to mere δόξαι, that is, to more or less imperfect portrayals of a mysterious and ever-fleeting *world in itself*; as a result, the notion of truth was detached from the alleged truthfulness of this or that human belief and placed in a higher and inaccessible plane. In like manner, once the value of human choices was made to depend on a rational and intelligible design, life itself was soon found lacking. Human existence emerged as a

faulty or downgraded existence and its contradictions led once more to the positing of a truer or higher form of existence. This shift from a subjective *here* to an objective *beyond* represents, for Nietzsche, a form of self-imprisonment or self-limitation. More than that, it constitutes the basis of a *slavish* form of cognition, directly opposed to human reason's original sovereignty. Socratic men are not in charge of their thoughts or judgements, nor are they able to create their own values. Their entire world is shaped and sustained by the constant adaptation to external forms of valuation, and this dependence amounts to a form of weakness: it betrays the will to be 'ruled' or 'commanded', to be a creature rather than a creator.[623]

At the hands of Socrates and his followers, truth became not only an epistemic imperative, but also a *moral* one. The historical and conceptual link between scientific truth and moral truth is a central Nietzschean theme, and the key to understanding Nietzsche's insistence on the relationship between Socrates, Plato and the Jewish and Christian traditions. In his 1886 preface to the *Birth of Tragedy*, Nietzsche suggests that his first book, which offered his first approach to the 'problem of science', is already dedicated, albeit indirectly, to the issue of morality. More specifically, it already puts forth a philosophy set 'against the *moral* interpretation and significance of existence.' His aesthetic characterization of ancient Greek wisdom allows only for an 'artist's meaning behind all that happens': 'a "god", if you will, but certainly only an utterly unscrupulous and amoral artist-god who frees himself [*der sich löst*] from the dire pressure of fullness and over-fullness, … and who wishes to become conscious of his autarchic power and constant delight and desire, whether he is building or destroying, whether acting benignly or malevolently.'[624]

With Socrates, however, this power is denied. Human autarchy is sacrificed to a new divinity, i.e. the notion of truth, and reason acquires for the first time an overtly moral dimension. The parallel between truth and morality constitutes one of the main

[623] *KSA* X, 61: 'Ihr meint, ihr suchtet "die Wahrheit"! Ihr sucht einen Führer und wollt euch gerne commandiren lassen!'; *KSA* III, 583: 'Wo ein Mensch zu der Grundüberzeugung kommt, dass ihm befohlen werden *muss*, wird er "gläubig"; umgekehrt wäre eine Lust und Kraft der Selbstbestimmung, eine *Freiheit* des Willens denkbar, bei der ein Geist jedem Glauben, jedem Wunsch nach Gewissheit den Abschied giebt.'

[624] *KSA* I, 17; *BT*, 8

conceptual bridges between Nietzsche's early writings, primarily focused on the value of truth (and the possibility of an aesthetic redemption of scientific reason) and his later writings, increasingly concerned with Jewish and Christian morality (and the possibility of a complete 'transvaluation' of moral values). Nietzsche's interpretation of the 'will to truth' as a religious or ascetic symptom marks a shift of direction from Athens to Jerusalem, from his early Dionysian worldview to his fierce critique of Christian morality.

As Nietzsche writes in the *Genealogy of Morality*, 'there is no 'presuppositionless' knowledge, the thought of such a thing is unthinkable, paralogical: a philosophy, a 'faith' always has to be there first, for knowledge to win from it a direction, a meaning, a limit, a method, a right to exist.' Therefore, whoever, like Socrates and Plato, 'tries to place philosophy "on a strictly scientific foundation", must first stand on its head not just philosophy, but also truth itself.'[625] Although Plato's scientific project aims to identify and suspend all of reason's dogmatic assumptions, it is itself supported by a dogmatic faith in the referential value of truth. In this respect, its formal structure is similar to that of many other moral or religious systems: by ranking every idea or judgement below the stage of ἐπιστήμη, Socratism takes the form of a negative theology, where the value of human beliefs is determined by their degree of imperfection. In this negative scheme, ideas and judgements are mere *imagines scientiae*, bound to emulate a truth they are unable to convey, and philosophy's task is to raise these images from the negative terrain of self-conceit (viz. *dogmatic* δοξάζειν, or οἴεσθαι εἰδέναι ἃ οὐκ οἶδεν) to the neutral terrain of ignorance (viz. *philosophical* δοξάζειν, or οἴεσθαι οὐκ εἰδέναι), and from then on to the positive terrain of science.

This elevation requires, therefore, that everything whatsoever be tested with regard to its scientific worth. Reality can no longer be deemed beautiful, or just, or pleasurable, or useful: it must be shown to be such. Contradiction and error are no longer allowed and the world is divested of its usual reliability. Wherever Socratism 'directs its probing gaze, it sees a lack of insight and the power of delusion, and it concludes from this lack

[625] *KSA* V, 400; *GM*, 112

that what exists is inwardly wrong and objectionable.'[626] As a consequence, the phenomenal world is reduced to the more or less distant or distorted echo of a scientific or noumenal world, concealed behind or beyond the doxastic realm.

> …Those who are truthful in that audacious and ultimate sense which faith in science presupposes *thereby affirm another world* than that of life, nature, and history; and insofar as they affirm this "other world", must they not by the same token deny its counterpart, this world, *our* world? […] It is still a *metaphysical faith* upon which our faith in science rests — that even we knowers of today, we godless anti-metaphysicians, still take *our* fire, too, from the flame lit by the thousand-year old faith, the Christian faith which was also Plato's faith, that God is truth; that truth is divine…[627]

Nietzsche's engagement with Plato's philosophy is usually broad-scoped, and rarely comes down to the actual argumentative detail of his dialogues. Plato and Platonism are often used as general historical labels, which tend to downplay the complexity and heterogeneity of the Platonic corpus. In this particular point, Nietzsche's attitude is hardly an exception. Nevertheless, the previous passage should not be too quickly construed as an example of the 'two-world reading' usually associated with Plato and his so-called Theory of Ideas.[628] Nietzsche's aim, here and elsewhere, is to show that a scientific project such as Plato's, grounded in a wholesale criticism of phenomenal reality, is naturally open to this kind of interpretation. More than that, Nietzsche points out that Plato's philosophy has indeed been interpreted in this way, as it constituted *de facto* one the main philosophical sources of Christian thought. And here, once again, his argument is not particularly original. Insofar as Christianity is based on a two-world cosmology, which echoes the fundamental contrast between an earthly and a divine standpoint, it does amount to a historical development of Platonism — a 'Platonism for the "people"'[629], elevated to a world religion.

But this dualism is not the only parallel between the Platonic and the Christian traditions. In the preface to *Beyond Good and Evil*, Nietzsche credits Plato not only with

[626] *KSA* I, 89; *BT*, 66. The original reads: '[Er] schliesst auf diesem Mangel auf die innerliche Verkehrtheit und Verwerflichkeit des Vorhandenen.'

[627] *KSA* III, 577; *GS*, 201

[628] See above section 4.1.

[629] *KSA* V, 12; *BGE*, 4

having invented 'the pure spirit' (*der reine Geist*), but also 'the good in itself' (*das Gute an sich*), thereby paving the way for 'millennia of Christian-ecclesiastic pressure.' [630] Concerning this second claim, we have seen earlier, when discussing the main features of Plato's diagnosis of human cognition, that the Platonic notion of good is primarily a *formal* notion, placed at the top of reason's conceptual alphabet.[631] Since human beings are by nature self-interested beings, human life is continually oriented towards what is perceived as the greatest or highest good available. Whether in a conscious or unconscious manner, all human judgements and choices are endorsed in view of what is deemed, at each given moment, the best or more advantageous alternative or course of action. But although 'the good', in this generic sense, is indeed 'the end of all our actions'[632], it is still a formal concept, liable to accommodate an infinity of different contents. Although every one of us aims at the greatest good, not all of us agree on what the greatest good actually is, and this disagreement is precisely what fuels many of Plato's dialogues.

Socrates' enquiries are often debates about what is *best* — namely, what is the best course of life, what is the best conduct in a given situation, what is the best kind of education, and so on. More specifically, the enquiries often revolve around whether it is best, for instance, to be just or unjust, to pursue philosophy or rhetoric, to embrace a life of pleasure or one of discipline, to take on a lover or to refrain from doing so, etc. In these disputes, Socrates' usual role is that of a referee: his first aim is to show that the answers to these queries are always specific declensions of τὸ μέγιστον ἀγαθόν, that is, specific *versions* of what the greatest good, in these situations, actually is; his second aim is to show that these versions, when subjected to a closer examination, are usually problematic. When questioned about their reasons for endorsing this rather than that alternative or course of action, Socrates' interlocutors are led to contradict themselves and to lose faith in their initial stances. Their idea of what is good and bad, desirable or undesirable, advantageous or disadvantageous, etc., is found to rest on a series of shaky and ultimately unintelligible assumptions.

[630] *KSA* V, 12
[631] See above section 1.3., I
[632] *Go* 499e8-9

What emerges from this procedure is once again the need for a philosophical re-education of natural consciousness. Even though human life is naturally oriented towards the good, one's idea of good is usually misguided and self-contradictory. As a consequence, what is perceived as good or advantageous may turn out to be neither, and lead to evil or harmful consequences. For the good to be actually beneficial, it must be accounted for *scientifically*: viz. it must be *known* to be good, rather than merely *claimed* or *believed* to be so. In other words, the good must be elevated from the doxastic to the epistemic realm, from οἴεσθαι εἰδέναι to actual εἰδέναι. And this elevation is philosophy's most important task: the continual effort to correct the discrepancy between the version of good adopted by consciousness and the good itself.

In Nietzsche's words, 'Plato wanted to use all his strength (the greatest strength a philosopher had ever had at his disposal!) to prove to himself that reason and the instincts converge independently on a single goal, on the Good, or "God"; and, ever since Plato, all theologians and philosophers have been on the same track.'[633] Yet insofar as the Platonic idea of good is a formal idea, it is still very far from the Christian idea of good and the more concrete imperatives of love, piety or charity. According to this line of interpretation, τὸ ἀγαθόν is above all a riddle, lacking a definitive solution, and Platonic morality is likewise an interrogative, rather than an affirmative or dogmatic system. It is primarily aimed at exposing the spuriousness of natural or doxastic morality, and introducing the possibility of a true or scientific form of morality. The transition from Plato to Christian theology is therefore not as inevitable or as straightforward as it might appear. For the Christian believer, the greatest good is no longer a question, but a certainty; although its true face may remain hidden, unfathomable, or outside human reach, it is nonetheless irreducible to a dialectical or philosophical demonstration. Whereas for Socrates and Plato intelligibility and coherence are basic conditions of truth, Christianity subordinates both of these requirements to the immediacy of religious faith. For the Christian, one must 'believe in order to understand'[634]: truth is not a rational or scientific conquest, but the content of a divine revelation, open to philosophers and non-philosophers alike.

[633] *KSA* V, 112; *BGE*, 81
[634] Isa 7:9

If Plato's idea of good is construed in a purely formal way, Nietzsche's claim that 'the Christian faith is also Plato's faith' is no better than a sweeping generalization. But this formal approach is not the only possible approach to Platonic morality, nor indeed the one usually favoured by Nietzsche. In the *Gorgias*, for example, as the different contenders try to reach an agreement concerning the best or most advantageous way of life, Socrates strives to show that their proposed solutions — namely, the life of rhetors, the life of despots or the life of hedonists — are not as good or as advantageous as they appear. Yet Socrates' role is not confined to this negative procedure. Apart from his immanent critique of the different views set forth by his interlocutors, he also advances a doctrine of his own. After having argued, against Polus, that the despot is not necessarily powerful, he goes on to claim, quite vehemently, that to do evil is worse than to suffer it, and even worse when the evildoer is left unpunished;[635] and after having argued, against Callicles, that the hedonist is not necessarily happy, he goes on to claim, with no less conviction, that the just man is the happiest of all men, and justice the greatest of all goods.[636] If Plato's aim were purely elenctic, or dialectical, the dialogue would consist simply in a systematic *reductio ad absurdum* of his opponents' standpoints. In the end, the greatest good, or the best way of life, would remain a formal category, lacking a satisfactory or coherent content. The reader would be told what the best way of life *is not* — namely, rhetoric, despotism, or hedonism — and encouraged to go on looking. Besides, if Plato's intention were to expose the doxastic nature of human cognition in a more radical or complete manner — as he does, for example, in the *Sophist*, the *Parmenides* or the *Theaetetus* — his critique would have to move beyond despotism and hedonism, good and evil, and extend to the entire alphabet of human reason. He would be led to establish, in a more general manner, that human beings are 'of no account in respect to wisdom'[637] — either of what is best or, indeed, of *anything else*. In the *Gorgias*, however, Plato's aim appears to be different. He is concerned not only with disproving

[635] *Go* 468e-469b
[636] *Go* 527b-e
[637] *Ap* 23a

his opponents' immoral claims, but also with introducing a moral worldview that is both intelligible and liable to be put into practice.[638]

Plato's departure from the strictly *negative* or *immanent* strategy usually associated with Socrates is also noticeable, for example, in the *Republic*, where the Platonic Socrates is equally categorical about the merits of a just and virtuous life, or in the dialogues' eschatological myths, where the exaltation of virtue, the immortality of the soul and the idea of a Final Judgement play an important part. All of these texts denote a moralizing tendency which outweighs the logical or dialectical element of a pure ἐκβολὴ τῆς δόξης. However, none of these passages is a definitive proof of Plato's moral dogmatism. Indeed, the very form of his dialogues should caution exegetes against such categorical conclusions — either because Plato's supposed seriousness is often hard to distinguish from Socrates' irony, or because what appears to be final in his dialogues is often only provisional, open to new and more subtle forms of refutation. In any case, it is fairly clear that Nietzsche's Plato is not a purely dialectical or elenctic Plato, bent on turning every human belief against itself. Nietzsche's critique of Platonic morality appears to be focused, rather, on the dogmatic attitude featured in the *Republic*, the *Gorgias* or the *Phaedo*.[639] His Plato is first and foremost an *edifying* Plato: the champion of justice and selflessness, of honesty and resignation. He is the main vehicle for the historical and cultural affirmation of a new and less ambiguous brand of dogmatism, to which Nietzsche will dedicate most of his later writings.

[638] The *Gorgias* ends with a practical exhortation, which encapsulates its entire pedagogic programme: 'Let us therefore take as our guide the doctrine now disclosed, which indicates to us that this way of life is best, to live and die in the practice alike of justice and of all other virtues. This then let us follow, and to this invite everyone else.' (*Go* 527e2ff). Of course this exhortation is still merely formal, insofar as the concepts of justice (δικαιοσύνη) and virtue (ἀρετή) in question are not simple or self-evident, but Plato seems to be appealing here to a kind of life that is nonetheless familiar to his readers, albeit hard to put into practice.

[639] In *Beyond Good an Evil*, Nietzsche contrasts Socrates' scepticism with Plato's moral naïvety. See *KSA* V, 112; *BGE*, 81, already quoted above: whereas Socrates 'had seen through to the irrationality of moral judgments', 'Plato ... was more innocent in such matters and lacked Socrates' plebeian craftiness.' His goal was to affirm the divine nature of the Good, thereby anticipating the task of 'all [future] theologians and philosophers.'

11.3. *Masters and Slaves*

As Nietzsche's main philosophical focus shifts from a critique of scientific reason to a critique of morality, his genealogical account of freedom is likewise transformed. In his earlier writings, freedom and slavery are chiefly associated with the contrast between an arbitrary or aesthetic mode of cognition and a logical or conceptual mode of cognition. In the former case, human beings were somehow abler to create their own values, as artists creating a work of art: their life resembled a conscious dream, where illusion was embraced and celebrated as such, and where logical and conceptual oppositions were somehow less binding than they are today. In the latter case, human life was confined to the 'prison walls' of science[640] and subordinated to an increasingly complex web of concepts, rules and regulations. Valuation became a passive, rather than an active process, and the real world was depreciated in favour of a new, *ideal* world, where truth reigns supreme.

To the later Nietzsche, mankind's original or primordial freedom is more directly linked to the notions of authority, strength and power. In the *Genealogy of Morals*, for example, the problem of freedom is once again framed by an evolutionary account of human reason. Primitive men are once more portrayed as naïve and capricious creatures, incapable of the intricate calculations of conceptual reason. They live above all for the present moment, moved by strong feelings rather than coherent or persuasive truths. Their value judgements 'are based on a powerful physicality, a blossoming, rich, even effervescent good health that includes the things needed to maintain it, war, adventure, hunting, dancing, jousting and everything else that contains strong, free, happy action.'[641] However, instead of emphasizing the playful or artistic aspects of this primitive existence, Nietzsche highlights the violence and the cruelty it necessary entails. In this primitive world, life is a permanent struggle and every action is simultaneously a display of strength and domination. Reason is not yet distinguished from the drive to satisfy one's most

[640] See *KSA* I, 883, 760.
[641] *KSA* V, 266; *GM*, 17

basic instincts, to sate one's immediate needs and appetites. Therefore, prediction and calculation are still overshadowed by a primitive scorn for safety and comfort, by a 'shocking cheerfulness and depth of delight in all destruction, in all the debauches of victory and cruelty.'[642]

Yet this savage innocence was short-lived. Just like the reckless artists celebrated in Nietzsche's previous accounts, these wild beings soon found themselves imprisoned within the confines of social life and led to conform to a series of rules and customs. Whereas their original freedom hinged precisely or their utter unaccountability, their new, 'cultured' existence demanded the limitation of their most basic instincts. As their power encountered the stronger power of a community, individual strength and cruelty bowed to measure and reasonableness; the call of nature was bound by the shackles of coherence and peace, of what is expected and of what is allowed. And this transformation led to an important change in the very physiology of human life: instead of a being made for action, to whom every resolution was simultaneously spiritual and physical, men were led to detach the element of thought from the element of action, and to grant it a new independence. They were forced to rely on concepts and inferences, on causes and effects — in a word, on their 'consciousness', the most fragile and undeveloped of all their organs.

This transformation led to a painful self-division. Instincts could no longer be vented freely, but their call was still as audible and as pressing as before. The original drive for power and domination, part of the very essence of physical life, was not suppressed, but merely held in check. And the free animal became a caged animal, torn between his own power and the power of others, between freedom and punishment; trapped in a hostile and reproving world, no longer shaped in his image. As a consequence, his instincts were internalized, forced to find new and secret forms of satisfaction; they were concealed, dematerialized and translated into spiritual phenomena. More importantly, however, this inward movement was accompanied by a feeling of inadequacy and self-repudiation: social censorship became self-censorship, and freedom itself became a source of shame.

[642] *KSA* V, 275; *GM*, 23

> This fool, this prisoner consumed with longing and despair, became the inventor of 'bad conscience'. With it, however, the worst and most insidious illness was introduced, one from which mankind has not yet recovered; man's sickness of man, of himself: as the result of a forcible breach with his animal past, a simultaneous leap and fall into new situations and conditions of existence, a declaration of war against all the old instincts on which, up till then, his strength, pleasure and formidableness had been based.[643]

This pivotal moment is the primary source of the moral worldview that will come to dominate human reason and human life. By turning against their primitive nature, human beings were led to draw a contrast between animal life and an ethereal or spiritualized life, placed on the opposite end of the ontological scale. This impulse of self-denial paved the way for new and increasingly spiritualized notions, such as selflessness, abnegation, asceticism and self-sacrifice. With time, it also laid the ground for the more sophisticated notions of a redeeming God and a transcendent or heavenly existence. For Nietzsche, these ideas are the main forces behind the increasing moralization of the modern world. In his genealogical account, the emergence of bad conscience and all of its derivatives sets the stage for the 'moral sickness' redefined and aggravated by Platonism, Judaism and Christianity.

Nietzsche's account is evocative of a key Hegelian motif, namely the *Phenomenology*'s so-called 'unhappy consciousness', equally torn between the concrete element of natural existence and the ideal element of Spirit. Nietzsche's 'bad conscience' is also triggered by a denial of life's intrinsic value and the attempted flight to a transcendent beyond. Moreover, Nietzsche's account is also associated with the asceticism of Judaism and primitive Christianity. In Hegel's case, however, this self-division is both necessary and liberating. The self-imposed wretchedness of the unhappy consciousness is merely the preparation to a new and more complete phenomenological stage, grounded in a dialectical reconciliation of self-identity and self-alienation. Although the unhappy consciousness is concerned 'to save and maintain itself at the expense of the *world*, or of its own actuality', its efforts will lead to the discovery or Reason, where 'the *existence* of

[643] *KSA* V, 323; *GM*, 57

the world becomes for self-consciousness its own *truth* and *presence*.'[644] For Nietzsche, on the other hand, the evolutionary turn from the element of strength and worldliness to that of humility and self-denial is a self-diminishing movement. It leads humankind away, rather than towards, the actualization of its full ontological potential.

Nietzsche's genealogical approach to morality amounts to yet another instantiation of the 'problem of valuation' (*das Problem vom Werthe*), already explored with regard to truth and science. Just as Nietzsche sought to look beyond truth's referential value, by asking about the actual *value* of the notions of truth and untruth, he is now set on looking beyond the moral presuppositions embedded in modern reason, by questioning the actual *value* of the notions of good and evil, right and wrong, selflessness and selfishness, justice and injustice, etc. In order to do so, he resorts to physiology, etymology and natural history, in an attempt to free philosophy and psychology from the moral presuppositions they have come to incorporate and to unveil a new and wider perspective on the issue of morality. A 'physiology of morality', as Nietzsche understands it, is meant to expose the contingency of moral concepts and to retrace their natural origin. Under this new optic, what is deemed good or bad, moral or immoral, is interpreted in light of the concrete demands of human life and human history, of physical health and physical growth. Instead of asking whether something is good or bad in *itself*, the physiologist must determine whether it is favourable to the development of the human type, whether it is a sign of health or a symptom of weakness and degeneration, whether it strengthens the will to live and to affirm oneself, or whether it undermines one's confidence and creativity.

All of these interrogations are primarily linked, Nietzsche argues, to human beings' spiritual emancipation and to the rise of bad conscience. This should not be taken to mean, however, that primitive men were purely physical or mechanical beings, devoid of spiritual life. The difference lies rather in their understanding of the relationship between the physical and spiritual components of life. Nietzsche aims to show that bad conscience is not a necessary trait of self-consciousness, but the result of a specific kind of self-consciousness — and, what is more, a *pathological* kind of self-consciousness. As soon as

[644] *HW* III, 178f.; *PS*, 139f.

the mind was detached from the body, and turned against it, spiritual life acquired an independence it did not possess. Instead of fuelling the body's natural strength, its immediate cry for self-fulfilment and self-affirmation, it channelled its activity to an internal, disembodied world, increasingly devoid of physical expression.

This separation took its toll, firstly, on the primitive notions of good and evil. In opposition to contemporary theories regarding the origin of morals, which located the primary source of good and evil in the notions of utility and custom, Nietzsche argues for the *aristocratic* foundation of these notions.[645] In the primitive world, good and evil were not prescriptive or utilitarian notions, but *descriptive* ones. The word 'good' was usually applied to 'good men', that is, to powerful or superior men, whose talent for domination was immediately recognized and celebrated. The epithets 'bad' or 'evil', on the contrary, were the negative counterparts of this primitive nobility, applied to weakness and lack of power. Nietzsche's point, however, is not simply that good was synonymous with power or strength: he argues, rather, that the concrete meaning of this qualification was unilaterally established by the stronger or nobler class — that is, by those who possessed the actual power to make others conform to their definition of good. In its primary form, then, the good was the immediate reflection of a dominant and autocratic worldview. Good things were the attributes of a strong and noble life: viz. the noble man's belongings, his physical complexion, his concrete achievements, his specific preferences, his way of life, etc. If these classifications possessed a moral content, it consisted simply in the noble man's appreciation of his own intrinsic 'goodness'.

The noble class was necessarily faced with a weaker class, against which it was able to exercise and display its sovereignty. But although weaker men were forced to comply with the values and practices dictated by stronger men, their obedience was counterweighed by the internalization of their own unfulfilled instincts. Their repressed power was the first root of bad conscience — the psychological 'swamp', as Nietzsche sometimes puts it, where the first seeds of morality were allowed to grow. As the power of these primitive masters increased, so did that of their slaves, but in a negative or inverted way: initially, their submission was the source of shame and self-loathing, leading to an

[645] The authors Nietzsche has in mind in the opening chapters of the *Genealogy of Morality* are Paul Ree and Herbert Spencer. See *KSA* V, 257ff.

increasingly severe condemnation of their bodily existence and to an emancipation of life's mental or spiritual element; soon afterwards, this growing asceticism was coupled with a secret resentment against nobility and domination, which grew in direct proportion to the hardships and humiliations they were forced to endure. Their inner strength was fuelled by the decrease of their actual strength, and a new kind of power — an indirect, spiritual, or 'subterranean' power — started taking shape.

Once again, Nietzsche's account displays important similarities with Hegel's characterization of self-consciousness. In the *Phenomenology of Spirit*, and later on in the *Encyclopaedia*, primitive subjectivity was also defined in terms of power, strength and recognition. Although Hegel's master-slave dialectic is part of a different philosophical project, with different aims and a very specific methodology, it can also be construed, generally speaking, as an alternative approach to the 'problem of valuation'. In Hegel's account, already discussed above, the most basic form of self-consciousness consists in the recognition of another self-consciousness and its own drive for recognition. In this initial stage, each self-consciousness is bound to look for itself outside of itself, and to try and reduce the other to its own self. In this symmetrical struggle, subjectivity starts out as a pure will to domination, whose *terminus ad quem* is the complete suppression of otherness.

This confrontation takes the form of a mortal struggle, where each recognizing self seeks to eliminate the other. But since the suppression of one of the contenders cancels his ability to recognize the other, and thus the latter's very source of identity, this primitive scenario is replaced by a more complex dialectical arrangement. Instead of life and death, the struggle becomes a matter of authority and obedience: the victor strips the vanquished of his independence, but is careful enough to ensure his own self-identity; in other words, he does not kill his victim, but turns him into a slave. Henceforth, master and slave are the protagonists of a new dialectical cycle, whose formal structure anticipates, to some extent, Nietzsche's account of the emergence of bad conscience. Initially, while the master is free to exhibit his power and to reap the fruits of his sovereignty, the slave is bound to a life of servitude and dispossession. However, his will to domination has not disappeared; it was merely thwarted by the master's triumph and forced to turn inwards. Gradually, nourished by work and self-discipline, it will reemerge as a new form of independence, enabling the slave to supersede both his own

enslavement and the master's sovereignty, and to strive for new and more sophisticated modes of emancipation.

Again, Hegel regards this whole process as logically necessary. Just as the master's sovereignty, due to its inner contradictions, is bound to be short-lived, so too the slave's submission is merely a stopping point in the path towards reason's self-fulfilment. Mastery, slavery and all the other phenomenological stages are necessary evils, as it were, without which human consciousness could not achieve its highest truth. In Nietzsche's account, on the contrary, the master's immediate power is qualitatively superior to the slave's internalized power. The replacement of the former with the latter represents, therefore, a qualitative loss.

A closer look at Nietzsche's basic argument will help clarify this difference. Once human beings learned to sublimate their natural drives, turning the struggle against nature and each other into a struggle *against themselves*, self-consciousness was torn in two. Its physical half was neglected in favour of its spiritual half, and the world became the outer setting of an inner and more essential form of reality. The pain and privation derived from weakness were imputed to the concrete side of existence and transformed into arguments against life and domination. As a result, the *natural* superiority of power, cruelty, nobility, war, selfishness and strength was replaced by the *spiritual* superiority of impotence, love, humility, peace, selflessness and compassion. This inversion of values was morality's first conceptual victory. Whereas stronger men did not need a moral worldview — their 'morality', if one can call it such, was immediately determined by the concrete demands of health, vigour and domination — weaker men were not as privileged. To make up for their natural inferiority, they were led to question the strong man's values and to dislocate them into a new and unnatural region. The noble or aristocratic conception of good was turned into its opposite, and exposed as *evil* — that is, morally reproachable; conversely, the noble or aristocratic conception of evil was likewise turned into its opposite, and revealed as *good* — that is, morally commendable. Moral precepts were born, therefore, from the artificial distinction between a 'slave morality' (*Sklavenmoral*), contrasted with mankind's primitive nature, and a 'master morality' (*Herrenmoral*), which is ultimately but the result of an internalized will to domination.

Given this basic diagnosis, Nietzsche sets out to expose the fundamental contradiction between morality's self-appointed meaning and its *natural* or *pre-moral* meaning. Although the first advocates of morality exalted the merits of humility and selflessness, these very concepts were born out of greed and selfishness. Although they drew away for the physical world, they did so out of an intense, but frustrated desire to be part of it. According to Nietzsche, 'slave moralists' were the inventors of a new and undeclared selfishness, which would prove far more violent than the *innocent* selfishness of primitive masters. Whereas, 'the noble method of valuation' 'acts and grows spontaneously, seeking out its opposite only so that it can say 'yes' to itself even more thankfully and exultantly'[646], the 'base method of valuation' is grounded in 'the *ressentiment* of those beings who, denied the proper response of action, compensate for it only with imaginary revenge. Whereas all noble morality grows out of a triumphant saying 'yes' to itself, slave morality says 'no' on principle to everything that is 'outside', 'other', 'non-self': and this 'no' is its creative deed.'[647]

Because of the slave's usual cunning, of his calculating nature, of his capacity to endure pain and humiliation, his power ends up surpassing that of the master. For the latter, cleverness is 'nowhere near as important as the complete certainty of function of the governing *unconscious* instincts, nor indeed as important as a certain lack of cleverness (*eine gewisse Unklugheit*), such as a daring charge (*das tapfre Drauflosgehn*) at danger or at the enemy.'[648] Nobility is not about reason, but about instinct, spontaneity, audacity. For the slave, on the contrary, power hinges on prudence and patience, on the willingness to wait and to compromise. From a historical or evolutionary perspective, the Nietzschean contrast between the immediate or uncultured power of masters and the spiritual or educated power of slaves evokes the general educational path followed by primitive mankind: the evolution leading from natural ignorance (ἀπαιδευσία) to *political* education (primitive παιδεία), to *technical* and *sophistical* education (παιδεία τῶν πολιτῶν, παιδεία τῶν τεχνιτῶν), to the scientific revolution advocated by Socrates and Plato. In the introduction to this work, we have seen that each of these leaps entailed the

[646] *KSA* V, 271; *GM*, 20
[647] *KSA* V, 271f.; *GM*, 20
[648] *KSA* V, 273; *GM*, 21

conviction of an increase in power, and therefore an increase in human freedom. What Nietzsche points out, however, is precisely the opposite, namely the idea that education is liable to bring about a loss of human beings' original freedom and the obligation to conform to dogmatic standards of valuation. Although Nietzsche's historical characterization of the transition from primitive mastery to primitive slavery is wide, heterogeneous, and hence not easy to locate within the educational progression I have sketched earlier, his main concern is to expose the *debilitating*, rather than the *empowering* effects of the previous forms of paideia. Primitive sovereignty, on his view, 'is the point where feeling reaches the opposite of the low temperatures needed for any calculation of prudence or reckoning of usefulness (*jede berechnende Klugheit*, *jeder Nützlichkeits-Calcul*) — and not just for once, for one exceptional moment, but permanently.'[649] In other words, primitive sovereignty is not grounded in technical or scientific achievements: its power, as well as the freedom it provides, are due to its independence from every rational or moral rule.

For this very reason, however, primitive masters are simultaneously the most fragile of beings. Their childlike inability to plan and to calculate, to sacrifice their immediate goals to more distant or more abstract rewards, renders them particularly vulnerable to the world's arbitrariness. Their life is usually more pleasurable and more intense than a life of submission, but it is also shorter and more perilous. They fight violence with violence, cruelty with cruelty, and shun every form of compromise. To borrow one of Nietzsche's favourite metaphors, noble men resemble those rare and exuberant plants born in extreme natural environments and incapable of surviving in milder conditions. Weaker men, on the other hand, are amphibious and agile, solicitous and malleable, prone to imitation and adaptation. In this respect, Nietzsche's characterization of sovereignty is indeed similar to Hegel's, differing only in his appreciation of it. In the *Phenomenology*, masters are also doomed to extinction, but in favour of a more complete form of self-consciousness; for Nietzsche, this extinction amounts to a degeneration, a disease, a downgrade — it is the movement whereby excellency gives in to mediocrity, and life is forced into the dogmatic moulds of science and morality.

[649] *KSA* V, 259; *GM*, 11f.

11.4. *Asceticism and nihilism*

Whereas Nietzsche's earlier contrast between aesthetic freedom and moral servitude was chiefly associated with the cultural transition from the ancient Greek world to Socratism (and Christianity), the opposition between a master and a slave morality is granted a wider historical background. The first kind of morality is generally linked to the pagan or 'barbaric' world, whether it be represented by the cruelty of the primitive Germanic peoples, by the nobility of Homeric heroes or by the strength of Roman generals. The second kind of morality is linked to the Jewish and Christian traditions. Judaism, in particular, represents for Nietzsche the slavish religion *par excellence*: the source of a vengeful hatred against all things noble and the root of the 'negative' or 'selfless' ideals which will come to dominate the Christian world.

This long genealogical progression is grounded in the basic fact of human suffering. For primitive masters, pain and violence were necessary and stimulating; they helped fuel and intensify 'all the debauches of victory and cruelty'. For weaker or unhealthier men, they were the cause of fear and unhappiness. Less equipped to face the world's natural hostility, weaker men were less able to protect themselves. More importantly, they were less able to satisfy their own craving for victory and satisfaction, which led them to turn inwards and against themselves. Starting from this basic diagnosis, Nietzsche re-reads the history of human culture as a history of human impotence: he regards the moral, philosophical and spiritual transformations underwent by the primitive, medieval and modern man as increasingly refined developments of the same defensive strategy; on his view, these transformations amount, for the most part, to the different strategies devised by powerless men to escape suffering and to find the satisfaction they crave.

By turning their back on the outer world, primitive slaves are led to endorse a series of 'ascetic ideals' liable to assume very different forms. The first one, already mentioned earlier, is the feeling of shame or self-loathing which leads weak men to repress and revile their most basic instincts. Culturally, this idea is linked to the habits of fasting, abstinence and self-sacrifice, dating back to the very dawn of human civilization. Secondly, and perhaps simultaneously, this feeling is extended to the whole world and to everyone else: the condemnation of one's actions and desires becomes a condemnation of

acting and desiring *as such*. For the first time, life itself is put on the balance and found wanting; natural wisdom gives way to the 'wisdom of Silenus' and to a wholly new perspective on human existence. As a result, the embodied individual retreats into his 'consciousness', his 'soul', his 'inner citadel', and embraces the life of a hermit, a monk, a Stoic. Thirdly, the world's former protagonism is projected into a new, spiritual world, along with the slave's desires and aspirations. With this new inversion, the earthly element wins back its value: the physical world becomes a *sanctified* world, the product of a *divine* will. Since every earthly reality is an image or a sign of God, the hermit and the Stoic are no longer forced into isolation. The atonement for their weakness is now dependent on their moral conduct, on the selflessness required to earn a place in the God's heavenly kingdom.

The idea of divinity, open to various meanings and interpretations, is the object of a series of important transformations. In the Greek world, for example, the gods are not yet endowed with moral or saintly qualities, nor are they the negative image of human beings' frustrations. According to Nietzsche, they are the 'reflections of noble and proud men in whom the *animal* in man felt deified.'[650] The Greeks 'used their gods expressly to keep 'bad conscience' at bay so that they could carry on enjoying their freedom of soul.'[651] With the Jewish and the Christian traditions, all of this changes. The Jewish God is the avenger of human weakness, the source of power which the slave failed to find in himself. God is the ultimate master, but he is also a reminder of human impotence, of men's fallen condition, of their intrinsic wretchedness. With his negative conception of divinity, the Jew 'seizes upon the ultimate antithesis he can find to his actual and irredeemable animal instincts, he reinterprets these very instincts as a debt [or guilt, *Schuld*] before God.' In doing so, 'he emits every "no" which he says to himself, nature, naturalness and the reality of his being as a "yes", as existing, living, real, as God, as the holiness of God, as God-the-Judge, as God-the-Hangman, as beyond, as eternity, as martyrdom without end, as hell, as the immeasurability of punishment and debt [*Schuld*].'[652]

[650] *KSA* V, 333; *GM*, 64
[651] *KSA* V, 333; *GM*, 64f.
[652] *KSA* V, 333; *GM*, 64. Translation altered.

The Jewish faith is the direct antithesis of the freedom enjoyed by Nietzsche's superior men. Its strength derives precisely from the condemnation of all the things that made great men great: the violence and the cruelty that enabled them to flourish and assert their power are deemed immoral and replaced by a new and unnatural fixation on stillness and renunciation; the overflowing health they enjoyed is countered by a spiritual indifference to the human body, to its strength and vigour, to its needs and appetites; the Olympian carelessness of their deeds, and indeed of their whole life, is replaced by a morbid obsession with guilt, debt, human accountability and divine justice. All of these contrasts attest, once more, to the radical incompatibility between the aristocratic mode of valuation praised by Nietzsche and the slavish mode of valuation employed by conventional morality. The first one speaks the language of health and physiology; its chief concern is to lead man upwards, onwards, towards the fulfilment of his 'highest potential power and splendour.'[653] The second one speaks the language of spirit and sentiment, and aims to detach man from the concrete element of earthly life.

With the rise of Christianity, this contrast is further aggravated. The emphasis is no longer placed on revenge and punishment, but on love, forgiveness and compassion. Life is no longer cursed, but sanctified, and the natural world is no longer shunned or reviled, but praised as a divine gift. For Nietzsche, however, this change of focus does not amount to a real transformation of the Jewish standpoint. Even though the world is accorded a new importance, its value is still entirely negative. Like the Jews, the Christians say 'no' to 'nature and naturalness', but do so in a more sophisticated way. To them, earthly life is not meaningful or desirable *in itself*, but only as a vehicle for 'redemption', 'salvation', or 'absolution', notions whose meaning is once again directly dependent on the acknowledgement of the fallen or sinful condition of humankind. Whereas Judaism denounces earthly strength and earthly power as spurious and shameful achievements, praised by the wicked, rather than the good, Christianity calls for an actual celebration of weakness, poverty and compassion. In order to 'redeem' themselves, sinners must not only recognize their natural helplessness, but also elevate it to a moral imperative. Since salvation is primarily destined to the poor-spirited, the humble, the

[653] *KSA* V, 253; *GM*, 8

soft-hearted and the destitute, true Christians must strive to become all of these things: they must turn their natural slavishness into a *deliberate* or *voluntary* slavishness. Not surprisingly, this attitude represents, for Nietzsche, the most dangerous attack on life, health and nobility. 'By taking the side of everything weak, base, failed, [Christianity] has made an ideal out of whatever *contradicts* the preservation instincts of a strong life.'[654] Accordingly, the Christian God is but a 'hybrid creature of ruin [*dies hybride Verfalls-Gebilde*], made from nullity, concept and contradiction, in which all the instincts of *décadence*, all the cowardices and exhaustions of the soul find their sanction!'[655]

While Judaism is rigid and sectarian, Christianity is merciful and cosmopolitan. Its moral worldview is not the prerogative of a chosen people, but an appeal sent out to humankind as a whole. In his later writings, Nietzsche charts Christianity's historical legacy, and is often engaged in detecting traces of the 'Christian attitude' in different philosophical traditions. He sometimes praises the occasional resurgence of the aristocratic values embodied by primitive masters — namely during the Italian Renaissance, 'a brilliant, uncanny reawakening of … the noble mode of valuation'[656]; or under the sway of Napoleon, in whom 'the problem of the *noble ideal itself* was made flesh'[657] — but is invariably led to recognize its exceptional nature. Modern European history is the history of a downfall, a domestication, a cooling down of man's primal instincts. The raw innocence of warriors and aristocrats was soon surpassed by the spiritual cunning of priests, prophets and zealots; the words 'God', 'sin' and 'salvation' were translated into every language and absorbed by every philosophical system. Morality cast its web across the ages and left its sting in different regions and schools of thought. Everywhere the noble element was forced into new and hybrid mixtures, often at the price of extinction.

[654] *KSA* VI, 171; *A*, 5

[655] *KSA* VI, 185; *A*, 16. Translation altered.

[656] *KSA* V, 287; *GM*, 33. Cf. *KSA* VI, 171: 'Die Menschheit stellt *nicht* eine Entwicklung zum Besseren oder Stärkeren oder Höheren dar, in der Weise, wie dies heute geglaubt wird. Der "Fortschritt" ist bloss eine moderne Idee, das heisst eine falsche Idee. Der Europäer von heute bleibt, in seinem Werthe tief unter dem Europäer der Renaissance; Fortentwicklung ist schlechterdings *nicht* mit irgend welcher Nothwendigkeit Erhöhung, Steigerung, Verstärkung.'

[657] *KSA* V, 288; *GM*, 33. Cf. *KSA* V, 120, 185, 199; *KSA* XII, 357.

This declining tendency is well visible, Nietzsche argues, in the 'theologian element' which pervades modern German philosophy. 'Kant's success', like that of Luther, or Leibniz, 'is a theologian's success.'[658] As regards Kant's epistemology, the distinction between a world of appearance and a world in itself, placed beyond the reach of human experience, represents for Nietzsche the new clothing of a much older religious ideal. In his eyes, all forms of idealism, by dislocating the essence of human life from a concrete here to an abstract beyond, amount to variations of the same theological deception. From a physiological point of view, their basic motivation is a feeling of weakness and their main result is a slavish existence. Likewise, as regards Kantian morality, the ideas of virtue, duty, selflessness or a 'good in itself' are all part of morality's decadent vocabulary, favoured by weak and ascetic men. Here, once again, the philosopher is at odds with the physiologist, and an obstacle to the attainment of human being's 'highest potential power.'

Another important name in this historical survey is Schopenhauer, in whom the later Nietzsche sees one of the greatest and most dangerous enemies of the aristocratic ideal. Seduced by the Christian instincts of pity and compassion, by the Buddhist emphasis on silence and self-denial, Schopenhauer is the leading modern pessimist. His enthronement of the will is really an enthronement of the *negation* of the will[659]: a philosophical indictment against the 'penal servitude of volition'[660], in which human man's natural attachment to the world is deemed illusory and burdensome. Hostile to life, to the senses, to the tumultuous joy of health and virility, Schopenhauer put life itself on the balance and concluded against its intrinsic value. His philosophy is one of the main sources of the 'will to the desert' or the 'will to nothingness' which Nietzsche considers the defining trait of modern European culture.

Finally, the aristocratic mode of valuation was further compromised by the modern dissemination of democratic and socialist ideas, whose first main surge dates back to the French Revolution. Once again, Nietzsche's reaction to this phenomenon is not hard to anticipate: in direct contrast to his *Herrenmoral*, reserved to masters and warriors, the

[658] *KSA* VI, 176; *A*, 9. Cf., for ex., *KSA* III, 13ff., 185ff.; *KSA* V, 24ff., 404f.; *KSA* VI, 177f., 250ff.; 358ff.; *KSA* XI, 215, 261f., 442.

[659] Cf. *KSA* V, 68, 74f.; *KSA* VI, 173f.

[660] *Welt als Wille und Vorstellung* I, 231, cit. in *KSA* VI, 348.

democratic ideal advocates what he often describes as a *Herdenmoral*, that is, a 'moral of the herd', or the 'populace'. Broadly construed, socialism and democracy amount to a re-enactment, in the political sphere, of the Christian ideal of redemption, destined to the weak and the disenfranchised. Like Christian morality, this new attitude leads to 'a mediocritization and depreciation of humanity in value'; to the 'degeneration and diminution of man to the perfect herd animal [*zum vollkommenen Heerdenthiere*]', to 'a dwarf animal [*zum Zwergthiere*] of equal rights and claims.'[661] Moreover, like Christian morality, socialism forces the strong and the powerful to restrain their natural might and to settle for a slavish mode of existence.

These cultural and historical forces, strengthened by the increasing popularity of 'modern ideas', lead men away form the fulfilment of their true potential, towards a life of comfort, consensus and 'reasonableness'.[662] As the 'instinct of life' dies down, so do the conflicts and hierarchies that fuelled the ancient world. Morality, asceticism and democracy all lead to the triumph of weakness and exhaustion, to the nihilistic tendency which dominates the modern world:

> Today we see nothing that wants to expand, we suspect that things will just continue to decline, getting thinner, better-natured, cleverer, more comfortable, more mediocre, more indifferent, more Chinese, more Christian — no doubt about it, man is getting 'better' all the time . . . Right here is where the destiny of Europe lies — in losing our fear of man we have also lost our love for him, our respect for him, our hope in him and even our will to be man. The sight of man now makes us tired — what is nihilism today if it is not *that*?[663]

[661] *KSA* V, 126ff.; 91f.

[662] *KSA* XIII, 17: 'Gesamt-Anblick des zukünftigen Europäers: derselbe als das intelligenteste Sklaventhier, sehr arbeitsam, im Grunde sehr bescheiden, bis zum Excess neugierig, vielfach, verzärtelt, willenschwach — ein kosmopolitisches Affekt- und Intelligenzen Chaos.'

[663] *KSA* V, 323; *GM*, 57. As with pessimism, despite Nietzsche's more frequent definition of nihilism as a sign of exhaustion, degeneration, weakness, etc., he also distinguishes between a *positive*, *active* or *strong* nihilism and a *negative*, *passive* or *tired* nihilism. What is at stake in both cases is the devaluation of previously accepted values; but whereas the first kind of nihilism can be 'a sign of strength', or a call for the endorsement of new and stronger values, the second kind is 'a sign of weakness', a symptom of *décadence* (Christianity, Buddhism). See *KSA* XII, 350ff ., 367.

12.

PHILOSOPHY UNCHAINED

As regards the issue of human freedom, Nietzsche's philosophical project can be divided into two main moments, corresponding to a descriptive and a prescriptive approach to the 'problem of valuation'. In the previous chapters, we have dealt mainly with the first of these moments: namely, with Nietzsche's critique of human beings' usual mode of cognition. His critical strategy often consists a) in exposing human thoughts and actions as the products of a global system of valuation, liable to assume different hierarchical configurations; and b) in arguing that the mode of valuation usually relied upon by human consciousness is dogmatic and misguided; it is not enabling of, but rather an obstacle to, the full actualization of freedom.

Nietzsche's criticism is centred on two complementary themes: he argues, firstly, that the referential value usually ascribed to the *epistemic* alternative between truth and untruth — whether inexplicitly, throughout the course of everyday life, or explicitly, in philosophical and scientific debates — betrays an error in valuation, which must be questioned and corrected; as he sometimes puts it, human beings' usual concern with truth and truthfulness speaks to their 'lack of perspective' and to the need for a wider and freer outlook on reality. Secondly, Nietzsche argues that the referential value usually ascribed to the *moral* alternative between good and evil — again both inexplicitly, in the course of everyday life, and explicitly, in religious and philosophical contexts — is the product of a misevaluation, which must also be corrected. The 'will to truth' and the 'will to the good' constitute, for Nietzsche, arbitrary modes of valuation automatically

regarded as necessary and self-evident. They are the standards with which all things are measured, but their own value is hardly ever called into question.

Furthermore, both of these standards are incompatible with what Nietzsche takes to be the real driving force behind human life: i.e. a pre-epistemic and pre-moral (or *trans*-epistemic and *trans*-moral) *will to power*, to domination, to victory and self-enhancement. Whenever human existence is subordinated to epistemic and moral imperatives, the resources needed to fuel and intensify the 'instinct of life' are unduly limited. Consciousness is denied the freedom to create its own values, or to change them at will, and the different possibilities of valuation are artificially reduced to a more or less narrow set of fixed ideas, which may lead to a decrease, rather than an increase in freedom and vitality. In certain cases, whenever the concern with truth and the good is elevated to a dogmatic creed, science and morality enter into direct contradiction with life: it is the case of asceticism and nihilism, favoured by the Judaism, Christianity and the main trends in modern philosophy and politics.

This contradiction brings us to the second, *prescriptive* moment of Nietzsche's project. Besides denouncing the self-limiting or slavish nature of human beings' usual standpoint, Nietzsche calls for what he names a 'revaluation' or 'transvaluation of all values' (*Umwerthung aller Werthe*), starting with the ideals of truth and morality. His aim is to transform the hierarchical structure of human values, in order to eliminate the usual discrepancy between reason and life.

12.1. *The will to power*

Throughout the previous sections, Nietzsche's reform of human cognition was generally associated with the idea of *power*. Instead of a life measured in terms of truth and untruth, right and wrong, good and evil, Nietzsche advocates a kind of life where personal choices are endorsed or rejected depending on whether and how far they contribute to an actual increase of personal power. This basic formula, however, is open to different interpretations: 'power' (*Macht*), 'powerfulness' (*Mächtigkeit*) and the 'will to power' (*der Wille zur Macht*) are not self-evident notions; moreover, given their central

role they play in Nietzsche's writings, determining their exact meaning amounts to nothing less than determining the exact nature of the global transformation he advocates.

In the opening pages of *The Antichrist*, Nietzsche offers an explicit description of his new 'table of values':

> What is good? — Everything that enhances people's feeling of power, will to power, power itself.
>
> What is bad? — Everything stemming from weakness.
>
> What is happiness? — The feeling that power is *growing*, that some resistance has been overcome.
>
> *Not* contentedness, but more power; *not* peace at any price [*überhaupt*], but war; *not* virtue, but capacity.[664]

In this description, as in many others of the same nature, the new way of life proposed by Nietzsche is not simply about power, or the 'will to power', but about 'growth', 'war' or 'capacity' (*Tüchtigkeit*). These different concepts are employed to qualify a general phenomenon which is associated, throughout Nietzsche's writings, with many other different terms and expressions, some of which used almost interchangeably — namely 'life' ('will to life', 'instinct of life', 'the ascending life', 'vitality', 'liveliness', 'health', 'energy', 'fullness', etc.), 'strength' ('will to strength', 'might', 'vigour', 'virility', 'hardness', etc.), 'self-affirmation' ('saying yes to oneself', 'saying yes to life', 'joy', 'delight', 'egoism', 'pride', 'self-glorification', etc.), 'self-enhancement' ('self-elevation', 'self-overcoming', 'expansion', etc.), 'domination' ('mastery', 'sovereignty', 'nobility', 'aristocracy', 'superiority', 'independence', 'command', 'victory', 'will to victory', 'audacity', 'valour', etc.), 'strife' ('conflict', 'combat', 'tension', etc.), and the list could go on.[665] All of these expressions refer to an existential regime whose exact contours are not entirely clear, not only because Nietzsche resorts to such a wide range of vocabulary, but

[664] *KSA* VI, 170; *A*, 57. Translation altered.

[665] Nietzsche's most frequent formulas include *Leben*, *Lebendigkeit*, *aufsteigendes Leben*, *Lebenskraft*, *Instinkt des Lebens*, *Wille zum Leben*, *Wachstum*, *Selbstüberwindung*, *Gesundheit*, *Fülle*, *Überfülle*, *Überfluss*, *Fruchtbarkeit*, *Reichtum, Überreichtum*, *Gewalt*, *Überwältigung, Stärke*, *Verstärkung*, *Kraft*, *Überkraft, aufsteigende Kraft*, *Wille zur Kraft*, *Härte*, *Tüchtigkeit*, *Spannung*, *Sieg*, *Wille zum Sieg*, *Grösse*, *Übergewicht*, *Übermacht*, *Übermenschheit*, *Mehr-Menschheit*, *Vornehmheit*, *Vorrang*, *Aristokratie*, *Herrschaft*, *Vorherrschaft*, *Herrschaftlichkeit*, *Unabhängigkeit*, *Für-sich-sein-wollen*, *Tapferkeit*, *Stolz*, *Ja-sagen*, *Ja-schaffen*, *Selbstbestimmung*, *Selbstbejahung*, *Selbstherrlichkeit*, *Selbstverherrlichung*, *Fröhlichkeit*, *Rausch*, etc.

also because he chooses to highlight, in different occasions, different aspects of the same basic ideas.

Generally speaking, Nietzsche's alternative to the standpoints of truth, science, morality and all other dogmatic or metaphysical modes of valuation, is a philosophy of power and domination; a philosophy that replaces the intellectual and moral deliberations that are typical of other systems of valuation with the *direct* experience of power and domination. But what exactly does this entail? And how does it relate to, or differ from, say, a peculiar form of utilitarianism, or evolutionism, or hedonism? In order to make out, as clearly as possible, the meaning and scope of Nietzsche's standpoint, it is important to avoid a few hermeneutical pitfalls which he himself calls attention to in his writings.

To begin with, a mode of valuation that has moved 'beyond' the dichotomies true/untrue or good/evil can be construed as a purely instrumental or Machiavellian mode of valuation: that is, as the outlook of a wholly selfish human being, entirely focused on his or her own self-interest. In this case, valuation would hinge solely on the maximization of personal advantage (whether in the form of power, gain, utility, etc.). Depending on the context, truthfulness and compassion would be just as valuable, if not more so, than deceit and cruelty, for both attitudes might lead to useful or empowering results. But although this line of interpretation seems to agree with Nietzsche's insistence on the 'selfish', 'proud' and 'pitiless' demeanour of superior men, it contradicts his general definition of nobility. As we have seen earlier, utilitarian concerns are unknown to, and unworthy of, a noble consciousness: they represent, for Nietzsche, a form of enslavement or self-abasement.[666] The noble man is innocent and sincere, hot-blooded and childish, wasteful and exuberant; his joy and his power are instantaneous, and not the result of a plan or a calculation, nor the indirect reward in view of which something else is pursued or undertaken. Scheming and economizing are left to the weak and to the vengeful, to logicians and moralists. Their virtue is patience, resignation, the laborious deduction of causes and effects. The noble man lives for the present moment — which is why Nietzsche is sometimes careful to distinguish between power in general and the

[666] See, for ex., *KSA* XI, 74: 'Zur Signatur des Sklaven: die Werkzeug-Natur kalt, nützlich, — ich betrachte alle Utilitarier als unwillkürliche Sklaven. Menschen-Bruchstücke — das zeichnet die Sklaven.'

actual feeling or *sensation* of power: viz. the immediate feeling that 'power is *growing*, that some resistance has been overcome.'[667] It is power in this concrete, physiological sense that he places at the heart of human existence.

Alternatively, Nietzsche often speaks of the 'instinct of conservation', of human beings' natural defences and of an in-bred drive to protect one's life against danger and death. He distinguishes between modes of behaviour that favour the prolongation of the human species and others that lead to degeneration and extinction. Often enough, things are deemed true, right or good that are unhealthy and pose a threat to human life; conversely, many of the things usually considered untrue, wrong or evil are in fact beneficial and life-enhancing. In light of these ideas, Nietzsche's philosophy of power could be construed as a philosophy of self-preservation, where values are measured according to their power to extend and protect human life. In *Beyond Good and Evil*, however, Nietzsche explicitly rejects this line of interpretation, which he includes in a list of 'philosophical prejudices' to be gotten rid of.[668] Against Spinoza, Darwin and modern physiologists, he argues that the instinct of conservation is not life's main driving force, but merely one of the most frequent indirect consequences of the will to power. This same point is also made, for example, in *The Gay Science*, where Nietzsche claims that the 'the wish to preserve oneself' is 'a limitation of the truly basic life-instinct, which aims at the *expansion of power* and in so doing often risks and sacrifices self-preservation.'[669] Furthermore, also in opposition to this idea, he often speaks of the superior man's 'unconcern and scorn for safety, body, life, [and] comfort'[670], as well of the joys of danger and the honour of dying fighting, as a hero in a battlefield.[671]

A third possibility is to interpret the 'feeling of augmented power' pursued by superior men as a peculiar form of pleasure or gratification, and the 'will to power' itself as a peculiar form of hedonism. To be sure, Nietzsche does not encourage the avoidance of pain or conflict, nor does he prescribe a life of comfort and ease. These are the very ideals he goes after in his critique of morality, democracy and nihilism. On the contrary,

[667] *KSA* V, 167; *KSA* VI, 170
[668] See *KSA* V, 27f.
[669] See *KSA* III, 27f; *GS*, 207f, but also *KSA* XII, 260ff.
[670] *KSA* V, 275; *GM*, 23
[671] See, for ex., *KSA* IV, 93ff., or *KSA* XIII, 21.

he often speaks of hardness, discipline and the endurance of pain as signs of strength and nobility. However, these adversities are not valued out of a sense of duty or abnegation. Pain and strife provide the resistance whose overcoming rekindles the feeling of power. They are tonic or stimulating agents, whose intensity determines the intensity of the pleasure they are able to procure.

Yet despite the close connection between power, pleasure and pain, Nietzsche rejects, once again, this kind of reading. In his own words, 'hedonism, pessimism, utilitarianism, [and] eudaemonism', insofar as they are all 'ways of thinking that measure the value of things according to pleasure and pain', are focused on 'secondary states and matters [*Begleitzustände und Nebensachen*]. They are all foreground ways of thinking and naïvetés, and nobody who is conscious of both *formative* powers [*gestaltende Kräfte*] and an artist's conscience will fail to regard them with scorn as well as pity.'[672] The will to power is thus something other than the 'will to pleasure', or the 'will to happiness', or the 'will to self-advantage', or the 'will to self-preservation'. These are all symptoms or consequences, but not the real thing.

What, then, does one will when one *wills power*? A definitive answer to this question is not easy to find in Nietzsche's writings, not because he forgot to tackle or clarify such an important issue, but because his main aim is precisely to overcome the philosophical prejudice underlying this kind of formulation. As he repeatedly argues, to ask about the specific *end* of the will to power — which for Nietzsche is synonymous with the will *in general* — is to look for one of the 'superfluous teleological principles'[673] he sought to eliminate. Whereas scientific and moral modes of valuation are grounded in a finalist approach to human life — whether focused on logical coherence, self-preservation, well-being, selflessness or self-satisfaction — a 'philosophy of the future' must look beyond this cognitive scheme. Power, like the 'true' or the 'good', is not something fixed, to be craved or pursued, but 'something *to be created*, which designates a *process*, or what is more a will

[672] *KSA* V, 160; *BGE*, 116. Translation altered. Cf. *KSA* XIII, 30: 'Lust und Unlust sind Nebensachen, keine Ursachen; es sind Werthurtheile zweiten Ranges, die sich erts ableiten von einem regierenden Werth; ein in Form des Gefühls redendes "nützlich" "schädlich", und folglich absolut flüchtig un abhängig. Denn bei jedem "nützlich" "schädlich" sind noch hundert verschiedene Wozu? Zu fragen.' Cf. 33ff.

[673] *KSA* V, 27f.

to overpower [*einen Willen der Überwältigung*] that has in itself no end.' It is 'a *processus in infinitum*, an *active determining*, not a becoming-conscious of something [that is] fixed and determined "in itself."'[674]

The will to power is not a means to an end, but an end in itself. It is something like a *pure physiological drive*, a primal overflow of health and energy which requires no aim or justification, but merely the occasion to be released and renewed. This will is not subordinated to, nor channelled towards, a higher truth or a moral imperative — it is itself the primary source of meaning and agency. Each of its choices is the result of a momentary relation of forces, maintained or abandoned according to whether it favours or obstructs the release and expansion of power. In this dynamic context, fixed ideas are mere obstacles: every categorical 'yes' or 'no', 'true' or 'untrue', 'good' or 'evil', is a sign of weakness, a trace of disease and sclerosis, a limitation to the endless array of possibilities open to the human will.

Nietzsche's conceptual shift from a teleological to a physiological mode of valuation (viz. from a *teleological* physiology to a *pure* or *creative* physiology) is made especially clear in a central passage from the *Genealogy of Morality*, where he criticizes the teleological bias underlying most modern accounts of the origin of legal institutions, political usages, social customs or religious rites. Once again, he compares all of these processes to the growth of a living organism, in which the notions of purpose or telos play a secondary and largely accidental role. What determines the evolution and growth of living things, Nietzsche argues, is their will to power and domination. Meaning and valuation are the cognitive translations of this primary will, and not the other way around:

> Every purpose and usefulness is a mere *sign* that a will to power has achieved mastery over something less powerful, and has impressed upon it its own idea [*Sinn*] of a function; and the whole history of a "thing", an organ, a tradition can to this extent be a continuous chain of signs, continually revealing new interpretations and adaptations,

[674] *KSA* XII, 385. Cf. 386: '— Daß die anscheinende "*Zweckmäßigkeit*" ... bloß die Folge jenes in allem Geschehen [sich] abspielenden *Willens zur Macht* ist; — daß das *Stärkerwerden* Ordnungen mit sich bringt, die einem Zweckmäßigkeits-Entwurfe ähnlich sehen; — daß die anscheinenden *Zwecke* nicht beabsichtigt sind, aber, sobald die Übermacht über eine geringere Macht erreicht ist und letztere als Funktion der größeren arbeitet, eine Ordnung des *Rangs*, der Organisation den Anschein einer Ordnung von Mittel und Zweck erwecken muß.'

> the causes of which need not be connected even amongst themselves, but rather sometimes just follow and replace one another at random.[675]

The meaning and ends favoured over time by a given organism, person or tradition are the records of different relations of power, domination, subjugation and resistance. Like the rings on a tree trunk, or the lines on a seismogram, they document the different forms taken on by the will to power, as well as the degree of their success.

> The form is fluid, but the "meaning" [*Sinn*] even more so… It is no different inside any individual organism: every time the whole grows appreciably, the "meaning" of the individual organs also shifts. […] True *progressus* [always appears] in the form of a will and way to *greater power* … and [is always accomplished] at the cost of countless smaller powers … The extent of each "progress" can actually be measured according to how much has had to be sacrificed to it.[676]

This physiological mode of valuation is Nietzsche's alternative to the in-grained finalism of science and morality. In his eyes, once philosophy and psychology manage to overcome their usual prejudices, they will be free to develop a '*doctrine of the evolution of the will to power*.'[677]

This novel task, as defined and attempted throughout Nietzsche's writings, can once again be divided into a descriptive and a prescriptive moment. The first one corresponds to Nietzsche's recurring claim that human life *in its entirety* — including every form of cognition, volition and agency — is guided by a will to power. This claim also extends, *a fortiori*, to those modes of cognition or valuation that result in a *decrease* of power, or in an outright negation of life. Indeed, insofar as these examples pose the greatest challenges to Nietzsche's physiological strategy, they are the ones he returns to more often.

Firstly, Nietzsche points out that the *epistemic* bias of philosophers and scientists is nothing other than a peculiar form of will to power, typical of overeducated or overspiritualized beings. Like every other animal, the '*bête philosophe* … also strives instinctively for an optimum of favourable conditions in which to fully release its power

[675] *KSA* V, 314; *GM*, 51. Translation altered.
[676] *KSA* V, 315; *GM*, 51f. Translation altered.
[677] *KSA* V, 38; *BGE*, 23

and achieve its maximum of power-sensation'[678]. Its real goal is not truth or science, not even happiness or well-being. These are secondary aims, superposed to a primary drive 'to power, to action, to the mightiest deeds.'[679] What distinguishes philosophers from other animals is thus the devious form usually assumed by their will to power: instead of a direct display of health and strength, 'philosophical power' is achieved by means of silence, seclusion, chastity, an abnormal interest in logic and coherence, a penchant for 'big questions' and 'essential problems'.

Secondly, Nietzsche argues that the *moral* bias of priests, saints and theologians is likewise a peculiar form of will to power, usually found in weak and delicate beings. In this case, his critique is less about the naïvety of this attitude and more about its outright hypocrisy. Like philosophers, priests are primarily driven by a *Wille der Überwältigung*. Their life is not really about virtue, selflessness or compassion: once again, these 'noble' values are either the products of self-delusion (in the most sincere cases) or deliberate ruses designed to conceal their real foundation. Physiologically, the *bête morale* or *pieuse* differs from other animals in its natural weakness, which translates into the need for new and negative means of self-affirmation. Its struggle against life and health is nonetheless a struggle for power, and its situation 'is therefore the precise opposite of what the worshippers of this ideal hold [*meinen*].'[680] The priest, the theologian and God itself are all born out of a will to power — yet their power is secured and strengthened by the exaltation of weakness or powerlessness.[681]

Finally, *nihilism* itself, as the corollary of the two previous tendencies, is ultimately but an absurd cry for domination. The final stage of human degeneration, it exhibits the will to power in its most negative or perverted form: a will born out of absolute exhaustion, whose preferred stimulant is the negation of all stimulants; a will at war against itself, fuelled by the madness of its own destruction. Even here, however, power is

[678] *KSA* V, 350; *GM*, 76. Translation altered.

[679] Ibid.

[680] *KSA* V, 366; *GM*, 88. C. Diethe translates *meinen* as 'imagine', but this solution fails to capture the ambiguity of Nietzsche's formulation: priests and moralists *hold* their ideal to be something different what it actually — whether out of naïvety or plain dishonesty.

[681] In the preface to *Ecce Homo,* Nietzsche speaks of prophets and founders of religions as 'jener schauerlichen Zwitter von Krankheit und Willen zur Macht.' (*KSA* VI, 259).

still the operating principle: as Nietzsche points out, this suicidal effort 'is and remains a will! ... for man still prefers to *will nothingness*, than *not* will…'[682]

If the epistemic, moral and nihilistic standpoints are indeed about power and domination, their actual results are not: according to Nietzsche, they lead to slavish and self-diminishing attitudes, which entail an overall decrease in health, power and freedom. Nietzsche's reaction to this state of affairs coincides with the *prescriptive* moment of his philosophy, focused on the actual transformation of human existence. Once the will to power is acknowledged as the founding principle of human life, philosophy must move beyond a mere critique of existing values. The will to power must be redirected towards a truly *empowering* form, which entails both a return to the primitive values of ancient aristocracy and the 'crossing over' into a new and radically different world.

12.2. *Return to Callicles*

Looking back on the genealogical path outlined along the previous chapters, Nietzsche's approach to the 'problem of valuation' can be interpreted as a response to an ancient philosophical debate, brought to the fore in Plato's dialogues: namely, the debate regarding the definition of the *greatest good* or the *best way of life* available to humankind. In the introduction to this work, and in the different sections dedicated to Plato's philosophy, I have called attention to the hierarchical structure of Plato's conception of human cognition. I have argued, with Plato, that all of our ideas and judgements are continually shaped by a self-interested assessment of the different situations we are faced with, and subordinated, whether consciously or otherwise, to what is perceived, at each given moment, to be the *greatest good* or the *best course of action* available. I have pointed out, furthermore, that this continual assessment is never confined to a specific event or situation, but is always somehow related to our whole existence. All of our ideas and judgements are ultimately subordinated to a global diagnosis of what is at stake in human

[682] *KSA* V, 412; *GM*, 120

life — to what matters most in life, to what is essential and what is secondary, to what is worth pursuing and what is not, etc.

This general theme, pervasive in Plato's writings, is explicitly dealt with in dialogues such as the *Republic*, the *Phaedo*, the *Philebus* or the *Gorgias*, to name only a few. In these texts, Socrates and his interlocutors are directly concerned with determining 'what course of life is best.'[683] The debate usually takes the form of a dialectical contest between different values and worldviews. Wisdom, pleasure or political domination are put forth as possible versions of the *greatest good*, and the different lives they prescribe — i.e. philosophy, hedonism or rhetoric — are likewise compared and discussed. In each case, the procedure amounts to the attempted translation of a formal principle, τὸ μέγιστον ἀγαθόν, into a concrete and consistent mode of valuation.

We find something very similar in Nietzsche's writings, where 'future philosophers' are charged with the task of 'determining the rank order of values.'[684] Once again, human cognition is conceived as a global hierarchical system, whose inner composition must be brought to light. Like Plato, Nietzsche calls attention to human beings' usual ignorance vis-à-vis their own valuation system: usually, their choices and actions are put down to a more or less immediate practical concern, wrongly perceived as an end in itself; alternatively, whenever science or religion step in, thoughts and actions are led back to more general or abstract ideals. But in all of these cases the higher term of the hierarchy remains out of sight. For Nietzsche, the whole system is about *power*: viz. about the *augmentation* of the 'feeling of power' that determines and motivates every single judgement or decision. Nietzsche's aim is not only to bring to light the real structure of human beings' system of valuation — that is, to expose every human intention as a more or less direct instantiation of the 'will to power' — but also to align human aims and actions with this newly discovered drive — that is, to inaugurate a mode of existence that is no longer unconsciously, but *explicitly* focused on the pursuit and accumulation of power.

Yet the resemblance between Plato's and Nietzsche's projects is even more profound. Nietzsche's proposed transition from a scientific and moral philosophy to a *physiology of*

[683] *Go* 500c4
[684] *KSA* V, 289; *A*, 34. Translation altered.

power is strongly evocative of the debate held throughout the *Gorgias*, already mentioned earlier. In particular, Nietzsche's views on power and domination are strikingly similar to those levelled by Callicles against Socrates. So much so, indeed, that Callicles and Nietzsche can almost be regarded as philosophical 'kindred spirits', united by their strongly anti-Platonic views.[685] Their doctrines coincide on a number of crucial points, which may be summed up as follows:

1. Firstly, like Nietzsche, Callicles bases his entire argument on the sophistical distinction between the laws of nature, which regulate human life, and the laws of custom or convention, which determine what is socially acceptable.[686] He calls attention to the contradiction between human beings' natural drives, responsible for the growth and evolution of humankind, and the norms which have come to be endorsed and upheld in modern, 'civilized' societies. Moreover, anticipating Nietzsche's approach, Callicles illustrates this distinction with a genealogical account of human nature, designed to contrast the power and freedom of primitive men with the impotence and the slavishness of modern men.

2. Secondly, like Nietzsche, Callicles argues that the laws of custom, along with the increasing domestication of humankind's primitive instincts, are due to a majority of weaker men, intent on protecting themselves from the danger posed by the strong. Their aim is to 'terrorize the stronger sort of folk who are able to get an advantage, and to prevent them from getting one over *them*'. In order to do so, 'they tell them that such aggrandizement is foul and unjust, and that wrongdoing is just this endeavour to get the advantage of one's neighbours.' [687] In Nietzschean terms, they replace nature's *Herrenmoral* with an artificially posited *Sklavenmoral*, whereby the aristocratic values of

[685] This closeness is both striking and enigmatic, inasmuch as Nietzsche, who knew Plato's work very well, does not refer once to Callicles, and only twice, in passing, to Plato's *Gorgias* (namely in *KSA* I, 790 and IX, 174). See MENZEL, op. cit., 80f.: 'Wenn daher Kallikles auch von Nietzsche nicht ausdrücklich genannt wird, so darf bei der weitgehenden, oft wörtlichen Übereinstimmung kein Zweifel darüber bestehen, daß eine Einwirkung stattgefunden hat. Sie kann ja auch eine unbewußte sein. Die Eindrücke, welche der junge klassische Philologe aus der Lektüre Platons erhalten hatte, sind haften geblieben, soweit sie mit der sich entfaltenden Gedankenwelt des Philosophen Nietzsche Beziehungen besaßen.'

On the rapport between Nietzsche and Callicles, see also Dodds' commentary on the *Gorgias*, ed. cit., 386-391; as well as KLOCH-KORNITZ, op. cit; LEYRA, op. cit.; and URSTADT, op. cit.

[686] Regarding the translation of νόμος as 'custom' or 'convention', see above fn. 175.

[687] *Go* 483c

primitive men are reviled and reversed: what was good according to nature (κατὰ φύσιν) — i.e. strength, virility, domination — becomes foul (αἰσχρόν), according to custom (κατὰ νόμον, or παρὰ φύσιν), and therefore evil (κακόν); conversely, what is evil according to nature — i.e. meekness, reasonableness, restraint — becomes socially accepted (καλόν), and therefore good (ἀγαθόν).

Interestingly, Callicles illustrates his argument with an image that is also profoundly Nietzschean: namely, the comparison between 'the best and strongest men' with 'young lions' taken from their infancy and 'told the while they must have but their equal share, and that this is what is fair and just.'[688] His point, once again, is that the physiological definition of power is incompatible with the moral or political one, and that the latter is conditional on the artificial repression of natural qualities and instincts. It is a power *contra natura*: not only because it is *unnatural*, but also because it is obtained and maintained through an actual struggle *against nature*.

3. In addition, although Nietzsche's and Callicles' critique is directed at the *moral* subversion of natural values, and consists therefore in an attack on conventional justice, social contractualism and democracy, it also extends to the *philosophic* or *scientific* subversion of natural values. Both of Socrates' opponents place the immediate power of the will above the intellectual concerns of scientific reason, and both regard Socrates' philosophical fixation with truth and meaning as a symptom of slavishness. Callicles is quite explicit in his warnings against the ridicules of philosophizing far on to life, in lieu of minding those things 'that ought to be familiar to the man who would be a thorough gentleman and make a good figure in the world.'[689] In his eyes, philosophy represents, furthermore, a source of danger, inasmuch as an overly minute consideration of one's thoughts and intentions may prove paralysing, and render one unable or even unwilling to defend one's life and honour.

[688] *Go* 483e-484a. Cf., for ex., *KSA* III, 548, where the lion is evoked by Nietzsche as a sign of personal strength: 'Ich will meinen Löwen und meinen Adler um mich haben, damit ich allezeit Winke und Vorbedeutungen habe, zu wissen, wie gross oder wie gering meine Stärke ist'; or 648: 'Du strenger Thurm, mit welchem Löwendrange / Stiegst du empor hier, siegreich, sonder Müh!'; *KSA* IV, 30, where 'future men' are compared to lions: 'Freiheit sich schaffen und ein heiliges Nein auch vor der Pflicht: dazu, meine Brüder, bedarf es des Löwen'; or 205: 'Noch war ich nicht stark genug zum letzten Löwen-Übermuthe und -Muthwillen'; or 531: '*lachende Löwen* müssen kommen!'; *KSA*, V, 334, where primitive Greek men are dubbed 'löwenmüthigen Kindsköpfe.'

[689] *Go* 484c9-d2

4. Finally, both Callicles and Nietzsche advocate a complete reform of this state of affairs. The hypocrisy of modern values must be acknowledged as such and replaced by a new mode of valuation, in which power and strength are granted their original protagonism. Once again, the freedom provided by this transformation is described by Callicles in terms that anticipate Nietzsche's own descriptions: in contrast to the image of the young lions forced to hold back their strength, he envisages the arrival of a slave capable of 'shaking all that he was taught', 'breaking his bonds' and rising to the role of master.[690]

The debate between Socrates and Callicles is centred on the notion of *power* (δυναστεία, *Macht*) — and, more specifically, on the determination of *what* power is, of *who* can be said to enjoy it and of *why* it is coveted in the first place. As usually, Socrates points out that none of these questions is simple or self-evident. His critique highlights a series of ambiguities that will reappear, in very similar ways, in Nietzsche's doctrine of the will to power. Like Callicles' νόμος τῆς φύσεως, Nietzsche's *Herrenmoral* is grounded in a constellation of notions whose definition is open to different and often contradictory interpretations.

The course of the argument was already outlined above: when asked about the exact nature of the δυναστεία of superior men, Callicles defines it as a form of strength, then wisdom, then courage and manliness. However, none of these definitions succeeds in capturing the exact phenomenon he has in mind, and Socrates' refutations lead him to settle on a fourth and final definition, i.e. power as πλεονεξία: that is, as the ability to 'let [one's] desires be as strong as possible and not chasten them', but rather 'minister to them when they are at their height, by reason of [one's] manliness and intelligence.'[691] This ability, he claims, is precisely what impotent men lack, and the reason why they praise temperance instead enjoyment, restraint instead of freedom, justice instead of force.

Again, the similarities with Nietzsche's thought are striking. Callicles' conviction that justice and temperance are born out of impotence and resentment anticipates

[690] *Go* 484a3-4
[691] *Go* 491e8-492a2

Nietzsche's insistence on the 'morality of *ressentiment*', which he associates with Platonism, Judaism, Christianity and democracy. Moreover, as noted by Menzel, Callicles' notion of πλεονεξία is directly translated by Nietzsche as the 'will to wanting-to-have-more' (*der Wille zum Mehr-haben-wollen*), which amounts to a variation of the will to power.[692] Finally, and more importantly, Callicles' difficulty in defining the exact nature of the power held by superior men also seems to anticipate Nietzsche's unwillingness to reduce the will to power to a specific end or ideal. For Callicles, as for Nietzsche, power is not merely about strength, wisdom or courage, as though powerful men would simply have to submit to these values and pursue them in the best or most effective way possible. Πλεονεξία seems to refer, rather, to power or volition in a more basic or primordial sense: not a movement oriented towards a specific end, but a movement which is itself its own end; not a matter of achieving this or that particular goal, but the ability to renew and intensify a physiological drive for self-enjoyment and self-overcoming. This idea is made clearer from 491d onwards, when Callicles condemns the ideals of personal autarky and self-mastery (ἄρχειν ἑαυτοῦ, ἐγκράτεια ἑαυτοῦ) as the life-goals of slaves and simpletons, and compares the alleged happiness of 'those who want nothing' to the happiness of corpses or stones. Whereas 'a man who has taken his fill can have no pleasure any more, […] a pleasant life consists rather in the largest possible amount of inflow.'[693]

Callicles is not focused on satisfaction as a *nomen rei actae*, but as a *nomen actionis*. He puts the emphasis not on the state being sated, but on the very movement of satisfaction. Insofar as his aim is to perpetuate and intensify a feeling of gratification, his standpoint can be interpreted as a conventional form of hedonism. Alternatively, however, and more interestingly, his standpoint can also be read as a preliminary attempt at the kind of 'transvaluation' championed more than two thousand years later by Nietzsche. In this light, the remainder of the *Gorgias* may likewise be read as a challenge to the cogency of Nietzsche's project and to the possibility of a *physiological* shift from a cognitive outlook, grounded in truth and morality, to a *metacognitive* outlook, grounded in a pure and all-encompassing 'will to power.'

[692] *KSA* XIII, 251. Cf. *KSA* IX, 284 and *KSA* XI, 586.
[693] *Go* 494a6-b2

Unlike ordinary men, 'overmen' must be able to create their own values, and to do so they must liberate themselves from the normative framework of science and morality. In the *Gorgias*, Socrates' final debate with Callicles revolves precisely around the question of whether this liberation is possible. Whilst Nietzsche aims to move beyond the oppositions truth/untruth and good/evil, Socrates reaffirms his doctrine of δοξάζειν, anchored in the *transcendental* value of these two oppositions. As in many other dialogues, Plato's views amount to the restatement of a global diagnosis of human cognition, based on two major observations: namely a) that every human belief is necessarily supported both by a conviction of truthfulness and by a complex web of 'silent' beliefs also supported by such a conviction; and b) that this doxastic web is not linear, but hierarchical, and in fact *universally* oriented towards the pursuit of what is perceived to be, consciously or otherwise, the greatest good or the best course of life available. In short, truth and goodness, as well as untruth and evil, are the formal standards underlying the endorsement of every human belief, choice, or course of action. They can be filled with different contents, or given different names — but never truly eliminated.

If the good is indeed 'the end of all our actions; if 'it is for its sake that all other things [are] done'[694], then every kind of life — be it the life praised by Callicles (and Nietzsche) or the one praised by Socrates — represents a *version* of the best possible life, endorsed and pursued for this very reason. As all versions, however, it may be *true* or *untrue* — viz. *illusory*, and therefore potentially harmful, or *genuine*, and actually beneficial. It is upon this disjunction that Plato grounds his entire attack on Callicles (and Nietzsche): if every life is doomed to be a version of a true and good life, the determination of its value is ultimately dependent on whether one can actually demonstrate, rather than merely assert, that it is in fact so. In other words, an accurate mode of valuation, capable of ensuring the attainment of power, strength, wisdom, or any other quality deemed beneficial or self-enhancing, is necessarily dependent on a *scientific* outlook on reality — which can only be provided, if at all, by the practice of philosophy.

[694] *Go* 499e7-500a2

For Nietzsche, on the contrary, everything hinges on the possibility of refuting this line of argument and showing that human beings' attachment to the good and the true is not as natural or as essential. As we have seen earlier with regard to the 'will to truth', the assessment of whether and to what extent this move is possible depends on how radically we choose to interpret Nietzsche's intentions. If, on the one hand, the 'transvaluation of all values' is made to mean that human beings' standard of valuation is no longer epistemic or moral, but physiological, this may simply amount to claiming that what is usually deemed good and true will cease to be so. Truthfulness and goodness will henceforth be determined by one's virility, rank, strength, health, etc., and therefore applied to things which were previously deemed evil or untrue. The main tenets of truth and morality will no longer be dictated by slavish or pious beings, but by strong and powerful ones. Yet although this version of the argument does reflect some of Nietzsche's specific formulations, it is still fairly superficial, as it affects only the *content*, and not the *form* of valuation. The oppositions truth/truth and good/evil have not been overcome, but merely translated in different ways.

This kind of reading is the one usually at work in most historical and political appropriations of Nietzsche. Indeed, many of the misunderstandings surrounding the notions of 'overhumanity', 'power' or 'nobility' have arisen from a superficial interpretation of Nietzsche's philosophical project, subordinated to very concrete political ends. In these cases, the 'transvaluation of all values' has been rather shy: the overcoming of truth and morality was motivated by very specific, albeit unconventional versions of the good and the true; 'new values', just like the old ones, were oriented towards a specific identification of the greatest good, or the best possible life.

A second, much more radical interpretation of Nietzsche's project consists in the attempt to overcome the attachment to the true and the good *altogether.* Only this version is faithful to Nietzsche's explicit intention of moving beyond teleology and purposiveness *as such*, and to free humankind from the slavishness of the 'will to truth' and the 'will to the good'. But here the stakes are of course much higher and the possibility of such a transformation is less easily fathomable. If every human judgement or decision entails, if only unconsciously or momentarily, the validation of a specific *version of things* (viz. a specific account of who one is, of where one stands, of what the world looks like, of what is at stake in a given situation, etc.), a life no longer supported

by such validation — or, which amounts to the same thing, a life where such version were lived as a mere version, *and nothing more* — would be stripped of its most basic practical references, and faced with the possibility of a complete existential paralysis.

Zarathustra voices this contradiction in the following terms:

> That I must be struggle and becoming and purpose and the contradiction of purposes — alas, whoever guesses my will guesses also on what *crooked* paths it must walk! Whatever I may create and however I may love it — soon I must oppose it and my love, thus my will wants it.[695]

In a new, 'overhuman' life, consciousness is faced with the same kind of abyss found at the edge of Plato's critique of human cognition: the pivotal moment when all human beliefs are perceived *as mere beliefs*, and thereby rejected. Afterwards, what lays in store for consciousness can only be guessed. What would life be like beyond the realm of belief? What form would knowledge and experience take? How free and powerful would one really be? For Nietzsche, as for Plato, these questions cannot be answered beforehand. They can only emerge as the result of a radical ontological transformation, whereby one's identity, the identity of others and the meaning of the whole world are changed into something else.

[695] *KSA* IV, 148; *TSZ*, 89f.

13.

CONCLUSIONS

The itinerary followed in the previous chapters has led all the way from the ancient Greek world to the dawn of postmodernity. Starting with Socrates and Plato, we have looked at some of the key moments of a long and complex philosophical debate, with concrete implications in every dimension of human life. What emerged from this survey, however, was not a clear or straightforward progression, leading to a clear or straightforward result. The examination of Plato's, Hegel's and Nietzsche's views on freedom and truth revealed an intricate web of theses, judgements and perspectives, irreducible to a single set of conclusions, or to a single line of interpretation.

The three projects we have considered are all grounded in the same general criticism. In all of them, human consciousness' usual account of reality is shown to be unreliable and in need of transformation, and in all of them this transformation is construed in a negative way, as the release from ignorance and self-deceit. According to Plato, Hegel and Nietzsche, only by eliminating all of the unwarranted assumptions underlying one's usual cognitive standpoint is one able to grasp reality's actual meaning. However, this task is dependent on a series of requirements that are anything but easy or self-evident. In order to break free from ignorance, one must be able not only to track down all of the assumptions active within one's consciousness, but also to cancel all of their effects. Moreover, for this operation to be successful, one must be able to guarantee that all assumptions have indeed been identified and cancelled, which leads to a challenging paradox: on the one hand, if this critical survey is not thorough enough, one's release from ignorance might simply amount to a new form of self-deceit, grounded in a

subsisting assumption or set of assumptions; on the other hand, since the latter are characterized precisely by their appearance of truth, there seems to be no way of making sure that ignorance and self-deceit have indeed been overcome.

None of the projects we have discussed was able to provide a full or definitive solution to this paradox. The release from ignorance envisaged, in different ways, by Plato, Hegel and Nietzsche, was haunted by the possibility of a doxastic or presuppositive residue, immune to philosophy's critical charges. Though all three authors directed their criticism at the entire edifice of human cognition, and though their methods succeeded in identifying and eliminating many of the unwarranted assumptions that guide human life, they were ultimately unable to ensure, in a final or unequivocal way, that no other assumption or set of assumptions was left in place.

Moreover, apart from the examination of each individual project, the comparison between Plato's, Hegel's and Nietzsche's critical approaches also revealed a number of ambiguities. Instead of a clear philosophical progression, or a clear argumentative path, we have come across a wide and complex labyrinth, where every thesis can be viewed from different angles, and every answer changed round into a question. Each project set up a global philosophical map, as it were, where the other two count as regional locations or possibilities. From Plato's perspective, Hegel's and Nietzsche's standpoints can be seen as sophisticated forms of δοξάζειν, anticipated, up to a point, by Socrates' enquiries; from Hegel's perspective, the Platonic and Nietzschean standpoints correspond to intermediate stages on the road to science and truth; from Nietzsche's perspective, Plato's and Hegel's philosophies amount to extreme forms of the 'will to truth', which must be criticized and overcome. In this complex dialogue, every interlocutor points in a different direction, guided by a different set of references, and the interpreter is left to decide which way to follow.

Finally, adding to these difficulties, the present survey was confined to a small operating field, amid a much wider philosophical universe. The three projects in question represent important conceptual landmarks, on which many others have drawn inspiration, but they are not immune to criticism. Although they propose global cognitive reforms, designed to replace all other methods and solutions, their approach is still a particular approach, whose universal aspirations are liable to be questioned and refuted.

For all of these reasons, the two main challenges raised in the beginning of this work are left unfulfilled. Firstly, the enquiry into the possibility of a truly insightful or educated standpoint, inspired by Plato's, Hegel's and Nietzsche's philosophies, has led to a negative result. The standpoint that has emerged from the enquiry is still exposed to ignorance and self-deceit, and still unable to guarantee the soundness of its judgements. Secondly, as a result of this, the enquiry into the possibility of a truly free or independent outlook on reality has also led to a negative result. The standpoint that has emerged from the enquiry is not yet a free or self-determining standpoint, but an imprisoned or enslaved one, still caught up in its own contradictions.

As we have seen earlier, in the first sections of the introduction, human freedom can be defined in different ways and depend on a wide array of requirements, including one's physical strength, material wealth, political autonomy or good fortune. But apart from these practical or contingent factors, freedom is also dependent on a cognitive dimension, which is no less important. In order to be free, one must be capable of choosing and acting according to one's interests, and this capacity requires, in turn, the ability to make sense of reality and to determine what is at stake in each practical situation. If one lacks these skills, one's freedom may easily become self-defeating: in the absence of a clear and reliable account of reality, one may end up mistaking what one wishes for what one really wants, or choosing against, rather according to, one's real interests.

Without the power to make informed and meaningful choices, there can be no real freedom. And yet this power is precisely what the present enquiry has found to be missing. If human consciousness is not in possession of a clear and reliable account of reality — or rather, if human consciousness, despite Plato's, Hegel's and Nietzsche's efforts, is unable to vouch for the clarity and reliability of its ideas and judgements — one's freedom is not real, but illusory. And what is more, if the power to make informed and meaningful choices underlies every other human power, this illusion is not partial, but complete. One may well be strong, or rich, or emancipated, or fortunate, but the freedom afforded by these qualities, unaccompanied by a suitable understanding of reality, cannot be taken for granted. For the former to be real and empowering, they must be recognized as such, accorded a specific value and put to a specific use. And since all of these operations are based on specific judgements, liable to be proven wrong, all of

these qualities are prone to misinterpretation and misuse. Instead of confirming one's freedom, they may amount to new sets of shackles, put in place by one's own hands.

Since human knowledge is unreliable, human life is not a free life. With this damning conclusion, we are brought back to the starting point. Philosophy's liberating efforts, in all their breadth and complexity, have failed to deliver, and Plato's initial diagnosis is once again confirmed. In light of this result, one might be tempted to follow Callicles' advice to Socrates and abandon the issue altogether. Instead of probing the foundations of one's ideas and judgements, in search of a solution that remains out of reach, one might simply accept their limitations and move on. But this attitude, as Socrates keeps pointing out, does not solve anything. Moreover, it can only be maintained at a price. Although one may disregard the contradictions of one's usual cognitive standpoint, these are no less real and no less serious. And though one may base one's entire life on instinct and common sense, all of one's choices and actions are still liable to contradict one's real interests. The price to pay for this voluntary ignorance is thus much higher than might be expected: in leaving behind philosophy's riddles, one may be gambling away what one values the most: namely, the possibility of a free and meaningful life.

SELECTED BIBLIOGRAPHY

1. *Introduction*[1]

AESCHYLUS, *Persians, Seven Against Thebes, Suppliants, Prometheus Bound*. Cambridge (MA), Loeb Classical Library, Harvard University Press, 2009

ARISTOTLE, *Physics* (2 vols.). Cambridge (MA), Loeb Classical Library, Harvard University Press, 1957-1934

———, *Poetics. Longinus: On the Sublime. Demetrius: On Style*. Cambridge (MA), Loeb Classical Library, Harvard University Press, 1995

EPICTETUS, *Discourses* (2 vols.). Cambridge (MA), Loeb Classical Library, Harvard University Press, 1956-1959

FICHTE, J. G., *Fichtes Werke*. Berlin, W. de Gruyter, 1971

FRÄNKEL, H., 'A Thought Pattern in Heraclitus', *American Journal of Philology* 59 (1938), 309-337

HADOT, P., *La Citadelle Intérieure*. Paris, Fayard, 1992

HORACE, *Satires, Epistles and Ars Poetica*. Cambridge (MA), Loeb Classical Library, Harvard University Press, 1966

JAEGER, W., *Paideia: The Ideals of Greek Culture* (3 vols.). Oxford, Oxford University Press, 1946

KANT, I., *Gesammelte Schriften*. Ed. Berlin-Brandenburgische Akademie der Wissenschaften. Berlin, Reimer/de Gruyter, 1900-

———, *Critique of Pure Reason*. Cambridge, Cambridge University Press, 1998

———, *Theoretical Philosophy*. Cambridge, Cambridge University Press, 2007

MARCUS AURELIUS, *Works*. Cambridge (MA), Loeb Classical Library, Harvard University Press, 1916

SCHILLER, F., *Sämtliche Werke* (eds. Fricke, G.; Göpfert, H.). München, Hanser, 1987

———, *Friedrich Schiller. Poet of Freedom*, vol.1. New York, Schiller Institute, 2015

SCHOPENHAUER, A., *Sämtliche Werke* (ed. Löhneysen, W.). Stuttgart/Frankfurt am Main, Cotta-Insel, 1965

[1] For the sake of coherence, Plato's, Hegel's and Nietzsche's writings, as well as all works directly related to them, will be referenced in sections 2, 3 and 4 of the bibliography, respectively.

SENECA, L. A., *Epistles 1-65*. Cambridge (MA), Loeb Classical Library, Harvard University Press, 1917

SOPHOCLES, *Ajax, Electra, Oedipus Tyrannus.* Cambridge (MA), Loeb Classical Library, Harvard University Press, 1994

———, *Antigone. The Women of Trachis. Philoctetes. Oedipus at Colonus*. Cambridge (MA), Loeb Classical Library, Harvard University Press, 1994

2. *Plato's Scientific Project*

ADAM, A. M., *Plato: Moral and Political Ideals.* Cambridge, Cambridge University Press, 2011

ALESSE, F., 'Alcuni esempi della relazione tra l'etica stoica e Platone', in Bonazzi, M.; Helmig C. (eds.), *Platonic Stoicism, Stoic Platonism*. Leuven, Leuven University Press, 2007, 23-40

ALEXANDRE, M., *Lecture de Platon*. Paris, Bordas/Mouton, 1968

ARISTOPHANES, *Birds, Lysistrata, Women at the Thesmophoria*. Cambridge (MA), Loeb Classical Library, 2000

———, *Clouds*, *Wasps*, *Peace*. Cambridge (MA), Loeb Classical Library, Harvard University Press, 1998

———, *Frogs, Assemblywomen, Wealth.* Cambridge (MA), Loeb Classical Library, Harvard University Press, 2002

———, *Fragments.* Cambridge (MA), Loeb Classical Library, Harvard University Press, 2008

BARNEY, R., 'Callicles and Thrasymachus', in Zalta, E. (ed.), *The Stanford Encyclopedia of Philosophy*, Winter 2011

BLUMENBERG, H., *Die Lesbarkeit der Welt*. Frankfurt am Main, Suhrkamp, 1986

———, *Das Lachen der Thrakerin. Eine Urgeschichte der Theorie*. Frankfurt am Main, Suhrkamp, 1987

BRISSON, L.; PRADEAU, J., *Dictionnaire Platon*. Paris, Ellipses, 2007

BROZE, M., 'Calliclès et Thrasymaque: sur la ruse et la violence chez Platon', *Revue de Philosophie Ancienne*, 9-1 (1991), 99-115

BURCKHARDT, J., *Grieschiche Kulturgeschichte*, Berlin/Stuttgart, Spemann, 1898-1902, re-ed. Darmstadt, Wissenschaftliche Buchgesellschaft, 1977

CARVALHO, M. J., 'Μέθοδος e ὑπόθεσις. O problema do pressuposto na fundação platónica da filosofia', in Ferrer, D. (ed.), *Método e Métodos*. Coimbra, Edições Universitárias de Coimbra, 2006

———, *Sobre a Prescindibilidade ou Imprescindibilidade do φρονεῖν. Um* Gedankenexperiment *no* Filebo *e no* Protréptico. Porto, Fundação Engenheiro António de Almeida, 2013

———, *Problemas Fundamentais de Fenomenologia da Finitude* – Tese de Doutoramento. Lisboa, 1996, 2 vols., XXIX + 1832 + 449 pp

CARVALHO, M. J. *et al.* (eds.), *In the Mirror of the Phaedrus*. Sankt Augustin, Academia Verlag, 2013

CICERO, *Orations: Pro Milone, In Prisonem, Pro Scauro, Pro Fonteio, Pro Rabirio Postumo, Pro Marcello, Pro Ligario, Pro Rege Deiotaro*. Cambridge (MA), Loeb Classical Library, Harvard University Press, 1989

CORNFORD, F. M., *Before and After Socrates*. Cambridge, Cambridge University Press, 1999

———, *Plato's Theory of Knowledge. The Theaetetus and the Sophist*. New York, Dover, 2003

DEUSCHLE, J., *Die Platonischen Mythen, insbesondere der Mythus im platonischen Phädrus*. Hanau, F. König, 1854

DIOGENES LAERTIUS, *Lives of Eminent Philosophers* (2 vols.). Cambridge (MA), Loeb Classical Library, Harvard University Press, 1991

EBERT, T., *Meinung und Wissen in der Philosophie Platons: Untersuchungen zum 'Charmides', 'Menon' und 'Staat'*. Berlin, W. de Gruyter, 1974

EURIPIDES, *Fragments. Aegeus-Meleager*. Cambridge (MA), Loeb Classical Library, Harvard University Press, 2008

FOGELIN, R. J., *Hume's Skepticism in the 'Treatise of Human Nature'*. London, Routledge and Kegan Paul, 1985

GOETHE, J., *Sämtliche Werke. Münchner Ausgabe*. München, btb, 2006

GOLDSCHMIDT, V., *Platonisme et Pensée Contemporaine.* Paris, Aubier-Montaigne, 1970

GOMPERZ, H., *Sophistik und Rhetorik. Das Bildungsideal des εὖ λέγειν in seinem Verhältnis zur Philosophie des V. Jahrhunderts*. Berlin, Teubner, 1912

GOMPERZ, T., *Grieschiche Denker. Eine Geschichte der Antiken Philosophie in 3 Bänden*. Frankfurt, Eichborn Verlag, 1996

GUTHRIE, W. K. C., *A History of Greek Philosophy*. Cambridge, Cambridge University Press, 1977

JOLY, R., *Le Thème Philosophique des Genres de Vie dans l'Antiquité Classique*, *Académie Royale de Begique, Memoires*. Classe des Lettres et des Sciences Morales et Politiques LI, 3, 1956

HAMANN, J. G., *Sokratische Denwürdigkeiten*, *Aesthetica in nuce.* Stuttgart, Reclam, 1986

HARTMANN, N., *Platos Logik des Seins*. Berlin, W. de Gruyter, 1965

HAVELOCK, E., *Preface to Plato*. Cambridge (MA), Harvard University Press, 1963

HEINIMANN, N., *Nomos und Physis. Herkunft und Bedeutung einer Antithese im Griechischen Denken des 5. Jahrhunderts.* Darmstadt, Wissenschaftliche Buchgesellschaft, 1965

HIRSCH, W., *Platons Weg zum Mythos*. Berlin, W. de Gruyter, 1971

HOMER, *Odyssey* (2 vols.). Cambridge (MA), Loeb Classical Library, Harvard University Press, 1919

HUME, D., *A Treatise of Human Nature*. Oxford, Oxford University Press, 2000

KERFERD, G. B., *The Sophistic Movement*. Cambridge, Cambridge University Press, 1981

KRISCHER, T., 'Die logischen Formen der Priamel', *Grazer Beiträge*, 2 (1974), 88-91

LAFRANCE, Y., *La Théorie Platonicienne de la Doxa.* Paris, Les Belles Lettres, 2014

LATONA, M., 'This Tale is Not My Own (οὐκ ἐμὸς ὁ μῦθος): Myth and Recollection in Plato', *Apeiron*, vol. 37, no. 3 (September 2004), 181-210

LEIBNIZ, G. W., *Die philosophischen Schriften von Gottfried Wilhlem Leibniz* (hrsg. C. I. Gerhradt). Berlin, Georg Olms Verlag, 1965

———, *Nouveaux Essais sur l'Entendement Humain.* Paris, Flammarion, 1990

———, *Correspondance avec Arnauld*. Paris, Vrin, 1993

———, *Principes de la Nature et de la Grâce. Monadologie*. Paris, Flammarion, 1996

———, *Discours de Métaphysique et Autres Textes (1663-1689)*. Paris, Flammarion, 2001

LIVINGSTON, D., *Philosophical Melancholy and Delirium: Hume's Pathology of Philosophy*. Chicago, University of Chicago Press, 1998

LONG, A. A., *Epictetus. A Stoic and Socratic Guide to Life*. Oxford, Oxford University Press, 2002

MENZEL, A., *Kallikles: Eine Studie zur Geschichte der Lehre vom Rechte der Stärkeren*. Wien, F. Deuticke, 1922

MULLER, R., *La Doctrine Platonicienne de la Liberté.* Paris, Vrin, 1997

NATORP, H., *Platos Ideenlehre*. Leipzig, Felix Meiner, 1903

NEHAMAS, A., *Virtues of Authenticity. Essays on Plato and Socrates*. Princeton, Princeton University Press, 1999

NOACK, K., *Platon und der Immoralismus: die Prototypen des extremen Naturrechts, Kallikles und Thrasymachos in der Darstellung Platons*. Bautzen, Lausitzer Druck- und Verglashaus, 2010

OHLERT, K., *Rätsel und Rätselspiele der alten Griechen.* Berlin, Mayer & Müller, 1912

PHILONENKO, A., *Leçons Platoniciennes*. Paris, Les Belles Lettres, 1997

PIEPER, J., *Über die Platonischen Mythen*. München, Kösel, 1965

PINDAR, *Olympian Odes*, *Pythian Odes*. Cambridge (MA), Loeb Classical Library, Harvard University Press, 1997

PLATO, *The Dialogues of Plato translated into English with Analyses and Introductions by B. Jowett, M.A., in Five Volumes.* 3rd edition revised and corrected. Oxford, Oxford University Press, 1892

———, *Platonis opera* (5 vols.). Burnet, J. (ed.). Oxford, Clarendon Press, 1900-1907 (repr. 1967-1968)

———, *Euthyphro, Apology, Crito, Phaedo, Phaedrus*. Cambridge (MA), Loeb Classical Library, Harvard University Press, 1914

———, *Theaetetus, Sophist.* Cambridge (MA), Loeb Classical Library, Harvard University Press, 1921

———, *Laches*, *Protagoras*, *Meno, Euthydemus.* Cambridge (MA), Loeb Classical Library, Harvard University Press, 1924

———, *Lysis, Symposium, Gorgias.* Cambridge (MA), Loeb Classical Library, Harvard University Press, 1925

———, *Statesman. Philebus. Ion.* Cambridge (MA), Loeb Classical Library, Harvard University Press, 1925

———, *Cratylus. Parmenides. Greater Hippias. Lesser Hippias.* Cambridge (MA), Loeb Classical Library, Harvard University Press, 1926

———, *Laws* (2 vols.). Cambridge (MA), Loeb Classical Library, Harvard University Press, 1926

———, *Charmides. Alcibiades I and II. Hipparchus. The Lovers. Theages. Minos. Epinomis.* Cambridge (MA), Loeb Classical Library, Harvard University Press, 1927

———, *Timaeus. Critias. Cleitophon. Menexenus. Epistles.* Cambridge (MA), Loeb Classical Library, Harvard University Press, 1929

———, *Gorgias. A Revised Text with Introduction and Commentary by E. R. Dodds.* Oxford, Oxford University Press, 1959

———, *Philèbe.* Paris, Les Belles Lettres, 1978

———, *Republic* (2 vols.). Cambridge (MA), Loeb Classical Library, Harvard University Press, 2013

PLUTARCH, *Plutarch's Lives*, vol. 5. Cambridge (MA), Harvard University Press, 1917

RIGINOS, A. S., *Platonica. The Anecdotes Concerning the Life and Writings of Plato.* Leiden, Brill, 1976

ROBIN, L., *La Pensée Grècque et les Origines de l'Esprit Scientifique.* Paris, Albin Michel, 1973

ROSEN, S., 'The Problem of Sense Perception in Plato's *Philebus*', in *Metaphysics in Ordinary Language*. New Haven, Yale University Press, 1999, 81–101

SCHMID, U., *Die Priamel der Werte im Griechischen von Homer bis Paulus*. Wiesbaden, Harassowitz, 1964

SCHMITZ, H. 'Physis versus Nomos. Platons politiktheoretische Auseinandersetzung mit Kallikles, Thrasymachos und Protagoras', *Zeitschrift für philosophische Forschung*, Bd. 42, H. 4 (Oct-Dec 1988), 570-596

SCHUHL, P., *La Fabulation Platonicienne*. Paris, Vrin, 1968

STEWART, J. A., *The Myths of Plato*. London, Macmillan, 1905

STRIKER, G., 'Plato's Socrates and the Stoics', in Waerdt, Paul A. Vander (ed.), *The Socratic Movement*. Ithaca, Cornell University Press, 1994, 241–251

TAYLOR, A. E., *Plato: The Man and His Work*. New York, Dover Publications, 2001

TRABATTONI, F., *Scrivere nell'Anima. Verità, Dialettica e Persuasione in Platone*. Firenze, La Nuova Italia (Pubblicazioni della Facoltà di Lettere e Filosofia dell'Università degli Studi di Milano, 154), 1994

TERZAGHI, N., 'La Scelta della Vita'. *Studi Italiani di filologia classica*, N. S. 1, 1920

VOEGELIN, E., *Plato.* Baton Rouge, Louisiana State University Press, 1981

VVAA., *Socrate. Tutte le Testimonianze, da Aristofane e Senofonte ai Padri Cristiani*. Roma, Biblioteca Universale Laterza, 1981

WINDELBAND, *Platon*. Stuttgart, Fromanns Verlag, 1910

WOLFSDORF, D., 'The Method ἐξ ὑποθέσεως at 'Meno' 86e1-87d8', *Phronesis*, vol. 53, no. 1 (2008), 35-64

———, *Trials of Reason. Plato and the Crafting of Philosophy*. Oxford, Oxford University Press, 2008

WOOLF, R., 'Callicles and Socrates: Psychic (Dis)harmony in the *Gorgias*', *Oxford Studies in Ancient Philosophy*, 18, 2000, 1-40

XENOPHON, *Memorabilia*, *Oeconomicus, Symposium, Apology*. Cambridge (MA), Loeb Classical Library, Harvard University Press, 2013

ZELTNER, H., 'Ideologie und Idee. Zum Kallikles-Gespräch (Platon, Gorgias 481 C ff.)', *Zeitschrift für philosophische Forschung*, Bd. 28, H. 1 (Jan-Mar 1974), 3-16

3. *Hegel's Phenomenological Project*

ADORNO, T., *Hegel. Three Studies.* Cambridge (MA), MIT Press, 1993

AMERIKS, K., 'The Hegelian Critique of Kantian Morality', in Moen, Marcia and Ouden, Bernard den (eds.), *New Essays on Kant*. New York, Peter Lang, 1987

BAUDELAIRE, C., *Oeuvres Complètes* (vol. 2). Paris, Michel Lévy Frères, 1868

BECKER, W., *Hegels Phänomenologie des Geistes. Eine Interpretation.* Stuttgart, Kohlhammer, 1971

BECKETT, S., *L'Innommable*. Paris, Minuit, 1953

BELAVAL, Y., *Études Leibnizieness: de Leibniz à Hegel.* Paris, Gallimard, 1993

———, *Leibniz Critique de Descartes.* Paris, Gallimard, 1960

BOER, K., *On Hegel: The Sway of the Negative*. New York, Palgrave Macmillan, 2010

BROWNING, G., *Hegel's Phenomenology of Spirit. A Reappraisal.* Dordrecht, Kluwer, 1997

BURZIO, P., *Lettura della Fenomenologia dello Spirito di Hegel.* Torino, UTET, 1996

CARVALHO, M. J., '"Verse Actuar a Sí Mismo". La Commedia en la Fenomenología del Espíritu', in Alonso, A.; Jiménez Redondo, M. (eds.), *Figuraciones Contemporáneas de lo Absoluto. Bicentenário de la 'Fenomenología del Espíritu' de Hegel (1807-2007)*. València, Universitat de València, 2009

CHIEREGHIN, F., *La 'Fenomenologia dello Spirito' di Hegel: Introduzione alla Lettura.* Roma, NIS, 1994

DILTHEY, W., *Hegel et Leibniz* (trans. Jean-Christophe Merle). Paris, Éditions du Cerf, 2002

DUDLEY, W., *Hegel, Nietzsche and Philosophy. Thinking Freedom*. Cambridge, Cambridge University Press, 2007

FERRO, B., 'De Leibniz a Hegel: Lei, Força e Infinitude na 'Fenomenologia do Espírito'', *Kairos: Revista de Filosofia e Ciência*, 7 (2013), 61-85

———, 'A Compreensão Hegeliana do Desejo (Begierde)', in Ferro, N.; Lima, P. (eds.), *Categorias Existenciais: Fim, Sentido, Ipseidade*, Coimbra, Instituto de Estudos Filosóficos, Universidade de Coimbra, 2016, 169-195

———, 'The Dialectics of Domination: Hegelian Echoes in Plato's 'Gorgias'', in Cervalho, M. J. Carvalho; Fidalgo, T. (eds.), *Plato's Gorgias: Labyrinth and Threads*. Coimbra, Instituto de Estudos Filosóficos, Universidade de Coimbra, 2016, 231-257

FLAY, J. C., 'Hegel's Inverted World', *Review of Metaphysics*, 23 (1970), 662-678

FORSTER, M. N., *Hegel's Idea of a Phenomenology of Spirit.* Chicago, Chicago University Press, 1998

FULDA, H. F., 'Geist, VIII: Der Begriff des Geistes bei Hegel und seine Wirkungsgeschichte', in Ritter, J. (ed.), *Historisches Wörterbuch der Philosophie*, Bd. III. Basel/Stuttgart, Schwabe & Co., 1971, 191-199

FULDA, H. F.; HEINRICH, D., *Materialien zu Hegels 'Phänomenologie des Geistes'*. Frankfurt am Main, Suhrkamp, 1973

GADAMER, H. G., *Hegels Dialektik. Sechs Hermeneutische Studien.* Tübingen, Mohr, 1980

GILBERT, B., *The Vitality of Contradiction. Hegel, Politics and the Dialectic of Liberal Capitalism*. Montreal, McGill-Queen's University Press, 2013

HARRIS, H. S., *Hegel's Development. Vol. I: Towards the Sunlight. 1770-1801*. Oxford, Oxford University Press, 1972

———, *Hegel's Development. Vol. II: Night Thoughts. Jena 1801-1806.* Oxford, Oxford University Press, 1983

———, *Hegel's Ladder*. Cambridge, Hackett Publishing Company, 1997

HEGEL, G. W. F., *Gesammelte Werke*. Edited by the Rheinisch-Westfälischen Akademie der Wissenschaften. Hamburg, Felix Meiner Verlag, 1968

———, *Werke in Zwanzig Bänden* (eds. E. Moldenhauer; K. M. Michel). Frankfurt am Main, Suhrkamp, 1986

———, *Phenomenology of Mind* (trans. J. B. Baillie). New York, Harper & Row, 1967

———, *Phenomenology of Spirit* (trans. A. V. Miller). Oxford, Oxford University Press, 1970

———, *Early Theological Writings* (trans. T. M. Knox). Philadelphia, University of Pennsylvania Press, 1971

———, *Faith and Knowledge* (trans. Walter Cerf and H. S. Harris). Albany, State University of New York Press, 1977

———, *Philosophy of Mind* (trans. W. Wallace and A. V. Miller). Oxford, Oxford University Press, 2007

———, *Outlines of the Philosophy of Right* (trans. T. M. Knox). Oxford, Oxford University Press, 2008

HEINRICHS, J., *Die Logik der 'Phänomenologie des Geistes'*. Bonn, Bouvier, 1974

HENRICH, D., 'Hegel und Hölderlin', in *Hegel im Kontext*. Frankfurt am Main, Suhrkamp, 1967, 9-40

HOBBES, T., *De Cive: The English Version* (ed. Howard Warrender). Oxford, Oxford University Press, 1984

———, *Leviathan* (ed. Noel Malcolm). Oxford, Oxford University Press, 2012

HOFFMANN, T. S. (ed.), "Hegel als Schlüsseldenker der Modernen Welt". *Beiträge zur Deutung der "Phänomenologie des Geistes" aus Anlaß ihres 200-Jahr-Jubiläums*. Hamburg, Felix Meiner Verlag, 2009

HONNETH, A., *Kampf um Anerkennung. Zur moralischen Grammatik sozialer Konflikte*. Berlin, Suhrkamp, 1994

———, *Das Ich im Wir: Studien zur Anerkennungstheorie*. Berlin, Suhrkamp, 2010

HORN, J. C., *Monade und Begriff. Der Weg von Leibniz zu Hegel*. Wien/München, Oldenburg, 1965

HOULGATE, S., *Hegel, Nietzsche and the Criticism of Metaphysics*. Cambridge, Cambridge University Press, 1986

HYPPOLITE, J., *Structure et Génèse de la Phénoménologie de l'Esprit.* Paris, Aubier, 1946

JAESCHKE, W., *Hegel-Handbuch.* Stuttgart, J. B. Metzler, 2003

JAMESON, F., *The Hegel Variations: On the Phenomenology of Spirit*. London, Verso, 2010

JANICAUD, D., *Hegel et le Destin de la Grèce*. Paris, Vrin, 1975

KAEHLER, K. E.; MARX, W., *Die Vernunft in Hegels Phänomenologie des Geistes.* Frankfurt am Main, Klostermann, 1992

KAINZ, H., *Hegel's Phenomenology of Spirit: Not Missing the Trees for the Forest*. Lanham, Lexington Books, 2008

KÖHLER, D; PÖGGELER, O., *G. W. F. Hegel: Phänomenologie des Geistes*. Berlin, Akademie-Verlag, 2006

KOJÈVE, A., 'Hegel, Marx et le Christianisme' [Compte rendu de Niel, H., 'De la Médiation dans la Philosophie de Hegel' (1945)], *Critique*, nº 3-4 (août-sept 1945), 339-366

———, *Introduction à la Lecture de Hegel. Leçons sur la Phénoménologie de l'Esprit professées de 1933 à 1939 à l'École des Hautes-Études.* Paris, Gallimard, 1971

LABARRIÈRE, P., *Structures et Mouvement Dialectique dans la "Phénoménologie de l'Esprit" de Hegel*. Paris, Aubier-Montaigne, 1968

———, *La Phénoménologie de l'Esprit de Hegel. Introduction à une Lecture.* Paris, Aubier-Montaigne, 1979

LABARRIÈRE, P.; JARCZYK, G., *Les Premiers Combats de la Reconnaissance. Maîtrise et Servitude dans la 'Phénomenologie de l'Esprit' de Hegel; Texte et Commentaire.* Paris, Aubier-Montaigne, 1987

———, *Hegel: Le Malheur de la Conscience ou l'Accès à la Raison. Liberté de l'Autoconscience; Stoïcisme, Scepticisme et la Conscience Malheureuse; Texte et Commentaire.* Paris, Aubier-Montaigne, 1989

LEBRUN, G., *La Patience du Concept. Essai sur le Discours Hégélien.* Paris, Gallimard, 1972

LUKÁCS, G., *Der junge Hegel.* Frankfurt am Main, Suhrkamp, 1973

LÖWITH, K., *Von Hegel zu Nietzsche. Der revolutionäre Bruch im Denken des 19. Jahrhunderts.* Hamburg, Felix Meiner, 2015

MARX, W., *Hegels Phänomenologie des Geistes: die Bestimmung ihrer Idee in 'Vorrede' und 'Einleitung'*. Franfurt am Main, Klostermann, 2006

———, *Die Selbstbewußtsein in Hegels Phänomenologie des Geistes.* Frankfurt am Main, Klostermann, 1986

MONETI, M., *Hegel e il Mondo alla Rovescia. Una Figura della Fenomenologia dello Spirito*. Florence, La Nuova Italia, 1986

NANCY, J., *Hegel: L'Inquiétude du Négatif.* Paris, Hachette, 1997

NEGELE, M., *Grade der Freiheit: Versuch einer Interpretation von G. W. F. Hegels 'Phänomenologie des Geistes'.* Würzburg, Königshausen & Neumann, 1991

NORTON, R., *The Beautiful Soul: Aesthetic Morality in the Eighteenth Century*. Ithaca, Cornell University Press, 1995

O'NEILL, J. (ed.), *Hegel's Dialectic of Desire and Recognition.* Albany, State University of New York Press, 1996

PAHA, B., *Die schöne Seele Hegels und die Literatur der Frühromantik*. München, Grin, 1992

PARRY, D., *Hegel's Phenomenology of the "We"*. New York, Peter Lang, 1988

PATTEN, A., *Hegel's Idea of Freedom*. Oxford, Oxford University Press, 1999

PINKARD, T., *Hegel's Dialectic: The Explanation of Possibility.* Philadelphia, Temple University Press, 1988

———, *Hegel's Phenomenology: The Sociality of Reason*. Cambridge, Cambridge University Press, 1994

PIPPIN, R., *Idealism as Modernism. Hegelian Variations*. Cambridge, Cambridge University Press, 1997

———, *Hegel on Self-consciousness: Desire and Death in the 'Phenomenology of Spirit'*. Cambridge, Cambridge University Press, 2011

PÖGGELER, O., *Hegels Idee einer Phänomenologie des Geistes.* Freiburg, Alber, 1973

PROUST, M., *À la Recherche du Temps Perdu*. Paris, Bibliothèque La Pléiade, Gallimard, 1954

RAUCH, L.; SHERMAN, D., *Hegel's Phenomenology of Self-consciousness. Text and Commentary.* Albany, State University of New York Press, 1999

RÖTTGES, H., *Der Begriff der Methode in der Philosophie Hegels.* Meisenheim am Glan, Hain, 1981

———, *Dialektik und Skeptizismus: die Rolle des Skeptizismus für Genese, Selbstverständnis und Kritik der Dialektik.* Frankfurt am Main, Athenäum, 1987

ROUSSEAU, J.-J., *Du Contrat Social, ou Principes du Droit Politique*. Paris, Garnier Frères, 1954

SCHEIER, C., *Analytischer Kommentar zu Hegels Phänomenologie des Geistes. Die Architektonik des erscheinenden Wissens.* Freiburg/München, Alber, 1980

SICKENDIECK-MEYER, B., *Affektpoetik: Eine Kulturgeschichte litterarischer Emotionen*. Würzburg Königshausen & Neumann, 2005

SIEP, L., *Anerkennung als Prinzip der praktischen Philosophie. Untersuchungen zu Hegels Jenaer Philosophie des Geistes*. Freiburg/München, Alber, 1979

———, 'The *Aufhebung* of Morality in Ethical Life', in Stepelevich, L. S.; Lamb, D. (eds.), *Hegel's Philosophy of Action*. Atlantic Highlands, Humanities Press, 1983

———, *Der Weg der 'Phänomenologie des Geistes'. Ein einführender Kommentar zu Hegels 'Differenzschrift' und zur 'Phänomenologie des Geistes'*. Frankfurt am Main, Suhrkamp, 2001

SMITH, S. B., *Hegel's Critique of Liberalism*. Chicago, University of Chicado Press, 1991

SOLOMON, R., *In the Spirit of Hegel.* Oxford, Oxford University Press, 1985

STERN, R., *Hegel and the Phenomenology of Spirit. Routledge Philosophy Guidebook*. New York, Routledge, 2002

STEWART, J., *The Unity of Hegel's Phenomenology of Spirit*. Evanston, Northwestern Universtity Press, 2000

TAMINIAUX, J., *La Nostalgie de la Grèce à l'Aube de l'Idéalisme Allemand. Kant et les Grecs dans l'Itinéraire de Schiller, de Hölderlin et de Hegel*. The Hague, Martinus Nijhoff, 1967

TAYLOR, C., *Hegel*. Cambridge, Cambridge University Press, 1975

VIEWEG, K., *Hegels Phänomenologie des Geistes: ein kooperativer Kommentar zu einem Schlüsselwerk der Moderne*. Frankfurt am Main, Suhrkamp, 2008

WAHL, J., *Le Malheur de la Conscience dans la Philosophie de Hegel.* Paris, Presses Universitaires de France, 1951

WALSH, W. H., *Hegelian Ethics.* London, Macmillan, 1969

WILLIAMS, R. R., *Recognition: Fichte and Hegel on the Other.* Albany, State University of New York Press, 1992

———, 'Hegel and Nietzsche: Recognition and Master/Slave', *Philosophy Today*, vol. 45, issue supplement (2001), 164-179

———, *Tragedy, Recognition, and the Death of God: Studies in Hegel and Nietzsche*. Oxford, Oxford University Press, 2012

WOOD, A., *Hegel's Ethical Thought*. Cambridge, Cambridge University Press, 1990

4. *Nietzsche's Metacognitive Project*

ABEL, G., *Nietzsche: Die Dynamik der Willen zur Macht und die ewige Wiederkehr*. Berlin, W. de Cruyter, 1984

ALLISON, D., *The New Nietzsche. Contemporary Styles of Interpretation*. Cambridge (MA), M.I.T. Press, 1995

BEARDSWORTH, R., *Nietzsche*. Paris. Les Belles Lettres, 1997

BITTNER, R.. 'Ressentiment', in Schacht, R. (ed.), *Nietzsche, Genealogy, Morality. Essays On Nietzsche's Genealogy of Morals*. Berkeley, University of California, 1994, 127-138

BLANCHOT, M., *L'Entretien Infini*. Paris, Gallimard, 1969

BOTET, S., *Le Zarathoustra de Nietzsche: Une Refonte du Discours Philosophique?*. Paris. Klincksieck, 2006

CAMUS, A., *Essais*. Paris, Bibliothèque de la Pléiade, Gallimard, 1965

CLARK, M., *Nietzsche on Truth and Philosophy*. New York, Cambridge University Press, 1990

COLLI, G., *Dopo Nietzsche*. Milano, Adelphi, 1974

CONSIGNY, S., 'Nietzsche's Reading of the Sophists'. *Rhetoric Review*, vol. 13, no. 1 (Autumn 1994), 5-26

CRÉPON, M., *Nietzsche: L'Art et la Politique de l'Avenir*. Paris, Presses Universitaires de France, 2003

DANTO, A. C., *Nietzsche as Philosopher*. New York. Columbia University Press, 2005

DELEUZE, G., *Nietzsche*. Paris, Presses Universitaires de France, 1977

———, *Nietzsche et la Philosophie*. Paris, Presses Universitaires de France, 1962

DIELS, H.; KRANZ, W. (eds.), *Die Fragmente der Vorsokratiker*. Berlin, Weidmann, 1952

GEMES, J.; MAY, S. (eds.), *Nietzsche on Freedom and Autonomy*. Oxford, Oxford University Press, 2009

GREEN, M. S., *Nietzsche and the Transcendental Tradition*. Urbana, University of Illinois Press, 2002

GRIMM, R. H., *Nietzsche's Theory of Knowledge*. New York, W. de Gruyter, 1977

HALÉVY, D., *Nietzsche*. Paris, Bernard Grasset, 1944

HEIDEGGER, M., *Nietzsche* (2 Bde.). Pfullingen, Günther Neske, 1961

HILL, R. K., *Nietzsche's Critiques: The Kantian Foundations of his Thought*. Oxford, Oxford University Press, 2003

JASPERS, K., *Nietzsche und das Christentum*. München, Piper, 1952

———, *Nietzsche. Einführung in das Verständnis seines Philosophierens*. Berlin, W. de Gruyter, 1981

JENSEN, A. K., *Nietzsche's Philosophy of History*. Cambridge, Cambridge University Press, 2013

KAUFMANN, W., *Nietzsche: Philosopher, Psychologist, Antichrist*. Princeton, Princeton University Press, 2013

KLOCH-KORNITZ, P., 'Der Gorgias Platons und die Philosophie Friedrich Nietzsches', *Zeitschrift für philosophische Forschung*, Bd. 17, H. 4 (Oct-Dec 1963), 586-603

KLOSSOWSKI, P., *Nietzsche et le Cercle Vicieux*. Paris, Mercure de France, 1978

LEYRA, I. P., 'Calicles y Nietzsche: Dos Formas Contrapuestas de Entender el Derecho del Más Fuerte', *Nexo. Revista de Filosofía*, núm. 2 (2004), 167-176

LONG, A.; SEDLEY, D. N., *The Hellenistic Philosophers* (2 vols.). Cambridge, Cambridge University Press, 1990

LOW, R., *Nietzsche, Sophist und Erzieher: Philosophische Untersuchungen zum systematischen Ort von Friedrich Nietzsches Denken*. Weinheim, Acta Humaniora, 1984

MIGOTTI, M., 'Slave Morality, Socrates, and the Bushmen: a Reading of the First Essay of the Genealogy of Morals', *Philosophy and Phenomenological Research*, vol. LVIII, nr. 4 (Dec 1998), 745–779

MITCHESON, K., *Nietzsche, Truth and Transformation*. New York, Palgrave Macmillan, 2013

MÜLLER-LAUTER, W., 'Nihilismus als Konsequenz des Idealismus', in Schwan, A. (ed.), *Denken im Schatten des Nihilismus. Festschrift für Wilhelm Weischedel zum 70. Geburtstag am 11. April 1975.* Darmstadt, Wiss. Buchges, 1975, 113-163

———, *Über Werden under Wille zur Macht: Nietzsche-Interpretationen I*. Berlin, W. de Gruyter, 1999

———, *Über Freiheit und Chaos: Nietzsche-Interpretationen II*. Berlin, W. de Gruyter, 1999

NEHAMAS, A., *Nietzsche: Life as Literature*. Cambridge, Harvard University Press, 1985

NIETZSCHE, F., *The Will to Power* (trans. Walter Kaufmann and R. J. Hollingdale). New York, Vintage, 1967

———, *Kritische Studienausgabe* (eds. G. Colli; M. Montinari). Berlin, W. de Gruyter, 1988 (2nd revised ed.)

———, *Human, All-too Human*: *A Book for Free Spirits* (trans. R. J. Hollingdale). Cambridge, Cambridge University Press, 1996

———, *The Birth of Tragedy and Other Writings* (trans. Ronald Speir). Cambridge, Cambridge University Press, 1999

———, *The Gay Science: With a Prelude in German Rhymes and an Appendix of Songs* (trans. Josefine Nauckhoff). Cambridge, Cambridge University Press, 2001

———, *Beyond Good and Evil* (trans. Judith Norman). Cambridge, Cambridge University Press, 2002

———, *Sämtliche Briefe. Kritische Studienausgabe in acht Bänden* (eds. G. Colli; M. Montinari). Berlin, W. de Gruyter, 2003

———, *The Antichrist, Ecce Homo, Twilight of the Idols* (trans. Judith Norman). Cambridge, Cambridge University Press, 2005

———, *On the Genealogy of Morality* (trans. Carol Diethe). Cambridge, Cambridge University Press, 2006

———, *Thus Spoke Zarathustra* (trans. Adrian del Caro). Cambridge, Cambridge University Press, 2006

PÜTZ, P., *Friedrich Nietzsche*. Stuttgart, Metzler, 1967

RICHARDSON, J., *Nietzsche's System*. New York, Oxford University Press, 1996

SCHABERG, W., *The Nietzsche Canon: A Publication History and Bibliography*. Chicago, University of Chicago Press, 1995

SCHRÖDER, W., *Moralischer Nihilismus: Radikale Moralkritik von den Sophisten bis Nietzsche*. Stuttgart, Reclam, 2005

SEDGWICK, P. R. (ed.), *Nietzsche: A Critical Reader*. Oxford, Blackwell. 1995

SIMMEL, G., *Schopenhauer und Nietzsche*. Leipzig, Duncker & Humblot, 1907

STEGMAIER, W., *Nietzsches Befreiung der Philosophie: Kontextuelle Interpretation des V. Buchs der 'Fröhlichen Wissenschaft'*. Berlin, W. de Gruyter, 2012

URSTAD, K., 'Nietzsche and Callicles on Happiness, Pleasure, and Power', *Kritike*, vol. 4, nr. 2 (Dec 2010), 133-141

VINCENZO, J., 'Socrates and Rhetoric: The problem of Nietzsche's Socrates', *Philosophy and Rhetoric*, 25-2 (1992), 162-182

WALLACE, R. J., 'Ressentiment, Value, and Self-Vindication: Making Sense of Nietzsche's Slave Revolt', in Leiter, B.; Sinhababu, N. (eds.), *Nietzsche and Morality*. Oxford, Oxford University Press, 2007, 110–137

WHITE, R. J., *Nietzsche and the Problem of Sovereignty*. Urbana, University of Illinois Press, 1997

WILCOX, J. T., *Truth and Value in Nietzsche: A Study of his Metaethics and Epistemology*. Ann Arbor, University of Michigan Press, 1974

WILLIAMS, B., *Truth and Truthfulness: An Essay in Genealogy*. Princeton, Princeton University Press, 2004

INDEX NOMINUM

(Socrates, Plato, Hegel and Nietzsche were not included in the following list, as they are continually mentioned throughout this work.)

Abraham, 273
Achilles, 158
Aeschylus, 19, 305n.478
Aesop, 373
Adorno, Theodor, 204n.306, 248n.383
Alcibiades of Scambonidae, 116-118, 117n.149, 164
Antigone, 305, 319
Apollo, 371, 374-376, 379-380, 383-384
Archimedes of Syracuse, 110-111, 111n.138, 111n.140
Aristophanes, 111, 111n.140, 113, 323
Aristotle, 3-4, 11-13, 13n.12, 320
Augustine of Hippo, 282n.437

Baudelaire, Charles, 323
Baumgarten, Alexander Gottlieb, 2n.1
Beckett, Samuel, 271n.421

Callicles of Acharnae, 112-114, 114n.146, 116-119, 124, 126n.169, 127-131, 128n.175, 153, 159, 263-264, 392, 418, 420-424, 420n.685, 430
Camus, Albert, 382
Chrysippus, 29
Creon, 305, 305n.478, 319

Danton, Georges, 311
Darwin, Charles, 343, 343n.543, 413
Descartes, René, 226
Diderot, Denis, 308n.480
Diogenes Laertius, 131n.186
Dionysos, 52, 335, 374-376, 375n.604, 379, 388

Engels, Friedrich, 256
Epictetus, 29, 263n.407
Eteocles, 305n.478
Euripides, 114n.146, 159, 159n.229, 376

Fichte, Johann Gottlieb, 54-55, 238n.365, 247n.382, 248n.382, 287
Fränkel, Hermann, 20, 134, 342, 343n.543

Gadamer, Hans-Georg, 243
Gall, Franz Joseph, 290
Glaucon, 127-128, 170
Goethe, Johann Wolfgang von, 75, 78, 243, 315n.496
Gorgias of Leontini, 106, 119-121, 123, 131,

Habermas, Jürgen, 248n.383
Heidegger, Martin, 59, 89, 252n.387, 334
Heraclitus of Ephesus, 20, 28, 135, 138, 154n.221, 155, 272, 342, 343n.543
Herder, Johann Gottfried von, 303n.475
Hippias of Elis, 106
Hobbes, Thomas, 252n.388, 301, 354, 358n.577

Hölderlin, Friedrich, 304n.475
Homer, 158-159, 403
Honneth, Axel, 247, 248n.383
Horace, 7
Hume David, 165-166, 165n.236, 175
Husserl, Edmund, 59, 89, 334

Isaiah (prophet), 282n.437

Jesus Christ, 273, 275-277, 282n.437, 326
Job, 274
Jocasta, 12n.8

Kant, Immanuel, 2n.1, 35, 54-58, 55n.72, 73, 89, 143n.205, 181, 192-193, 200, 229-230, 287, 298-299, 299n.465, 300n.465, 304n.475, 311-313, 312n.487, 336, 347, 360n.578, 407
Kierkegaard, S., 252n.387, 334
Kojève, Alexandre, 248, 248n.383

Lavater, Johann Kaspar, 290
Leibniz, Gottfried Wilhelm, 80n.97, 81n.98, 145n.207, 208, 226-227, 228n.346, 282n.437, 407
Löwith, Karl, 335
Luther, Martin, 407

Marcellus, Gaius Claudius, 111
Marcus Aurelius, 29, 263n.407
Marx, Karl, 248n.383, 256, 258-259, 334n.534, 335n.534
Meier, Georg Friedrich, 2n.1
Moses, 273
Münchhausen, Baron of, 3

Napoleon Bonaparte, 406
Natorp, Paul, 145n.207

Odysseus, 117, 157-158
Oedipus, 10-14, 12n.8, 13n.8, 24-25, 305n.478

Parmenides, 141n.198, 236, 272
Pascal, Blaise, 123n.169
Penelope, 176
Pericles, 117
Pindar, 157-158
Plutarch, 110, 111n.138
Polus of Acragas, 119, 123-127, 126n.169, 130, 392
Polynices, 305, 305n.478
Prodicus of Ceos, 106
Prometheus, 19-20, 25, 35
Protagoras of Abdera, 19, 106
Protarchus of Athens, 69, 87
Proust, Marcel, 189
Pygmalion, 359
Pyrrho, 200, 353

Robespierre, Maximilien, 311, 294n.453
Rousseau, Jean-Jacques, 293, 293n.453, 301, 310-311, 315n.496

Saint-Just, Louis Antoine de, 311
Schelling, Friedrich Wilhelm J. von, 54-55
Schiller, Friedrich von, 4-5, 5n.5, 7, 303n.475, 371n.595
Schopenhauer, Arthur, 28, 57-58, 334, 373, 375, 375n.604, 407
Seneca, Lucius Annaeus, 29, 176
Shakespeare, William, 125n.168
Sophocles, 10-14, 12n.8, 19, 25, 114n.146, 157-158, 305
Spinoza, Baruch, 413
Stirner, Max, 334
Strepsiades, 111n.140
Suetonius, 125n.168

Thales of Miletus, 110, 111n.140, 113, 274

Theodorus of Cyrene, 104
Thomas Aquinas, 282n.437
Thomas, Saint, 276
Thrasymachus of Calchedon, 106, 127-128, 128n.175
Timothy, Saint, 282n.437
Tiresias, 13, 158
Tisias of Sicily, 106

Vitruvius, 110n.137

Xenophon, 125n.168

Zarathustra, 52, 341-342, 344-345, 350, 353, 370, 426
Zeus, 145n.208, 157