Kairos

Revista de Filosofia & Ciência
Journal of Philosophy & Science

Julho | July 2015

13

Artigos
Papers

Scientific Realism, Adaptationism and the Problem of the Criterion
Fabio Sterpetti

Ortega y Gasset on Natural Selection and the Vocation of Man: A Rejection or Elaboration of Darwinism?
Rabi Lior

Les présupposés théologiques de la conscience historique moderne : une lecture löwithienne du sens de l'histoire
Laura P. Kassar

Kant e a fundamentação da moral
Sérgio Mascarenhas

Dossier:
Imagética Atomística

Is there such a thing as 'Greek Atomism'?
Zbigniew Kotowicz

Models and Molecules: Representation in the Work of John Dalton
Rachel Dunn

La physique laplacienne dans la seconde moitié du XIXe siècle: Joseph Boussinesq – la pratique et la réflexion autour de l'atomisme en France vers 1875
João Príncipe

How to Conceive the Atom: Imagery vs. Formalism
Giora Hon & Bernard R. Goldstein

Observing stars, representing atoms: images and objectivity in the physical sciences
Eduardo Zubia

Kairos. Revista de Filosofia & Ciência

Kairos. Journal of Philosophy & Science

ISSN: 1647-659X

Edição: Centro de Filosofia das Ciências da Universidade de Lisboa

Índice | Index

ARTIGOS | PAPERS

Scientific Realism, Adaptationism and the Problem of the Criterion 7
Fabio Sterpetti

Ortega y Gasset on Natural Selection and the Vocation of Man: A Rejection or Elaboration of Darwinism?
Rabi Lior 47

Les présupposés théologiques de la conscience historique moderne : une lecture löwithienne du sens de l'histoire
Laura P. Kassar 83

Kant e a fundamentação da moral
Sérgio Mascarenhas 111

DOSSIER: Imagética Atomística – Atomistic Imagery 143
Organizado por Lídia Queiroz

Is there such a thing as 'Greek Atomism'?
Zbigniew Kotowicz 145

Models and Molecules: Representation in the Work of John Dalton
Rachel Dunn 157

La physique laplacienne dans la seconde moitié du XIXe siècle: Joseph Boussinesq – la pratique et la réflexion autour de l'atomisme en France vers 1875
João Príncipe 179

How to Conceive the Atom: Imagery vs. Formalism
Giora Hon & Bernard R. Goldstein 213

Observing stars, representing atoms: images and objectivity in the physical sciences
Eduardo Zubia 237

Artigos | Papers

Scientific Realism, Adaptationism and the Problem of the Criterion

Fabio Sterpetti
(Sapienza University of Rome)
fabio.sterpetti@uniroma1.it

Abstract Scientific Realism (SR) has three crucial aspects: 1) the centrality of the concept of truth, 2) the idea that success is a reliable indicator of truth, and 3) the idea that the Inference to the Best Explanation is a reliable inference rule. It will be outlined how some realists try to overcome the difficulties which arise in justifying such crucial aspects relying on an adaptationist view of evolutionism, and why such attempts are inadequate. Finally, we will briefly sketch some of the main difficulties the realist has to face in defending those crucial aspects, and how such difficulties are deeply related: they derive from the inability of SR to satisfyingly avoid the sceptical challenge of the criterion of truth. Indeed, SR seems not to be able to fill the so-called 'epistemic gap' (Sankey 2008). In fact, the epistemic gap cannot be filled in no way other than obtaining a criterion of truth, but such a criterion cannot be obtained if the epistemic gap obtains.

Resumo O Realismo Científico (RC) tem três características nucleares: (1) a centralidade do conceito de verdade; (2) a ideia de que o sucesso é um indicador fiável de verdade; e (3) a ideia de que a inferência para a melhor explicação é uma regra segura de inferência. Neste artigo mostraremos como alguns realistas tentam superar as dificuldades suscitadas pela justificação daquelas três características, à luz de uma concepção adaptacionista da evolução, e por que razão tais tentativas nos parecem ser inadequadas. Finalmente, descreveremos brevemente algumas das principais dificuldades que os realistas enfrentam quando defendem as três características mencionadas, e como tais dificuldades estão intimamente relacionadas: elas derivam da incapacidade do RC em evitar, de um modo satisfatório, o desafio cepticista do critério de verdade. O RC parece-nos ser, de facto, incapaz de superar o chamado 'hiato epistémico' (Sankey 2008). Na verdade, o hiato epistémico não pode ser superado sem a assunção de um critério de verdade, mas um tal critério não pode ser obtido se o hiato epistémico se verificar.

1. Scientific Realism

It seems to be fair to say that Scientific Realism (SR) has, at least in its mainstream formulations, three crucial aspects: 1) the centrality of the concept of truth, 2) the idea that success is a reliable indicator of truth, and 3) the idea that the Inference to the Best Explanation is a reliable inference rule. It is impossible to account here for all the realist positions which have been elaborated so far, but it is important to underline that there are several formulations of SR which are cast in different terms, and which are not based on all (or, any of) the aspects outlined above. For example, Devitt denies that SR has to be understood in terms of truth.[1] In what follows we will focus on what can be labeled a 'standard' realist position, i.e. a position which is based on all the three aspects described above, and our considerations will be limited to such kind of SR. Thus, throughout this article by 'SR' we will mean 'standard scientific realism'.

1.1. Scientific Realism, Truth, and The No Miracle Argument

SR can be briefly described as the claim that our best scientific theories are true. As Saatsi and Vickers state: "scientific realists seek to establish a link between theoretical truth and predictive success."[2]

The concept of truth is central for SR. For example, Giere states that: "virtually every characterization of scientific realism I have ever seen has been framed in terms of truth."[3]

The most shared view of truth among the realists[4] is that of *truth as correspondence*.[5] For example, Sankey states that: "correspondence

[1] Devitt, 1997.
[2] Saatsi and Vickers, 2011, 29.
[3] Giere, 2005, 154.
[4] In what follows, for brevity, 'realist' will be used in place of 'scientific realist'.

theories which treat truth as a relation between language and reality are the only theories of truth compatible with realism."[6, 7]

The main argument to sustain SR is the No Miracle Argument (NMA).[8] The central idea of the NMA is that the truth of a scientific theory is the best, or the only scientifically acceptable, explanation of its empirical success.

The problem is that success seems not to be an adequate indicator for truth: it is not easy to support the idea that only truth can lead to success.[9] In fact, claiming that the success of a theory is due to its being true would imply that such theory should not be radically modified over time or ever considered false. If only truth implies success, then there could not be a theory which is empirically successful and false. But the history of science seems not to allow us to support such a claim.[10] Saatsi summarizes this line of reasoning in the form of a *reductio ad absurdum* as follows:

---

[5] On the necessity for the realists to aim at a substantive theory of truth, and on the relation between such a kind of theory of truth and the correspondence view of truth, see Sher, 2004.

[6] Sankey, 2008, 17.

[7] Many positions have been elaborated on the issue of truth (see for a survey Burgess and Burgess, 2011). Thus, even if truth as correspondence seems to be the most widespread view among the realists, not any realist adopts such view. Here we will focus on correspondence, but some of the objections we will deal with can be formulated even with respect to other conceptions of truth.

[8] Putnam, 1975; Psillos, 1999.

[9] On why success is not a reliable indicator of truth see Wray, 2013. Against the claim that only true theories can account for novel predictive success see Stanford, 2000.

[10] See, e.g., Laudan, 1981. Cf. also Worrall, 2008, 287: "The chief obstacles to this view [SR] are precisely those posed by the facts about theory-change in science. If we accept that earlier theories in the history of science were quite radically false and yet enjoyed striking predictive success, then it can scarcely be claimed that it would be a miracle if present theories enjoyed the success they do and yet were not even approximately true." For a detailed analysis of a historical case of successful and false theory see Saatsi and Vickers, 2011. On the relation between success and truth see also Held, 2011.

> (1) Assume that success of a theory is a reliable test for its truth. (2) So most current successful scientific theories are true. (3) Then most past scientific theories are false, since they differ from current successful theories in significant ways. (4) Many of these past theories were also successful. (5) So successfulness of a theory is not a reliable test for its truth.[11]

It has been argued that it is unfair to claim that it is necessary to show that *only* the truth implies success, if we want to rely on success in order to claim for the truth of our successful theories. Indeed, success should be understood as a *good* indicator of truth in the same sense in which any indicator is taken to be 'good' in scientific practice, i.e. accordingly to its 'rate of success'. In this perspective, the reliability of an indicator has to be cast in statistical terms.[12] If this approach is right, then it would be pointless to underline that there are some cases in which success has not been a good indicator of truth, because, on average, we nevertheless could claim that success is a good indicator of truth.

The problem with this approach is that in order to show that an indicator is 'good' we have to show that the rates of 'false positives' and 'false negatives' that it produces are both small.[13] But to do that, we should construct four sets: that of the false positives, that of the false negatives, that of the true positives, and that of the true negatives. Comparing the dimension of each set with the totality of the theories considered, we could derive the different rates we are interested in. In the context we are dealing with, false positives are 'successful but false theories', false negatives are 'true but unsuccessful theories', true positives are theories that are both 'true and successful', and true negatives are theories that are both 'false and unsuccessful'.

[11] Saatsi, 2005, 1089.
[12] Lewis, 2001.
[13] Lewis, 2001, 374-375.

There are two main problems when trying to construct such sets: 1) the very notions of 'false negative' and 'true negative' are not clear at all in this context. For example, Saatsi states that: "I can make no sense of the idea of delineating a non-arbitrary, well-defined collection of *both false and unsuccessful* theories (…) or true yet unsuccessful;"[14] 2) in order to construct the four sets mentioned above, we have to determine in which set each theory has to be put, i.e. we should already know which theory is true. The fact is that in the ordinary clinical practice, the context from which Lewis takes the analogy between the measure of the reliability of the empirical success of the theories and the measure of the reliability of the medical tests (think, e.g., to the pregnancy tests), this can easily be done, because there are several ways of assessing the rate of success of the indicator we want to evaluate which are independent from that indicator (we can easily check how many times a pregnancy test has been successful without having to rely on that test). But when dealing with the debate on the truth of our best scientific theory, there is not a way of assessing the correctness of all the theories we are evaluating independent from our *best* theory. But we take our *best* theory to be *true* exactly because it is *the most successful*. As Wray states, in order to infer that our theories are likely true it is not sufficient to state that "*if* the false positive and false negative rates for our test are low, and *if* most of our current theories are successful," then our best theories will likely be true, we "also need to know that most successful theories are true."[15] Thus, there is not an independent way of evaluating our indicator, i.e. there is not a way of justifying the claim that success indicates truth without circularly assuming that very claim.

This is a very debated issue, and cannot fully presented here, but this brief digression is just aimed at underlining that there is not an

[14] Saatsi, 2005, 1096.
[15] Wray, 2013, 1727.

easy way to coherently develop and defend the intuitive idea that there is no need to defend the claim that only truth leads to success if we want to maintain the realist claim that success is a good indicator of truth.

1.2. The NMA and the Inference to the Best Explanation

The importance of the Inference to the Best Explanation (IBE)[16] for the realists is due to the fact that IBE is the kind of inference that, if valid, would allow the realists to derive the truth of the confirmed theory from the empirical success of such theory, i.e. to support the core tenet of the realist position.[17] Whether an author accepts IBE or not has been considered to be roughly equivalent to whether she embraces realism or not.[18] Indeed, the NMA can be described as an IBE.[19]

---

[16] Lipton, 2004. IBE is a kind of inference introduced by Harman, 1965, but already considered by Peirce (Kapitan, 1992, 2; Psillos, 2011a, 128), which can be described as: "one infers, from the premise that a given hypothesis would provide a 'better' explanation for the evidence than would any other hypothesis, to the conclusion that the given hypothesis is true" (Harman, 1965, 89).

[17] The relevance of IBE for the realists is widely recognised. Cf., e.g., Douven, 2002, 355, who states that: "defenses of scientific realism typically rely on Inference to the Best Explanation."

[18] Cf. Hitchcock, 1992, 151: "The principal link between the scientific realism debate and competing theories of scientific explanation is provided by a family of inference patterns which march under the slogan 'inference to the best explanation'."

[19] Psillos, 1999. Frost-Arnold, 2010, 45, given that the basic form of IBE is:
(1)*q* is the best explanation of *p*
(2)*p*
(3)*q*
describes the NMA as an IBE in the way that follows:

> (1)The (approximate) truth of mature scientific theories is the best explanation of their empirical success.
> (2)Mature scientific theories are empirically successful.
> (3)Mature scientific theories are (approximately) true.

So, if to support SR the realists rely on the NMA, they have to face the problems related to the nature of IBE, i.e. to the nature of abduction, given that Harman states that the "'inference to the best explanation' corresponds approximately to what others have called 'abduction',[20] and that IBE can be seen as a generalization of abduction.[21, 22]

Another relevant feature of this abductive realist line of reasoning is that it seems able to secure the reliability of the IBE abductively, relying on a sort of benign circularity. For example, Psillos states that: "It is transparent that the NMA has two conclusions," the first is that "we should accept as (approximately) true the theories that are implicated in the (best) explanation," and the second is that "since,

[20] Harman, 1965, 88.

[21] See Cellucci, 2013, § 18.7. Different positions on the relation between abduction and IBE have been adopted. Mackonis sketches the situation and presents his personal view as follows: "Some researchers do not conceptually discriminate between IBE and abduction or use the term 'abduction' as standing for IBE (Barnes 1995; Carruthers 2006; Douven 2011; Fodor 2000; Josephson and Josephson 2003; Niiniluoto 1999; Psillos 2002), but this stance is wrong: there is more to IBE than mere abduction. Some others argue that IBE and abduction are conceptually distinct (Campos 2009; Minnameier 2004; Hintikka 1998; McKaughan 2008), however, this stance is also an exaggeration: two concepts are indeed related. The most accurate description of the relation between IBE and abduction is to state that they overlap to some degree" (Mackonis, 2013, 976). We will follow Cellucci's view (2013).

[22] It may be objected that Peirce stated the conclusion of abduction in the form: "Hence, there is reason to suspect that *A* is true" (Peirce, CP 5.189) and not "Hence, *A* is true." But it is important to underline that here we are dealing with abduction in relation to SR and to the realist view of abduction, and realists' aim is to claim for the truth, or the approximate truth, of the scientific theories. If realists consider abduction to be the inference rule normally used by scientists in their work (cf. Psillos, 2002, 605: "abduction, suitably understood as Inference to the Best Explanation, offers the best description of scientific method") and consider the product of scientists' work, i.e. the scientific theories, to be true, we can affirm that realists tend to see abduction and IBE as inferences to the truth. For example, Psillos states that: "It should be taken to be implicit in the realist thesis that the ampliative-abductive methods employed by scientists to arrive at their theoretical beliefs are reliable: they tend to generate approximately true beliefs and theories" (Psillos, 1999, xviii).

typically, these theories have been arrived at by means of IBE, IBE is reliable (truth-conducive)."[23]

To sum up, SR is centered on the concept of truth, and in order to support their main argument realists rely on success as an indicator of truth, and on a kind of abductive inference to infer from the success of a theory the truth of such theory.

So, two main goals for the realist are: 1) to secure the link between success and truth, and 2) to justify the reliability of such inference.

2. Scientific Realism and Adaptationism

Some realists, let's call them: evolutionary realists, try to overcome the difficulties which arise in attaining the goals just mentioned (and some of such difficulties will be analyzed in section 3), relying on an adaptationist view of evolutionism. In this section it will be outlined why such attempts are inadequate.

2.1. Scientific Realism and Evolutionary Adaptationism

It is important to analyse the centrality of some realist assumptions in the adaptationist view of evolution, in order to underline the circularity of the realist's attempt to rely on such view to naturalize the human ability to produce knowledge. This attempt is in fact intended to secure the realist's confidence in the link between success and truth relying on evolutionary considerations. On the one hand, such an approach is motivated by the idea that a naturalistic

[23] Psillos, 2011b, 24.

stance should commit us with evolutionism,[24] and so, given that many realists view themselves as naturalists, that SR should be compatible with evolutionism. On the other hand, such an approach is motivated by the difficulties the realist faces in justifying her claims, and by the idea that evolutionism may provide a solution to such difficulties.[25]

2.2. Evolutionary Adaptationism

If a realist tries to naturalize knowledge through evolutionism, the position at hand is that of adaptationism.[26] Indeed, Evolutionary Adaptationism (EA) can be briefly described as the claim that natural selection is the only relevant cause of the evolution of a trait and that every relevant trait is an adaptation.[27] Being an adaptation means for a trait to be able to increase the fitness of the organisms which present such trait.[28] A strong correlation between adaptive traits and an increase of fitness is asserted by adaptationists: if a trait is individuated, then an increase of fitness due to such trait has to be measured, and if an increase of fitness is measured, then an adaptive trait has to be the cause of such increase.

[24] Cf. Giere, 2006, 53: "If evolutionary naturalism is understood to be a general naturalism informed by the facts of evolution and by evolutionary theory, then no responsible contemporary naturalist could fail to be an evolutionary naturalist in this modest sense."
[25] Thomson, 1995.
[26] De Cruz, 2007.
[27] Orzack and Forber, 2010; Lewens, 2009; Godfrey-Smith, 2001. Each of these claims has been contested, but the point here is to make clear why EA is so appealing for the realists.
[28] Cf. Resnik, 1989, 195: "An adaptation as a product is roughly anything that is the product of the process of adaptation (natural selection)."

2.3. Adaptationism and IBE

Adaptationism is deeply related to SR. First of all, who supports EA supports even the validity of IBE, given that abduction is the inference commonly used by adaptationists, and we have already seen that to accept the IBE almost amounts to embrace SR. For example, Griffith calls the IBE the "adaptationist abduction."[29] Secondly, IBE is evolutionarily justified by the adaptationists claiming that it is "plausible that the human brain was selected in part because of its capacity for and disposition to such inference [i.e. IBE]," given the "survival value of this disposition."[30] Thus, as the realist in the case of the NMA, the adaptationist uses abduction in her arguments, and justifies the reliability of such inference abductively.

So, combining selectionism and functionalism in order to naturalize SR seems to many realist authors a coherent strategy:

> The proper function of some mechanism, trait, or process in evolved organisms is ultimately relative to fitness, and the brain has as proper function the production of beliefs that are fitness-enhancing.[31]

But this is not sufficient to link adaptation and truth. Another step is necessary to achieve such goal:

> The evolutionary argument (…) contends that natural selection will form animal brains that tend to produce true beliefs (…). Cognitive faculties

[29] Griffiths, 1996, 521. Resnik states that the adaptationists search for adaptationist explanation, "an explanation that hypothesizes that a given trait is an adaptation (the product of natural selection)" (Resnik, 1989, 196), and that in order to try to confirm such hypothesis they "practice what philosophers of science have dubbed 'inference to the best explanation'" (Resnik, 1989, 201); Lewens states that "the method" of IBE is the "most favoured by many adaptationists" (Lewens, 2009, 179).
[30] Goldman, 1990, 39.
[31] De Cruz and De Smedt, 2012, 413.

> that are widely off the mark would seriously compromise a creature's ability to survive and reproduce.[32]

So, in order to naturalize realism, to produce beliefs that are *fitness-enhancing* has to mean to produce *true* beliefs.[33]

We have reached the realist core of this view: the crucial premise of such view is that only true beliefs can be useful. Thus, given that only useful beliefs producers will be selected, because useful beliefs increase fitness, selection will retain only true beliefs producers.

2.4. Adaptationism and Truth

Stich has sketched this position as follows: "the argument seems to be that natural selection favors true beliefs, (…). So if an organism is the product of natural selection, we can safely assume that most of its beliefs will be true."[34]

This means to connect adaptation to success. And given that for the realist the success is related to the truth, this amounts to connect adaptation to truth.[35]

---

[32] De Cruz and De Smedt, 2012, 416–417.

[33] There is a deep relationship between the evolved structure of the human brain and scientific knowledge, in fact "from an evolutionary perspective, science is a recent development in our species. Thus, scientists have to draw on the same cognitive resources as other people, and they are subject to the same cognitive limitations" (De Cruz, 2011, 205). So, those cognitive resources have to be able to produce true beliefs in order to give us humans the ability to produce true scientific theories as SR claims.

[34] Stich, 2011, 83.

[35] The analogy between the concept of adaptedness and that of correspondence has been analyzed in Godfrey-Smith, 1998. Philosophers who have tried to naturalize intentionality have often proposed a theory of meaning based on the idea of a relation of correspondence between organic inner states and states of the world: "Correspondence as a property of thought, and adaptedness as a property of biological structure and behavior, are both apparently used in explanations of success" (Godfrey-Smith, 1998, 174); in such a view, correspondence is seen as "*a general-purpose fuel for success* in dealing with the world" (Godfrey-Smith, 1998, 172).

The problem is that if we commit ourselves to claim that human cognitive structures are true beliefs producers because they have been selected for, and so that adaptedness implies truth, then we should be able to demonstrate that every cognitive adaptation is able to produce nothing else than (or at least mostly) true beliefs.

But this is not an easy task, because, as Stich, among others, has stressed,

> it is simply not the case that natural selection favors true beliefs over false ones. What natural selection does favor is beliefs which yield selective advantage. And there are many environmental circumstances in which false beliefs will be more useful than true ones.[36]

The problem of connecting selection and truth is analogous to that viewed above of linking the truth of a theory to its empirical success. As the success cannot guarantee the inference to the truth of its (hypothesized) cause, so the survival is not able to discriminate among its possible causes.[37] So both success and survival are not reliable indicator of truth.

The realists face a dilemma: either they reduce their explanation to a vacuous tautology,[38] or accept that success is caused not only by

[36] Stich, 2011, 83. See also McKay and Dennett, 2009. It may be objected that such kind of claims is self-defeating (e.g. De Cruz et al. 2011). This objection is discussed in Sterpetti, forthcoming.

[37] Goldman, 1990, 40, summarizes some of the difficulties that such a position has to face as follows: "That some cognitive capacity exists and serves useful functions does not show that it was selected; that it was selected does not show that it produces mostly true beliefs; that it produces true beliefs in one context does not show that it continues to do so in others."

[38] Cf. Downes, 2000, 435, who states that the risk is to claim that "successful action is guided by beliefs produced by a mechanism selected for to produce the kind of beliefs that produce successful action." Cf. also Godfrey-Smith, 1998, 188, who states that adaptiveness and fitness, like "correspondence (…) are relations that appear, *prima facie*, to have a role in explanation of success. However, when people have tried to say more clearly what these properties are, they have often drawn so heavily on the relation between fitness and success that success has become partially constitutive of fitness or

true beliefs, but this would amount to dismiss realism,[39] at least to the extent that the empirical success of a theory is considered to be reliably related to its being true, so that we can safely infer the truth from the success, which indeed is considered to be the main realist 'intuition'.[40]

It may be objected that the realists are not committed to the claim that only the truth leads to success, and that the evolutionary realists may easily account for the cases of selected false and useful beliefs. What they have to show is just that the *majority* of the selected beliefs are true. But, as Plotkin clearly states, the difficulty of linking 'truth' and 'survival' is not a matter of degree, because even "if it were only very rarely untrue, but the holders of the untruths survived and reproduced, that would be enough to nullify any foolish claim by evolutionary epistemology to overcoming the justification problem," i.e. to support the claim that success is a good indicator of truth: in fact, only "if survival and reproduction are absolutely correlated with knowledge could they be an infallible guide to true belief," but "this is not the case."[41] Thus we cannot safely rely on the success of some organism to assess the truth of such organism's knowledge.

In order to avoid misunderstandings, it is worth specifying that here we are assessing the epistemological attempts made to naturalize SR through evolutionism, i.e. to justify SR relying on selection. We are not dealing with the cognitive and evolutionary issue of evaluating the truthfulness or the falseness of our evolved beliefs relying on our

---

adaptiveness (…). So these properties have tended to lose their capacity to causally explain success."

[39] In fact, in an evolutionary scenario "the relevant notion of truth (…) is truth that is instrumental in aiding survival in the short run and contributes to reproductive fitness. But this is no way to reconstruct the notion of truth on the standard [correspondentist] account. For the standard account to work, it requires an independent conception of truth and then an account of how this is related to selection" (Downes, 2000, 435).

[40] Worrall, 1989b.

[41] Plotkin, 1997, 234.

present scientific knowledge, which we deem to be true because of its success. An 'evolutionary road' to SR tries to show that we are able to produce true scientific theories because evolution gives us some sort of truth-tracking ability. Thus, if we take the evolutionary road we have to show that selection leads to the truth without relying on the truth of our scientific knowledge, otherwise our attempt would be plainly circular. From this perspective, it is selection which has to secure our ability in producing true science; it cannot be the truth of our science that justifies our claim that the majority of the beliefs produced by our evolved cognitive systems are true.[42]

It is important to take those two different approaches distinct, because it is easy to confound them. This kind of confusion occurs when the evolutionary realists try to defend their position relying on some account of the 'false but useful' beliefs which tries to explain the evolutionary origin of our false beliefs taking for granted the truth of our knowledge, while they (the realists) should provide a justification of the fact that we are inclined to take for granted the truth of our science. For example, McKay and Dennett state that in

> many cases (perhaps most), beliefs will be adaptive by virtue of their veridicality. The adaptiveness of such beliefs is not independent of their truth or falsity. On the other hand, the adaptiveness (or otherwise) of *some* beliefs is quite independent of their truth or falsity.[43]

To give an example of an adaptive misbelief they refer to a supernatural belief, which may be advantageous even if it is *clearly* false. In this case, we can easily tell the true from the false from the

[42] On a similar point cf. Worrall, 1989a, 384: "Those who make the right inductions, those who base their actions on generalizations that have enjoyed predictive success had, and have, a selective advantage. This argument no doubt needs careful handling. But however carefully handled and however persuasive it can be made to seem, it is clearly circular. It is based on our belief in the correctness (or essential correctness) of Darwinian theory. But this in turn (...) is based on our methodological principles."
[43] McKay and Dennett, 2009, 507.

point of view of our present science. Thus, we can easily determine which beliefs are useful and true, which beliefs are useful and false, and then try to give an evolutionary explanation of the latter. Moving along this line, McKay and Dennett are not compelled to justify their assumption that true beliefs are more conducive to usefulness and success than false beliefs, they simply take it for granted, and construct their examples from the vantage point of our present science. This is clearly a licit stance in trying to evolutionarily explain misbeliefs, but has not to be confused with a licit epistemological justification of the evolutionary realist claim that true beliefs are more adaptive than false ones.

The fact is that also in the context of an evolutionary justification of SR, exactly as in the case of the inference from the empirical success of the theories to the truth we have seen above, we cannot reach an Archimedean point of view and assess whether the hypothesis that 'only the truth leads to success' is true or not. We can only rely on success. And we can safely compare some measures of success. But does this make the hypothesis that is the truth which explains the success more confirmed? The answer is in the negative. The point is that it is easy to mistake an increase in the success for an increase in the confirmation of the hypothesis that it is only the truth which can explain the success, exactly because we tend to implicitly assume that it is only the truth that leads to success. But as many other implicit assumptions we make, and notwithstanding how much this assumption about the truth seems evident to us, such assumption may be unreliable.[44]

This point is related to the fact that referring to 'truth approximation' instead of referring to 'truth' is often considered to be

[44] Think, e.g., to Nozick's (2001) position, according to which if we accept evolutionism we cannot state neither if what appears to us as self-evident and necessary is instead contingent, nor if it is (or has ever been) even true.

sufficient to secure SR against the objections inspired by the theory change or by the falseness of some of our beliefs. According to this approach, some aspects of a theory may not be strictly true, but we can nevertheless claim that we are approximating the truth. The evidence for the fact that we are approximating the truth should be given by the increase in the success of our theory. But this approach rests on the assumption that an increase in success may be due only to an increase in the quantity of truth present in our theory. But, as already noted, an increase in success may be due to an increase of truth only if only truth can lead to success. In other words, the more cautious formulation of SR, i.e. the 'approximation-to-the-truth' formulation, rests exactly on the same claim that 'only truth can leads to success' on which rests the straightforward formulation of SR, i.e. the 'truth' formulation. Thus claiming that our theories, or beliefs, are just approximation to the truth, does not solve the problem of justifying SR.[45]

Let's turn to the fact that many evolutionary realists face the objection that not only truth leads to success claiming that it is not difficult for them to account for some few false beliefs which are also useful, if the majority of the useful beliefs are true. The problem is that in order to construct the set of the false beliefs, that of the true beliefs, and then compare them, the realists have to already know which are the false beliefs and which the true ones. That is, the realists can avoid the objection based on the difficulty of inferring the truth from the success by claiming to be able to show that the set of the useful-and-true beliefs is greater than the set of the false-and-useful beliefs, and so that an inference from the success to the truth is justified at least in the majority of the cases. But to construct those

[45] We are not denying that our ability in coping with the world and in elaborating theories which are more successful than the previous ones is something that we have to try to explain. The point is assessing if the traditional solution is satisfying.

sets they need to use exactly what they instead should ground. But if this would be possible, then we should admit that we already possessed a justified way to tell the true from the false beliefs. A realist approach should justify the tool it uses in discriminating the true from the false beliefs, and it should not use that very tool in doing it.

So, the only way to maintain an *evolutionary* realist perspective, and secure that success is implied *only* by true beliefs, is *to presuppose* such a realist connection.[46] For example, Millikan states:

> *Assuming that* the capacity to form and to use beliefs has survival value mainly in so far as the beliefs formed are true (…), and *assuming that* humans currently have this capacity in part because, historically, having it had survival value, the mechanisms in us that produce beliefs, (…) all have in common at least one proper function: helping to produce true beliefs.[47]

Thus, realists can succeed in naturalizing knowledge through EA only because EA assumes realism among its premises, and adding to EA another crucial realist premise, i.e. that only true beliefs can increase fitness.

3. Abduction, Truth, and the Problem of the Criterion

In this section, we will be briefly sketch some of the main difficulties the realist has to face in relation to each of the crucial aspects of SR outlined in section 1, and how such difficulties are related. Such difficulties may in fact be re-conducted to the inability of SR to satisfyingly avoid the problems deriving from the sceptical challenge of the criterion of truth.

[46] On the fact that in such kind of argument the claim that only true beliefs can be useful may only be assumed, cf. Sage, 2004.
[47] Millikan, 1984, 317, italics mine.

3.1. Abduction and the Production of True Hypotheses Problem

As stated above, IBE is considered by realists a valid inference and is also commonly seen as an ampliative inference, in some way similar to induction. IBE is also generally considered to be related to abduction.

The problem is that some author has (convincingly) claimed that abduction is neither ampliative, nor truth-preserving.[48]

For example, analyzing the way normally inference rules are classified,[49] Cellucci states that abduction is: 1) neither ampliative; 2) nor truth-preserving:

$$\text{(ABD)}\ \frac{B \rightarrow A \quad A}{B}$$

In fact, with regard to 1): "conclusion B is a subformula of the major premise B→A, and so is already contained in it." With regard to 2): "if A is true and B is false, then B→A is true, so both premises of (ABD) are true but the conclusion B is false;"[50] so abduction cannot be considered neither ampliative nor truth-preserving.

Indeed, "what generates new information is not (ABD), but rather the process that yields its major premise, B→A, thus something prior

[48] We put aside here the other criticisms that have been already moved to IBE and focus on Cellucci's one for two reasons: 1) there is an enormous and well known literature on the former; 2) Cellucci's criticism to IBE is different from the others criticisms already moved to IBE because these are, normally, at their turn committed to some concept of truth and to the deductivist view, which considers deductions as justified. Cellucci's work, instead, not only denies that abduction is ampliative, and denies a pivotal role to the concept of truth, but also shows that the asymmetry between deductive inferences/justified and non-deductive inferences/unjustified is untenable (Cellucci, 2006).

[49] Cellucci sees abduction as a counterexample to the standard classification of inference rules, given that "(ABD), on the one hand, like deductive rules, is non-ampliative, and, on the other hand, like non-deductive rules, is not truth preserving" (Cellucci, 2013, § 18.2).

[50] Cellucci, 2011, 124.

to (ABD)."[51] This line of thought goes back at least to Frankfurt, who states:

> Clearly, if the new idea, or hypothesis, must appear in one of the premisses of the abduction, it cannot be the case that it originates as the conclusion of such an inference; it must have been invented before the conclusion was drawn.[52]

So, what really contributes to the ampliation of knowledge occurs *before* abduction and cannot be described as an abduction, because it is beyond our rational control, i.e. it is beyond logic:

> Our first premisses (…) are to be regarded as an extreme case of abductive inferences, from which they differ in being absolutely beyond criticism. The abductive suggestion comes to us like a flash. It is an act of *insight*.[53]

So, the ampliative process of hypothesis production is distinct from abduction and is not describable by abduction as an inference rule, because "it is sub-conscious and so not amenable to logical criticism,"[54] while "reasoning is deliberate, voluntary, critical, controlled, all of which it can only be if it is done consciously."[55] For example, Kapitan states that in "strict Peircean terms, the emergence of hypotheses is not a matter of inference and, therefore, not a matter of a unique form of inference."[56] Many authors try to solve this puzzle by adding *something* to the Peircean scheme of abduction, in order to account for what they call 'the creative side' of abduction,[57] and try to

[51] Cellucci, 2013, § 18.2.
[52] Frankfurt, 1958, 594. Peirce himself writes that "*A* cannot be abductively inferred, or if you prefer the expression, cannot be abductively conjectured until its entire content is already present in the premiss" (Peirce, CP 5.189) [note that in Peirce's text '*A*' corresponds to '*B*' in the formula given above].
[53] Peirce, CP 5.181.
[54] Peirce, CP 5.181.
[55] Peirce, CP 2.182.
[56] Kapitan, 1992, 8. Cf. Hoffmann, 1999, 278: "the logical form for itself leaves the question unanswered how to get the hypothesis."
[57] Cf., e.g., Aliseda, 2006; Magnani, 2009; Hoffmann, 1999.

avoid that abduction may be considered nothing more that the fallacy of affirming the consequent.[58] For example, Schurz analyzes and classifies different kinds of abduction.[59] When he comes to the more creative rather than selective patterns of abduction, it is easy to see that what he is doing is incorporating some other ampliative inference rule, such as induction or analogy, in the supposed first stage of abduction, i.e. that of the hypothesis production. This is transparent, for example, in § 6.2, where he analyzes what he calls 'analogical abduction', i.e. an abduction in which the hypothesis has been produced by analogy. But at this point it is not clear why we should continue to talk of 'abduction', and refer to Peirce, if the process of hypothesis production can be described by different rules and it is itself not abductive. The so-called 'creative' part of abduction is clearly not an abduction. This confirms that abduction is not ampliative.

So, following Peirce, we should *presuppose* that humans are able to (unconsciously) produce hypotheses which in some way are likely to be true.[60]

How could we account naturalistically for the human ability to (unconsciously) produce true hypotheses? The answer could seem to be, *prima facie*, referring to evolution:

> A *naturalistic* basis means that Peirce likens abductive instinct to those instincts that animals possess for getting food and reproducing (…). If animals have innate tendencies that help them to survive in their environments, why not to assume that we as human beings have analogously innate tendencies for finding correct theories? This kind of an instinct would obviously have strong adaptive value for us.[61]

[58] Magnani, 2009, 15.
[59] Schurz, 2008.
[60] Cf. Peirce, CP 7.220: "The existence of a natural instinct for truth is, after all, the sheet-anchor of science." Cf. also Anderson, 1986, 152: "For Peirce, however, because abduction is not a matter of pure chance, science is understandable. As Rescher says, 'Peirce insists that trial and error cannot adequately account for the existing facts'."
[61] Paavola, 2005, 134.

The problem, as already seen above, is exactly that that adaptive beliefs have to be true beliefs can be either postulated or supported by an abductive argument.

So, trying to naturalize abduction and show that abduction is reliable in finding true hypotheses relying on an evolutionary account of abduction which endorses an adaptationist view of evolution, which at its turn assumes the validity of abduction, would be circular.

3.2. Truth and the Problem of the Criterion

As we have already seen, the concept of truth realists usually are committed to is that of truth as correspondence. This is the most widely accepted concept of truth, but nevertheless it is elusive. For example, Bunge states that "everyone uses the correspondence concept of truth, but nobody seems to know exactly what it is."[62]

The problem for the realist derives from the divide that exists between the conception of truth she adopts and the epistemic optimism that characterizes SR. In fact, given that: 1) the usual definition of knowledge realists assume is that of 'justified true belief';[63] and that 2) truth is intended as a mind-independent correspondence to the world, it can be stated that the "realist conception of truth is a non-epistemic conception of truth, which enforces a sharp divide between truth and rational justification."[64]

[62] Bunge, 2012, 74.

[63] Cf. Sankey, 2008, 14, fn. 2: "The traditional justified true belief account of knowledge is a minimal condition for a realist conception of knowledge." The problem of providing an effective criterion, which emerges in relation to the concept of truth, emerges also in relation to the concept of knowledge, cf. Sankey, 2008, 101: "The justified true belief analysis of knowledge provides a set of conditions, satisfaction of which qualifies a subject as having knowledge. It does not provide criteria which enable a subject to recognize that those conditions obtain, and is thereby in possession of knowledge."

[64] Sankey, 2008, 112. Cf. also Sankey, 2008, 16: "It is important to note that there are a number of alternative theories of truth which contrast with the

But this means that the concept of truth normally endorsed by realists is intrinsically unable to provide a *criterion* of truth, i.e. a tool to determine the truth of a given statement: given "the non-epistemic nature of truth, there is no logical relation between method and truth."[65] This is obviously a problem for SR, given that SR maintains that our best scientific theories are true, and that it is through the scientific method that scientific theories are developed, evaluated, and selected.

In fact, SR, while denying the relation between method and truth, is at the same time "a position of epistemic optimism, which holds against the sceptic that humans are able to acquire knowledge of the world."[66] So, the problem for the realist is to show how the scientific method can lead to the truth and fill the epistemic gap, i.e. to solve what Sankey has called 'the problem of method and truth'.[67]

The main difficulty is that to solve such problem amounts to solve what has been called since antiquity 'the problem of the criterion of

---

correspondence theory of truth. (…). According to such theories of truth, truth is a property which a belief or statement may have in virtue of some epistemic property of the belief or statement. (…). Because such theories of truth identify truth with an epistemic property of belief, they are sometimes called 'epistemic theories of truth'."

[65] Sankey, 2008, 112. Cf. also Sankey, 2008, 112: "On the realist conception of truth, truth is a relation of correspondence that obtains between statements and mind-independent states of affairs that obtain in the world. (…). Thus, truth depends solely on the way the world is (…). As such, no epistemic condition enters into the realist conception of truth."

[66] Sankey, 2008, 112. In fact, those authors, as Laudan and van Fraassen, who support a non-epistemic concept of truth and deny that we can fill the epistemic gap between method and truth, are labeled by realists 'scientific sceptics' (Sankey, 2008, 116). On the realists' epistemic optimism, cf. Sankey, 2008, 34: "It would be perfectly consistent for the scientific realist to refrain from any positive epistemic commitment to the truth or progressiveness of science. (…). But realists typically do not adopt such a sceptical attitude toward science. They typically support a stronger epistemic thesis to the effect that science has made progress toward the truth, and, in so doing, has produced genuine knowledge about the objective world."

[67] Sankey, 2008.

truth'.[68] A simple argument which brings a very hard challenge, at least for the realist.[69]

The main feature of the argument of the criterion is that it is built up in such a way to force the opponent in one of three alternatives,[70] each of which is considered to be able to show that the opponent's claim is untenable.[71] The idea is that through such argument is possible to show how a claim is not really justified in the way its proponent pretends that it is.

Indeed, the argument of the criterion forces the opponent to assent that its claim is justified: 1) circularly; 2) just assumed and not justified; 3) justified referring to some other claim, and that this leads to an infinite regress.[72]

---

[68] See Cling, 1997, 1994; Floridi, 1994, 1993; Chisholm, 1982; Sankey, 2012, 2011, 2010. The problem of the criterion of truth is the ancient sceptical paradox of the wheel: "in order to know any proposition we must first know a criterion, but in order to know a criterion we must already know some proposition" (Cling, 1997, 109).

[69] See Cellucci, 2013, § 18.19; Chisholm, 1982. On the relevance of such problem, cf., e.g., Chisholm, 1982, 61: "'The problem of the criterion' seems to me to be one of the most important and one of the most difficult of all the problems of philosophy" and Rescher, 2003, 22: "It is difficult to exaggerate the significance of this extremely simple line of reasoning."

[70] Cf. Floridi, 1993, 207: "It is a combination of Agrippa's second, fourth and fifth tropes, i.e. *regressus*, *hypothesis* and *diallelus*."

[71] This is a debated issue, which cannot be developed here, see, e.g., Cling, 2003, on circularity, Cling, 2004; Alston, 1989; Klein, 1999, on infinite regress, Betz, 2010, on *petitio principii*.

[72] Cf. Sextus Empiricus, 1976, II.2: **"**[...] in order to decide the dispute which has arisen about the criterion, we must possess an accepted criterion by which we shall be able to judge the dispute; and in order to possess an accepted criterion, the dispute about the criterion must first be decided. And when the argument thus reduces itself to a form of circular reasoning the discovery of the criterion becomes impracticable, since we do not allow them [the dogmatics] to adopt a criterion by assumption, while if they offer to judge the criterion by a criterion we force them to a regress *ad infinitum*. And furthermore, since demonstration requires a demonstrated criterion, while the criterion requires an approved demonstration, they are forced in circular reasoning."

Here we cannot account for the attempts which have been made to solve the problem of the criterion,[73] but it will suffice to underline how SR seems not to be able to offer a justification of its claims about truth and knowledge which can face the challenge of such argument.[74]

The reason lays in the concept of knowledge that SR adopts. If knowledge is related to the truth, then SR cannot avoid or face the problem of the criterion, given that the concept of truth as correspondence does not provide an adequate criterion of truth.

Indeed, the 'standard' realist approach to the problem is making reference to Tarski's work. For example, Ruttkamp states that truth "is a semantical relation between language and reality. Its meaning is given by a modern (Tarskian) version of the correspondence theory."[75] But the inadequacy for SR of a merely semantic treatment of the issue of the concept of truth is clearly stated by Sankey, when he analyzes the relation between the Tarskian T-scheme and the most widely adopted theories of truth.[76] For Sankey, the correspondence theory of truth is obviously compatible with the T-scheme, but the latter is not sufficient to clarify the concept of truth that SR needs.[77] Something

[73] The issue has been set in its actual form by Chisholm, 1982. Cf. Cling, 1994, 232: "Chisholm says that there are only three possible responses to this problem: (i) adopt skepticism, (ii) claim to have an answer to 'how are we to decide whether we know?' and use it to answer 'what do we know?' (methodism), or (iii) claim to have an answer to 'what do we know?' and use it to answer 'how are we to decide whether we know?' (particularism)."

[74] This does not mean that we have to embrace scepticism, but that SR is probably not the right way of accounting for scientific knowledge. In fact, until SR is maintained, and so it is the traditional account of truth and knowledge, knowledge is unattainable by humans, and so the sceptical arguments are unavoidable. On this, cf. Cellucci, 2013.

[75] Ruttkamp, 2002, 177.

[76] Cf. Sankey, 2008, 49: "the T-scheme is common ground to all the standard theories of truth. The disquotational, pragmatic, coherence, internalist, verificationist and correspondence theories of truth all agree that truth, *whatever it is*, must conform to the T-scheme."

[77] Cf. Sankey, 2008, 111: "Since a statement is true just in case the state of affairs to which it corresponds obtains, the correspondence conception satisfies the equivalence condition specified by Tarski's T-scheme:

more than a mere *definition* of truth is needed to provide a criterion of truth. In fact, Tarski's scheme does not give us a criterion to tell the true from the false,[78] nor it is able to connect "language and reality," because, as Bunge states,

> It does not contrast language, or rather its epistemic designatum, with extralinguistic reality – which is what 'correspondence' is supposed to mean. Indeed, Tarski's formula just bridges a bit of language (…) to a bit of metalanguage.[79]

Moreover, Cellucci has convincingly shown that not only the concept of truth as correspondence cannot provide a realist criterion of truth, but that the most important conceptions of truth that have been proposed until now are inadequate as a criterion of truth,[80] and that to provide a criterion of truth is indeed necessary in order to support a realist conception of science:

> If the goal of science is truth and, on the other hand, a concept of truth does not provide a criterion of truth, we will generally be unable to determine whether a given sentence is true or false. Then, (…) the goal of science cannot be generally achieved because it transcends human capacities.[81]

(T) 'P' is true if any [*sic* = and] only if P.
While the T-scheme is not a definition of truth, it provides a minimal condition of adequacy that must be satisfied by any account of truth."

[78] It is worth noting that Tarski states that a "criterion" of truth "will never be found" (Tarski, 1944, 363–364), and that there is no point in complaining that the concept of truth as correspondence does not provide a criterion of truth, since the concept of truth as correspondence "is not designed at all for this purpose" (Tarski, 1969, 69). On this point see Cellucci, 2014.

[79] Bunge, 2012, 66.

[80] Cellucci, 2014, 2013. More precisely, Cellucci, 2014, analyses the following concepts of truth: 1) truth as correspondence, 2) truth as intuition of the essence, 3) truth as consistency, 4) truth as systematic coherence, 5) truth as possession of a model, 6) truth as provability, and shows that they are all inadequate as a criterion of truth.

[81] Cellucci, 2013, 152–153.

But if "truth is humanly transcendent, we will generally be unable to recognize a truth when we reach it, therefore we will be unable to say that we have acquired knowledge of the world through it."[82]

But this would obviously be at odds with realist's epistemic optimism. In fact, SR needs not only a definition of truth as correspondence, but a way of showing that the correspondence such defined between the theory and the world actually occurs, it needs a tool to tell the true from the false,[83] i.e. a criterion of truth.[84]

It may be argued that we have not to be realist about a whole theory, and that we could *select* the parts of a theory we intend to be committed to. The problem is how to 'carve' the theory we are interested in and determine which are its ontologically relevant parts. We should be able to discriminating which elements were indispensable in order for the theory to give accurate predictions and

[82] Cellucci, 2014, 4.

[83] For example, Rescher states that in "characterizing a claim as true, we indicate that what it states corresponds to the facts, so that its assertion is in order. But while this factually ('stating which is the case', 'corresponding to the facts') is what truth is all about, we cannot apply or implement it as such: it does not provide a basis on which the truth of claims can be determined" (Rescher, 2003, 146–147). Cf. also Munz, 1993, 177: "it is comparatively easy to arrive at a concept of truth. But as soon as one seeks to define the precise criteria by which one can determine whether any particular statement is true or not, one is confronted with an array of possibilities, none of which are satisfactory."

[84] On the features such a criterion should have, cf., e.g., Cling, 1997, 110–111: "An ideal criterion would express a reliable and complete touchstone of truth: a mark or sign by which we could distinguish true from false propositions on any topic. Having such a criterion is not the same as grasping the meaning of 'true', for we might understand what it would be for a proposition to be true without being able to tell the true from the false. To have a criterion of truth is to grasp that some detectable property other than the property of being true would correctly distinguish true from false propositions. Thus we may think of an ideal criterion of truth as a principle according to which a specified property *C* is such that *C* is not part of the meaning of 'true', but a proposition *P* would have *C* (in the appropriate circumstances) if, and only if, *P* were true. Ideally, then, a criterion of truth would provide us with a perfectly reliable indicator of truth and of falsehood."

take only those indispensable elements as existing. The problem is that we can do this kind of evaluation only in relation to our *past* theories and from the point of view of our *present* best theory.[85] The risk is that our explanation of the success of the past theories makes explicit the positive contribution to the empirical success of the past theories only of those elements which have been retained and passed in the new theory. There could be other parts of the theory which were 'indispensable' to reach the predictions but that now we judge as irrelevant, and there is the possibility that some of the elements which we now judge as indispensable and relevant will be discarded in the future. Thus, our judgement on what exists could be based only on the similarity between certain parts of the past theory and certain parts of the new theory. In order to claim that this similarity tells us what exists we should be able to defend the claim that our present theory *is* true. But we can do that only relying on success. The problem, again, is how to justify the claim that success is due to the truth.

It may also be objected that we should intend a criterion not as a tool able to tell the true from the false, but just as an *epistemic norm* aimed at justifying a belief, whose satisfaction does not entail the truth of the justified belief. Thus, a criterion may be satisfied and the belief which satisfies such a criterion may nevertheless be false. Our criterion would be fallible, and we could account for the distinction between truth and justification.

The problem is that in any case SR aims at the truth. But if we do not have a tool to state with certainty whether something is true or not, we will never be able to know whether we have reached the truth or not. Moreover, normally the realists maintain that science arrives at truth thanks to its *method*. Scientific method is what makes a belief *justified*. Thus, if method is related to the justification, and the realists

[85] Stanford, 2000.

take the method to be truth-conducive, this means that, in a realist perspective, justification cannot be taken to be *really* so distinct from the truth. For example, Sankey states that the questions "about the relation between method and truth divide scientific realism from anti-realism in the philosophy of science," since they are "questions about the truth-conduciveness of method. While they relate directly to the epistemic status of method, they bear indirectly on the nature of rational justification. For if use of method conduces to truth, then, given the relation between method and justification, the warrant provided by method is warrant with respect to truth."[86] In this perspective, an epistemic norm is equivalent to a criterion of truth:

> Sextus Empiricus (…) speaks of a 'criterion of truth' that is used to 'judge of reality and non-reality' (…). Since an epistemic norm is used to justify belief, and since belief involves belief in the truth of the content of the belief, an epistemic norm plays the same role as a 'criterion of truth'.[87]

But if a criterion of truth does not entail the truth of the belief which satisfies it, then in what sense should it be defined as a *criterion* of truth? If something can satisfy a criterion of truth and at the same time not to be true, then such a criterion is simply not a real criterion of truth. Thus, it is not easy for the realists to avoid to face the problem of the criterion.

In order to solve such a puzzle,[88] Sankey adopts what he calls 'abductive realism', i.e., while "the success argument is usually employed to argue for the approximate truth of theories," he would "extend the argument to the truth-conduciveness of rules of method." But the difficulty of providing a criterion of truth cannot be avoided or

[86] Sankey, 2008, 109.
[87] Sankey, 2012, 5, fn. 1.
[88] Cf. Sankey, 2008, 118: "The attempt to combine a naturalistic account of epistemic warrant with the realist view of truth as the aim of science must face the problem that no empirical evidence may show directly or conclusively that use of a methodological rule yields theoretical truth."

solved by such abductive approach: "The point is not that satisfaction of methodological rules constitutes truth. The point, rather, is that satisfaction of methodological rules is best explained by truth."[89]

Again, what at its most SR can offer to support its strong claims about truth and reality is a sort of NMA. The truth cannot reliably be inferred from the success of a theory, but the truth of such theory can be proposed as the best explanation for its success. But this move does not really explain anything at all, given that such an explanation is based on the *assumption* that 'only truth can explain success', but no ground is given for such assumption.

This line of reasoning seems to be symptomatic of the fact that SR cannot ground its claims on other than success. And cannot do this if not abductively. There is no way of justifying the belief in the truth of the theoretical statements other than relying on success. And there is no way of inferring the truth of theoretical statements from success other than relying on abduction. And there is no way of justifying abduction other than abductively.

But: 1) success is not an adequate criterion to discriminate *only* the true theories; 2) abduction is neither an ampliative, nor a truth-preserving inference, thus it cannot be said to be a *reliable* inference; 3) a circular justification of abduction is not sufficient to make abduction a reliable inference, because rule circularity is not a benign sort of circularity.[90, 91]

---

[89] Sankey, 2008, 107.

[90] Many realists (e.g., Psillos, Sankey, Ellis) claim that some circularity is not dangerous and can be accepted. Contrary to such position, at least in the context of the justification of the inference rules, cf. Cellucci, 2006, 210–211: "That there is something basically wrong with rule-circularity appears also from the fact that, if to prove the validity of a rule of inference of deductive logic one is entitled to use that very same rule, then some invalid rule can be proved to be valid. For example, consider the *abduction rule*, that is, the rule:

$$(\text{Abd})\ \frac{B \quad A \to B}{A}\ .$$

So it seems that there are just two ways to try to warrant the realist's epistemic optimism.[92] The first is to justify our epistemic abilities referring to evolution, but this leads to a sort of circularity. The second is to take for granted a metaphysical framework which allows to warrant the realist's claims, but this amounts to presuppose what should instead be derived.

The first way is that proposed, e.g., by Sankey. Indeed, Sankey is aware that even "if it is granted that realism is the best explanation of the success of science, it does not follow that it is to be accepted as true." So, he tries to provide a naturalized account of our epistemic

---

Using Abd one can give the following formal proof of the validity of Abd:

$$\text{(Abd)}\ \frac{\text{TT}\,(\rightarrow)\dfrac{\mathrm{T}(A\rightarrow B)\quad \mathrm{T}(B)}{\mathrm{T}(A\rightarrow B)\rightarrow \mathrm{T}(B)}\qquad \dfrac{\text{TT}\,(\rightarrow)}{\overbrace{\mathrm{T}(A)\rightarrow\big(\mathrm{T}(A\rightarrow B)\rightarrow \mathrm{T}(B)\big)}}}{\mathrm{T}(A)}\ .$$

It is no use to object that this proof does not provide a justification of Abd since, while the proof of the validity of MP uses MP, which is valid, this proof uses Abd, which is invalid. For to justify the validity of MP is just what is at issue."

[91] The problem of the criterion of truth and that of justifying and classifying inference rules are related: there is no way to justify an inference rule avoiding circularity (Cellucci, 2006, *contra* Howson, 2000, shows that not evenin the case of *Modus Ponens* an inference rule can be demonstrated to be non circularly justified), and there is no way to state an adequate criterion of truth: "If we interpret 'criterion of truth' as 'inference rule', Sextus Empiricus' argument becomes: Those who profess to validate an inference rule are bound to have some inference rule to validate it. Now this inference rule either is not validated or has been validated. If it is not validated, then it cannot be trusted, for no matter of dispute is to be trusted without being validated. If it has been validated, the inference rule used to validate it, in its turn, either has been validated or has not been validated, and so on *ad infinitum*" (Cellucci, 2013, 310).

[92] Another way may be to follow the 'particularist' approach to the problem of the criterion sustained by Chisholm, 1982, and claim that the optimistic assumption needs not to be justified, and that it is sufficient to provide some case of knowledge we know to be true to support the realist optimism. But such attempt has been already analysed and criticized, and seems not so promising. On this see Cling, 1994.

abilities.[93] In fact, he states that "the naturalistic realist may treat the problem of knowledge as the broadly empirical problem of explaining how cognitive agents embedded in the natural world are able to use their epistemic capacities to promote their survival."[94]

The problem is that Sankey *assumes* the idea that only true beliefs may promote the survival, and so that if our epistemic abilities have been selected, they cannot but produce true beliefs, otherwise we would have become extinct.[95]

Thus, for him the empirical problem of explaining how our epistemic capacities may promote the survival is solved relying on the concept of truth.

The problem is that we started from the difficulty of giving a criterion of truth and justify the adoption of a correspondence theory of truth. So, it is the concept of truth that should be secured empirically, and not an empirical problem that should be explained referring to the truth.

Moreover, in supporting his claim that only true beliefs may promote the survival, Sankey explicitly refers to the proposals of Rescher (1977) and Kornblith (1993). Details of such proposals are not relevant here, but what is worth noting is that both these authors connect our human epistemic abilities to evolution, and that in their proposals it is implicitly assumed the key assumption of

[93] Sankey, 2010, explicitly refers to the problem of the criterion and tries to combine the two traditional responses to the problem, i.e. the 'particularist' and the 'methodist' (see above fn. 73). In fact, he maintains a particularist stance following Chisholm, 1982. But, given that he acknowledges that Particularism begs the question against the sceptic and that it may be not extended to face the relativist's challenge, he tries to solve this problem giving "a naturalistic account of epistemic warrant," which is very similar to the methodist proposal given by Rescher, 1977.

[94] Sankey, 2008, 8.

[95] Cf. Sankey, 2010, 14: "Epistemic norms which lead us systematically astray in our beliefs about the surrounding environment will inevitably give rise to frustration, harm or even death."

adaptationism we have already seen above, i.e. the connection between adaptive success and truth, which allows them to derive from the selective process the justification of the truth-conduciveness of the products of our evolved abilities. In other words, such attempts, as that of Sankey, implicitly assume exactly what they are supposed to be able to naturalize. So, Sankey's attempt is circular.

The second way is that proposed, e.g., by Ellis (2009). If the relation of correspondence cannot be secured epistemically, it can be metaphysically postulated. Ellis adopts a correspondence theory of truth, and tries to connect it to a truth-makers theory. He, as Sankey, strongly underlines the importance of the mind independence of truth for SR. Ellis doesn't think that the epistemic gap can be filled by the scientific practice. For him, to support metaphysical claims such as realist's, what is needed is a metaphysical framework. Such a metaphysical framework should define what exists in the world in order to decide if a statement is made true by the world or not. This obviously amounts to presuppose what a scientific realist should instead derive.[96]

Indeed, there is no problem in defining the truth as a correspondence to the world, and in stating that a statement is made true by the world. The problem is *how* to provide a criterion which is able to ascertain *if* the world actually makes true a given statement, especially in the context debated here, i.e. SR, whose characterizing

[96] On the circularity of his approach, cf. Ellis, 2009, 19: "Metaphysical necessitation is the relation that holds between things in the world and the things they make true. That is, it is what is usually called a 'semantic relation'. (…). The concept of plausibility that we require to define semantic relations all depend on our general metaphysics, that is, our theory of the ultimate nature of reality. And, for such a theory to be adequate, it must be consistent with our best understanding of the world, and able to accommodate all of the things we truly believe in. This is a logical circle, of course. But it is inescapable. A postulated existent is ontologically plausible if and only if it fits into an adequate metaphysical theory. And a metaphysical theory is adequate if and only if it accommodates all of the things that we truly believe in."

claim is the belief in the truth of the theoretical, i.e. unobservable, parts of scientific theories. How could we identify the truth-maker of a theoretical statement, in order to state that it makes true such theoretical statement, if what we rely on to believe in the existence of what such theoretical statement refers to is just the success of the theory such theoretical statement belongs to?

The scientific realist can try to avoid the ambiguities related to the linguistic formulation of the theories of truth, and support a truth-makers conception of the relation between theories and the world. The problem, again, is that the relation of correspondence cannot be shown to actually occur between the theory and the world, but just between the theory and a 'plausible' metaphysical framework which tries to describe the world. Moreover, the way in which this framework is developed is abductive, as clearly emerges from Ellis' own words: "A necessitation relation is metaphysical if the proposition *p* whose truth is to be explained is made true by some objectively existing thing or state of affairs X." But how to identify or determine such X? We have to "specify an X such that (a) X could plausibly exist, and (b), if X were to exist, then *p* would have to be true." Obviously, in order to specify such X, "*we must have some prior views about the nature of reality*."[97]

So, there is no way to avoid to support realist's claims other than circularly or postulating exactly what she is wanting to demonstrate, and both these ways are unsatisfactory. In other words, the realists seem not be able to avoid the problem of the criterion of truth, given that they rely on truth, nor their attempts seem to be able to satisfyingly face such challenge, given that they lead to circularity or to a *petitio principii*.

[97] Ellis, 2009, 19–20, italics mine.

4. Conclusions

We have pointed out that three crucial aspects of SR are: 1) the centrality given to the concept of truth, 2) the idea that success is a reliable indicator of truth, and 3) the idea that IBE is a reliable inference rule.

We have then outlined how some realists try to naturalize such crucial aspects relying on an adaptationist view of evolutionism, and why such attempts are inadequate.

Finally, we have briefly sketched the main difficulties the realist has to face in relation to any of such crucial aspects, and how such difficulties are deeply related. Such difficulties may in fact be re-conducted to the inability of SR to satisfyingly avoid the problems deriving from the sceptical challenge of the criterion of truth.

This is due to the tension between what the realist maintains about truth and what she maintains about knowledge. In other words, SR seems not to be able to fill the epistemic gap. In fact, the epistemic gap cannot be filled in no way other than obtaining a criterion of truth, but such a criterion cannot be obtained if the epistemic gap obtains.

Acknowledgements

I would like to thank two anonymous reviewers for their helpful comments and suggestions.

References

Aliseda, A., 2006, *Abductive Reasoning*, Dordrecht, Springer.

Alston, W., 1989, Epistemic Circularity. *Philosophy and Phenomenological Research*, 47, 1–32.

Anderson, D.R., 1986, The Evolution of Peirce's Concept of Abduction. *Transactions of the Charles S. Peirce Society*, 22, 145–164.

Betz, G., 2010, Petitio principii and circular argumentation as seen from a theory of dialectical structures. *Synthese*, 175, 327–349.

Bunge, M., 2012, The correspondence theory of truth. *Semiotica*, 188, 65–75.

Burgess, A.G. and Burgess, J.P., 2011, *Truth*, Princeton, Princeton University Press.

Cellucci, C., 2006, The Question Hume Didn't Ask: Why Should We Accept Deductive Inferences? In: *Demonstrative and Non-Demonstrative Reasoning*, Cassino, Edizioni dell'Università degli Studi di Cassino, 207–235.

-- 2011, Classifying and Justifying Inference Rules. In: *Logic and Knowledge*, Newcastle Upon Tyne, Cambridge Scholars Publishing, 123–142.

-- 2013, *Rethinking Logic. Logic in Relation to Mathematics, Evolution, and Method*, Dordrecht, Springer.

-- 2014, Knowledge, Truth and Plausibility. *Axiomathes*, 24, 517–532.

Chisholm, R.M., 1982, *The Foundations of Knowing*, Minneapolis, University of Minnesota Press.

Cling, A.D., 1994, Posing the Problem of the Criterion. *Philosophical Studies*, 75, 261–292.

-- 1997, Epistemic Levels and the Problem of the Criterion. *Philosophical Studies*, 88, 109–140.

-- 2003, Self-supporting Arguments. *Philosophy and Phenomenological Research*, 66, 279–303.

-- 2004, The trouble with infinitism. *Synthese*, 138, 101–123.

De Cruz, H., 2007, Innate Ideas as a Naturalistic Source of Mathematical Knowledge. Dissertation, Vrije Universiteit, Brussel.

De Cruz, H., 2011, *Through a Mind Darkly*. PhD thesis. Groningen: University of Groningen.

De Cruz, H. et al., 2011, Evolutionary Approaches to Epistemic Justification. *Dialectica*, 65, 517–535.

De Cruz, H. and De Smedt, J., 2012: Evolved cognitive biases and the epistemic status of scientific beliefs. *Philosophical Studies*, 157, 411–429.

Devitt, M., 1997, *Realism and Truth*, Princeton, Princeton University Press.

Douven, I., 2002, Testing Inference to the Best Explanation. *Synthese*, 130, 355–377.

Downes, S.M., 2000, Truth, Selection and Scientific Inquiry. *Biology and Philosophy*, 15, 425–442.

Ellis, B., 2009, *The Metaphysics of Scientific Realism*, Durham, Acumen.

Floridi, L., 1993, The Problem of the Justification of a Theory of Knowledge. Part I: Some Historical Metamorphoses. *Journal for General Philosophy of Science*, 24, 205–233.

-- 1994, The Problem of the Justification of a Theory of Knowledge. Part II: Morphology and Diagnosis. *Journal for General Philosophy of Science*, 25, 17–49.

Frankfurt, H.G., 1958, Peirce's Notion of Abduction. *The Journal of Philosophy*, 55, 593–597.

Frost-Arnold, G., 2010, The No-Miracles Argument for Realism: Inference to an Unacceptable Explanation. *Philosophy of Science*, 77, 35–58.

Giere, R.N., 2005, Scientific Realism: Old and New Problems. *Erkenntnis*, 63, 149–165.

-- 2006, Modest Evolutionary Naturalism. *Biological Theory*, 1, 52–60.

Godfrey-Smith, P., 1998, *Complexity and the Function of Mind in Nature*, Cambridge, Cambridge University Press.

-- 2001, Three Kinds of Adaptationism. In: *Adaptationism and Optimality*, Cambridge, Cambridge University Press, 335–357.

Goldman, A., 1990, Natural Selection, Justification, and Inference to the Best Explanation. In: *Evolution, Cognition and Realism*, Lanham, University Press of America, 39–46.

Griffiths, P.E., 1996, The Historical Turn in the Study of Adaptation. *The British Journal for the Philosophy of Science*, 47, 511–532.

Klein, P.D., 1999, Human Knowledge and the Infinite Regress of Reasons. *Noûs*, 33, 297–325.

Harman, G.H., 1965, The Inference to the Best Explanation. *The Philosophical Review*, 74, 88–95.

Held, C., 2011, Truth does not explain predictive success. *Analysis*, 71, 232–234.

Hitchcock, C.R., 1992, Causal Explanation and Scientific Realism. *Erkenntnis*, 37, 151–178.

Hoffmann, M., 1999, Problems with Peirce's Concept of Abduction. *Foundations of Science*, 4, 271–305.

Howson, C., 2000, *Hume's Problem*, Oxford, Oxford University Press.

Kapitan, T., 1992, Peirce and the Autonomy of Abductive Reasoning. *Erkenntnis*, 37, 1–26.

Kornblith, H., 1993, Inductive Inference and its Natural Ground, Cambridge (MA), MIT Press.

Laudan, L., 1981, A Confutation of Convergent Realism. *Philosophy of Science*, 48, 19–49.

Lewens, T., 2009, Seven types of adaptationism. *Biology and Philosophy*, 24, 161–182.

Lewis, P., 2001, Why the Pessimistic Induction Is a Fallacy. *Synthese*, 129, 371–380.

Lipton, P., 2004, *Inference to the Best Explanation*, London, Routledge.

Mackonis, A., 2013, Inference to the best explanation, coherence and other explanatory virtues. *Synthese*, 190, 975–995.

Magnani, L., 2009, *Abductive Cognition*, Berlin, Springer.

McKay, R.T. and Dennett, D.C., 2009, The evolution of misbelief. *Behavioral and Brain Sciences*, 32, 493–510.

Millikan, R., 1984, Naturalist Reflections on Knowledge. *Pacific Philosophical Quarterly*, 65, 315–334.

Munz, P., 1993, *Philosophical Darwinism*, London, Routledge.

Nozick, R., 2001, *Invariances*, Cambridge (MA), Harvard University Press.

Orzack, S.H. and Forber, P., 2010, Adaptationism. In: The Stanford Encyclopedia of Philosophy, http://plato.stanford.edu/archives/fall2010/entries/adaptationism/.

Paavola, S., 2005, Peircean abduction: Instinct or inference? *Semiotica*, 153, 131–154.

Peirce, C.S., 1931–1958, *Collected Papers of Charles Sanders Peirce* (CP), Cambridge (MA), Harvard University Press.

Plotkin, H., 1997, *Darwin Machines*, Cambridge (MA), Harvard University Press.

Psillos, S., 1999, *Scientific Realism*, New York, Routledge.

-- 2002, Simply the Best: A Case for Abduction. In: *Computational Logic: Logic Programming and Beyond. Essays in Honour of Robert A. Kowalski. Part II*, Berlin, Springer, 605–625.

-- 2011a, An Explorer Upon Untrodden Ground: Peirce on Abduction. In: *Handbook of the History of Logic. Volume 10. Inductive Logic*, Amsterdam, Elsevier, 117–151.

-- 2011b, The Scope and Limits of the No Miracles Argument. In: Explanation, Prediction, and Confirmation, Springer, Dordrecht, 23–35.

Putnam, H., 1975, *Mathematics, Matter and Method*, Cambridge, Cambridge University Press.

Rescher, N., 1977, *Methodological Pragmatism*, Oxford, Blackwell.

-- 2003, *Epistemology*, Albany (NY), SUNY Press.

Resnik, M., 1989, Adaptationist Explanations. *Studies in History and Philosophy of Science*, 20, 193–213.

Ruttkamp, E., 2002, *A model-theoretic realist interpretation of science*, Dordrecht, Kluwer.

Saatsi, J. and Vickers, P., 2011, Miraculous Success? Inconsistency and Untruth in Kirchhoff's Diffraction Theory. *The British Journal for the Philosophy of Science*, 62, 29–46.

Saatsi, J., 2005, On the Pessimistic Induction and Two Fallacies. *Philosophy of Science*, 72, 1088–1098.

Sage, J., 2004, Truth-Reliability and the Evolution of Human Cognitive Faculties. *Philosophical Studies*, 117, 95–106.

Sankey, H., 2008, *Scientific Realism and the Rationality of Science*, Burlington, Ashgate.

-- 2010, Witchcraft, Relativism and the Problem of the Criterion. *Erkenntnis*, 72, 1–16.

-- 2011, Epistemic relativism and the problem of the criterion. *Studies in History and Philosophy of Science*, 42, 562–570.

-- 2012, Scepticism, relativism and the argument from the criterion. *Studies in History and Philosophy of Science*, 43, 182–190.

Schurz, G., 2008, Patterns of abduction. *Synthese*, 164, 201–234.

Sextus Empiricus, 1976, *Outlines of Pyrrhonism*, Cambridge (MA), Harvard University Press.

Sher, G., 2004, In Search of a Substantive Theory of Truth. *The Journal of Philosophy*, 101, 5–36.

Stanford, P.K., 2000, An Antirealist Explanation of the Success of Science. *Philosophy of Science*, 67, 266–284.

Sterpetti, F., forthcoming, Formalizing Darwinism, Naturalizing Mathematics. *Paradigmi*.

Stich, S., 2011, *Collected papers. Volume 1*, Oxford, Oxford University Press.

Tarski, A., 1944, The semantic conception of truth and the foundations of semantics. *Philosophy and Phenomenological Research*, 4, 341–376.

-- 1969, Truth and proof. *Scientific American*, 220, 63–77.

Thomson, P., 1995, Evolutionary Epistemology and Scientific Realism. *Journal of Social and Evolutionary Systems*, 18, 165–191.

Worrall, J., 1989a, Fix It and Be Damned: A Reply to Laudan. *The British Journal for the Philosophy of Science*, 40, 376–388.

-- 1989b, Structural Realism: The Best of Both Worlds? *Dialectica*, 43, 99–124.

-- 2008, Theory-change in science. In: *The Routledge Companion to Philosophy of Science*, New York, Routledge, 281–291.

Wray, K.B., 2013, Success and truth in the realism/anti-realism debate. *Synthese*, 190, 1719–1729.

Ortega y Gasset on Natural Selection and the Vocation of Man: A Rejection or Elaboration of Darwinism?

Rabi Lior
rabilior@gmail.com

Abstract In many of his writings and lectures Ortega y Gasset criticized Darwinism on numbers of issues. In this paper I aim to reexamine his critique and to prove that in 1916 the harsh critique was designed to hide the similarity between his ideas and Darwin's idea of Natural Selection. The origins of Ortega's idea on the vocation of man can be traced in his dialogue with Darwin in the year 1916. In the historiography his Philosophy of Life is conceived as a metaphysics that stands beyond biological and any other natural explanation. My contention on the other hand is that Ortega's concept of the vocation of man is an elaboration of Darwin's Natural Selection. I will try to show that in 1916 Ortega conceived the human character and the vocation of man as a product of Natural Selection. Instead of following Ortega's critiques on Darwin and his effort to show that there is no connection between his Philosophy of Life and Darwinism the article proposes to reconsider Ortega's thought as a constant dialogue with Darwin's theory of Natural Selection. The Change in Ortega's conception of Darwin's evolution came rather late in his thought and in the first section we will see how Ortega described technology in a manner that allows the human subject to change the environment in his favor.

Resumo Em muitos dos seus escritos e lições, Ortega y Gasset criticou diversos aspectos do darwinismo. Neste artigo procuro reexaminar tais críticas, mostrando que, em 1916, a crítica mais dura foi dirigida para ocultar a semelhança entre as suas ideias e a teoria darwinista da selecção natural. As origens da concepção orteguiana da vocação do Homem podem ser rastreadas no seu diálogo com Darwin em 1916. De acordo com a historiografia, a filosofia orteguiana da vida é concebida como uma metafísica que se situa para além de qualquer explicação biológica ou naturalista. Eu defenderei, pelo contrário, que a concepção de Ortega acerca da vocação do Homem foi, na verdade, elaborada a partir da teoria darwinista da selecção natural. Procurarei mostrar que, em 1916, Ortega concebia a natureza e a vocação humanas como produtos da selecção natural. Ao invés de seguir as críticas de Ortega a Darwin, bem como a sua tentativa de mostrar a inexistência de qualquer ligação entre a sua filosofia da vida e o darwinismo, eu

proponho-me reconsiderar o pensamento de Ortega como um diálogo permanente com a teoria da seleção natural. A mudança de perspectiva de Ortega relativamente à teoria darwinista da evolução veio bastante tarde no seu pensamento, e na primeira secção deste artigo procurarei mostrar como Ortega descreveu a tecnologia de uma forma que permite ao sujeito humano mudar o meio ambiente em seu favor.

I. The Way the Subject Reforms the Natural World.

In his work "The Animal that therefore I am" Derrida distinguishes between an animal's reaction and response. Do animals really respond to our gestures or just react? Derrida wonders why man is the only animal that feels the need to get dressed or in his own words why "clothing would be proper to man, one of the "properties of man?"[1] Can an animal feel it is naked?! Derrida concludes that an animal can never really be naked because it neither feels nor sees itself naked. He mentions that "clothing derives from technology".[2] It is technology that allows man to create fashionable clothing.

In his lectures on technology Ortega also focused on the differences between man and animal in relation to man's technical capabilities. Ortega's main argument was that technology is a manner in which man tries to reform nature in his favor. Man does so in order to fulfill his needs. However, man does not only work to fulfill his biological needs, but rather tries to go beyond them. For the human being it does not suffice only 'to be' (estar), but to live well (bienestar)[3].The animal does not need technology because for the animal it is enough to be able to take from nature only what it is necessary in order to survive. But man has a different task, creative

[1] Derrida, 2002, 373.
[2] Derrida, 2002, 374.
[3] "El bienestar y no el estar es la necesidad fundamental para el hombre, la necesidad de las necesidades", Ortega, 1933, 33.

task. This creative task is the main difference between Ortega and Darwin conception of man. By using this idea Ortega aims to go beyond the biological depiction of man and to differentiate between his philosophy and Darwin's evolution whose ideas as we shall see later in this paper fascinated him and influenced on the articulation of his own thought.

In his late lectures we can notice Ortega's subtle dialogue with Darwin and his critiques on Darwin's evolution. In 1951 in the conference in Darmstadt Ortega refers to Natural Selection and examines a situation in which certain individuals don't possess the natural abilities to adapt to a certain natural or environmental problem. Unlike Darwin, his argument is that in such extreme situation these individuals will try to adapt environment in their favor. Technology for Ortega means the ability of the subject to adapt nature in its favor. This interpretation is a way to criticize what Ortega conceives as the mechanical explanation of Darwin: "The technology is contradictory to the idea of the subject's adaptation to the medium; it is rather the adaptation of the medium to the subject. This is enough to allow us to suspect that we see a movement to the opposite direction that the biologists refer "[4]. In his lectures on technology Ortega aimed to overcome the biologist conception of man and mainly the Darwinist idea of Natural Selection. Therefore, he also argued that technology is not a natural response to biological needs.[5] Technology is a product of man's "fantasy": "As a species, man's desires do not have much in common with the instincts, with nature". Technology is a product of our "desired fantasizes".[6] Ortega mentions that man for example has a desired fantasy- to follow what he fantasizes as justice[7].

[4] Ortega y Gasset, 1933, 31.
[5] Ortega, 1933, 26.
[6] Ortega, 1951, 51. Cerezo contends that there is a strict parallelism between the conference in Darmstadt and the lectures from 1933: "Hay así un estricto paralelismo entre las dos conferencias, la del 1933 y la de 1951: el 'centauro

Human being is always looking to find things he sees in his fantasies, things he cannot always find in nature. Technology is a way the human tries to achieve these desired fantasies[8]. Man is not satisfied with his circumstance. In this manner he is different from the other animals and he will try to adapt the circumstances to his fantasies. Ortega distinguishes between three possibilities:

1. The circumstance does not present any difficulty to the human.
2. The circumstance present unbearable difficulties.
3. The circumstance present difficulties together with conformabilities.

He rejects the first two options and accepts the last one. For Ortega, humans are never satisfied with their circumstances and this can explain how they try to adapt nature to their own fantasies. It is not only Natural Selection that affects man, but it is rather man who tries to adapt nature to his wishes. Therefore, the human has different task from these of the other animals. In this manner Ortega distinguishes between his own philosophy and Darwin's Natural Selection.

In his later work Ortega also refers to the idea of "instinct of preservation" and contends that humans live because they want to live. The Instinct of preservation is a narrow definition of man. In this aspect Ortega's later thought is not so different from Nietzsche's thought. In her article Pearson writes that "in his theory of life Nietzsche sharply criticizes the view that the aim and goal of life is self-preservation and places all the emphasis on the enjoyment a living thing gets out of simply discharging its force"[9]

---

ontológico' de la primera es 'el ser excéntrico' de la segunda". Cerezo, 2011, 243.

[7] Ortega y Gasset, 1951, 107.

[8] "Las necesidades biológicamente objetivas no son, por sí, necesidades para él", Ortega, 1933, 34.

[9] Pearson, 1997, 91

Focusing mainly on this biological definition might cause us to forget that man can control his instincts by the use of his will. Human life cannot be understood when it is strictly based on the biological interpretation. Ortega differentiates between man and animal by the ability of the first to decide to end its life, man has the capability to suicide. Humans do not have to focus only on their basic needs, on their survival or their being in the world. The difference between man and animal lies in man's ability to have desired fantasies, desires that aim to change the environment to his needs and comforts. Man's task is not only to live, but mainly to live well. This difference between man and animal, the ability to choose, to look forward, to feel tension between his fantasies and nature, are the main features in Ortega's thought that differentiate it from the biological explanation. Ortega's later work aims to overcome any understanding of man in biological manner. Already in 1924 Ortega published an article about vitalism in which he tries to explain how his concept of vitality differs from the biological concept. In this year he argued that he used many biological terms mainly in order to criticize the idealistic version of rationalism[10]. When he referred to Goethe's concept of life in his paper from 1932 Ortega argued that Goethe's concept of life is biological, botanical.[11] The aim of Ortega was to overcome what he conceived as the conception of life itself from outside.

The dialogue with Darwin existed through all of Ortega's thought. The explicit critique appeared mainly in "Meditation on Technology", but also in his writing from the twenties or thirties[12]. In his paper on

[10] Ortega, 1924.

[11] "Claro es que Goethe nos desorienta porque su idea de la vida es biológica, botánica", Ortega, 1932, 405.

[12] Marías argues that the concept of human life in Ortega's thought does not refer to the biological aspects of life: Marías, 2000, 66. In his later work Ortega went beyond the biological conception of human life, but as we shall see in the next sections the young Ortega was influenced by Darwin and biological concept of life.

Goethe from the year 1932 Ortega presented what he conceived as the main difference between his own philosophy and the evolutionists. His contention was that human life is a drama not an evolution.[13] Human life is about choices, decisions and opportunities and humans' main focus is on the future. Humans' focus is on the future, on the plans they should make for their future. Man's focus is to change the circumstances, making them different according to his plans. In this manner Ortega understands the present dimension in human's life only in relation to his future, his circumstances[14]. Animals do not feel that there exists a big difference between them and their circumstances. In contrast, man experiences the difference because he has desired fantasies and his focus is on the future.

II. How Can a Darwinian Explain the Creativity of the Human Mind?

In his first book "Meditations on Quixote" published in 1914 Ortega dedicates two short paragraphs designed to reconsider what he conceives as the most relevant or central conclusion that can be deduced from Charles Darwin's idea of Natural Selection. In these two short paragraphs (that appear at the end of the book) Ortega tries to show the difference that exists between his and Darwin's concept of life. The critique of Ortega looks at odds with his own articulation of human life that is not based on the traditional philosophical ideas of soul or cogito. Hence, reading these paragraphs we ask ourselves if

[13] "Para la planta, el animal o la estrella, vivir es no tener duda alguna respecto a su proprio ser. Ninguno de ellos tiene que decidir ahora lo que va a ser el intante inmediato. Por eso su vida no es drama, sino...evolución", Ortega, 1932, 407.
[14] "Mi presente no existe sino merced a mi futuro, bajo la presión de mi futuro", Ortega, 1930, 131-132.

there really is a big difference between Ortega and Darwin's concepts of life. It may be so that what seems as an anti-Darwinist attack is only a conscious exaggeration that intends to disguise the similarity. In other words, stressing only the supposed difference might be intended to disregard the similarities. In these paragraphs Ortega wonders where the source of human creativity comes from. Furthermore, he raises serious doubt that the concept of adaptation to the environment can explain the creativity of cultural heroes. By concentrating on the creative ability of the heroes of human culture Ortega is aiming to doubt the relevance of the Darwinist description. He argues that human life should be understood as freedom and not as deterministic adaptation to the material surrounding:

> The natural sciences that are based on determinism have conquered the biological camp. Darwin believes that he succeeds to imprison the vital- our ultimate hope, within the physical necessity. Life descends to be no more than matter. The physiology turns to mechanics. The organism that appears as independent unity, able to work for itself, is inserted in the physical medium like a figure in a carpet. Our actions are not reactions. There is no liberty, originality. To live is to adapt: to adapt means to leave the material surrounding to penetrate us, we are expelled from ourselves. Adapting is renouncing. Darwin moves aside the heroes from their doing.[15]

Here we can see that Ortega argues that if we adopt the Theory of Evolution we would not be able to understand human ability to articulate original ideas, for example we would not be able to properly conceive how there are humans (los héroes: the heroes) like Cervantes or Flaubert whose work and ideas are so creative and original. This puzzle is further elaborated in the work of Daniel Dennett who asks in one of his articles the following question: "How can an arrangement of a hundred billion mindless neurons compose a creative mind, an I?"[16] This dilemma as we know does not prevent Dennett from suggesting a

[15] Ortega y Gasset, 2010, 244.
[16] Dennett, 2009, 396.

plausible solution to the concept of the mind. In his book Ortega does not try to find a solution to the problem he raised. Hence, it seems that here one can either follow the idea of Natural Selection or Ortega's metaphysics of human life.[17] However, I believe that to deal with this issue is much more intricate than it may seem by focusing only on these two paragraphs. In the following sections I will try to show how the main ideas in Ortega's Philosophy are actually an effort to elaborate much further the conclusions that can be derived by analyzing Darwin's Theory of Evolution.

The main critique on Darwin in these two paragraphs is aimed to show that the mind has a potential of originality. When it is conceived according to a biological point of view it will be hard to understand the nature of human creativity since "Darwin pushed aside the heroes from their work on earth".[18] Ortega concludes that Darwinism cannot explain the existence of heroes' creative human mind. This critique of Ortega might be responsible for the fact that researchers usually did not put any serious effort to examine the close similarity between Ortega and Darwin's thought. In the following paragraphs I will try to demonstrate the similarities between the young Ortega and Darwin's concept of human life. Ortega's critique may be explained by an intended effort to hide the similarities between his and Darwin's concept of life. I believe that Ortega's concept of human character and its vocation is actually an elaboration of Darwin's thought and an effort to go much further philosophically. This elaboration leads to an articulation of a metaphysical concept of human life.

[17] Ortega's original philosophy is mainly interpreted as an effort to go beyond both realism and idealism. His 'categories of human life' are presented in many of his books that appear after 1928. See for example: Rodríguez Huescar, 2002.
[18] Ortega, 2010, 244.

In his nine important and albeit forgotten lectures at the University of Buenos Aires from 1916 Ortega focuses mainly on the philosophical implications of the advancement in biology. These lectures appeared on print only in the year 1995. This late appearance may also be responsible for the fact that none of Ortega's researchers focuses on Ortega's alternative or elaboration to Darwinian Theory.

In these lectures from 1916 Ortega maintains that human preferences, predilections and tendencies are gradually formed during evolution. However, since these tendencies become so natural their evolutionary formation is forgotten. Each person has different tendencies and human vocation is examined according to the ability of the individual to follow or not to follow his own tendencies. This ability to choose to obey or not to the personal tendencies is also a product of evolution; a process in which human attitude towards the surrounding went beyond the interest in immediate dangers and the human started to be interested in finding the truth for itself. This ability to go beyond the immediate necessities is also connected, as we will see, to the ability to choose in favor or against our personal tendencies, to obey the truth that is supposed to be located beyond the subjective point of view.

III. The Objectives of Philosophy and Biology.

In his youth writings Ortega maintains a constant dialogue with Nietzsche in relation to his Theory of Perspectives, Elitism and critique on scientific assumptions.[19] During August-October 1916 Ortega gave nine long lectures at the University in Buenos Aires. In his third lecture

[19] See for example the following articles: Ortega, "Glosas", in: Vida Nueva, 1 de diciembre de 1902, or Ortega, "El sobrehombre", in: *El Imparcial*, 13 de Julio 1908. On Ortega and Nietzsche see Sobejano, 1967.

from that year he refers specifically to Nietzsche's argument according to which human truth is an illusion and that therefore "science, moral and culture are unstable delusions".[20] According to Ortega, Nietzsche's argument is 'an elegant blasphemy', to be applauded by the 50-year-olds of the previous generation. Ortega does not underestimate Nietzsche's argument regarding human consciousness or truth, and feels the need to justify carefully the alternatives he subsequently offers for them.

In his nine lectures in Argentina Ortega does not focus only on Nietzsche's Theory of Perspectives but also refers to Nietzsche's interpretation and critique on Darwinism. In these lectures we can find more than one occasion in which Ortega examines explicitly Nietzsche's interpretation and alternative to Darwin's Theory of Evolution.[21] The dialogue with the philosophical implications of Darwinism is common to both Nietzsche and Ortega. However, while Nietzsche's critique on Darwin is more explicit, it might seem that Ortega's own interpretation and alternative to Darwinism is far more implicit.[22] What is Ortega's own attitude towards Darwinism and/or the way in which Nietzsche reads Darwin? Focusing on this question can surprisingly lead to a new interesting insight regarding the relations between Philosophy of Life and Darwinism and it further leads to another question: Must a Metaphysical attitude towards human life as freedom necessarily contradict Darwin's 'dangerous idea'?

In his books Nietzsche calls Darwin a 'mediocre' thinker and tries to stress the differences between his own concept of life and Will to

[20] Ortega, 1996, 84.
[21] Ortega, 1996, 68, 78-9, 84.
[22] Contrary to what is assumed in the historiography I believe that Ortega's dialogue with the biologists' interpretation of human life and Darwin's theory of Natural Selection does not sum up with his argument from 1914 that Darwinian will find it hard to explain human acts of creativity. Here I will try to show the real extent of the debate that Ortega had with Darwin and the biologist perspective on human life.

Power and Darwin's idea of Natural Selection.[23] On the contrary Ortega never speaks of himself as being "anti-Darwinian" and his dialogue with Natural Selection is more implicit in comparison to Nietzsche's, who explicitly expresses a severe critique. Although Ortega follows some aspects of Nietzsche's interpretation of Darwin it seems that he refrains himself from criticizing him directly. Ortega's attitude derives as I believe from the distinction he makes between the objectives of biology and philosophy. On two occasions in his lectures Ortega argues that although biologists and philosophers try to understand human life both the philosophers and biologists need to remember that there is a difference between their objectives and methods. The biologist analyzes what he considers as the real facts (hechos), while the philosopher on the other hand would define the biologist's analysis as a confined examination of what we perceive through our senses. The biologists must also adopt philosophical assumptions. What are these assumptions? According to Ortega the biologists' assumptions rely on the positivist Theory of Knowledge. The biologist treats human consciousness as a biological fact. Declaring that facts are all there is in the world is an exposition of a philosophical argument. Unlike the biologist the philosopher is responsible to go beyond the pure analysis of facts. The philosophers' questions are: 1. Does truth exist? 2. What is the correct method to arrive towards truth: rational or empirical? Ortega maintains that by choosing the empirical method the biologists appropriate a philosophical truth according to which truth derives from the sense data:

> Natural science departs from the facts and these facts are not more than particular class of truth. The biological facts are the affirmation that such phenomenon takes place. Something is not a truth because it is a fact; On the contrary, something is a fact because we convinced

[23] See for example where Nietzsche writes in *Will to Power* a section "against Darwin's theory", Nietzsche, 1968, 343.

> ourselves through pre-biological reasons, prior to our observation, that it is such.[24]

For Ortega the question here is whether to conceive natural facts as the ultimate truth as the biologists are doing or to follow the philosopher that is more skeptical regarding human assumptions. A philosopher can decline the assumption that facts are all there is. He can argue for example that truth lies behind the facts. Ortega suggests that the philosophy focuses on the ultimate truth. Therefore, the philosopher has to define what the method to arrive towards the truth is. Ortega concludes that the philosopher asks questions which cannot be answered by relying only on the medium of experience.[25]

Hence, Ortega suggests that there is another central difference between biology and philosophy. Biology is based on the assumption that the ultimate reality is to be found in the facts. Philosophy on the other hand cannot be based on any assumptions. Philosophy is the effort to reflect on the ultimate truth and it begins by presenting a doubt regarding all of our unreflective assumptions.[26] As we will further see the difference between the objectives of philosophy and biology is central for our understanding of Ortega's analysis of both Darwinism and the critique of Nietzsche on Darwin. The difference between the biologist and the philosopher's work derives from their different objectives. The different definitions of what the ultimate purpose of an organ, consciousness or human life as a whole really is derives from a different definition of life's objective. Ortega's main argument is that the philosopher is not allowed to determine for the

[24] Ortega, 1996, 85.
[25] Ortega, 1996, 86.
[26] "Toda ciencia particular parte de algún supuesto, por lo menos, de suponer que el conocimiento verdadero es posible y consiste en proceder de esta o aquella manera determinada. Pero una ciencia cuyo problema está en averiguar si es posible el conocimiento verdadero, la teoría y en qué consiste, claro es que no puede partir de ningún supuesto", Ortega, 1996, 73-74.

biologist what the objective (el fin) of human life is.[27] Since there is a difference between the objectives of philosophy and biology there is also an obligation for the philosopher to be careful when he reflects on the conclusions of the biological analysis. Therefore, there might be two presentations of what human life or consciousness (as we will see later) is. According to Ortega, both of these presentations may be equally correct. On the one hand, the biologist can argue that the aim of human life is to preserve itself. How can we decide that an individual succeeds in preserving itself? Darwin for example argues that the criterion is to be traced in the ability to have predecessors. The struggle for existence is not the survival of the individual but the survival of his characteristic and his ability to pass those to a next generation.[28] On the other hand, there is the interpretation of human life from the philosophical viewpoint. A philosopher like Nietzsche for example can argue that the aim of living organisms and human life is not survival or self-preservation. Rather, the goal of human life is to increase its power: "It can be shown most clearly for every living thing, that it does everything, *not* in order to preserve itself, but to become *more*".[29]

Therefore, Ortega assumes that the differences between the objectives of philosophy on the one hand and biology on the other hand, a difference that many philosophers like the positivists and naturalist tended to disregard, is actually central for the philosophical debate of what life's objectives are. Considering this difference between the objectives of philosophy and those of biology we can understand why Nietzsche's discussion on Darwinism is wrong "the

[27] "El filósofo ha de guardarse muy bien de pretender enseñar a los biólogos qué cosa sea el fin de la vida", Ortega, 1996, 68.
[28] "I should premise that I use the term Struggle for Existence in a large and metaphorical sense, including dependence of one being on another, and including (which is more important) not only the life of the individual, but success in leaving progeny". Darwin, 1998, 50.
[29] Nietzsche, 1968, 366.

philosopher has to be really careful not to pretend to decide for the biologist what the objective of human life is".[30] This conclusion is also applicable for the biologists who might tend to decide for the philosopher what the objectives of human life/consciousness/nature are. In order to examine the plausibility of Ortega's analysis of Darwin's and Nietzsche's interpretation of Natural Selection it is important first to present Nietzsche's arguments and then to reconsider whether Ortega had a real alternative of his own to Darwinism.

IV. Naturalism versus Teleology.

One of Darwin's earliest critics, Robert MacKenzie, conceived the idea of Natural Selection as 'an inversion of reasoning'. The central idea of Darwin is that the process of natural selection is not guided by a divine or absolute wisdom; there is not any intentional guidance by an intelligent designer.[31] During the course of time the critique on Darwinism represented mainly a focus on the conception of Natural Selection as a blind law. What are the arguments of those who criticize Darwin's presentation of natural selection as blind process? The critique is based on three main lines of thought: the idea of intelligent design, the 'anthropic principle' and the teleological argument:[32]

> 1. Those who favor the idea of intelligent design argue that the organic world is too tightly functioning to be a product of the blind law of Natural Selection.
>
> 2. Some physicists believe that had the laws of nature not been exactly as they are then life could never have evolved. Since the exact form of

[30] Ortega, 1996 68.
[31] Dennett, 2009, 393.
[32] Since in this section I examine mainly the teleological argument that stood at the center of the interest of Ortega and Nietzsche alike I will focus mainly on the mentalist and naturalist explanations of Natural Selection.

> natural laws could be infinite range, the only conceivable explanation is that there was an involvement of a designer.
>
> 3. The teleological argument criticizes the idea of Natural Selection as blind law. Its adherents point out that the world is not just thrown together randomly, but functions in a harmonious way towards an end. The human eye for example seems to have been designed for the purpose of enabling sight.[33]

The idea of Natural selection as a blind law leads to the insight that there exists a process in which absolute ignorance gradually leads to the appearance of wisdom and human consciousness. Different authors refer specifically to the question of teleological argument in Darwin's writings. Phillip R. Solan argues that the development of Darwin's conception of nature teaches us that in his maturity Darwin never adopted a "belief in a purposeless nature that undercuts all teleological explanations". According to Solan, "there is a persistence of philosophy of nature, indebted to certain strands within German reflections on nature that underpins the concept of natural laws".[34]

Jeff Wallace argues that what may seem as teleology in Darwin's expressions is no more than a metaphor. If we were to follow Darwin's expressions some of us would be able to assume that he describes the process of Natural Selection as an intentional process. In the third chapter he writes. “Every selected character is fully exercised by her; and the being is placed under well-suited conditions of life”.[35] Here we might argue that Darwin means that it is Mother Nature that chooses or 'selects' for us. We might assume that Natural Selection is not a blind process. In other words, metaphorically the process of Natural Selection can be interpreted in this context as an intentional agency; the readers might be tempted to assume that it is nature that selects, rejects, preserves, adds up or works. However, Wallace suggests that

[33] Ruse, 2003, 168-173.
[34] Solan, 2005, 159.
[35] Darwin, 1998, 65.

Natural Selection should not be conceived as an intentional process.[36] In the same manner the researcher, Dieter Wandschneider argues that Darwinism is limited to "casual patterns of explanation" and the rejection of every teleological interpretation.[37]

In the 'Origin of Species' the argument is that the gradual appearance of different species and the gradual divergence of character was not guided by any intelligent designer; it was rather guided by what Darwin defines as unintentional, unguided or blind law of nature. The different species were not independently created but appeared gradually during history from other very small number of species.[38] Therefore, since the process of Natural Selection is not intentional we cannot conceive the appearance of the human eye according to the teleological argument. Dawkins argues that what we might mistakenly conceive as the 'watchmaker' is actually a blind process.[39] Hence, Darwin's naturalism means that the blind and unaided process of Natural Selection should not be interpreted in any mentalist manner. Natural selection does not imply an element of conscious choice. There is no intervention by a creative designer in the process and the difference between man and other animals is a difference of degree. Humans do not have a different kind of mental powers but the difference with other animals is a difference of degree.[40]

For our purposes it is important to reconsider the idea of Natural Selection as an intentional process or teleology since this kind of

[36] Wallace, 1998, VII- XXIII.
[37] Wandschneider, 2005, 196-215.
[38] "Natural Selection leads to the divergence of character…On the view that each species has been independently created, I can see no explanation of this great fact in the classification of all organic beings; but, to the best of my judgment, it is explained through inheritance and the complex action of natural selection", Darwin, 1998, 100.
[39] Dawkins, 1986.
[40] Darwin, 2010, 17-39.

interpretation was adopted at least apparently by Nietzsche. Nietzsche's harsh critique on Darwin was well known by Ortega. In his lectures Ortega appropriated few features of Nietzsche's critique of Darwin although he also acknowledged the importance of distinguishing between naturalism and the mentalist explanations. In his third lecture in Buenos Aires from 1916 Ortega argues that human life is not to be understood as adaptation to the surrounding. He suggests that life should rather be conceived as an 'artificial creative act that is consisted of the being enhancement and progress'.[41] Here we are encountering a puzzle that can be considered as a contradiction or can also point out that we have to examine this argument within a larger frame of Ortega's thought. The idea of progress might suggest the existence of teleology, but I believe the apparent contradiction can easily be overcome.

We have seen that Ortega argues that there is a necessity to separate between the objectives of philosophy and biology. This separation allows him to argue that a philosopher like Nietzsche should refrain himself from suggesting the biologist what the ultimate objective of life is.[42] Nietzsche suggests the idea of Will to Power as might be understood as nature's intentional goal. How then should we interpret Ortega's argument that life is consisted of self enhancement and progress that is not natural but is rather 'artificial'? Can it lead us to assume that Ortega does not maintain the necessary boundaries between philosophy and the naturalism, boundaries that he himself argued that other philosophers should always take into consideration?

The teleological argument according to which the organic world is not just thrown together randomly, but works and functions towards a certain end was well known to Nietzsche. The idea of an intelligent designer is adopted in theology and philosophy, and therefore it

[41] Ortega, 1996, 79.
[42] Ortega, 1996, 68.

seems implausible that Nietzsche himself who is a renowned critic of rationalist philosophy and theology adopted this point of view. In section 13 of his book 'Beyond Good and Evil' Nietzsche writes that we have to be careful and not to adopt any teleological explanations when we are to discuss human life.[43] In this manner it seems that Nietzsche follows the naturalistic explanation of human life and that he does not add up any teleological explanation to the natural processes. Nietzsche looks here as a philosopher within the boundaries of naturalism. However, in many places in his writings Nietzsche describes the Will to Power as a will 'towards' (zu, auf) an intentional goal. Furthermore, Nietzsche takes Darwin to claim that the organism' has its own goal- survival.[44] Is the Darwinian idea of Natural Selection interpreted here beyond the limits of naturalism? In the 'Will to Power' the idea of Natural Selection seems as an idea that is not interpreted on a purely natural basis. In his *Will to Power* Nietzsche assumes that there exists a conscious intentional process in the natural world according to which everything is aiming towards enhancing its own power.[45]

In his lectures Ortega implicitly argues that Nietzsche's teleological interpretation of Darwin and the proposition of a teleological explanation to nature can lead to a misinterpretation of Darwinism. Furthermore, Ortega maintains that it can also lead to what he defines as the inability to remember the difference between the objectives of philosophy and biology. But as we have seen Ortega

[43] "In short, here, as elsewhere else, let us beware of *superfluous* teleological principles! – one which is the instinct of self-preservation (we owe it to Spinoza's inconsistency). Nietzsche, 1997, 10.

[44] Anti-Darwin. As for the celebrated "struggle for life, it seems to me, in the meantime, to be more asserted than proved. It occurs, but only as an exception; the general aspect of life is *not* a state of want or hunger; it is rather a state of opulence, luxuriance, and even absurd prodigality, - where there is a struggle, it is a struggle for power", Nietzsche, 2004, 46.

[45] More on the tension between naturalism and teleology in Nietzsche's concept of *Will to Power* can be read in Richardson, 2004.

himself also argued in his third lecture that the Darwinian concept of adaptation to nature is a 'secondary operation' (operación secundaria) while the enhancement of the self should be understood as the first and prior operation. Life's changes are not a passive response to a physical medium but rather a process guided towards self enhancement:

> During the second half of the 19th century the Darwinian theory governed the laboratories. Darwinism supposes that the regulated objective of each and every vital phenomenon is adaptation to the medium. Living is to adapt, living is understood as an act, secondary operation; respect for the medium, this law together with its variations or its capricious variations rules the forms of the organic being… In Front of this point of view, comes another old and new vigor. According to this, life is not a secondary activity that only replies to the agent of the physical medium. Life is an artificial act of creativity that consists in expanding and growing of the being (ser).[46]

In this section we can notice that Ortega contradicts his own effort to stress the differences between the goals of biology and philosophy. This kind of interpretation of Ortega's thought is possible. However, I believe that when Ortega argues that life is moving towards self enhancement he means that this might be the way the philosopher wants (or more correctly he as a philosopher wants) to explain life. The philosopher cannot adopt the natural explanation of life. Since the philosopher is not restricted to naturalism he is allowed to present what he conceives as the real or the ethical goal of human life. This goal should not necessarily be restricted to the naturalistic explanation of the human life. In order to elaborate this way of interpretation of Ortega's argument we need first to closely reconsider Ortega's attitude towards the Darwinian description of life's goal and then to reexamine Ortega's own perspective regarding the evolution of human mind.

[46] Ortega, 1996, 78-79.

V. The Darwinians and the Goal of Life.

In his third lecture in Buenos Aires Ortega examines the biologists' analysis of the purpose of each organ in the human body. He starts by asking what the biologist's or the philosopher's definition of the goal of life is? This is a question common to both the philosopher and the biologist and each one tends to present another objective. Ortega argues that the Darwinians conceive the objective of life as an ability to adapt to the physical surrounding. In his opinion the philosopher cannot follow the manner in which the Darwinist defines life's goal as an adaptation to nature. This explanation seems for Ortega too naturalistic for a philosopher to be able to adopt it. According to Ortega the definition of life's objective is in the realm of ethics and therefore the Darwinist explanation cannot be applicable for the philosopher. Since the definition of life's objective is "the biggest problem of man" the philosopher cannot leave this question to be solved by the biologist.[47] Philosophy deals with the ultimate questions of ethics and metaphysics. Since the definition of life's goal is ethical the philosopher's duty is to suggest what the real life goal is. Hence, the philosopher's duty is to contemplate on the problem according to ethical premises and to suggest another definition of life's goal. The philosopher's definition cannot stem from biology but rather has to be based on ethical arguments. In this section we will reconsider how important it is for Ortega to maintain the separation between the goals of philosophy and those of biology. This separation is central in his thought and allows him to argue that both the Darwinist and the philosopher's definitions of life are plausible and legitimate. Furthermore, the separation between ethics and biology

[47] "Este es un problema de ética, el más grande problema del hombre", Ortega, 1996, 76.

will enable Ortega to reconsider the human character and its vocation along the lines of both biology and ethics.

First Ortega concentrates on explaining the difference between instrument and objective and he begins by presenting the naturalistic argument. The naturalistic analysis of the human organs focuses on the medium that each organ is responsible to serve. The human organs have different and certain functions. Since the human mind is also an organ, the naturalist argues that it is also designed in order to serve for a certain function. In the same manner as all the other organs in the human body the human mind is responsible to assist the human preservation; the human mind has to be analyzed according to the function it has to serve. A biologist who tries to explain what the function of the organ is usually focuses on showing the organ's utility for the entire body. For the biologist the form of our body and the way the organism moves are also conceived as tools that are aimed to serve a certain goal. For example, the heart does not beat for its own sake. The heart beat is a function designed to bring blood in order to assist the other organs to function effectively; the heart has a certain function. Each organ is a medium that is supposed to serve a purpose. The organ is designed in a manner that will enable it to function effectively. The organ's ability to serve well as a medium for the goal of adaptation of the human body to the physical surrounding is examined according to utilitarian criterions. This is the explanation for the function of each organ and it seems that Ortega tends to accept this argument as a plausible explanation.

However, while Ortega accepts the biological description of the function of each organ he argues that the real problem is to explain what the final objective of all the organs is. He argues that the biologists (and the philosophers) debate on this question: "What is the objective for which all of life serves as a medium. This is the question

that the biologists debate".[48] According to Ortega, the biologists cannot reach an agreement and cannot decide what the objective of life is; an objective is that all the organs together are supposed to serve. This argument of Ortega may seem implausible since we know that Ortega could have presented the idea of survival or self-preservation, ideas that were familiar to him at least through Nietzsche's critique from which he himself had cited in his previous lecture. We can assume that Ortega refrains at this point from presenting the idea of self-preservation because of a central reason: the necessity to separate between the objectives of ethics and those of biology. Ortega assumes that it is not the role of the philosopher to decide for the biologists: "It is not urgent to solve this question now because it is a subject for biology".[49]

At this point we can see again that the separation between the objectives of philosophy or ethics and biology is central for understanding Ortega's attitude towards Darwinism. The repeated argument is that the philosopher cannot assign goals for the biologist's description of life. Furthermore, as we have seen before in the third lecture Ortega also refrained himself from presenting or reexamining the goals that the biologists themselves assigned to life. At that point in his lecture Ortega only mentions what the Darwinists consider as life's goal and does not consider other viewpoints. Why is it important for him to mention only the Darwinian point of view during this part of his lecture? Ortega does not give any explicit reason, but his choice to present the Darwinist point of view (in a manner that he conceives as unsettled in biology) might suggest that Ortega recognizes the social and scientific importance of Darwinism. Therefore, we need to ask how according to Ortega the Darwinians

[48] "Cuál sea el fin a quien todo lo vital sirve como medio es cosa que discuten los biólogos", Ortega, 1996, 76.
[49] "Pero aún reducidos a la finalidad biológica, no nos urge ahora resolver la cuestión entre otras porque es asunto para el biólogo", Ortega, 1996, 76.

explained the goal of life. Ortega asserts that for the Darwinians the goal of all the organs is to serve as a medium for another ultimate goal. For the Darwinians the ultimate goal is the ability to adapt to the physical environment.[50]

Ortega's argument is that not all the biologists concur with the definition regarding the ultimate goal of the biological organism; however he does not intend to examine the other biological definitions apart from the Darwinian one. The Darwinists conceive the ability of the biological organs to adapt to nature as a goal while others may present other goals. However it is very important to stress the central point that for the philosopher all the biologists are basing their explanation on naturalism. Therefore, Ortega concludes that the philosopher should not ascribe any goals of life for the biologist since for the philosopher the real goal of life is ethical and it stands beyond the naturalistic explanation. Hence, here we see again how Ortega separates between the ethic's way of defining what life's goal is and the biological definition of the goal of the human organism. According to Ortega, the ultimate question of the life' goal is an ethical problem that should be handled by the philosophers and not by the biologists:

> What is the objective that all the vital (organs) serve as a medium is a thing that the biologist debate. I refer strictly to the biological objective, the objective of the individual organism. I am not referring here to the objective or the ultimate meaning that is given to life in the universe: here it is an ethical problem, the biggest problem of the human being.[51]

This explanation of Ortega obligates us to reconsider his interpretation of Darwinians explanations of life's goal. It seems that Ortega is aware of the importance of ethic and cannot accept the narrow naturalist description of the goal of human life. This separation between two kind of human goals- one naturalistic and the other

[50] "Para el darwinismo triunfante de la pasada centuria el fin de la vida, el resorte esencial de la vida, la adaptación al medio físico", Ortega, 1996, 76.
[51] Ortega, 1996, 76.

ethical, suggests the understanding of the necessary separation between what there is and what there ought to be. This approach to the objectives of the philosopher and the biologist does not necessarily suggest any implicit intention to deny or diminish the importance of Darwinism. I believe that this attitude is aimed to acknowledge the limits of the Darwin's theory. This separation is not based on an argument that considers the 'struggle for survival' as dangerous or totally unethical. On the contrary, Ortega is not aiming to disrespect the Darwinians. The only serious interest that we can learn from the lecture is that there are two realms: the realm of what there is and the realm of what there ought to be. In the following sections we will see how this analysis relates to Ortega's central ideas regarding the mind and the human vocation.

VI. Human Reason: From Vital Utility towards Looking for the Truth for Itself.

The dialogue with Darwin and the biological explanation of human mind and character appears in almost all of Ortega's lectures in Argentina. The discussion in these lectures revolves around psychology or more specifically around the behaviorist theory and the biologist explanation of human mind. The discussion suggested by Ortega is aimed to explain two central new ideas:

1. A theory of knowledge that aims to present an alternative to the traditional debate between idealism and realism. This important theory that considers human life as a 'radical reality' is examined in the historiography about Ortega.[52]

[52] The researchers are basing their arguments on ideas from Ortega's other writings. The nine lectures in Argentina are hardly examined.

> 2. Another central idea is the concept of human vocation. I believe that Ortega is interested here in examining the idea of human vocation along and beyond the lines of the biologist's explanation. On the one hand, the vocation cannot be conceived metaphysically, religiously or mystically since it has its roots in evolution. Since Ortega does not conceive evolution in a mentalist manner he does not believe that the human vocation is part of a designed plan. On the other hand, human vocation is a 'born' (nativo) system of preferences. Therefore, these human tendencies and preferences might be conceived as if they were meant intentionally to be designed in a special way.[53]

In this section I would like to concentrate on one central concept of Ortega that I believe is developed as part of the dialogue he maintains with the biologists. The human character and its vocation is an idea that is aimed to overcome the limited Darwinian description while being able to stay within the Darwinian discourse. My argument here is that we need to reconsider Ortega's description as an idea that originates from a dialogue with Darwin. In order to present this argument we need first to examine the way in which Ortega describes the evolution of human cognition and the evolution of the human objective of knowing.

In biology we describe the different human organs as a system of utile functions. The legs, ears, eyes have different functions and their utility is examined according to how well they serve for certain needs. Human consciousness (consciencia) has also a function; it is supposed to serve the human for certain needs he has to accomplish. Hence, in order to understand these needs we have to analyze the mind’s function according to the principle of vital utility. Ortega's argument is that the process of perception of reality serves also a certain function:

[53] Ortega, 1996, 128.

"In principal there is nothing that exists in the individual that cannot be understood according to its vital utility. One of these functions is what we define as consciousness".[54] This argument of Ortega implies that in the same manner as all the human biological organs the human mind cannot be described in a spiritual manner. Here it seems that Ortega follows the guiding lines of naturalism. It seems that Ortega who will later crystallize a metaphysical conception of human life does not try to refute the naturalistic assumption according to which all events in the universe are natural. The act of knowing the immediate surrounding is the same as all the natural phenomena that occur in the human body. This argument that follows the naturalist theory helps Ortega avoid dealing with the metaphysical puzzle of how the mental process can causally influence the non-mental process.

Although he adopts a naturalistic attitude towards the mind we can notice that in his other lectures Ortega is presenting arguments that can be understood as a way to refute any interpretation of the human act of knowing in a pure naturalist manner:

> 1. In his seventh lecture Ortega criticizes those who make equivalence between the human psyche and the brain. He claims that understanding the 'brain' (cerebro) as an equivalent to the human psyche signifies an effort to return to the traditional 'myth' of the human soul. Identifying the brain with the psyche is a reduction of psychology to physiology. Hence, Ortega stresses the necessity to overcome what he conceives as a kind of reductionism. His effort to stress the difference between psychology and physiology is evident and symbolizes an intention to go beyond the naturalization of the mind.[55]

[54] Ortega, 1996, 48.
[55] Ortega, 1996, 130-31.

2. Another important argument is to be found in his eighth lecture. Here, Ortega argues that there is a necessity to separate between two conceptions of the mind: the mind as something that is composed of things on the one hand or as an entity that performs acts on the other hand. Human mind is to be described according to the acts it makes; our psyche is pure activity (actividad). Ortega does not elaborate what he exactly means by this description of the psyche as activity. It seems as an idea that resembles to Brentano and Husserl's intentionality since he also explicitly mentions the importance of their ideas: "The reference to objects, intentionality is the character or the plasma that differentiates and distinguishes the psyche. We are separating here from the modern doctrines; with this (idea) we are trying to overcome modernity".[56] It is important to stress that this argument according to which the brain should not be conceived as a 'thing', can also be conceived as part of the general critique on traditional metaphysical speculations regarding the existence of the human soul or pure spiritual entity. However, in the same manner it might also be understood as part of a critique on the conception of the human mind in a naturalist manner, a conception that Ortega presented in the first and fourth lectures.[57]

In his first lecture Ortega presented a naturalistic conception of the human mind. However, when we reexamine it closely we can see that there is also a simultaneous effort to avoid from referring to the mind as a pure natural process. One possible way to solve this puzzle can be made when we consider how Ortega perceives the descent of man and the evolution of cognition. In his lectures Ortega tries to

[56] Ortega, 1996, 148.
[57] Ortega, 1996, 148.

present two opposite descriptions of the human mind. I believe that what separates between these two different descriptions of the same phenomena can be explained by the process of evolution.

The fact that Ortega conceives the brain as a natural organ that has a certain function does not prevent our ability to focus on another way in which he describes the mind. I believe that Ortega tries to crystallize a plausible argument that might allow us to settle between these two opposite functions of the mind. We can understand his choice to reconcile between the two opposite conceptions by trying to perceive how he depicts human character and its vocation as a product of Natural Selection.

The application of the principle of vital utility (utilidad vital) to the human process of gaining knowledge on the surrounding means that truth or the things in the world are conceived according to the restricted subjective point of view. This principle leads to a separation between things as they are conceived by us and things as they really are in the world without our subjective utile perception of them. Applying the concept of vital utility to the act of knowing leads Ortega to argue that at the beginning of the evolution of the human species the brain functioned as a "utilitarian deformation of reality".[58] However, it seems that Ortega believes there is also much evidence that this historical period has gradually changed: "the biological evolution is long, but it is important to (show) its manifestation. Therefore I have to make this summary".[59] The summary is presented explicitly in the first and the fourth lectures although it is also shortly represented in other lectures. Ortega's argument is that at the beginning of evolution to know meant to enable the human to act in a useful manner without any interest in finding the truth as it is for itself. However, later in human evolution to know becomes a goal for its own sake. At a certain

[58] Ortega, 1996, 49.
[59] Ortega, 1996, 48.

point the process of knowing 'abandons' (abandona) the principle of utility. Here we find the central point in Ortega's understanding of the implications of the gradual evolution; at this point we can also notice the importance of Darwin's Theory of Evolution in Ortega's description of the act of knowing. However, Ortega does not try to explain what he conceives as the causes of the historical change. We can only see that for him the historical change is probably most notable in the appearance of philosophy and metaphysics in Greece. Ortega says that "it is very interesting that there is a moment in human history", in which few people in Greece started to be surprised by the nature of the things around them and started to reflect on their real nature.[60] The causes of the change are not analyzed by Ortega and he only argues that the birth of rationality appeared when the Asiatic people immigrated to Greece. At the arrival to Greece these immigrants under the leadership of Tales started to ask what the nature of things is: "Tales de Mileto is the first person who abandons the mythology... he starts asking "what the thing is?"[61]. This change as it is depicted by Ortega is not only another small historical change but a radical change in the gradual and long evolution of the human being. Ortega implicitly suggests that this change is one of the most important changes in human evolution. Philosophy and the quest for ultimate truth, and its call to re-examine our basic assumptions about reality, represents an inflection point in evolution of human attitudes towards the world—before, we had mythological thinking, after we had philosophy. The results of this ostensibly gradual change have been historically dramatic.

This historical change is represented in extreme manner by the modern rationalist. The rationalist believes that human thought can set us free from any utilitarian attitudes and can help us know the real

[60] Ortega, 1996, 49.
[61] Ortega, 1996, 47.

nature of things. In his lectures Ortega has a rather complicated attitude towards rationalism. On the one hand, he criticizes the rationalists for forgetting the origin of the process of knowing: the long process of evolution. On the other hand, Ortega believes that the consequences of gradual evolution and the overcoming of the constraint of utility can allow us freedom, a freedom to follow or not to follow our vocation.[62]

VII. Human Character and its Vocation.

Each human being has different preferences, attitudes and predilections. As some people do it is of course possible to treat the existence of these personal preferences as a wonder. The mystic will define it as a vocation or destiny designed for each human being. Since for the mystic the experience of human vocation is a wonder, he is not troubled to give a detailed explanation for the causes of the phenomena. The mystic's main effort is to explain the possible consequences of the phenomena on a person's life. Ortega adopts a concept that has mystical resonance; however he aims to go further than the mystic. In his lectures he argues that each person or people have different human tendencies and characteristics. The perspective we have is part of our tendencies. Ortega's effort is to analyze the process in which the human character has been shaped. His main argument is that each human has a "system of preferences that are produced by nature".[63]

Ortega argues that one person might be practical while the other is more theoretical, there is one who is interested in art while other person may be even totally blind to the art. Is there a possibility to

[62] Ortega, 1996, 94.
[63] Ortega, 1996, 128.

explain human tendencies and attitudes without using mystical or metaphysical explanation? I believe that Ortega implicitly tries to present an argument that is based on the Theory of Evolution. By describing the human characteristics as products of natural evolution Ortega also implicitly intends to present the philosophical implications of Natural Selection.

Theoretically each human can have the same interest towards each and every thing surrounding him. Attention cannot be given to all the things that exist around us. In order to concentrate on one object or one problem we need to distract ourselves from focusing on another object: "As I was saying, we can only focus ourselves on a single point when we manage to divert our attention from all the others".[64] There is a natural limit to the potential of human attention: a person who is troubled solving a mathematical problem must first distract himself from other problems. Getting to know each person better means we have to try consider towards what objects he gives his attention: "Tell me towards what you attend to and I will tell you who you are".[65] Following this description of Ortega it may seem as if each human has an active role in choosing his predilections and tendencies. Each human is free to choose his own vocation. Furthermore, we have seen that Ortega criticizes Darwin and argues that the evolutionary process of adaptation cannot explain human creativity. Hence we might even argue that Ortega's idea of free will can be described as a capability of wanting to be different in our preferences and purposes. However, I believe that this kind of reading is evidently mistaken. In his lecture he argues that each consciousness has its frame of objects and that frame of mind or what he calls "tessitura" is a product of natural evolution.

[64] Ortega, 1996, 127.
[65] Ortega, 1996, 128.

The idea of human vocation is very important for understanding how Ortega conceives the human will. By analyzing Ortega's concept of free will we can see that it should be understood as an ability to choose to follow or to refuse to follow our characteristic. A person which has a tendency towards mathematics can choose to distract his attention from other objects around him in order to concentrate on solving a certain problem. He can also choose to refrain from solving the problem. However, this person is not able to choose to have the same tendency towards art or music. The human characteristics or the mind frame are natural products of evolution. In his lecture Ortega refrains himself from describing the origin of mind frame (tessitura): "From where it (the mind frame- tessitura) came from or how it has been shaped, I cannot tell it now and here".[66] Ortega refrains himself from explaining the process of evolution however it is definitely conceived a thing acquired naturally.

The tendencies we have are not actively chosen by us but are rather chosen for us during the long gradual process of evolution. Ortega defines human characteristics as part of the "frame of mind". He suggests that characteristics are "produced by nature" (nativo): "Each individual has a system of preferences produced by nature".[67] Furthermore, since the tendencies and the frame of mind are a product of nature it will be impossible to try and change them: The Spanish thinker Juan de Valdés was right in saying that to change the customs is the same as choosing to die".[68] We can conclude by saying that human vocation is the ability to follow or not to follow our tendencies and preferences. We cannot change our characteristic but we can indeed choose when we want to follow them and when we do not. Liberty exists but within the limits of the frame of our mind,

[66] Ortega, 1996, 127.
[67] Ortega,1996, 128.
[68] Ortega, 1996, 129.

perspective and character. These are products of Natural Selection. Ortega implicitly follows Darwin and his critique is rather a conscious intention to hide the deep connection between his and Darwin's thought.

VIII. Summary.

In his writings Ortega referred to the differences between his own philosophy of life and the biological conception of life. The biological conception of Lamarck or Darwin interested him and in his early thought he was influenced by Darwin's concept of Natural Selection. In his later work Ortega presented human life in a manner different from Darwin from many aspects: the focus on the future dimension, the inner desired fantasies of the subject and especially the dissatisfaction from the circumstances. The argument of Ortega was that it is the subject who reforms the environment to its favor. But early in his writings Ortega followed Darwin's Natural Selection and the biological interpretation of the human character stands at the center of his thought. Therefore, we are not allowed to conclude that Ortega's mature philosophy is a Darwinist philosophy, on the contrary. But we also cannot ignore the huge influence the ideas of Darwin had on the thought of the young Ortega. Ortega's conception of vocation of man is an elaboration of Darwin's Theory of Evolution.

Acknowledgements

I would like to thank my family for their support and understanding and for allowing me to dedicate my free time to individual research. I would also like to thank Kairos' three anonymous reviewers whose comments helped me present Ortega's analysis of Darwin's concept of Natural Selection in a much wider historical context.

References

Cerezo, P., 2011, *José Ortega y Gasset y la razón práctica*, Biblioteca Nueva, Madrid.

Darwin C., 1998, Struggle for Existence, *The Origin of the Species*, Wordsworth.

-- 2010, *The Descent of Man*, Abridged and with an Introduction by Michael Ghiselin, New York, Dover Publications.

Dawkins R., 1986, *The Blind Watchmaker*, Harmondsworth, Penguin

Dennett D., 2009, In Darwin's Wake, Where am I?, In: *The Cambridge Companion to Charles Darwin*, Second Edition, Edited by Jonathan Hodge and Gregory Radick Cambridge: Cambridge University Press

Derrida, J., 2002, *The Animal That Therefore I am*, Translated by David Wills, Critical Inquiry, 28.

Marías, J., 2000, *Acerca de Ortega*, Espasa libros.

Nietzsche, F., 1968, *The Will to Power*, Edited by Walter Kaufmann, New York.

-- 1997, *Beyond Good and Evil. Prelude to a Philosophy of the Future*, New York, Dover Publications.

-- 2004, *The Twilight of the Idols*. New York, Dover Publications.

Ortega y Gasset J., 1902, "Glosas", *Vida nueva*. In: *Obras completas*, Tomo I, Madrid, Alianza Editorial, 1983.

-- 1908, "El sobrehombre", *El Imparcial*. In: *Obras completas*, Tomo I. Madrid, Alianza Editorial, 1983.

-- 1924, Ni vitalismo ni racionalismo, *Obras completas*, Tomo III, Madrid.

-- 1932, Goethe desde dentro, in *Obras Completas*, Tomo III, Madrid, pp. 395-420.

-- 1984, Qué es conocimiento? (1930), *Revista de Occidente*, Alianza Editorial, Madrid.

-- 1996, *Meditación de nuestro tiempo. Las conferencias de Buenos Aires, 1916 y 1928* (1916, 1928), Edición de José Luis Molinuevo, Primera edición, Fondo de Cultura Económica.

-- 2004, Meditación de la técnica y otros ensayos sobre ciencia y filosofía (1933), *Revista de Occidente*, Alianza Editorial, Madrid.

-- 2004, Conferencia en Darmstadt (1951), in: *Meditacion de la técnica y otros ensayos sobre ciencia y filosofía*, Alianza Editorial, Madrid, pp. 99-108.

--2010, *Meditaciones de Quijote* (1914), Edición de Julián Marías, Octava Edición. Madrid, Cátedra- letras Hispánicas.

Pearson, K. A., 1997, Nietzsche contra Darwin, in: *Viroid Life: Perspectives on Nietzsche and the Transhuman Condition*. London, Routledge, pp. 85-122.

Richardson J., 2004, *Nietzsche's New Darwinism*, Oxford University Press.
Rodríguez Huéscar A., 2002, *La innovación metafísica de Ortega. Crítica y superación del idealismo*, Madrid, Biblioteca Nueva.

Ruse M., 2003, Belief in God in a Darwinian Age, in: *The Cambridge Companion to Darwin*.

Sobejano, G., 1967, *Nietzsche en España*, Madrid, Gredos.

Solan P., 2005, It Might be Called Reverence, In: *Darwinism and Philosophy*, Edited by Vittorio Hösle and Christian Illies. Notre Dame, Indiana, University of Notre Dame Press.

Wallace, J., 1998, Introduction, In: Darwin, The Origin of the Species. Wordsworth,

Wandschneider D., 2005, On the Problem of Direction and Goal in Biological Evolution, In: *Darwinism and Philosophy*. University of Notre Dame Press.

Les présupposés théologiques de la conscience historique moderne : une lecture löwithienne du sens de l'histoire

Laura P. Kassar
Université d'Ottawa
lpyco081@uottawa.ca

Abstract What Karl Löwith describes in *Meaning in History* is first and foremost the conflictual nature of modernity. By carefully unveiling throughout his work what appears to be the fragile, even seemingly "illegitimate" architecture upon which modernity's relation to history and to philosophy of history itself is built, Löwith reveals the profound dependency that ties philosophy of history to its theological presuppositions. According to him, for philosophy of history to be able to fully come to grasp with its own meaning, it must be understood within the framework of Christian theological notions: indeed, for Löwith, modernity has to be interpreted in terms of a secularization of the Christian idea of eschatology. The study of the main concepts found in *Meaning in History* that we propose to undertake here aims at a better understanding of the problem of modernity's relation to "thinking historically" and its problematic entanglement in the Judeo-Christian tradition.

Resumo O que Karl Löwith descreve em *Meaning in History* é, em primeiro lugar e acima de tudo, a natureza conflitual da modernidade. Revelando cuidadosamente, através de toda a sua obra, aquela que parece ser a frágil, e mesmo a aparentemente "ilegítima", arquitectura sobre a qual foi erguida a relação da modernidade com a história e com a filosofia da história, Löwith revela a profunda dependência que vincula a filosofia da história às suas pressuposições teológicas. Segundo Löwith, só se poderá compreender o significado da filosofia da história, se esta for concebida com base em noções teológicas cristãs. Com efeito, para o filósofo alemão, a modernidade deve ser interpretada à luz de um processo de secularização da noção cristã de escatologia. A nossa análise dos principais conceitos de *Meaning in History* tem como objetivo fornecer uma melhor compreensão do problema da relação da modernidade com o "pensamento histórico" e do seu problemático emaranhamento na tradição judaico-cristã.

C'est dans son ouvrage intitulé *Histoire et Salut – les présupposés théologiques de la philosophie de l'histoire* que Karl Löwith formule son projet de montrer que la philosophie moderne de l'histoire s'enracine dans les dogmes chrétiens de la rédemption et que cette même philosophie de l'histoire résulte d'une sécularisation du modèle de l'eschatologique biblique. À l'époque où Löwith écrit, la thèse qu'il soutient n'est pas complètement neuve. Avant lui, le travail sur la sécularisation était déjà entamé chez des auteurs tels que Max Weber et Carl Schmitt, de même que chez Nietzsche. L'originalité de Löwith sur la question théologico-politique se situe sans doute au niveau de sa méthode : en retraçant un parcours dans l'ordre chronologique *inversé*, il va examiner différents moments importants de la constitution de la pensée historique moderne, partant de Burckhardt et Marx, rebroussant jusqu'à Augustin et Orose, et passant bien entendu par Hegel, auprès duquel se cristallise le terme « philosophie de l'histoire » dans l'ampleur de sa signification moderne. Au fil du livre et à l'envers du temps chronologique, Löwith travaille à développer dans les détails cette notion de philosophie de l'histoire, et cela de façon à mener vers une prise de conscience véritable des sources qui animent et posent réellement son héritage.

La thèse de Löwith est la suivante : la philosophie de l'histoire s'ancre profondément dans des concepts d'abord théologiques, et plus précisément ; la pensée historique moderne ne prend pleinement son sens que lorsqu'interprétée dans le cadre d'une eschatologie chrétienne sécularisée, c'est-à-dire *mondanisée,* traduite dans un monde terrestre dépouillé de Dieu comme altérité radicale. Cette idée soutenue par Löwith ouvre la voie aux questions suivantes : après la sortie du christianisme, – où l'expression « sortie du christianisme » est entendue à la manière de Marcel Gauchet – comment faut-il désormais penser l'histoire ? Est-il même possible d'envisager un dépassement du modèle de la pensée historique qui nous est légué

par la religion chrétienne ? Ou du moins, serait-il envisageable de libérer la modernité d'un *télos* chrétien qui lui semble décidément anachronique ? Ces interrogations sont laissées en suspens dans *Histoire et salut,* où Löwith n'offre pas de réponse claire. La conclusion de l'ouvrage ainsi que le scepticisme qu'on devine chez lui traduisent d'ailleurs son hésitation à poser un diagnostic définitif. La réflexion qu'il entreprend dans *Histoire et Salut* permet toutefois de démêler des notions indispensables à l'étude du problème moderne du sens de l'histoire et de son inscription dans la tradition judéo-chrétienne. D'une certaine façon, tout se trouve déjà dans le sous-titre : ce sont les *présupposés théologiques de la philosophie de l'histoire* qui font l'objet de la réflexion de Löwith, et c'est avec lui que nous nous proposons d'examiner quels sont effectivement ces présupposés, de même que les implications de ceux-ci. L'objectif de l'enquête qui se fera ici ne sera pas de répondre aux questionnements laissées en suspens par l'auteur d'*Histoire et salut*, ni non plus de nous situer dans le débat autour de la question de la sécularisation, mais bien d'approfondir les indices que Löwith nous lègue pour une compréhension de la modernité postchrétienne qui caractérise la conscience historique qui marque encore notre époque, à l'aube du XXI^{e} siècle. Avec l'intervention d'auteurs tels que O. Cullmann, H. Blumenberg et J-C. Monod, nous chercherons par le fait même à explorer les différentes thèses qui se dessinent autour de la réflexion sur la philosophie de l'histoire au XX^{e} siècle, de façon à révéler, en fin de parcours, la position paradoxale qu'endosse Löwith sur la question de la conscience historique moderne, position paradoxale qui le maintien non seulement dans le sillage des catégories judéo-chrétiennes desquelles il ne semble pouvoir échapper, mais qui l'enserre plus précisément dans les présupposés de la tradition idéaliste allemande et de son achèvement dans la pensée de Hegel.

La conscience historique moderne : grecque ou chrétienne ?

Dans son ouvrage, Löwith décrit puis oppose ce qu'il présente comme les deux différentes possibilités quant à la manière dont les hommes sont entrés en rapport avec l'histoire, qu'il classifie en deux modèles de conscience historique : d'une part, le modèle des Anciens (c'est-à-dire le modèle grec), et d'autre part le modèle chrétien. Ces deux modèles sont qualifiés par Löwith comme se situant à l'antithèse l'un de l'autre, et c'est en partie à partir de cette antithèse qu'il explique les contradictions et la confusion de la conscience historique moderne. D'abord, l'Antiquité grecque pensait essentiellement le temps du monde comme un cycle, donc comme issu d'un cosmos ordonné dont l'orientation était fondamentalement « anhistorique », autrement dit : détachée de tout contexte historique ou de toute succession d'évènements inter-reliés. L'expérience historique des Grecs était celle de la fiabilité d'une nature humaine reproduisant certaines tendances de façon cyclique, dans une alternance de «phases de bonheur et de phases d'échec».[1] De même en politique, où les victoires et les échecs se succédaient selon un mouvement naturel, circulaire, en harmonie avec celui même du cosmos dans lequel ils se déroulaient. Une philosophie de l'histoire aurait donc été un non-sens pour les Grecs, qui n'avaient de rapport à l'avenir que celui d'être relatif au présent. Il n'y avait, d'ailleurs, pas de mot dans le vocabulaire grec pour désigner le terme « histoire » au singulier, terme qui apparaît si évident à la connaissance moderne, qu'il semble presque impensable qu'il n'ait pas toujours fait partie du langage. Toutefois, les Grecs n'avaient effectivement de notion que *des* histoires (*historiae*) au pluriel, ce qui signifie qu'ils ne se rapportaient qu'à des évènements particuliers, non pas à une histoire du monde

[1] Löwith, 2002, 29-30.

comme pourvue d'un sens et d'un lien, encore moins d'une direction. Selon Löwith, cette absence de philosophie de l'histoire chez les Grecs n'est toutefois pas due à un aveuglement face aux grands évènements historiques, mais plutôt à la lucidité qu'ils conservaient sur un monde soumis au hasard et au changement : il était possible d'en dégager un récit, mais non pas des certitudes. L'historien classique questionnait donc davantage les *causes* derrière les événements, sans avoir à se soucier ni se demander, comme le fera plus tard l'historien moderne, aussi bien que Karl Löwith lui-même: « comment cela va-t-il se poursuivre ? ».[2]

C'est avec l'avènement de la tradition judéo-chrétienne que s'opère une première rupture profonde d'avec le schéma de la conception ancienne du rapport humain au temps. Pour les juifs et les chrétiens, l'histoire devient comprise avant tout comme une histoire du Salut. La déchirure radicale qui s'effectue entre Dieu et le monde, alliée à la croyance eschatologique en « une fin ultime de l'histoire »,[3] c'est-à-dire en la fin des temps sous forme du Jugement dernier, est ce qui fait, pour la majorité de penseurs de la question de la sécularisation, la spécificité de cette tradition. Les premiers chrétiens, en reconnaissant en Jésus la figure du Christ, ont pensé que la venue du Royaume de Dieu était imminente, que la réalisation de l'eschatologie était directement associée à la mort corporelle de Jésus, et que c'est précisément la mort du Christ qui ferait sonner l'heure du Jugement dernier.[4] De toute évidence, le Royaume s'est fait attendre, et à l'époque des premiers chrétiens, l'interprétation de l'eschatologie en fut bouleversée. Nonobstant, c'est la reconnaissance de l'événement christique propre à la religion chrétienne qui situe, selon Löwith, le point de départ à notre conception linéaire de l'histoire. En

[2] Löwith, 2002, 40.
[3] Löwith, 2002, 25.
[4] Cullmann, 1966, 25-27.

effet, le calendrier occidental dont nous avons hérité ne part pas d'un point initial, mais bien d'un point *central*, la naissance de Jésus ; et à partir de ce point, fusent deux numérotation inversées, l'une remontant indéfiniment dans le passé (av. J.-C.), et l'autre projetant indéfiniment dans l'avenir (apr. J.-C.). Le théologien Oscar Cullmann explique :

> La théologie affirme que c'est à partir de cet événement central que l'histoire dans son ensemble doit être comprise et jugée. Cet événement auquel on assigne la date de l'an 1 constitue le sens dernier et le critère de toute l'histoire qui l'a précédé et qui l'a suivi.[5]

Désormais, « fin » (*Zweck*) de l'histoire et « sens » (*Sinn*) de l'histoire viennent à se confondre dans un rapport nouveau des évènements historiques à une *fin*, qui pose un but (*Ziel*) à venir[6] : les évènements isolés prennent tout à coup *sens* de par leur inscription nécessaire dans ce déploiement linéaire qui culmine dans l'accomplissement effectif de l'eschatologie, mettant ainsi fin au temps du monde humain, et au monde lui-même. Nous assistons, par rapport à la vision du monde qu'avaient les Grecs, à un rétrécissement de ce dernier. De cette notion de rétrécissement tel qu'exposé par Löwith, on retrouve une interprétation intéressante chez Habermas : le monde, dit-il, « qui, par nature, est, se tient en lui-même, s'éteint et se forme à nouveau, se voit soumis à une histoire du Salut qui le ravale au rang de création transitoire, advenant en vue de l'homme et non plus en vue de soi ».[7] À partir du moment où apparaît la croyance en un monde « d'après », en un monde réuni avec Dieu, la conscience se tourne vers l'avenir, vers un ultime futur où se situe l'essentiel de l'expérience humaine. La croyance au Salut replace le monde humain au milieu de deux autres temps propres au Nouveau Testament : soit,

[5] Cullmann, 1957,13.
[6] Löwith, 2002, 273.
[7] Habermas, 1974, 158.

celui qui précède la création, où le plan divin se trouvait déjà préparé, et celui qui vient, celui qui contient les évènements « finaux ». En ce sens, le monde prend véritablement la forme d'une transition qui prépare un avenir attendu et espéré. Passé et présent préparent l'avènement d'un Royaume futur, alors que l'histoire de ce monde ci est celle de la souffrance humaine, de l'espérance, de la crainte. Avec l'arrivée du christianisme et impensable sans elle, une philosophie de l'histoire devient désormais possible, et c'est précisément cette question de la philosophie de l'histoire qui nous intéressera principalement ici.

Il est important de noter que si, pour Löwith, le questionnement sur le sens de l'histoire est possible grâce au caractère singulier de la religion chrétienne, son interprétation peut aussi devenir dangereuse, surtout lorsque déracinée de son contexte proprement théologique. Par là s'explique en partie la position inconfortable dans laquelle se trouve l'Occident postchrétien. La conscience historique moderne s'est certes débarrassée de l'organisation chrétienne du monde social et du monde sacré, mais elle demeurerait prise, selon Löwith, dans certains de ses schèmes de pensée, notamment celui d'une historicité linéaire basée sur l'idée du présent comme préparation à un avenir compris comme accomplissement dans l' « au-delà ». D'un autre côté, la conscience scientifique moderne renoue en plusieurs points avec la vision que les Anciens avaient du cosmos, c'est-à-dire celle d'un regard rationnel qui observe et mesure les récurrences des phénomènes de la nature. Ainsi, l'esprit moderne est indécis, il « regarde le monde de deux yeux différents, celui de la foi et celui de la raison ».[8] En effet, la modernité tente de réconcilier les deux pôles constitutifs mais contradictoires qui l'animent, toutefois sans parvenir à rester pleinement cohérente ni avec l'un, ni avec l'autre de ces pôles, et sans non plus arriver à atteindre un nouvel

[8] Löwith, 2002, 255.

équilibre. L'interprétation de la modernité ne saurait donc se faire que selon des variations sur les deux modèles de la conscience historique que propose Löwith. L'une des thèses fortes d'*Histoire et Salut* consiste à exprimer le fait que le monde moderne reste effectivement à la fois chrétien sans y croire, étant dépouillé de toute perspective véritablement religieuse, ainsi que grec sans l'être réellement, parce que maintenu dans ce schéma d'historicité linéaire qui aurait été insensé pour les penseurs de l'Antiquité. Et si le rapport de l'homme à sa représentation du monde a basculé avec la modernité, l'expérience humaine de l'Histoire comme constituée d'épreuves et d'échecs durables reste pourtant universelle.

D'ailleurs, à l'époque où écrit Löwith, les dangers liés à une absolutisation de l'histoire mondaine sont bien présents à l'esprit de l'auteur, et cela explique dans une certaine mesure la méfiance dont il fait preuve à l'égard du modèle chrétien de la conscience historique ; celle-ci demeure, selon lui, directement héritière de la structure eschatologique chrétienne. Löwith médite les conséquences possibles d'un tel « hybride » dans la manière de se rapporter à l'histoire, et affirme qu'« il est effrayant, mais conforme à l'esprit du Nouveau Testament, de penser que cette répétition d'actions et d'épreuves à travers toutes les époques est nécessaire pour parachever la passion du Christ ».[9] Le phénomène des « religions » dites séculières, qui sont apparues sous diverses formes au XX^e^ siècle et dont l'une des manifestations est le nazisme durant la Seconde Guerre Mondiale, doivent donc être considérés, explique Löwith, comme un résultat dangereux issu de la confusion de la modernité par rapport ses fondements.

[9] Löwith, 2002, 235.

Quel emploi de la terminologie théologique ?

Avant d'entrer plus avant dans la thèse de Löwith, il nous apparaît pertinent de prendre le temps de clarifier certains termes qu'il emprunte à la théologie, mais dont l'usage qu'il fait de ces termes dépasse bien souvent leur cadre spécifiquement théologique ; cadre qu'il ne décrit d'ailleurs que très sommairement, n'étant pas théologien avant tout, mais bien philosophe. Ainsi, nous pouvons questionner la légitimité du recours de Löwith aux concepts théologiques qu'il met en place dans *Histoire et salut* : l'usage qu'il en fait est-il un usage terminologique théologiquement et conceptuellement fidèle ? N'y a-t-il pas ce que nous pourrions supposer être un souci rhétorique chez Löwith, qui animerait son emploi de certains concepts et qui vaudrait la peine d'être éclairé ? Il nous semble donc important, pour bien saisir la valeur ou encore les lacunes des propos de Löwith, de replacer quelques concepts centraux d'*Histoire et salut* (tels que ceux d'histoire, d'eschatologie, de sécularisation, de progrès) dans leur contexte originel. Nous nous proposons donc de confronter l'appropriation löwithienne de la terminologie théologique employée dans *Histoire et salut* à d'autres appropriations faites par des auteurs traitant également du problème et du rôle de la sécularisation dans la genèse de la modernité. Nous trouvons d'une part dans les œuvres d'Oscar Cullmann, contemporain de Löwith et théologien suisse du vingtième siècle – d'ailleurs cité par Löwith dans *Histoire et salut* à multiples occasions – des pistes quant à l'origine des expressions d'« histoire du salut », ainsi que du terme d'«eschatologie». D'autre part, pour ce qui est du terme de «sécularisation» en tant que telle, nous en trouvons une bonne histoire retracée par Jean-Claude Monod dans son ouvrage *La querelle de la sécularisation – de Hegel à Blumenberg*. Nous faisons donc ici appel aux élaborations théoriques de Cullmann et Monod, auxquelles nous

ajouterons plus tard la contribution de Blumenberg, pour poser dans un premier temps le cadre de notre examen de la terminologie à partir de laquelle Löwith bâtit les thèses de son ouvrage.

D'abord, si l'on se fie à Cullmann dans *Le salut dans l'histoire*, le terme « histoire du salut » n'est pas biblique[10] : c'est-à-dire que l'on ne retrouve aucune occurrence de cette expression dans le Nouveau Testament. On pourrait donc remettre en question, tout comme le fait Löwith, la légitimité même du recours à l'expression «histoire du salut» tout court : celle-ci serait mal choisie, et de faire une lecture du Nouveau Testament à la lumière d'une « histoire du salut » serait aborder la Bible avec une catégorie (celle de l'*histoire*) qui n'est pas sienne de prime abord. Cependant, il ne serait pas forcément une erreur d'affirmer qu'il y a de fait dans les écritures du Nouveau Testament une « histoire du salut », dans la mesure où celle-ci est effectivement composée de certains évènements significatifs liés (par exemple, la naissance et la mort du Christ, etc.,) qui se déroulent *dans* l'histoire profane, donc à même l'histoire du monde. En ce sens, il est possible du moins de faire une *analogie* entre l'histoire du salut et l'histoire mondaine, dans la mesure où toutes les deux sont composées d'évènements imbriqués les uns dans les autres, qu'elles se déroulent toutes les deux dans le monde terrestre, et qu'elles partagent la même part de contingence propre à celui-ci. Toutefois, Cullmann souligne l'importance de distinguer l'histoire sacrée de l'histoire profane : l'histoire du salut, nous le répétons, comporte un nombre *restreint* d'évènements significatifs, et s'écarte de toute interprétation philosophique immanente à l'histoire du monde. Autrement dit, les évènements qui composent l'histoire du salut ont lieu dans le monde terrestre, mais ne l'expliquent pas *lui* ; ils sont interprétés en vue d'un *autre*. C'est là, dans cette promiscuité sans doute involontaire des domaines du sacré et du profane au sein de la

[10] Cullmann, 1966, 70.

langue courante, que se situerait l'erreur dans l'utilisation qui nous est parvenue du terme d'« histoire du salut ». Löwith soulèvera d'ailleurs à quelques reprises les problèmes qui sont nés de cette utilisation erronée de l'expression « histoire du salut », et surtout des malentendus qui adviennent de sa transposition dans le monde profane – c'est-à-dire de sa transposition dans le domaine de l'histoire terrestre – malentendus sur lesquels nous reviendrons en détail plus loin.

Ensuite, toujours selon Cullmann, le concept d'« eschatologie » a également subi des transformations dues à l'utilisation courante qui en a été faite et qui s'est pervertie, comme nous le verrons, dès le premier siècle. Il est supposé que la source de l'expression provienne de la phrase suivante, tirée d'un livre de l'Ancien Testament intitulé le *Siracide* (aussi appelé *l'Ecclésiastique*) : « Dans tout ce que tu fais, souviens-toi de la fin » (7.36).[11] L'étymologie des termes « eschatologie », et de son adjectif « eschatologique » renvoient donc directement au temps de la fin, c'est-à-dire aux événements finaux. Un tropisme s'est effectué graduellement, et non par hasard, de sorte que le sens du concept d'eschatologie s'est progressivement déplacé du « temps de la fin » au « temps de la décision » ; ce qui peut sembler banal, dans la mesure où le Jugement dernier peut représenter à la fois le moment de la *fin* – strictement temporel – et le moment de la *décision* de la part de Dieu. La différence s'aggrave cependant dans le langage courant, car devient « eschatologique » toute situation dans laquelle l'homme est placé devant une décision. Pour Cullmann, c'est l'utilisation libérale qu'en ont fait les auteurs de la pensée dite «existentialiste», ou du moins celle attribuée aux précurseurs de l'existentialisme du XX^{e} siècle (il faut penser à Kierkegaard, entre autres), qui aurait amplifié davantage l'établissement de ce deuxième sens du terme « eschatologie », fixant désormais son sens double. En

[11] Löwith, 2002, 73.

vérité, l'idée d'une « structure eschatologique [...] [suppose] la priorité de ce qui est à venir par rapport au donné présent. Cet avenir n'est pas lié au développement de virtualités présentes »,[12] mais à sa détermination en fonction de l'événement ultérieur, c'est-à-dire en fonction de l'évènement *ultime*. Il est donc nécessaire, lorsque l'on se réfère à l'eschatologie, de garder à l'esprit son caractère d'abord et essentiellement temporel, comme signifiant non pas en premier lieu l'instant de décision, mais son rapport toujours déjà lié avec ce qui est à venir.

En ce qui concerne le concept de sécularisation que nous avons déjà survolé, le mot est d'abord forgé sur le latin « *saeculum* », qui est issu de l'expression grecque « ce siècle »[13]. Dans la Vulgate, le terme « *saeculum* » se traduit de manière à exprimer « le monde présent », par opposition au « monde d'après », lequel correspond au royaume du Christ. Autrement dit, le mot « siècle » renvoie au monde terrestre ou profane, et est connoté de manière à insister sur la distance qui sépare le royaume de Dieu de toute forme de royaume temporel. Il est important de rappeler qu'avec le christianisme s'instaure cette rupture radicale entre Dieu et le monde, rupture qui n'avait pas encore eu lieu dans aucune des formes de religion précédentes. Nous pourrions presque dire qu'avec le christianisme, Dieu devient l'anti-monde[14], dans la mesure où Dieu et le monde ne sont plus simplement éloignés l'un de l'autre, mais bien étrangers l'un à l'autre, et pensés désormais dans un rapport de tension.[15] En plus de référer au monde temporel, « siècle » désigne aussi un « rapatriement vers l'ici-bas »,[16] un « rapatriement » qu'il est possible d'interpréter de deux manières. L'on peut comprendre cet *ici-bas* ou bien dans le sens péjoratif,

[12] Florival, 2001, 433.
[13] Monod, 2002, 17.
[14] Gauchet, 1985, 189.
[15] Taubes, 2009, 48.
[16] Monod, 2002, 17.

comme le lieu du péché, ou bien simplement comme l'ensemble des activités de la vie quotidienne qui sont non-orientées vers le salut. De cette deuxième compréhension s'est alourdie la définition de la sécularisation (rappelons qu'il y a deux termes dans la langue allemande pour exprimer l'expression de « sécularisation » : *Säkularisierung* et *Verweltlichung*), définition qui sera plus tard appliquée au processus spécifique consistant à faire passer des biens ou des institutions du domaine religieux au domaine profane.

Au sens le plus strict, l'action de séculariser revient à soustraire à quelque chose (un objet, une fonction, etc.,) sa signification ou son pouvoir religieux, et de la faire entrer « dans le siècle », c'est-à-dire dans le monde temporel, mondain. Avec Hegel, le terme *Verweltlichung* vient à désigner le processus par lequel il y a « réalisation » (*Verwirklichung*) des contenus véritables de la religion dans l'ici-bas : c'est-à-dire par *transfert* des contenus religieux abstraits dans l'histoire du monde terrestre.[17] Ainsi, lorsque Löwith décrit la sécularisation du schéma de l'eschatologie chrétienne dans *Histoire et salut,* c'est en vertu de la représentation hégélienne de sécularisation comme *Verweltlichung* (qui se traduit littéralement par *mondanisation*), qu'il pense la sécularisation de l'eschatologie comme la transposition d'un certain rapport chrétien à l'événement *ultime* dans le monde profane, « déchristianisé ». Ici, la *mondanisation,* chez Löwith comme chez Hegel, « est comprise comme ''incarnation'', dans l'ordre temporel-politique, de principes jusqu'alors affirmés sous une forme religieuse (chrétienne) »[18] : c'est en ce sens que la modernité aurait, selon Löwith, effectivement vidé la croyance en l'eschatologie de son caractère strictement religieux, et donc abstrait, pour le faire advenir dans le « siècle », et plus précisément, pour le projeter au niveau de l'histoire concrète du monde.

[17] Monod, 2002, 29.
[18] Monod, 2002, 30.

Les fondements théologiques de la philosophie de l'histoire

Il a déjà été mentionné précédemment que la philosophie de l'histoire n'est devenue possible, pour Löwith, qu'avec l'apparition de la tradition judéo-chrétienne. Et que cela signifiait que l'idée même de s'interroger sur le sens ou l'absence de sens de l'histoire se trouvait à être un phénomène historiquement conditionné par la croyance chrétienne en l'eschatologie. Pour penser en termes de « philosophie de l'histoire », il faut que le regard puisse se hisser au-dessus des évènements particuliers : si la chouette de Minerve ne prend son envol qu'à la tombée de la nuit, c'est que l'histoire, et encore davantage la pensée de l'histoire comme *philosophie* de l'Histoire au sens hégélien, doit parvenir au moment où elle devient visible pour elle-même. Ce qui n'est envisageable qu'en liant les évènements entre eux et en les rapportant à un principe directeur ; et cela revient essentiellement à poser une fin ou un *eschaton*. Le sens des évènements historiques s'interprète dès lors en fonction d'une direction : l'histoire ne relate plus une simple succession d'évènements, mais exprime la progression vers un avenir qui culmine en son orientation achevée. Par rapport aux Grecs, le christianisme aurait produit une déviation du concept d'*historiae* vers le futur, vers un « en vue de quoi », et par conséquent aurait ouvert sur un nouveau modèle linéaire de l'historicité à partir duquel l'Occident postchrétien se conçoit toujours. La compréhension qu'a l'esprit moderne de lui-même est par conséquent chrétienne en ses fondements, et c'est l'un des arguments clés qu'apporte Löwith pour soutenir sa thèse selon laquelle l'Occident moderne est le produit de la sécularisation du schéma de l'eschatologie biblique.

Plus encore, Löwith avance que ces fondements sur lesquels repose la modernité représentent une construction artificielle, voire une construction illégitime : en effet, en tant qu'issu de la théologie, le

modèle chrétien de la conscience historique n'est pleinement dans son droit que lorsqu'il s'en tient à l'explication purement théologique du dogme, et non pas lorsqu'il s'aventure à prendre la forme d'une philosophie de l'histoire du monde. C'est-à-dire que pour le chrétien, il n'y a d'histoire vraiment authentique que celle du Salut, et celle-ci n'est décidément pas « histoire du monde ». La sortie de la religion, processus dont l'intention était celle d'une tentative antireligieuse et émancipatrice, n'aurait donc libéré l'Occident du christianisme qu'en *apparence*. Selon Löwith, si la modernité se conçoit comme entièrement libérée de la religion, c'est qu'en fait, elle est aveugle à l'illusion de cette émancipation : en réalité, la conscience moderne reste toujours en dette par rapport à la religion, et impensable en dehors d'elle. Ou autrement dit, «le monde moderne est moins un monde émancipé du christianisme qu'un monde chrétien sécularisé».[19] En outre, pour Löwith comme plus tard pour Gauchet, ce serait au sein même du christianisme que se trouverait le germe de sa propre « défaite »[20] : en radicalisant l'écart qui sépare le monde humain de son tout-autre, la religion chrétienne aurait rendu inopinément aux hommes leur autonomie sur le monde, et ce faisant, aurait procuré le « premier pressentiment du progrès humain ».[21] Le processus de sécularisation que décrit Löwith dans *Histoire et salut* est donc étonnamment à la fois fruit et héritage du christianisme, comme issu involontairement d'une « logique interne » à celui-ci. C'est en ce sens que Gauchet qualifiera le christianisme de la « religion de la sortie de la religion ». Nous assistons donc, après cette « sortie de la religion », à l'ère d'une eschatologie séculière postchrétienne, où Löwith affirme, à la manière de Hegel, que « l'histoire du Salut est projetée sur le plan

[19] Monod, 2002, 29.
[20] Gauchet, 1985, 284.
[21] Löwith, 2002, 105.

de l'histoire du monde, et cette dernière est élevée au rang d'histoire du Salut ».[22]

Comment cette eschatologie séculière postchrétienne se présente-t-elle à nous ? En d'autres mots, que signifie concrètement l'application du principe théologique de l'histoire du Salut à l'Occident moderne postchrétien ? Ce qui se produirait, après le déclin du christianisme, c'est un déplacement de la croyance en l'eschatologie, initialement attribuée au « monde à venir » vers une eschatologie de «l'avenir du monde»,[23] où la foi en l'histoire du Salut s'«immanentise» désormais dans un accomplissement terrestre de son sens, et ce sous l'emblème d'une nouvelle foi en l'Histoire mondaine. Le dogme chrétien de la fin des temps est transposé au monde humain, est sécularisé sous la forme de ce que nous pouvons appeler une « théorie profane du progrès » : celle-ci vient remplacer dès lors la fonction de Dieu en tant que Providence divine. Löwith affirme dans *Histoire et salut* que l'illusion, ou encore la *chimère*[24], qui caractérise l'Occident démocratique moderne, soit celle d'un progrès sans limites, constitue la forme profane de la foi en ce Dieu providentiel, qui représente le tout à fait Autre ; c'est-à-dire le tout-puissant, le pré-voyant (*pro-videre*) et le supra-mondain. Dans sa forme sécularisée, la « religion du progrès » conserve, quoique souterrainement et se manifestant autrement, l'esprit chrétien du Jugement dernier, en vertu duquel l'accomplissement de l'eschatologie symbolise la délivrance, pensée comme la résolution du problème du mal et de la souffrance, réminiscences du péché originel. En effet, l'idée moderne du progrès s'accompagne avec la confiance quasi-absolue en une élimination graduelle des maux humains, cette confiance étant principalement causée et sans cesse renforcée par le développement toujours plus

[22] Löwith, 2002, 87.
[23] Monod, 2002, 38.
[24] Barash, 1998, 72.

important des sciences naturelles et techniques. L'accroissement considérable des découvertes scientifiques ainsi que des connaissances empiriques qu'a connu l'Occident depuis la défaite du christianisme a eu pour résultats la consolidation du culte voué au progrès, de même que l'apparition des « religions » dites séculières, dont les discours se construisent à partir d'une « loi » du progrès, s'exprimant de façon semblable aux autres lois scientifiques, c'est-à-dire selon une logique de l'ordre du nécessaire et de l'inévitable.

Ce qui est espéré et promis par les religions séculières, c'est l'accomplissement de la communauté totalement libre et libérée[25] (on peut penser au « royaume terrestre de la liberté »[26] de Marx), et cette réalisation absolue de la liberté s'atteint à travers une émancipation de toutes les formes d'injustices humaines que reconnaissent l'esprit chrétien et son héritier sécularisé, le moderne. Löwith soulève plusieurs problèmes importants dans le schéma de la foi postchrétienne en l'eschatologie du progrès, notamment celui qui a lieu lorsque se transforme le rapport humain à l'espoir ; passant d'un espoir chrétien qui transcende le monde, à sa version moderne qui se vit à même le monde. Löwith juge que l'espérance d'une réconciliation terrestre, laquelle serait inconsciemment empruntée au christianisme, est en réalité trompeuse et marquée par la déraison, étant donné que l'homme se place dorénavant dans l'attente perpétuelle de quelque chose de meilleur. En conséquence, le temps présent est toujours perçu comme insatisfaisant, et la volonté de corriger la nature humaine escalade rapidement dans une poursuite incessante, voire vaine, d'une utopie inatteignable.

Plus encore, le danger des religions séculières est celui d'une absolutisation de l'Histoire, où « le regard se fixe hypnotiquement »[27]

[25] Perron, 2005, 9.
[26] Löwith, 2002, 278.
[27] Habermas, 1974, 156.

vers l'avenir, et produit un rétrécissement du monde humain au *télos* qui lui a été confié. Au nom du Progrès, il devient possible de tout sacrifier ; les régimes totalitaires du XX^e^ siècle sont la preuve des dommages irréparables qu'ont pu occasionner une telle interprétation de l'histoire (celle-ci se comprenant, répétons-le, à partir d'une version mondanisée du schéma de l'eschatologie biblique). Bien qu'il procède surtout par allusions, Löwith ne cache d'ailleurs pas sa méfiance par rapport aux conséquences politiques – ou plutôt aux *responsabilités* politiques – d'une telle philosophie de l'histoire : « En érigeant le temps, et donc l'histoire, au rang de source de la vérité, la philosophie s'est transformée en une instance d'action et d'intervention et, intentionnellement ou non, en ''instrument de légitimation de n'importe quelle forme de terreur politique'' ».[28] Témoin de l'Allemagne nazie des années 30 d'une part et de l'Italie fasciste de l'autre, les inquiétudes de Löwith se révèlent bien fondées en ce sens, et permettent de révéler l'architecture fragile sur laquelle sont édifiés à la fois les temps modernes en tant que tels, ainsi que la réflexion qu'ils opèrent sur eux-mêmes. Il constate d'autant plus la situation aporétique dans laquelle se trouvent les relations entre la philosophie et la politique qui marquent son époque : il s'avère, malheureusement, qu'autant la philosophie héritée de la tradition de l'historicisme (associée à Hegel et aux « théologies philosophiques » du XIX^e^ siècle) que son « calque négatif »[29] anti-historiciste (voire anti-métaphysique, à la manière de Heidegger) débouchent paradoxalement l'une et l'autre sur un décisionnisme historique « vide »[30] et dévastateur.

En ce sens, il n'est ainsi pas surprenant de sentir chez Löwith une certaine nostalgie pour le modèle grec de la conscience historique ainsi qu'une attitude sceptique par rapport à la modernité. Sans qu'il

[28] Donaggio, 2013, 146.
[29] Donaggio, 2013, 205.
[30] Donaggio, 2013, 204.

ne prenne explicitement position lui-même à un moment dans l'ouvrage, il ose parfois proposer un retour à une forme plus classique du rapport à l'histoire. Habermas, dans son essai sur Löwith intitulé « Karl Löwith ou le retrait stoïque hors de la conscience historique », indique une parenté d'esprit entre le point de vue de l'auteur d'*Histoire et salut* et celui du stoïcien : tous deux partagent la fidélité à une vision grecque du cosmos comme étant éternel, cyclique, et déjà complet en lui-même.[31] En effet, pour Löwith, l'histoire est essentiellement constituée d'évènements qui *arrivent*, et non pas d'évènements qui sont *agis* par l'homme,[32] d'où son désir de replacer l'être humain dans le mouvement naturel du cosmos, par opposition à au-dessus de celui-ci. En se faisant l'avocat subtil de la vision cosmologique grecque du monde et en osant souhaiter se débarrasser de la question du *sens* de l'histoire, Löwith n'insinue pas forcément que l'histoire n'a aucun sens du tout : l'absence de *sens* ici peut également signifier « que nous sommes détachés de la question du sens, que nous en sommes libérés, parce que nous n'attendons pas de l'histoire qu'elle puisse donner à la vie de l'homme un sens que celle-ci n'aurait pas pu avoir sans elle ou qui pourrait lui manquer ».[33] Löwith déplore ainsi le pervertissement moderne du rapport au temps présent et sa nostalgie pour le regard que portaient les Anciens sur le monde doit être comprise comme un désir de délier les errements de la conscience devenue « historicisée », c'est-à-dire de la conscience aveuglément centrée sur l'homme et confinée aux limites que posent une absolutisation de l'histoire du monde dans son rapport à un sens compris comme une *fin*.

Nous aimerions ici nous attarder sur quelques nuances, et peut-être soulever quelques lacunes dans la manière dont Löwith, et avec

[31] P. Riesterer, 1969, 91.
[32] P. Riesterer, 1969, 92.
[33] Löwith, 2002, 273.

lui d'autres penseurs de la sécularisation, ont pensé l'eschatologie mondanisée. D'un point de vue qui les situe eux-mêmes à la fin du processus achevé de la sécularisation de l'Occident chrétien, ces auteurs sont portés à assumer que dans le rapport à la pensée eschatologique, il y a une équivalence qui s'installe entre la foi chrétienne en le Jugement dernier et la foi immanentisée en la «religion du Progrès». Autrement dit – et c'est cette équivalence qui nous intéresse – suite à la sortie de la religion, c'est ce qui faisait l'*objet* de l'eschatologie qui se serait transformé sous les effets de la sécularisation du monde, mais la structure eschatologique elle-même aurait subsisté et survécu à ce processus, s'exprimant désormais de manière autre, c'est-à-dire de manière conforme à un monde dépourvu de transcendance. Ce que nous aimerions souligner ici, c'est qu'en réalité, l'adéquation qui se veut posée par Löwith entre l'eschatologie biblique et l'eschatologie séculière n'en est pas véritablement une. Par-là, nous voulons dire que lors du transfert du schéma eschatologique, le passage de son domaine propre, soit celui de la théologie, à celui du domaine mondain a en fait vidé l'eschatologie de son contenu eschatologique, de sorte qu'en bout de ligne, il ne reste plus d'« eschatologie » dans le véritable sens du terme. Il n'est donc pas entièrement exact de tenir pour acquis, comme semblerait le faire Löwith dans *Histoire et salut*, que la modernité connaît une eschatologie séculière du progrès, parce que ce n'est plus dans la structure d'une « eschatologie » qu'elle perdure, mais dans une nouvelle forme de linéarité historique. C'est que, comme nous l'avons énoncé, le terme « eschatologie » renvoie au moment de la *décision*, et nous avions plus tôt insisté sur l'importance du caractère *temporel* de l'expression. Nous voudrions maintenant insister sur le fait qu'au moment de la décision, celle-ci relève d'une intervention qui vient de l'*extérieur* (c'est-à-dire de Dieu en tant qu'altérité radicale), et que cette intervention vient mettre *fin* à une période de l'histoire divine ou

sacrée. En revanche, contrairement au Jugement dernier qui vient marquer un arrêt, le progrès implique d'une part l'idée d'un mouvement *continu*, et d'autre part, un mouvement qui provient de l'*intérieur* du monde.[34] En ce sens, l'histoire du progrès serait plutôt celle de l'autolégitimation du monde présent (pour prendre les termes de Blumenberg), alors que l'histoire du Salut, seule à être véritablement eschatologique, représente bel et bien l'histoire de la *fin* des temps.

C'est l'une des critiques que fera d'ailleurs Blumenberg à l'égard des thèses de Löwith. Contrairement à ce dernier, qui réduit la naissance du progrès à une simple sécularisation de l'eschatologie, Blumenberg explique dans *La légitimité des Temps modernes* qu'il « serait faux [de] déduire que l'idée de progrès devrait être mise sur le compte d'une provenance chrétienne »,[35] et cherchera conséquemment à défendre une vision du progrès qui n'est pas construite à partir de motifs eschatologiques chrétiens.[36] Nous avons vu que Löwith soutient, avec Hegel, l'idée du Progrès comme personnifiant la forme immanente de la Providence divine. Chez Blumenberg, au contraire, on retrouve une « plurivocité de l'expression de sécularisation »,[37] plurivocité qui le mènera à envisager une toute autre alternative pour expliquer la genèse de la modernité et de son rapport à la notion de progrès. Au lieu de proposer une sécularisation *de* l'eschatologie (c'est-à-dire par transfert, comme le fait Löwith), celui-ci considère l'idée d'une sécularisation *par* l'eschatologie.[38] La différence est importante : une sécularisation *par* l'eschatologie ne signifie pas que c'est l'objet de l'eschatologie en tant que tel qui est sécularisé, mais bien que l'eschatologie, pour parler dans les termes de J-C Monod, se

[34] Talay, 2011, 380.
[35] Blumenberg, 1999, 64.
[36] Talay, 2011, 380.
[37] Monod, 2002, 229.
[38] Monod, 2002, 228.

trouve « désamorcée ». C'est donc le fait que l'eschatologie chrétienne se soit vue obligée de s'adapter au monde, et qu'en outre elle ait été désamorcée par l'historicisation progressive de celui-ci, qu'elle s'est vue sujette à un *réel* processus de sécularisation. Il n'est donc pas le lieu ici de parler d'une « foi » immanentisée dans l'Histoire : en effet, chez Blumenberg, progrès et eschatologie ont dans les faits très peu à voir l'un avec l'autre. Ce qu'il appelle le « théorème de la sécularisation » (pour désigner la tendance du XX^e^ siècle à faire de la « sécularisation » la ritournelle d'une stratégie argumentative) occupe plutôt chez lui le statut de « métastase terminologique »,[39] et la philosophie de l'histoire elle-même lui apparaît plutôt comme une nouvelle forme de « mythe réconciliateur ».[40]

Pour Blumenberg, le progrès est principalement issu de la curiosité humaine en tant qu'elle est un facteur constant de la nature humaine, et c'est un progrès qui fonctionne selon une logique de l'essai et de l'erreur, dont la visée est celle de régler des besoins ou des désirs humains immédiats. Il affirme en ce sens que « tout événement, et ceci au sens le plus large, a un caractère de correspondance, va au-devant d'une question, d'une exigence, d'un malaise, enjambe une inconsistance, résout une tension ou occupe une place vacante »[41] au sein de la connaissance et dans l'éventail des besoins humains. Moins radical à ce propos que celui de Löwith, le point de vue de Blumenberg considère le progrès non pas comme un absolu métaphysique, mais comme un relatif accroissement des connaissances empiriques au fil du temps, accroissement porté par la recherche naturelle à l'homme de l'amélioration de ses conditions. Nous avons là, donc, deux postures possibles : celle, forte, de Löwith, qui condamne en grande partie l'illégitimité des fondements de la

[39] Blumenberg, 1999, 73.
[40] Monod, 2002, 243.
[41] Blumenberg, 1999, 435.

modernité ; et celle de Blumenberg, qui, dans un certain sens est plus encline à accepter le compromis et qui défend une version positive des temps modernes. Aux yeux de Blumenberg, en effet, les temps modernes ne demeurent décidément pas réduits simplement au résultat d'un processus de sécularisation du mode chrétien d'organisation du monde social et transcendant. Appuyant la thèse de Blumenberg, et allant encore plus loin dans ce sens, Monod explicite de son côté dans *La querelle de la sécularisation* ce que Blumenberg soutient de manière davantage sobre et nuancée dans *La légitimité des Temps modernes* : alors qu'il faut souvent dégager la thèse chez Blumenberg à partir de ses propos souvent implicites, il est indiscutable pour Monod que « la pensée moderne du progrès n'est pas réductible à une forme d'eschatologie sécularisée, pas plus que les concepts politiques modernes ne sont simplement des concepts théologiques sécularisés ».[42]

Conclusion

Il ne s'agit pas ici de donner tort à l'une ou à l'autre des genèses qui sont proposées respectivement par Löwith et Blumenberg quant au rapport qu'entretient la conscience moderne à l'histoire et à la pensée historique du progrès. D'ailleurs, nous aurions tendance à acquiescer lorsque Monod suggère qu'il y a, du moins dans une certaine mesure, un réel « dialogue de sourds »[43] entre Löwith et Blumenberg sur la question du problème de la sécularisation. Toutefois, il est clair pour Monod que « le devenir d'une religion n'est pas pensable comme le développement immanent de sa ''logique'' ou de ses ''contradictions''

[42] Monod, 2002, 282.
[43] Monod, 2002, 229.

internes ; il est tributaire d'un jeu de forces historiques »[44] et en ce sens, Monod s'éloigne des interprétations philosophiques telles que celles de Löwith ou de Gauchet, dont la tendance est de ramener les phénomènes de « la modernité » à une force motrice dont ils cherchent à révéler les voies souterraines ; force qui s'avère être, bien entendu, le christianisme. Il faut donc plutôt rapprocher la position de Monod de celle de Blumenberg, pour qui, comme nous l'avons souligné, le « théorème de la sécularisation » occupe le statut d'une ritournelle : à chaque époque sa chanson, pourrait-il dire. En ce qui nous concerne, si la posture de Blumenberg a ses attraits, elle ne discrédite pas pour autant le processus d'*enquête* mené par Löwith auprès des figures historiques qu'il met en scène dans *Histoire et salut,* de même que les constats qui ressortent de son analyse. La légitimité de l'inquiétude de Löwith par rapport à l'homme moderne, demeure en effet intacte à nos yeux, et sa sensibilité aux désarrois de la conscience moderne nous révèle son intime compréhension des enjeux qui marquent son siècle et qui guettent ceux à venir.

Pour faire un bref retour sur nos pas, nous avions posé en matière d'entrée la question suivante : après la sortie du christianisme, comment devons-nous encore penser l'histoire ? En parcourant les thèses soutenues par Löwith dans *Histoire et salut*, nous avons pu saisir l'ampleur de la problématique qu'engage cette interrogation. Le questionnement historique en lui-même s'est rapidement révélé être pour lui un legs de la tradition judéo-chrétienne, à partir de laquelle s'opère une rupture d'avec le modèle grec de la conscience historique. Pour la première fois, un bris se fit dans l'ancienne conception circulaire du temps, donnant lieu à un nouveau rapport à l'historicité : celui-ci se déployant désormais dans un temps linéaire, sous forme de la structure eschatologique biblique. Löwith défend que la pensée moderne de l'histoire demeure issue de ce schéma de l'eschatologie

[44] Monod, 2002, 286.

chrétienne, mais dans une version sécularisée, c'est-à-dire *mondanisée*, déposée dans l'immanence du monde profane. Et cette version sécularisée de la foi en l'histoire du Salut, Löwith la situe en modernité sous la forme d'une confiance quasi-absolue dans le progrès. Il explique que

> tout l'effort moderne pour parvenir sans cesse à de nouvelles améliorations et à de nouveaux progrès s'enracine dans le seul et unique progrès que connaît le chrétien, celui qui conduit à ce règne de Dieu dont la conscience moderne s'est émancipée et néanmoins demeure dépendante, comme un esclave fugitif l'est de son lointain maître.[45]

Blumenberg prend son contrepied : il présente de son côté une autre version de la genèse de la modernité, qui comme nous l'avons vu, repose sur des bases qu'il présente comme « légitimes », se tenant loin de tout « mythe réconciliateur » ou de toute analogie avec le schéma de l'eschatologie chrétienne. Par-delà les rapports respectifs qu'ils entretiennent face à la tradition philosophique qui les précède, l'intérêt que nous avons trouvé dans la confrontation des deux auteurs est surtout celui de comparer l'utilisation qu'ils font l'un et l'autre des termes d'« eschatologie » et de « sécularisation ». Nous avons vu que Löwith posait une adéquation parfaite entre la fonction de l'eschatologie biblique et celle de l'eschatologie du progrès, et que cette adéquation est en réalité contestable dès lors que l'on se colle de près à l'étymologie du mot « eschatologie », étymologie que nous appuyons ici sur les œuvres de Cullmann. Löwith aurait donc peut-être tort de parler d'une « eschatologie sécularisée du progrès » au sens littéral. Toutefois, si l'on considère la notion d'adéquation chez Löwith plutôt comme une *analogie* – comme une mise en rapport – son analyse devient fort intéressante et révélatrice des plusieurs tendances que semblent avoir pris la modernité, tendances qui rendent compte

[45] Löwith, 2002, 116.

des tensions qui habitent l'esprit moderne. Blumenberg, s'il est plus nuancé dans sa position, fait de son côté non pas une condamnation, mais bien une certaine éloge de la légitimité des temps modernes : il valorise en effet le progrès comme étant naturel, comme résultant de l'histoire de la curiosité humaine et consacre d'ailleurs plusieurs chapitres dans *La légitimité des Temps modernes* au thème de la curiosité.

Il reste que pour Löwith, l'Occident moderne est dans sa nature confusément à la fois chrétien et non-chrétien, dans la mesure où la tentative moderne d'émancipation de la religion par la sécularisation graduelle de ses structures a produit le renversement de la foi chrétienne, non pas le dépassement de celle-ci. Transposée sous la forme d'une religion du Progrès, l'eschatologie séculière postchrétienne est, pour Löwith, un emprunt illégitime de catégories qui ne lui appartiennent pas en vérité. C'est, selon lui, l'architecture illégitime de la modernité qui est à la source de la montée des formes de « religions séculières», ainsi que de l'occidentalocentrisme aveugle aux coûts humains dont fait preuve les temps modernes. La philosophie de l'histoire est donc un schème de pensée artificiel : elle représente, en quelque sorte, les œillères de l'homme de la modernité postchrétienne. Les thèses développées par Löwith dans *Histoire et salut* le mènent à souhaiter le retour à la lucidité grecque sur le monde et le temps : il n'y a, modestement, que *des* histoires, et non pas *une* Histoire. Cependant, autant qu'il cherche à se débarrasser de l'univocité et de l'absolutisation qu'entraînent la philosophie de l'histoire née de la sécularisation de l'eschatologie biblique, de fait, c'est précisément *une* histoire (et non pas *des* histoires) qui se voit retracée par Löwith dans *Histoire et salut.* Sa posture reste donc paradoxale, et sa critique de la philosophie de l'histoire souffre de ne savoir s'exprimer qu'à l'aide des catégories qu'elle redoute. C'est que Löwith, suggérons-nous, se situe au crépuscule de la pensée

systématique allemande du XIX[e] siècle et demeure en particulier très héritier de Hegel. En effet, explique Monod, « la thèse de Löwith paraît plus convaincante lorsqu'elle restreint son application aux philosophes allemands qui ont explicitement proposé une lecture de l'Histoire ''trinitaire'', comme Hegel » ;[46] et encore davantage, si Löwith semble patauger encore dans les spectres de l'idéalisme allemand, on peut supposer que c'est en raison de sa proximité historique avec les catastrophes de la Deuxième Guerre Mondiale : les répercussions directes qu'il a subies (*Histoire et salut* est publié pour la première fois en 1949, quelques années seulement après la fin de la guerre alors qu'il est toujours en exil aux États-Unis), l'inscrivent directement dans l'histoire de la pensée allemande.

[46] Monod, 2002, 222.

Références

Barash J.A., 1998, The Sense of History: On the Political Implications of Karl Löwith's Concept of Secularization, *History and Theory*, 37 (1), 69-82.

Blumenberg H., 1988, *La légitimité des Temps modernes,* Paris, Éditions Gallimard (Sagnol, Schlegel, Trierweiler, Dautrey. 1999)

Cullmann, O., 1957, *Christ et le temps : temps et histoire dans le christianisme primitif*, Neuchâtel, Delachaux & Nietslté.

-- 1966, *Le salut dans l'histoire : l'existence chrétienne selon le nouveau Testament,* Neuchâtel, Delachaux & Nietslté.

Donaggio E., 2004, *Karl Löwith et la philosophie. Une sobre inquétude,* Paris, Éditions Payot & Rivages (Audegean. 2013).

Florival G., 2001, Originaire et *eschaton*, *Laval théologique et philosophique,* 57 (3), 421-446.

Gauchet M., 1985, *Le désenchantement du monde*, Paris, Éditions Gallimard (2012).

Habermas J., 1971, *Profils philosophiques et politiques*, Paris, Éditions Gallimard (Ladmiral. 1974)

Löwith K., 1953, *Histoire et Salut : les présupposés théologiques de la philosophie de l'histoire,* Paris, Éditions Gallimard (Challiol-Gillet, Hurstel, Kervégan. 2002).

-- 1949, *Meaning in history : the theological implications of the philosophy of history,* Chicago, University of Chicago Press.

Monod J-C., 2002, *La querelle de la sécularisation : théologie politique et philosophies de l'histoire de Hegel à Blumenberg*, Paris, Librairie J. Vrin.

Perron L., 2005, Sécularisation, démocratie et eschatologie du politique, *Science et Esprit,* 57 (1), 59-69.

Riesterer B., 1969, *Karl Löwith's View of History : A Critical Appraisal of Historicism*, Netherlands, Martinus Nijhoff.

Talay Z., 2011, A dialogue with Nietzsche: Blumenberg and Löwith on history and progress, *History of European Ideas,* 37, 376-381.

Taubes J., 2009, *Eschatologie occidentale*, Paris, Éditions de l'Éclat.

Kant e a fundamentação da moral

Sérgio Mascarenhas
Instituto Superior de Gestão Bancária &
CEDIS, Faculdade de Direito, Universidade Nova de Lisboa
smascrns@gmail.com

Abstract Jean-Luis Bruch remarked that Law has a very special place in Kant's thinking. His language was to a great extent shaped by legal terminology, to the point that it displays a deep articulation between legal concepts and philosophical thinking, especially in the field of morals. We are thus led to ask to what extent Law in its theoretical and methodological dimensions underscores the construction of the Kantian system, more specifically on what concerns the domain of practical reason. The sources of Law (the modes of creation and expression of legal entities) included, in the XVIII century as today, the legislation, the court decisions, the customs and the jurisprudence. Yet, their relative importance deeply changed from then to today. In the present paper we will attempt to bring forward how these four alternative sources of Law shape the different formulations of the categorical imperative.

Resumo Anotou Jean-Luis Bruch que "o direito ocupa um lugar muito particular no pensamento do filósofo Kant", por isso a sua linguagem é marcadamente jurídica: há em Kant uma articulação profunda entre conceitos jurídicos e pensamento filosófico, mas é na moral que este "pano de fundo jurídico" mais se acentua. A questão que colocamos é a de saber em que medida o quadro teórico e metodológico do direito informa a construção do seu sistema, em particular no que respeita ao domínio da razão prática. No século XVIII os modos de formação e revelação do direito (as suas fontes) incluíam, tal como hoje, legislação, jurisprudência, costume e doutrina, porém com pesos relativos e formas de expressão marcadamente diferentes das presentes. No nosso estudo procuraremos explicitar a subjacência destes quatro momentos de juridicidade às várias formulações do imperativo categórico, pedra angular da conceção de moralidade de Kant.

1. Introdução

"O direito ocupa um lugar muito particular no pensamento do filósofo" Kant, anotou Jean-Luis Bruch[1], o mesmo Kant que ensinou direito de 1766-67 até praticamente ao fim da sua vida académica[2]. Não espanta assim que a sua linguagem seja marcadamente jurídica, mesmo quando se debruçou sobre questões aparentemente afastadas dos domínios do direito[3]: "o espírito jurídico comanda e imbui todo o sistema kantiano: desde o tribunal da razão até ao Deus-juiz, os seus esquemas de pensamento são essencialmente jurídicos"[4]. Não por acaso, o próprio Kant[5] diferenciou "os filósofos com formação jurídica" dos demais[6], posicionando-se, é evidente, entre os primeiros. Se Kant segue os modelos de problematização de questões jurídicas mesmo quando trata de domínios não jurídicos[7], é na moral como um todo que este "pano de fundo «jurídico»"[8] mais se acentua. Considere-se, a título de exemplo, a caraterização por Kant da *dedução transcendental* precisamente tomando por referência o conceito jurídico de dedução[9];

---

[1] Bruch, 1969, 18.
[2] Philonenko, 1988, 28.
[3] Gil, 1986, 35.
[4] Bruch, 1969, 18. Especifiquemos apenas que o direito é bem mais do que "uma aplicação doutrinal da *Crítica da Razão Prática*".
[5] Salvo quando se indique em contrário, citam-se as obras de Kant de acordo com a paginação do correspondente volume dos *gesammelte Schriften* editados pela Academia de Berlim. As principais obras são referidas na sequência do texto pelas seguintes abreviaturas:
CFJ *Crítica da faculdade do juízo*, 1992;
CRP*Crítica da Razão Pura*, 1985;
CRPr*Crítica da razão prática*, 1986(a);
FMC *Fundamentação da Metafísica dos Costumes*, 1986(b);
L *Lógica*, 1979b;
MC*A metafísica dos Costumes*, 2005.
[6] MC, 6:347.
[7] Gil, 1986, 27.
[8] Torralba, 2009, 393.
[9] CRP, 3:106; Guyer, 2002, 30.

a forma como modela a auto imputação do ato literalmente em termos de processo judicial[10], mesmo em resposta ao apelo emocionado de uma jovem senhora[11]. Daí, propõe Philonenko[12], poder-se defender que a expressão que traduz de forma mais literal *Urteilskraft* seja a expressão «faculdade judiciária», hoje inadequada por ter um sentido demasiado específico.

Não admira assim que o direito também tenha estado no centro das preocupações do filósofo de Königsberg e que tenha marcado a sua linguagem e as estruturas do seu pensamento, em particular no domínio da moralidade para onde Kant remete o direito como subdomínio especial. Aliás, a obra do filósofo alemão mostrou-se de leitura particularmente transparente para quem, de entre os seus contemporâneos, tinha formação jurídica, filósofo ou não filósofo. Tal poderá explicar, em boa medida, a receção que lhe deram os pensadores do direito, pois "não é por acaso que os primeiros discípulos de Kant, antes mesmo da publicação por este da sua própria *Doutrina do Direito*, manifestaram, por uma proliferação de escritos jurídicos, a fecundidade do criticismo neste domínio"[13]. Como terá assinalado o contemporâneo de Kant, Forberg, "no espaço de três anos … os kantianos importunaram o mundo com doze teorias do direito natural, nem uma a menos, e a décima terceira vai-se-lhes juntar pouco tarda"[14], doze teorias que assinalaram uma corrente ininterrupta que se alargaria por 108 obras jurídicas de inspiração kantiana nas décadas subsequentes[15], pelo menos em parte porque Kant se exprimia na mesma linguagem dos juristas, académicos ou práticos.

[10] MC, 6:438-440; ver também Torralba, 2009, 390.
[11] Kant, 1969, 11:320.
[12] Philonenko, 1979, 8.
[13] Renaut, 1986, 9 e nota 10, que aponta um outro fator a justificar esta proliferação de obras jurídicas: o impacto da revolução francesa.
[14] Philonenko, 1988, 30, nota 10.
[15] Lamego, 2005, XVI.

Temos, assim, que há em Kant uma articulação profunda entre conceitos jurídicos e pensamento filosófico. A questão que colocamos é a de saber em que medida o quadro teórico e metodológico do direito informa a construção do seu sistema, em particular no que respeita ao domínio da liberdade, da razão prática, do pensamento moral. No presente estudo avançamos com a hipótese de que essa articulação se observa nas várias formulações do imperativo categórico, pedra angular da sua conceção de moralidade. Proporemos que essas formulações são tributárias da estruturação de fontes do direito contemporânea de Kant.

2. A estruturação do direito contemporâneo de Kant

Como se exprimia o direito na época do nosso filósofo? No século XVIII os modos de formação e revelação do direito (as suas fontes) incluíam aqueles que encontramos nos sistemas jurídicos contemporâneos – legislação, jurisprudência, costume, doutrina –, porém o peso relativo e a forma como se exprimiam tinham diferenças marcadas para com o presente. Em particular, fora já ultrapassado o ponto de equilíbrio entre lei e costume no lento crescimento daquela e declínio deste, processo multisecular que levou à inversão do peso relativo dos mesmos[16]. A lei fora assumindo uma função cada vez mais importante, impulsionada pela difusão da escrita e pela centralização do poder[17], no entanto só progressivamente atingiu a depuração técnica que lhe daria a expressão quase industrial dos nossos dias. Ao

[16] Gilissen,1988, 237. Sendo certo que o costume ainda marcava o funcionamento da sociedade, quer porque persistia vivaz nas esferas mais afastadas dos centros de poder e de saber, quer porque, com a sua redução a escrito, o seu conteúdo tendia a gravitar para outras fontes do direito, Gilissen,1988, 274-282.
[17] Gilissen,1988, 302.

tempo de Kant assistia-se às primeiras experiências de codificação e de constitucionalização, ou seja, da produção das formas mais acabadas e sistemáticas de produção legislativa segundo princípios racionais. Pelo seu lado, a jurisprudência evoluía em sentidos diferentes nas várias regiões europeias. Se, por um lado, no meio anglo-saxónico ela se afirmava como uma fonte incontornável, nos países do sul da Europa era encarada com maior distanciamento[18]. Finalmente, florescia o direito natural a que hoje se chamaria teoria do direito e doutrina, impulsionado pela esperança de uma racionalidade jurídica objetiva, corrente que encontrou particular expressão onde se pode constituir como contrapeso às tendências de centralização de poder, precisamente o caso da Alemanha de Kant[19]. Na *Metafísica dos costumes* encontramos a delimitação expressa entre legislação natural e legislação positiva, entre direito natural e direito civil[20]. Quanto à primeira contraposição, diz-nos Kant que as leis vinculativas para as quais é possível uma legislação externa podem ser ou naturais quando se lhes "pode reconhecer vinculatividade mesmo sem legislação externa, *a priori*, mediante a razão", ou positivas quando "não obrigam de todo sem uma legislação externa efectiva"[21]. Acresce que "o que ordena (*imperans*) através de uma lei é o legislador (*legislator*). É autor (*autor*) da obrigatoriedade da lei, mas nem sempre autor da lei. Neste último caso, a lei seria positiva (contingente) e arbitrária"[22]. Finalmente, temos que "os Direitos, enquanto doutrinas sistemáticas, dividem-se em Direito natural, que assenta em puros princípios *a*

[18] Sobre a evolução da jurisprudência em França, cf. Hilaire, 1994.
[19] Gilissen, 1988, 364.
[20] Kant defende ainda a contraposição entre estado de natureza e estado civil, MC, 6:242 e 6:313.
[21] MC, 6:224. Kant acrescenta que se pode pensar "uma legislação exterior que contenha somente leis positivas; mas então deveria ser precedida por uma lei natural que fundamentasse a autoridade do legislador". Hoje diríamos que para o nosso filósofo o direito natural constitui o critério último de validade do direito.
[22] MC, 6:227.

priori e Direito positivo (estatutário), que dimana da vontade de um legislador"[23]. No respeitante ao costume e à jurisprudência, Kant não os considera expressamente no seu sistema moral, o que não quer dizer que estejam ausentes do seu pensamento, antes modelam-no de forma subterrânea e menos direta, como veremos em momento mais oportuno.

É este quadro jurídico que marca a conceção de moral do nosso pensador. Na *Crítica da Faculdade de Juízo* Kant enuncia o plano do seu sistema filosófico dividindo a filosofia em parte teórica e parte prática, em que a primeira tem por domínio a natureza[24], legisla mediante a faculdade de entendimento[25] e representa os seus objetos na intuição como fenómenos[26]. Já a parte prática da filosofia respeita ao conceito de liberdade[27], legisla mediante a faculdade de razão[28] e representa no seu objeto a coisa em si[29]. Uma e outra respeitam à determinação das leis ou regras a que a pessoa está sujeita nas circunstâncias concretas do seu agir[30]. Entre componente teórica e componente prática encontra-se a faculdade de juízo como *termo médio* sem domínio próprio[31] mas que constitui a pedra mestra ou pedra-angular que, para Kant, "procura reconciliar entre si os dois domínios da natureza e da liberdade, do conhecimento e da ação"[32]. A razão prática exerce-se independentemente do exercício da razão

[23] MC, 6:237 3 6:313.
[24] CFJ, 5:XII.
[25] CFJ, 5:XVII.
[26] CFJ, 5:XVIII.
[27] CFJ, 5:XI.
[28] CFJ, 5:XVII.
[29] CFJ, 5:XVIII.
[30] Deixamos por esclarecer até que ponto se pode encarar a repartição entre razão teórica e razão prática como correspondendo à divisão entre *questão de facto* e *questão de direito*.
[31] CFJ, 5:XXI.
[32] Morão, 1988, 469-470.

teórica[33] "dado que a razão prática não tem a ver com objectos para os *conhecer*, mas com a sua própria faculdade de *tornar reais* aqueles … isto é, com uma *vontade*, que é uma causalidade"[34]. A razão prática concretiza-se em *metafísica dos costumes*, sistema de leis *a priori* da moralidade[35], e esta tem por *contraponto* a antropologia moral[36] "como o outro elemento da divisão da filosofia prática em geral … doutrinas e preceitos fundados na experiência"[37].

Como faculdade intermédia, a faculdade de juízo opera mediante dois procedimentos ou dois exercícios judicativos, a *faculdade de juízo determinante* e a *faculdade de juízo reflexiva*. A primeira estabelece a relação entre a ação possível e uma lei dada. Opera por subsunção e no caso do juízo prático traduz-se na imputação do ato ao agente[38]. A faculdade de juízo reflexiva realiza a operação inversa, parte da situação dada para a determinação da lei, do particular para chegar ao universal. No juízo reflexivo a vontade, que não se refere senão à lei, não pode ser denominada de livre ou não livre, porque não se refere às ações mas directamente à legislação concernente às máximas das ações (a própria razão prática, portanto)"[39]. O exercício da razão reflexiva prática consiste assim na determinação da legislação que rege a ação[40], ou seja, "aqui não se trata do esquema de

[33] Por isso "à lei da liberdade (enquanto causalidade não sensivelmente condicionada), por conseguinte, também ao conceito do bem incondicionado, não se pode proporcionar como base nenhuma intuição, portanto, nenhum esquema, em vista da sua aplicação *in concreto*" (CRPr, 5:122).
[34] CRPr, 5:160.
[35] MC, 6:205.
[36] Ou *antropologia prática* (FMC, 4:388).
[37] MC, 6:217; Guyer, 1998, xv.
[38] MC, 6:438.
[39] MC, 6:226.
[40] Daí que "julgar se alguma coisa é ou não um objecto da razão *pura* prática é apenas a distinção entre a possibilidade ou impossibilidade de *querer* essa ação pela qual, se para ela tivéssemos o poder (acerca do que deve a experiência julgar), um certo objecto se realizaria" (CRPr, 5:100-101), sublinhado nosso. A dimensão fenoménica, empírica, da ação não é questão que se coloque ao juízo reflexivo, antes é questão que se coloca em sede de

um caso produzido segundo leis, mas do esquema (se é que o termo é aqui adequado) de uma lei em si mesma"[41].

3. As formulações do imperativo categórico

A determinação das leis morais pressupõe um critério de fundamentação das mesmas, "o sistema pressupõe a *Fundamentação da Metafísica dos Costumes*"[42]. A questão a que a *Fundamentação* procura dar resposta é a de saber qual o critério que permite identificar as proposições normativas que constituem leis da moralidade e que integram o sistema dos costumes. Uma tal lei é, na terminologia de Kant, um *imperativo categórico*, "um imperativo que, sem se basear como condição em qualquer outra intenção a atingir por um certo comportamento, ordena imediatamente este comportamento"[43], imperativo este que "não se relaciona com a matéria da ação e com o que dela deve resultar, mas com a forma e o princípio de que ela mesma deriva"[44]. A questão para o agente

juízo determinante. Os termos em que isso ocorre não cabem no âmbito da presente comunicação.

[41] CRPr, 5:121.

[42] CRPr, 5:13-14

[43] FMC, 4:416.

[44] FMC, 4:416. Kant dificulta a compreensão do seu pensamento ao utilizar em vários sentidos a expressão 'imperativo categórico', sentidos que convém distinguir para se evitarem mal entendidos:

- 'Imperativo categórico' como uma das modalidades possíveis de imperativo, por contraposição com os imperativos hipotéticos. Na sequência referiremos o conceito de imperativo categórico com maiúsculas, 'Imperativo Categórico'.
- 'Imperativo categórico' como imperativo categórico concreto. Referiremos os imperativos categóricos concretos com minúsculas, 'imperativo categórico'.
- 'Imperativo categórico' como lei ou norma universal moral. Utilizaremos a expressão 'lei moral' para referirmos o imperativo categórico neste sentido.

concreto nas circunstâncias concretas em que pretende empreender a ação é a de determinar qual o imperativo que lhe permite configurar o comportamento devido para, depois, agir em conformidade com essa configuração do comportamento devido. Temos aqui duas operações consistentes com as modalidades de juízo que referimos antes: primeiro, uma aferição do *mandamento*[45] que fornece o critério para a ação, ou seja, de apuramento de qual é a lei moral (mediante um juízo reflexivo) a que o agente deve obedecer no caso (mediante um juízo determinante). Interessa-nos em particular tomar este conceito no sentido de teste da imperatividade categórica, pois o que motiva o presente estudo é inquirir como se determinam as normas que regem a ação, como se fundamenta a normatividade destas.

Diz Kant que "as três maneiras indicadas de apresentar o princípio da moralidade são no fundo apenas outras tantas fórmulas dessa mesma lei, cada uma das quais reúne em si, por si mesma, as outras duas"[46]. Daí que as quatro formulações

> dão contributos distintos para a clarificação e especificação do imperativo categórico – enriquecem gradualmente o seu significado até que, por fim, o seu sentido completo pode ser entendido. E, uma vez completamente entendido, o imperativo categórico pode ser lido em retorno em cada uma das fórmulas de maneira a torná-las equivalentes, como exige Kant.[47]

Só a combinação das várias formulações da imperatividade categórica "apresenta as condições necessárias para mostrar que a ação de acordo com o IC é uma possibilidade real para qualquer agente racional"[48], só todas em conjunto "são suficientes para se conceber a possibilidade de um reino de agentes racionais que agem

---

- 'Imperativo categórico' como teste de determinação da lei moral. Referiremos este teste como 'teste da imperatividade categórica'.

[45] FMC, 4:416.
[46] FMC, 4:436.
[47] Pogge, 1998, 189.
[48] Guyer, 1998, 223.

em cumprimento de IC ou para a conceção de um objeto possível para o conceito de IC"[49]. Em suma, a existência de três fórmulas e quatro formulações permite a Kant enfatizar, em cada uma delas, uma componente diversa da imperatividade categórica.

De facto, o imperativo categórico "tem uma forma, expressa na exigência de universalidade. Mas também contém a especificação deste requisito relativamente ao domínio de seres (racionais) capazes de se comportar moralmente[50] e a exigência de completude dentro deste domínio"[51]. Ou seja, cada uma das três fórmulas explicita um *carater*[52] do Imperativo Categórico mas incorpora os outros dois carateres de forma implícita[53]. Temos assim[54]:

[49] Guyer, 1998, 224.

[50] Levanta-se aqui, é certo, a questão de saber quem são em concreto esses seres a quem devemos considerar como fim em si, membros da humanidade. O contacto com eles ocorre em termos empíricos, logo, como é que do fenómeno se pode chegar à identificação de quais os entes que são coisa em si? Como anota Alain Renaut, "para além da questão de saber a quem ou ao quê é que se aplica o imperativo categórico, trata-se sobretudo de determinar como se reconhece *no mundo sensível* a presença dum ser razoável, logo, da liberdade" (Fichte, 1984, 406 nota 32). "A questão que Kant não resolveu e que põe o *Direito Natural* é a de saber se podemos distinguir no mundo sensível, sem sair do mundo fenoménico (quer dizer, sem recorrer a uma solução do tipo da utilizada por Kant a propósito da terceira antinomia) o mecanismo (em direito universal) e «efeitos» de liberdade, e se uma tal distinção, supondo que é possível, é legítima sem se sair dos quadros de uma filosofia da finitude … a questão colocada parece dificilmente iludível: se não se pode distinguir entre mecanismo e liberdade *no seio mesmo dos fenómenos*, toda a visão moral do mundo e da história é desprovida de sentido" (Fichte, 1984, 406 nota 32).

[51] Pogge, 1998, 205.

[52] Sobre a noção de *carater* (L, p. 9:58 ss).

[53] O princípio supremo da doutrina da virtude da *Metafísica dos Costumes* constitui uma formulação muito próxima de explicitar de forma completa os carateres do Imperativo Categórico: "age de acordo com uma máxima dos fins tal que assumi-los possa ser para cada um uma lei universal" (MC, 6:395).

[54] FMC, 4:436. Allen Wood aborda as três fórmulas em termos diversos daqueles que aqui apresentamos mas chega a conclusões próximas das que enunciamos (Wood, s.d., 26). Considera que a segunda fórmula, precisamente

	Unidade da forma	Pluralidade da matéria	Totalidade do sistema
	Universalidade (máxima)	Humanidade (fins em si)	Legislação (intersubjetividade)
1.ª fórmula	Explícito	Implícito	Implícito
2.ª fórmula	Implícito	Explícito	Implícito
3.ª fórmula	Explícito	Implícito	Explícito

Temos que as três fórmulas do Imperativo Categórico não se distinguem entre si por constituírem outros tantos conceitos de Imperativo Categórico – só há um tal conceito – antes, cada uma delas enfatiza uma componente desse conceito[55].

A *Fundamentação* não se limita a definir o que são imperativos categóricos, ela centra-se antes na questão de saber em que termos se estabelecem imperativos, para o que propõe um conjunto de testes com base nos quais se pode verificar se uma dada proposição normativa ou *máxima*[56] constitui ou não uma lei moral.

Quais são esses testes que permitem determinar qual é a lei moral que rege a ação? Kant toma como ponto de partida o próprio conceito de Imperativo Categórico: "vamos primeiro tentar se acaso o simples conceito de imperativo categórico não fornece a sua fórmula, fórmula que contenha a proposição que só por si possa ser um imperativo categórico"[57]. A fórmula dá-nos *o que importa fazer*, o procedimento

na formulação que elegemos, constitui a mais acabada formulação do Imperativo Categórico.

[55] Significa isto que, ao contrário de José Torralba, consideramos que se podem "traduzir entre si todas as fórmulas do imperativo categórico" (Torralba, 2009, 279 nota 155).

[56] CRPr, 5:35.

[57] FMC, 4:420. Na *Crítica da Razão Prática* Kant especificou e clarificou o alcance desta etapa no desenvolvimento do seu sistema moral: "quem sabe o que para um matemático significa uma *fórmula*, que determina muito exatamente o que importa fazer para tratar uma questão e não a deixa falhar,

que permite responder à questão de saber se uma máxima se pode constituir em imperativo categórico[58]. Corrija-se, as fórmulas, pois na *Fundamentação* Kant não dá uma, dá três, de que a primeira se desdobra em duas formulações. Temos assim quatro proposições que correspondem às três fórmulas da imperatividade categórica[59]:

> **Primeira fórmula, primeira formulação**[60]: "*Age apenas segundo uma máxima tal que possas ao mesmo tempo querer que ela se torne lei universal*"[61].

---

não considerará como insignificante e dispensável uma fórmula, que faz o mesmo relativamente a todo o dever em geral" (CRPr, 5:14, nota 1).

[58] Vários intérpretes destacam o carater procedimental do Imperativo Categórico que faz dele um teste de imperatividade categórica: para Günter Elscheid ele "deve ser entendido como uma instrução sobre a forma de introduzir questões morais num certo processo intelectual" (Ellscheid, 2009, 246); Shelly Kagan considera que a *fórmula da lei universal* é um teste de máximas (Kagan, 2002, 122ss); de forma mais limitada Christine Korsgaard refere que "a Fórmula da Lei Universal é configurada como um processo de decisão" (Korsgaard, 1996, 39); Allen Wood também entende que a primeira formulação da primeira fórmula constitui um teste de normatividade (Wood, s.d., 11), porém previne contra o que há de excessivo e redutor na ênfase exagerada numa leitura da teoria moral de Kant como "um tipo qualquer de procedimento de decisão racional" (Wood, 2002, 167).

[59] Anote-se que a apresentação que aqui fazemos da diferenciação entre conceito e fórmulas e a própria identificação das fórmulas não é consensual. Por exemplo, Alexis Philonenko considera que a primeira proposição – a nossa primeira formulação da primeira fórmula – corresponde à noção de imperativo, noção essa separada das fórmulas, enquanto a nossa segunda formulação da primeira fórmula é, para este tradutor e intérprete de Kant, a expressão única da primeira fórmula (Philonenko, 1989, 114 e 115). Sem prejuízo disso, Philonenko marca bem que as fórmulas não são derivadas do Imperativo Categórico, princípios especiais decorrentes deste, mas que são apenas outras tantas expressões do mesmo. Já Paul Guyer entende que "as formulações adicionais do IC *[a segunda e a terceira, adicionais em relação à primeira]* definem condições que também são necessárias para tornar inteligível as duas maneiras diferentes por que pode ser possível a adoção do PLU *[princípio da legislação universal]* por *qualquer* agente racional"» (Guyer, 1998, 222). Gruyer não discute a diferença entre as duas formulações da primeira fórmula e desdobra a terceira fórmula em duas formulações. Apesar do mérito da sua abordagem, optámos, neste domínio, por nos mantermos fiéis à exposição de Kant.

[60] "Age de tal modo que a máxima da tua vontade possa valer sempre ao mesmo tempo como princípio de uma legislação universal" (CRPr, 5:54); "age

> **Primeira fórmula, segunda formulação**: "*Age como se a máxima da tua ação se devesse tornar, pela tua vontade, em* **lei universal da natureza**"[62].
>
> **Segunda fórmula** (princípio da humanidade e de toda a natureza racional em geral ou imperativo prático): "*Age de tal maneira que uses a tua humanidade, tanto na tua pessoa como na pessoa de qualquer outro, sempre e simultaneamente como fim e nunca simplesmente como meio*"[63].
>
> **Terceira fórmula** (princípio da autonomia da vontade)[64]: "*Age segundo máximas de um membro universalmente legislador em ordem a um reino dos fins somente possível*"[65].

Estas quatro formulações dão-nos outras tantas apresentações do conceito de Imperativo Categórico. Dão-nos igualmente quatro testes de imperatividade categórica, quatro maneiras de identificarmos *o que importa fazer*. É sobre estes que vai incidir a continuação do nosso estudo.

Antes de prosseguirmos para a análise dos testes da imperatividade categórica convém reiterar uma ideia chave que já mencionámos: estes testes não nos dão resposta à questão de se saber se a ação concreta é permitida ou proibida, dão antes resposta à questão de se saber se a máxima da ação corresponde ou não à lei

---

segundo uma máxima que possa valer simultaneamente como lei universal" (MC, 6:225).

[61] FMC, 4:421.

[62] FMC, 4:421.

[63] FMC, 4:429.

[64] Kant não apresenta uma formulação acabada na sua primeira abordagem à terceira fórmula, antes descreve esta como "a ideia *da vontade de todo o ser racional concebida como vontade legisladora universal*" (FMC, 4:431) "*por meio de todas as suas máximas*" (FMC, 4:432), "o conceito segundo o qual todo o ser racional deve considerar-se como legislador universal por todas as máximas da sua vontade para, deste ponto de vista, se julgar a si mesmo e às suas ações" (FMC, 4:433). Consideramos que ele acaba por dar uma formulação desta terceira fórmula na passagem antes citada. Anote-se também que Kant se dispensou de fornecer exemplos de aplicação deste terceiro teste, ao contrário do que fez no caso das duas primeiras fórmulas (FMC, 4:432 nota de Kant).

[65] FMC, 4:439.

moral que deve reger a ação. Se uma máxima passa o teste da imperatividade categórica, essa máxima conforma-se com a lei moral e constitui o imperativo categórico que rege a ação, a ação é permitida se for conforme com tal imperativo; se o teste falhar, a máxima não corresponde à lei moral, logo não nos dá o imperativo categórico que rege a ação, a ação não deve conformar-se com tal máxima. Neste caso é necessário refazer o procedimento a partir de outra máxima, por sua vez a sujeitar ao teste da imperatividade categórica. Este processo deverá ser repetido as vezes necessárias para se chegar à formulação do imperativo que rege a ação (sendo certo que um juízo bem orientado não procederá às cegas na formulação das máximas)[66]. De facto, os testes da imperatividade categórica consistem em juízos (reflexivos) que permitem verificar se a máxima corresponde à lei moral e constitui um imperativo categórico. Da sua realização nós não podemos retirar a conclusão de que a ação é permitida (ou proibida), apenas podemos concluir que uma dada proposição constitui ou não uma lei prática. Só assim "uma ação praticada por dever tem o seu valor moral, *não no propósito* que com ela se quer atingir, mas na máxima que a determina; não depende portanto da realidade do objecto da ação, mas somente do *principio do querer* segundo o qual a ação, abstraindo de todos os objectos da faculdade de desejar, foi praticada"[67].

Identificada a lei prática, coloca-se então a questão de saber se a ação que o agente pretende realizar é permitida ou não, mas a resposta a esta questão implica um juízo sobre a moralidade da ação, juízo este determinante, onde se opera a subsunção da ação (ou da

[66] É comum a confusão entre as duas operações. Por exemplo, afirma Christine Korsgaard: "em termos gerais, se uma máxima passa o teste do imperativo categórico a ação é permitida; se falhar, a ação é proibida e, nesse caso, o que é requerido é a ação ou omissão opostas" (Korsgaard, 1997, xxi).
[67] FMC, 4:399-400.

máxima da ação) à lei moral que rege esse juízo, tendo em vista a imputação do facto ao agente:

> a *moralidade* é pois a relação das ações com a autonomia da vontade, isto é, com a legislação universal possível por meio das suas máximas. A ação que possa concordar com a autonomia da vontade é *permitida*; a que com ela não concorde é *proibida*.[68]

As quatro formulações devem permitir testar proposições para determinar se estas são imperativos categóricos[69], a cada uma daquelas correspondendo um diferente teste da imperatividade categórica[70]. Cumpre, assim, analisarmos cada um desses testes. Na demonstração seguiremos a ordenação que nos foi dada pelo próprio Kant: a primeira fórmula corresponde ao *método rigoroso* a aplicar no juízo moral mas, para dar à lei moral «*acesso* às almas», convém proceder previamente à aplicação dos testes das demais fórmulas[71]. A nossa ordem será, pois, a seguinte: segunda fórmula; terceira fórmula; primeira fórmula, segunda formulação; primeira fórmula, primeira formulação.

---

[68] FMC, 4:439. Em consequência, "a função de um princípio fundamental nunca pode ser diretamente pôr um termo a questões morais difíceis; pode apenas servir para dar enquadramento geral adequado no qual regras morais e questões controversas devem ser colocadas e discutidas. Mesmo então qualquer formulação do mesmo deve ser vista como provisória – um objeto de constante reflexão crítica e de contínua reinterpretação e rearticulação" (Wood, 2002, 174).

[69] Que as fórmulas possam constituir testes decorre do próprio conceito de fórmula, entendido por Kant como consistindo em "regras de que a expressão serve de modelo à imitação" (L, 9:77).

[70] Do ponto de vista do caso concreto, o teste da imperatividade categórica "não é proposto como um algoritmo para decidir todas as questões morais com precisão. Ele reduz grandemente a indeterminação moral. E para além disto, a sua engenhosidade consiste em que facilita a decisão ao transformar a mesma de uma respeitante ao sujeito numa situação concreta (onde pode ser bastante difícil evitar a má fé e a desonestidade) numa respeitante ao mundo em geral. Aqui o imperativo categórico é, como deve ser, um procedimento geral para a construção de experiências mentais moralmente pertinentes" (Pogge, 1998, 206).

[71] FMC, 4:437.

4. Segunda fórmula: o imperativo categórico modelado no costume

Vimos que a segunda fórmula diz: *Age de tal maneira que tu e cada um usem sempre e simultaneamente, tanto na tua pessoa como na pessoa de qualquer outro, da humanidade como fim, conforme uma lei universal*. Para Kant os "seres racionais estão pois todos submetidos a esta *lei* que manda que cada um deles *jamais* se trate a si mesmo ou aos outros *simplesmente como meios*, mas sempre *simultaneamente como fins em si*"[72]. Para o nosso propósito as expressões chave são, «age», «usem sempre» e «se trate»[73].

Saliente-se, não há aqui invocação expressa da máxima da ação, a moralidade exprime-se no comportamento. Ora

> a aquisição de um hábito ou a sua perda consiste em estabelecer em si uma inclinação persistente sem a intervenção de qualquer máxima, através da satisfação reiterada dessa inclinação, e isso é não um princípio do modo de pensar, mas um mecanismo do modo de sentir.[74]

É no *uso*, na forma de *tratar*, *mecanismo do modo de sentir*, que se revela o imperativo sem a mediação da expressão do mesmo em proposição normativa, *princípio do modo de pensar*. De acordo com Kant, para o homem comum a moralidade exprime-se no comportamento concreto, não no entendimento abstrato[75].

---

[72] FMC, 4:434.

[73] Todas as traduções para língua inglesa que consultámos traduzem por «treat» a expressão vertida em português com «se trate» (Kant, 1997, 41; Kant, 2002, 51;Kant, 2008, 32; Kant, 2009, 35). Já quanto a «usem», não existe idêntico consenso. Mary Gregor e Allen Wood traduzem por «use» (Kant, 1997, 38; Kant, 2002, 47), enquanto Thomas Abbot e Jonathan Bennett traduzem por «treat» (Kant, 2008, 29; Kant, 2009, 32).

[74] MC, 4:479.

[75] "Se, porém, se perguntar – o que é, então, verdadeiramente a *pura* moralidade na qual, como pedra-de-toque, se deve ponderar o conteúdo moral de cada ação? – … na razão comum dos homens, ela [esta questão] está já há

O uso pode exprimir a moralidade e é necessário que o faça, tanto mais que a moralidade não se forma apenas com o exercício filosófico da razão[76]. Pelo contrário, ninguém mais do que o filósofo deve reconhecer que "diante de um homem de classe inferior, um burguês ordinário, no qual perceciono uma retidão de caráter de um grau tal que eu, no que me toca, não tenho consciência de possuir, *o meu espírito inclina-se*, quer eu queira quer não e por muito que eu levante a cabeça para que não lhe passe despercebida a superioridade da minha condição"[77]. Do uso ou forma de tratar conforme à lei, espera-se que seja constante e que não varie de ação para ação (no sentido amplo que inclui a omissão), que seja idêntico em todas as situações «sempre e simultaneamente». Por outras palavras, espera-se que seja *habitual*, tendo em atenção que o "hábito (*habitus*) é uma destreza para agir e uma perfeição subjetiva do arbítrio"[78].

Suscita-se aqui um problema:

> em toda a destreza desse tipo *[do hábito]* é um hábito livre (*habitus libertatis*); porque quando é costume (*assuetudo*) dessa liberdade, quer dizer, uma conformidade que se converteu em necessidade por repetição frequente da ação, não é um hábito que proceda da liberdade

---

muito resolvida, não certamente mediante fórmulas gerais abstractas, mas pelo uso habitual" (CRPr, 5:277).

[76] Menos ainda com esse exercício pelo filósofo Kant "como se, antes dele, o mundo estivesse totalmente na ignorância ou no erro acerca da natureza do dever" (CRPr, 5:14, nota 1).

[77] CRPr, 5:136. O caminho que levou Kant até uma tal posição não foi direto nem imediato. Como ele próprio referiu num momento de autoanálise, "por inclinação sou inquisidor. Sinto uma sede abrasadora de conhecimento, a agitação que acompanha o desejo de progredir no mesmo, e a satisfação em cada avanço nele. Houve um tempo quando cria que ele constituía a honra da humanidade e em que desprezava quem nada sabe. Nisto Rousseau corrigiu-me … Aprendi a honrar os homens e considerar-me-ia mais inútil do que um trabalhador comum se não cresse que esta minha forma de ver pode dar valor a todos os outros no estabelecimento dos direitos da humanidade" (Korsgaard, 1996, 37).

[78] MC, 6:407.

> e, portanto, não é um hábito moral. Deste modo, a virtude não pode ser definida como o hábito de praticar ações conformes à lei,[79]

daí que "as máximas morais, ao contrário das técnicas, não podem fundar-se no costume (pois que este releva da componente física da determinação da vontade), uma vez que, mesmo que a prática das máximas morais se tornasse costume, o sujeito perderia com isso a liberdade de adoptar as suas máximas, liberdade essa que carateriza a ação praticada por dever"[80]. Não é assim qualquer uso que permite revelar a existência de um imperativo categórico por detrás da ação que lhe é conforme.

A questão é pois a de se saber como e em que termos o uso, o hábito, podem ser práticas conformes à moralidade, tendo presente o "valor do caráter, que é moralmente sem qualquer comparação o mais alto, que consiste em fazer o bem, não por inclinação, mas por dever"[81]. Como passar do «mecanismo do modo de sentir» para «um princípio do modo de pensar»? Isso pode ocorrer "se se acrescentasse: 'determinar-se a agir pela representação da lei'; e, nesse caso, o hábito não é uma disposição do arbítrio, mas da vontade, a qual, com a regra que adopta, é ao mesmo tempo uma faculdade de desejar universalmente legisladora, e só um hábito semelhante pode ser considerado como virtude"[82], o que é, no fundo, possível porque "o entendimento mais vulgar pode discernir sem instrução qual a forma que, na máxima, se presta à legislação universal, e qual a que não"[83].

A moralidade pré-existe à expressão do seu conceito e das respetivas fórmulas a que só o filosofar pode dar corpo. Pré-existe porque a ação não está necessariamente dependente da razão discursiva, antes também se exprime na *razão comum*, *vulgar*, e nesta

[79] MC, 6:407.
[80] MC, 6:409.
[81] FMC, 4:398-399.
[82] MC, 6:407.
[83] CRPr, 5:49.

ela pode formar-se com base no exemplo[84]. É este o caso do homem reto cujo "exemplo apresenta-me uma lei que confunde a minha presunção quando a comparo com a minha conduta e o seu cumprimento, por conseguinte, a sua *praticalidade*, vejo-a demonstrada diante de mim através da ação", ou seja, de "a lei, tornada concreta através de um exemplo" (exemplo que, diga-se de passagem, "confunde sempre o meu orgulho")[85].

Mas a questão não fica por aqui. É que "no que se refere à força do exemplo … aquele que os outros nos dão não pode fundar nenhuma máxima de virtude. Pois que esta consiste precisamente na autonomia subjetiva da razão prática de cada homem, por conseguinte, em que não é a conduta de outros homens que nos há-de servir de móbil, mas sim a lei"[86], por isso "o bom exemplo (a conduta exemplar) não deve servir de modelo, mas tão-somente como prova de que é factível aquilo que é prescrito pelo dever"[87]. O exemplo não pode redundar na repetição mecânica. Como fundamenta ele a formação moral do homem? O ponto de partida é a imitação que "é para o homem ainda inculto a primeira determinação da vontade para aceitar máximas que subsequentemente faz suas"[88].

Vemos que aquilo que a *razão comum* faz sem reflexão, pode a filosofia adotar como um primeiro passo no método para testar

[84] Precisamente porque para Kant "o conhecimento do universal *in concreto* é conhecimento *comum*" (L, p. 9:27).
[85] CRPr, 5:136.
[86] MC, 6:480.
[87] MC, 6:480. O exemplo permite sedimentar a clareza subjetiva, da intuição, necessária para a distinção estética (L, 9:62) entendida no sentido kantiano de *sensível*. Esta distinção, caraterística da razão comum, é insuficiente e pode, com facilidade, entrar em conflito com o seu contraponto, a distinção lógica, por conceitos. Sem prejuízo disso, "é pela conjugação das duas, a distinção estética ou popular e a distinção escolástica ou lógica, que consiste a *lucidez* … o talento de apresentação luminosa, adaptada à faculdade de compreensão do *entendimento comum*, de conhecimentos abstratos e profundos" (MC, 6:246).
[88] MC, 6:479.

máximas para verificar se constituem imperativos categóricos, para que a pessoa se *oriente no pensamento*. Este primeiro passo consiste assim em "fazer do juízo segundo leis morais uma ocupação natural, de certa maneira um hábito, que acompanhe todas as nossas próprias ações livres como igualmente a observação das ações livres dos outros, e de o tornar o mais penetrante perguntando, primeiramente, se a ação é objetivamente *conforme à lei moral* e a que [lei]"[89], pelo que "o meio experimental (técnico) para educar na virtude reside no bom exemplo que o professor ele próprio possa dar"[90]. Sem prejuízo disto, o entendimento vulgar, concreto, baseado no exemplo empírico, não é suficiente para fundamentar a formação de leis morais pois pode induzir em erro, erro esse decorrente da "*influência despercebida da sensibilidade sobre o entendimento* ou, para melhor dizer, sobre o *julgamento*" da confusão "entre o que é simplesmente *subjetivo* com o que é *objetivo*" [91].

Torna-se assim patente porque é que a observação do uso constitui o primeiro teste de imperatividade categórica, teste que facilita o *acesso às almas*. Ora no direito o uso é a base do costume. Se a imperatividade categórica se revela nos usos e se apura pela observação dos mesmos, ela aproxima-se do costume jurídico que, também ele, se revela em termos semelhantes. Aparentemente Kant modela o teste da imperatividade da segunda fórmula no costume jurídico. Este teste é, porém, insuficiente pois, vimo-lo, falta-lhe rigor para fornecer o juízo definitivo da existência de um imperativo categórico[92]. São necessários outros testes para se poder extrair tal conclusão.

[89] CRPr, 5:284.
[90] MC, 6:479.
[91] L, 9:54.
[92] Não surpreende esta desconfiança relativamente ao costume da parte de Kant, ela é consonante com a evolução que esta fonte do direito sofreu a partir da Idade Média.

5. Terceira fórmula: o imperativo categórico modelado no direito civil

Acabamos de ver que podemos chegar à moralidade pela observação do comportamento, nosso e dos outros. Se dois sujeitos se observam mutuamente cada um deles pode destacar da ação concreta *todas as suas máximas*[93], esteja em causa a sua própria ação ou a ação dos demais. Uns e outros podem dar um passo adicional e, tendo por base a simples observação do comportamento próprio e alheio, tomarem consciência de uma máxima comum ao comportamento dos vários participantes na interação. Além disso, a observação do comportamento alheio coloca os agentes em relação entre si, o que lhes permite estabelecer comunicação tendo em vista um acordo relativamente à máxima das respetivas ações e ao estabelecimento da lei que as rege. Quer dizer, podem criar "uma ligação sistemática de seres racionais por meio de leis objetivas comuns"[94] e constituir um *reino* em tais termos que "somos certamente membros legisladores de um reino moral, possível mediante a liberdade, proposto ao nosso respeito pela razão prática, mas ao mesmo tempo, no entanto, somos os seus súbditos, não o seu soberano"[95]. Estamos aqui no âmbito da terceira fórmula do Imperativo Categórico, fórmula que abre a moralidade à intersubjetividade. Por esta via podemos chegar ao "conceito segundo o qual todo o ser racional deve considerar-se como legislador universal por todas as máximas da sua vontade"[96], máximas essas que assim são suscetíveis de ser agregadas numa totalidade, num sistema[97].

---

[93] FMC, 4:432.
[94] FMC, 4:433.
[95] CRPr, 5:147.
[96] FMC, 4:433.
[97] FMC, 4:436. Esta operação implica uma alteração de fundo no procedimento moral pois dispondo de um sistema de máximas tornadas «leis objetivas

De onde vem esta constituição de um sistema de leis morais no contexto de um reino composto por seres racionais em interação legisladora? O seu modelo é o processo legislativo que tinha progredido ao longo de séculos e estava em vida de Kant a atingir a maturidade com a emergência dos seus instrumentos paradigmáticos: os *estatutos* (leis formais), os códigos e as constituições do direito civil dos estados modernos. Para Kant o direito positivo é o direito legislado. A caraterização da autonomia, na *Fundamentação*, como sujeição à lei de que a pessoa é ela mesma autora[98], é consistente com o desenho que Kant dá do poder legislativo[99].

6. Primeira fórmula, segunda formulação: o imperativo categórico modelado no direito natural

Cumpre agora identificarmos a dimensão procedimental da segunda formulação da primeira fórmula do Imperativo Categórico que, recordemos, nos diz:

> *Age como se a máxima da tua ação se devesse tornar, pela tua vontade, em lei universal da natureza.*

Estamos neste caso a reportar-nos aos imperativos enquanto constituintes de "uma legislação das ações universalmente semelhante a uma *ordem natural*"[100], legislação que se forma por um ato de

comuns», o homem pode-se "julgar a si mesmo e às suas ações" por referência a esse sistema (FMC, 4:433).

[98] FMC, 4:431ss. "A vontade não está pois simplesmente submetida à lei, mas sim submetida de tal maneira que tem de ser considerada também como *legisladora ela mesma*, e exatamente por isso e só então submetida à lei (de que ela se pode olhar como autora)" (FMC, 4:431).

[99] MC § 46, 6:313-315.

[100] FMC, 4:431. Não está aqui em causa a natureza empírica, regida pelas leis do entendimento. Antes, trata-se agora da "ideia de uma natureza não

vontade, daí que "as máximas têm de ser escolhidas como se devessem valer como leis universais da natureza"[101]. Porém, é "graças à razão [que], somos conscientes de uma lei à qual, como se uma ordem natural houvesse de brotar da nossa vontade, estão sujeitas todas as nossas máximas"[102]. A essa lei "atribuímos, pelo menos no aspeto prático, uma realidade objetiva, porque a consideramos como objeto da nossa vontade, enquanto seres racionais puros"[103]. Como seres dotados de razão e pelo exercício dessa razão podemos configurar uma legislação suscetível de valer como ordem natural. Podemos fazê-lo porque "o entendimento mais vulgar [de um ser dotado de razão] pode discernir sem instrução qual a forma que, na máxima, se presta à legislação universal, e qual a que não"[104].

Fácil se torna observar que o referencial que subjaz a esta formulação no quadro das conceções jurídicas da época de Kant é o da noção de *direito natural*, para que apela a ideia de *lei universal da natureza*, noção que, no quadro das fontes do direito, se prolongou até hoje no conceito de doutrina ou "prática filosófica da ética académica"[105]. É este modelo de construção do direito com base num exercício científico ou sapiente, enfim, *racionalista*[106], que constitui o

empiricamente dada e, no entanto, possível através da liberdade; consequentemente, de uma natureza suprasensível" (CRPr, 5:76).

[101] FMC, 4:436.

[102] CRPr, 5:76.

[103] CRPr, 5:76.

[104] CRPr, 5:49.

[105] Brito, 2011, 38.

[106] "Ao uso dos conceitos morais é unicamente adequado o *racionalismo* da faculdade de julgar, o qual nada mais tira da natureza sensível do que o que a razão pura pode também pensar por si, isto é, a conformidade à lei, e nada transporta para a natureza suprasensível a não ser o que, inversamente, se pode realmente representar no mundo sensível por ações segundo a regra formal de uma lei natural em geral" (CRPr, 5:125).

terceiro procedimento de referência para a formação de imperativos categóricos[107].

7. Primeira fórmula, primeira formulação: o imperativo categórico modelado no precedente

Resta-nos a primeira formulação da primeira fórmula:

> *Age apenas segundo uma máxima tal que possas ao mesmo tempo querer que ela se torne lei universal.*

De acordo com esta formulação, o teste do imperativo categórico implica as seguintes etapas: configuração das circunstâncias da ação; configuração da máxima; configuração do imperativo. Kant precisa melhor o conjunto de operações que permitem a formulação do imperativo na *Crítica da Razão Prática* e na *Metafísica dos Costumes*. Na última especifica que "deves considerar as tuas ações primeiramente segundo o seu princípio subjetivo; mas podes reconhecer se esse princípio pode ter também validade objetiva apenas no seguinte: em que, submetido pela tua razão à prova de te pensares por seu intermédio como universalmente legislador, se qualifique para uma tal legislação universal"[108]. A chave está no procedimento de universalização, o raciocínio que permite transformar a configuração do caso individual numa máxima, proposição subjetiva e genérica, e depois transformar a máxima numa proposição objetiva

[107] Muito crítico de Kant, Michel Villey considera que este "é, numa larga medida, o continuador da escola do direito natural de que defende, reforça e leva à perfeição suprema as principais conclusões ao fundá-las numa filosofia nova" (Villey, 2002, 251-252), sem prejuízo de que, nesse processo, Kant cria uma cisão entre direito natural e direito racional, sendo igualmente certo que o filósofo alemão não teria retirado as consequências da mesma (Villey, 2002, 66-67).
[108] MC, 6:225.

e universal pela execução de um juízo reflexivo. Para Kant este procedimento intelectual respeita à forma da proposição pois "todas as máximas têm, com efeito: 1) uma *forma*, que consiste na universalidade"[109].

Ainda aqui a inspiração jurídica é evidente. O processo intelectual a que Kant faz apelo não é mais do que o raciocínio de construção do precedente judicial, a criação jurisprudencial de normas a partir do caso concreto submetido ao juiz para decisão quando falta uma regra adequada[110]. Avancemos um século, século marcado pela emergência de um paradigma do direito formalista e normativista contra o qual se levanta um dos arautos do *movimento do direito livre*, Herman Kantorowicz, em *A luta pela ciência do direito* de 1906[111]. Neste manifesto defende a adoção, como princípio norteador da criação livre do direito, duma disposição do primeiro artigo do projeto de código civil suíço, então em discussão e que viria a ser aprovado e publicado em 1910[112]. A referida disposição reza o seguinte: «na falta de uma disposição legal aplicável, o juiz decide segundo as regras que ele

[109] FMC, 4:436. Anote-se que Kant não considera que sejam exigidas quaisquer particulares capacidades intelectuais, congénitas ou adquiridas, para a realização desta operação, dado que "o entendimento mais vulgar pode discernir sem instrução qual a forma que, na máxima, se presta à legislação universal, e qual a que não" (CRPr, 5:49). A razão não é privilégio do filósofo.

[110] Reiteremos que esta criação de normas morais modelada na jurisprudência, exercício da faculdade de juízo reflexiva que permite a formulação de uma proposição universal partindo da máxima do caso, não se confunde com a aplicação do imperativo categórico, também ele um exercício da faculdade de juízo, neste caso determinante, que na sua dimensão ética se exprime como *consciência moral*, consciência esta que opera de acordo com um modelo construído sobre o exercício da judicatura (MC, Segunda Parte. § 13., 6:437-440). Para contraste veja-se esta contraposição, na perspetiva do juízo judicial sem referência ao pensamento de Kant em Hilaire, 1994, 181-182.

[111] Kantorowicz, 2011.

[112] O projeto foi da autoria de Eugen Huber (http://hls-dhs-dss.ch/textes/f/F4533.php), teórico do direito, neo-kantiano, e amigo de Rudolf Stammler, igualmente neo-kantiano.

estabeleceria se tivesse de atuar como legislador»[113], nem mais. Avancemos ainda mais algumas décadas e recordemos o art. 10º 3. do Código Civil português de 1966:

> Na falta de caso análogo, a situação é resolvida segundo a norma que o próprio intérprete criaria, se houvesse de legislar dentro do espírito do sistema.

Retenhamos o essencial: *segundo as regras que o juiz estabeleceria se tivesse de atuar como legislador*, no caso suíço; *a norma que o próprio intérprete criaria, se houvesse de legislar*, no caso português; na sua essência, estas duas disposições não dizem mais do que aquilo que nos diz Kant na primeira formulação da primeira fórmula do Imperativo Categórico[114]. Além disso, o direito positivo destes dois países remete o recurso ao procedimento universalizante para uma função supletiva, de último recurso. Mas não nos disse Kant para chegarmos ao teste da universalidade depois de passarmos pelo teste do uso e pelo teste da legislação?

8. Conclusão: a moralidade como *facto* da razão

Decorre do exposto que a definição do imperativo categórico pode implicar o recurso aos diversos testes de imperatividade categórica propostos por Kant[115]:

[113] http://www.admin.ch/ch/f/rs/210/a1.html. Deixámos de parte a referência inicial ao costume, compreensível no contexto da integração e harmonização dos diversos ordenamentos jurídicos da Confederação Helvética e que, aliás, também Kantorowicz omitiu na sua referência.

[114] E não se colocava já a hipótese deste juízo reflexivo de criação de normas no Código Civil francês de 1804? De facto, ao proibir aos juízes a criação de normas para a disposição dos casos que lhes eram submetidos (art. 5º), este diploma requeria já o Imperativo Categórico, nem que fosse para proibir o recurso ao mesmo (Hilaire, 1994, 181).

[115] Afloram aqui, bem vistas as coisas, as três condições universais que Kant identifica para evitar o erro: "1) pensar por si mesmo [teste da autonomia], 2)

- Nos termos da segunda fórmula, ela ocorre mediante um juízo de universalização da máxima da ação identificável no hábito, ou seja, por referência a outras ações, próprias ou alheias;

- Nos termos da terceira fórmula, ela resulta de um encontro de vontades dirigidas à universalização da regra geral intersubjetivamente estabelecida;

- Nos termos da primeira fórmula, ela tem lugar através de um juízo racional de universalização da máxima da ação racionalmente elaborada pelo agente, tendo por referência a natureza.

Estas vias são coerentes. Os dois primeiros testes da imperatividade categórica não são incompatíveis nem necessariamente independentes um do outro. Aplica-se aos imperativos categóricos em geral e à lei moral o que Kant diz do direito positivo: "as formas do Estado não são senão a letra (*litera*) da legislação ordinária do estado civil e podem, portanto, subsistir enquanto forem consideradas necessárias ao mecanismo da Constituição política por um costume antigo e longo (portanto, sob um ponto de vista meramente subjetivo)"[116]. O agente pode verificar a existência do uso no seu comportamento e no comportamento dos outros; o encontro de vontades pode, a partir daí, ficar implícito no uso coerente dos vários agentes ou ser explicitado em acordo e legislação. Entre as duas alternativas, várias são as combinações possíveis, daí que é possível por *modificações paulatinas* chegar-se à *Constituição de uma República pura*[117] ou, pelo menos, progredir nesse sentido (nos termos da *Fundamentação*, avançar no sentido da concretização do *reino dos fins*). Por outro lado, há igualmente uma ponte entre o teste da

pensar pondo-se no lugar do outro [teste do uso], 3) pensar sendo sempre consequente consigo mesmo [teste da universalidade]" (L, 9:57).
[116] MC, 6:340.
[117] MC, 6:340.

terceira fórmula e o teste do direito natural dado que "pode, pois, pensar-se uma legislação exterior que contenha somente leis positivas; mas então deveria ser precedida por uma lei natural que fundamentasse a autoridade do legislador"[118].

As várias formas de se determinar a lei moral, os vários testes de imperatividade categórica, surgem-nos assim como generalizações das grandes fontes de criação do direito que nos permitem reconhecer as normas que regem a ação, mas não mais, não nos dão, só por si, um critério último de moralidade. A *fundamentação da metafísica dos costumes* dá-nos a possibilidade da lei moral e do imperativo categórico mas não demonstra a sua necessidade[119]: "é-nos totalmente impossível a nós homens explicar como e porquê nos interessa a *universalidade da máxima como lei*, e, portanto, a moralidade"[120]. É certo que, procura Kant demonstrá-lo na terceira parte da *Fundamentação*, "se pode indicar o único pressuposto de que depende" a possibilidade do imperativo categórico, pressuposto esse que é a ideia de liberdade[121]. A *Crítica da Razão Prática* procura demonstrar a liberdade transcendental e demonstrar a possibilidade e necessidade do exercício prático da razão. É certo,

> a regra [prática] diz que se deve simplesmente proceder de um certo modo. A regra prática é, pois, incondicionada, por conseguinte, apresentada *a priori* como uma proposição categóricamente prática, mediante a qual a vontade é de modo absoluto e imediato objetivamente determinada (pela própria regra prática que aqui constitui, pois, uma lei),[122]

mas isto é uma possibilidade de que está por demonstrar a necessidade. O exercício prático *puro* da razão é colocado apenas de forma problemática, insuficiente para fundamentar a moralidade.

[118] MC, 6:224.
[119] FMC, 4:444-445.
[120] FMC, 4:460.
[121] FMC, 4:461.
[122] CRPr, 5:55.

A resposta de Kant a esta *coisa assaz estranha* que *não tem igual em todo o restante conhecimento prático* onde "o pensamento *a priori* de uma legislação universal possível, pensamento que, por conseguinte, é simplesmente problemático, é ordenado incondicionalmente como lei, sem nada tirar da experiência ou de qualquer vontade exterior"[123] é a afirmação do *facto da razão* que "não é um facto empírico mas o facto único da razão pura, que assim se proclama como originariamente legisladora *(sic volo, sic iubeo)*"[124], facto "indissoluvelmente ligado à consciência da liberdade da vontade, que até mesmo se confunde com ela"[125]. O facto da razão (prática), proporcionado pela lei moral, é "um facto absolutamente inexplicável a partir de todos os dados do mundo sensível e do âmbito global do nosso uso teorético da razão, facto esse que anuncia um puro mundo inteligível, o *determina* até *positivamente* e dele nos permite conhecer alguma coisa, a saber, uma lei"[126]. Por via dele "a lei moral transporta-nos, em ideia, para uma natureza em que a razão pura, se fosse provida de um poder físico a ela adequado, produziria o soberano bem, e determina a nossa vontade a conferir a sua forma ao mundo sensível enquanto conjunto dos seres racionais"[127]. Do facto da razão somos conscientes *a priori* donde ele ser *apoditicamente certo*[128], mesmo se "a realidade objetiva da lei moral não pode ser demonstrada por nenhuma dedução"[129].

[123] CRPr, 5:55.
[124] CRPr, 5:56
[125] CRPr, 5:72. *Facto*, "pois assim se pode chamar uma determinação da vontade, que é inevitável, embora não se baseie em princípios empíricos" (CRPr, 5:96).
[126] CRPr, 5:74.
[127] CRPr, 5:75.
[128] CRPr, 5:81.
[129] CRPr, 5:81.

Retenhamos o essencial: em Kant a fundamentação última da moralidade[130] é *questão de facto*, mesmo se o facto em causa não é equacionado com um facto empírico, fenoménico, natural, nem seja redutível a uma simplista contraposição entre *ser* e *dever ser*. Este *facto* desvela-se num processo que decorre em vários estágios: revelação no comportamento concreto dos agentes morais; acordo explícito ou implícito entre os agentes morais sobre o sentido da moralidade; reflexão sobre o domínio moral modelada sobre os princípios a que obedece o domínio da natureza; exercício pessoal da razão. Programa moral este que obriga a uma articulação total de natureza e liberdade, articulação que se manifesta, nomeadamente, na conformação à luz das mesmas estruturas lógico-formais.

[130] Reiteremos, "largamente dependente dum modelo jurídico, por muito cuidado que ele [Kant] ponha na distinção entre a avaliação moral e jurídica" (Longuenesse, 2005, 261).

Referências

Brito, J. de S., 2011, Falsas e verdadeiras alternativas na teoria da justiça. http://www.fd.unl.pt/docentes_docs/ma/jsb_MA_9652.doc, último acesso 2014-12-15.

Bruch, J.-L., 1969, Introduction. Kant et sa Correspondence. *In* Kant, 1969, 5-19.

Ellscheid, G., 2009, O problema do direito natural. Uma orientação sistemática. *In* Kaufmann, 2009, 211-280.

Fichte, J. G., 1984, *Fondement du droit naturel selon les principes de la doctrine de la science*, PUF, Paris.

Gil, F., 1986, *Provas*. Lisboa: Imprensa Nacional – Casa da Moeda.

Gilissen, J., 1988, *Introdução histórica ao direito*. Lisboa: Fundação Calouste Gulbenkian.

Guyer, P. (ed.), 1998, *Kant's Groundwork of the Metaphysics of Morals. Critical Essays*. Lanham : Rowmann & Littlefield Publishers, Inc..

-- 1998, «Introduction». *In* Guyer, 1998, xi-xlv.

-- 1998, «The Possibility of the Categorial Imperative». *In* Guyer, 1998, 215-246.

-- 2002, "Kant's Deductions of the Principles of Right". In Timmons, M. (ed.), 2002, Kant's Metaphysics of Morals. Interpretative Essays. Oxford, Oxford University Press, 23-64.

Hilaire, J., 1994, «Jugement et jurisprudence». *Archives de Philosophie du Droit* 39, 181-190.

Huber, E. s.d. *Dictionnaire historique de la Suisse*. http://hls-dhs-dss.ch/textes/f/F4533.php, ultimo acesso 2015-07-14.

Kagan, S., 2002, «Kantianism for Consequencialists». *In* Kant, 2002, 111-156.

Kant, E., 1969, *Lettres sur la morale et la religion*. Paris: Aubier.

-- 1979a, *Critique de la faculté de juger*. Paris : Vrin.

-- 1979b, *Logique*. 2ª ed. Paris : Vrin.

-- 1985, *Crítica da razão pura*. Lisboa: Fundação Calouste Gulbenkian.

-- 1986a, *Crítica da razão prática*. Lisboa: Edições 70.

-- 1986b, *Fundamentação da Metafísica dos Costumes*. Lisboa: Edições 70.

-- 1992, *Crítica da faculdade do juízo*. Lisboa: Imprensa Nacional – Casa da Moeda.

-- 1997, *Groundwork for the Metaphysic of Morals* (Tradução de Mary Gregor). Cambridge: Cambridge University Press (reimpressão de 2003).

-- 2002, *Groundwork for the Metaphysic of Morals* (Tradução de Allen Wood). New Haven: Yale University Press.
-- 2005, *A metafísica dos Costumes*. Lisboa: Fundação Calouste Gulbenkian.
-- 2008, *Groundwork for the Metaphysics of Morals* (Tradução de Jonathan Bennett). http://earlymoderntexts.com/kgw.html, ultimo acesso 2011-03-25.
-- 2009, *Groundwork for the Metaphysics of Morals* (Tradução de Thomas Kingsmill Abbott). ebooks@Adelaide. http://ebooks.adelaide.edu.au/k/kant/immanuel/k16prm/complete.html, ultimo acesso 2011-10-06.

Kantorowicz, H., 2011, «The Battle for Legal Science». *German Law Journal*, vol. 12, nº 11, 2005-2030, http://www.germanlawjournal.com/index.php?pageID=11&artID=1395, ultimo acesso, 2014-12-15.

Kaufmann, A., Hassemer, W. (orgs.), 2009, *Introdução à filosofia do direito e à teoria do direito contemporâneas*. Lisboa: Fundação Calouste Gulbenkian.

Korsgaard, C., 1996, *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press (reimpressão de 2000).
-- 1997, «Introduction». *In* Kant, 1997, vii-xxx.

Lamego, J., 2005, «A Metafísica dos Costumes : a apresentação sistemática da filosofia prática de Kant». *In* Kant, 2005, IX-XXXIV.

Longuenesse, B., 2005, *Kant and the Human Standpoint*, Cambridge, Cambridge University Press (reimpressão 2009).

Morão, A., 1988, «Crítica da Razão Prática. 1788-1988». *Revista Portuguesa de Filosofia*. Tomo XLIV, 465-474.

Philonenko, A., 1979, «Introduction». *In* Kant, 1979(a), 7-42.
-- 1988, *L'œuvre de Kant*, tomo 1. Paris: Vrin.
-- 1989, *L'œuvre de Kant*, tomo 2. Paris: Vrin.

Pogge, T., 1998, «The Categorial Imperative». *In* Guyer, 1998, 189-213.

Renaut, A., 1986, *Le système du droit. Philosophie et droit dans la pensée de Fichte*. Paris: PUF.

Torralba, J., 2009, *Liberdad, objeto práctico y acción. La faculdad del juicio en la filosofía moral de Kant*. Hildsheim : Georg Olms Verlag.

Villey, M., 2002, *Leçons d'histoire de la philosophie du droit*. 2ª ed. Paris : Daloz.

Wood, A., 2002, «What is Kantian Ethics?». *In* Kant, 2002, 157-181.
-- s.d., «The Supreme Principle of Morality». http://web.stanford.edu/~allenw/webpapers/SupremePrincipleMorality.pdf, ultimo acesso 2014-12-1.

Dossier

Imagética atomística

Organizado por Lídia Queiroz (CFCUL)

Is there such a thing as 'Greek Atomism'?

Zbigniew Kotowicz
(Centro de Filosofia das Ciências da Universidade de Lisboa, CFCUL)
z.kotowicz@yahoo.fr

Abstract In this article I look at the term 'Greek atomism' which is often used in a misleading way. 'Atomism' is a modern term, it was coined in the seventeenth century; the Greeks did not have a corresponding one, they referred to the philosophies of Democritus (and Leucippus) and Epicurus as very distinct doctrines, rather than the two representing the same philosophical system. And they had good reasons for doing so as there are considerable differences between Democritus and Epicurus. The presentation of some of these differences as well as the atomist doctrine introduced in the seventeenth century by Pierre Gassendi will form the core of the article. It will lead to the conclusion that the term 'atomism' often means Gassendi's version and it conjures up an image of billiard balls bouncing of each other which is inadequate to deal with the Epicurean doctrine and misleading when approaching Democritus.

Resumo Neste artigo, considero a expressão 'Atomismo Grego', frequentemente utilizada de modo ambíguo. 'Atomismo' é um termo moderno, cunhado no século XVII. Os Gregos não tinham um termo correspondente, referiam-se às filosofias de Demócrito (e Leucipo) e de Epicuro como doutrinas marcadamente distintas, mais do que enquanto representantes do mesmo sistema filosófico. E tinham boas razões para o fazer, pois existem diferenças consideráveis entre Demócrito e Epicuro. A apresentação de algumas dessas diferenças, bem como a doutrina atomista introduzida por Pierre Gassendi no século XVII, constituem o núcleo do artigo. Sustentam a conclusão segundo a qual o termo 'atomismo' significa frequentemente a versão de Gassendi e evoca a imagem de bolas de bilhar colidindo umas com as outras, uma imagem inadequada para lidar com a doutrina de Epicuro e enganadora no que diz respeito a Demócrito.

Although a number of studies show that atomism never disappeared, it is nevertheless true that since antiquity it occupied a marginal (to say the least) place in philosophical preoccupations and only regained some prominence in the 17th century mostly due to the work of Pierre Gassendi.

In his time, Gassendi had the reputation of a scientist of the highest rank. However, viewed from the historical perspective his contribution to science was unremarkable; he made no significant scientific discoveries, there is no Gassendi law, for example. However, his great merit was the introduction of atomism into scientific thought; it was an ontological hypothesis that the science of his time needed.[1] Gassendi's reconstruction of the ancient system (and what some of this entailed, I will return to later) has been accepted as what goes under the term 'atomism' and sometimes is referred to, wrongly, as 'Greek atomism'.

When one goes back to the early days of the doctrine, namely to the thought of Democritus, and begins to examine the matter closely, one quickly discovers that we are taught the atomist doctrine poorly with the result being that when we hear the word 'atomism' the image of billiard balls bumping off each other immediately springs to mind like a knee-jerk effect. This image is wrong. The thought of the Greeks was far more subtle and complex than the mechanistic world commonly associated with it.

The hopelessly meagre fragments of Democritus that have come down to our times set a limit to what can be said with confidence about his thought.[2] Nevertheless, it is clear that Democritus' system was coherent and consistent. This was at least Aristotle's view;

[1] Koyré, 1973, 321.

[2] Strictly speaking one should speak of Democritus and his predecessor Leucippus. But so little is known about the earlier man that apart from laying down the basics of the doctrine nothing of any detailed thought can be attributed to him.

although he makes his disagreements plain, he shows immense respect for the man, often singling him out as the most astute. Aristotle devoted a monograph to him, which has not, however, survived to our times either. Plato's reaction to Democritus also deserves attention. According to Diogenes Laertius a story circulated that he wanted to burn Democritus' writings.[3] Diogenes goes on to say that Plato was only dissuaded from this when it was pointed out to him that the circulation of the writings was so wide that he could not succeed in destroying them all. One can believe the story or not. On the one hand Diogenes takes as his source a certain Aristoxenus who was apparently notorious for spreading libellous gossip, and therefore not trustworthy; on the other hand, the sentiments that Plato expresses in the *Laws* make it plain that burning books would have been well within his range. (And since this was the fate of Protagoras' books[4] the Athenians would have tolerated such a measure.) What is, nevertheless, incontestable is that in all his writings Plato not even once mentions Democritus' name; and it is this that makes Diogenes think the story might be true.

What was it that Plato found so unacceptable? Clearly, a doctrine that lacked any sign of teleology and divine design must have been anathema to him. But if Democritus had argued it through a primitive mechanistic atomism Plato would have had no difficulty in destroying the arguments. We know from his treatment of the Sophists that when dealing with opponents he did not shirk from conflict and was capable of underhand tactics; why not do the same to Democritus? Diogenes thought that Plato feared having 'to match himself against the prince of philosophers'.[5] Maybe; but it still does not tell us anything concrete about Democritus' views (and Diogenes' own account is too brief and

[3] Diogenes Laertius, 1980, X 40.
[4] Diogenes Laertius, 1980, IX 52.
[5] Diogenes Laertius, 1980, IX 40.

sketchy to be of much help). Nevertheless, everything suggests that he was a philosopher of immense depth.

In post-Ancient times, atomism practically disappeared, in the intellectual climate dominated by the teachings of St Augustine there was no place for such a godless theory. When it was re-introduced by Pierre Gassendi into scientific thought in the 17th century it was based in principal on the later Epicurean version which in a number of respects differs from the system of Democritus. Some differences between Democritus and Epicurus are apparent quite immediately. The first, relevant in this context, concerns the conception of the atom. While Democritus did not accord the atoms the property of weight (at least that is the view of the majority of scholars), Epicurus did assign weight to atoms. As a consequence the natural movement of the atom is a downward vertical fall. However, no life would emerge if the atoms just fell in a straight line

> One further point in this matter I desire you to understand: that while the first bodies are being carried downwards by their own weight in a straight line through the void, at times quite uncertain and uncertain places, they swerve a little from their course, just so much as you might call a change of motion. For if they were not apt to incline, all would fall downwards like raindrops through the profound void, no collision would take place and no blow would be caused amongst the first beginnings: thus nature would never have produced anything.[6]

This 'swerve' is known by its Latin rendering the *clinamen* (from *clinare* – to incline). This swerve makes atoms collide, which, in turn, form vortexes from which various forms of life emerge. The *clinamen* was also the principle of indeterminacy in Epicurus' system,

> it break[s] the decrees of fate, that cause may not follow cause from infinity, whence comes this free will in living creatures all over the earth, whence I say is this will wrested from the fates by which we proceed

[6] Lucretius, 1982, 2.216-2.224.

> whither pleasure leads each, swerving also our motions not at fixed times and fixed places, but just where our mind has taken us?[7]

Gassendi 'cleaned up' the doctrine to make it acceptable to the Church (of which he was himself a loyal servant). He removed the *clinamen* and made the atoms move in all directions in conformity with Democritus' thought. He also removed the notion of multiple worlds which Epicurus argued and which had also been earlier advanced by Democritus.

The great merit of Gassendi was to argue for the legitimacy of the concept of the void, which he perceived as empty space in which atoms move. This new version of atomism was adopted by Newton, it was the kind of hypothesis that science of the time needed and it was on the whole accepted by the scientific community, although there was also some opposition (Ernst Mach and Wilhelm Ostwald, the most notable examples). At the beginning of the 20th century, the French experimentalist Jean Perrin confirmed the atomist hypothesis and it was practically universally accepted.

However, the atomism that re-emerged in the 17th century was a lifeless mechanics that could only be given some sense by positing an omniscient God who arranges that matter conform to a divine scheme, manifest in the immutable laws of nature. In time God was deemed unnecessary, 'Sire, I had no need for that hypothesis' declared the astronomer Pierre-Simon de Laplace, the laws of nature were so precise, he claimed, that in principle any past or future event could be deduced from these laws if it were possible to take all the factors into account. This was not the vision of either Epicurus or Democritus.

To see the problem more clearly let us begin with this sentence that opens Werner Heisenberg's reflections on the ancient doctrine:

[7] Lucretius, 1982, 2.254-2.260.

> The concept of the atom [...] has its origin in ancient Greek philosophy and it was in that early period the central concept of materialism as taught by Leucippus and Democritus.[8]

This statement may seem to be a fair reflection of the thought of the ancient sages but, as it happens, not altogether; in fact, it is somewhat misleading. This is not the way the Ancients presented the doctrine, the most common opening formula that they used was almost always the same: Aristotle stated that 'Leucippus and his associate Democritus say that the full and the empty are the elements'[9], Aëtius states: 'Leucippus of Miletus says that the filled and the void are principles and elements', Diogenes Laertius: 'The All includes the empty as well as the full' and finally Cicero repeats the same 'Leucippus admits two principles: the filled and the void'.

This way of presenting the doctrine suggests that these early Greeks did not begin with atoms but with a dichotomy, an oscillation between what is and what is not, that is, they begin with a dialectics of presence and absence. But there is more to it than just a dichotomy, in Democritean terms the atom and the void are presented in three different ways, all three stating them as opposites, aside the 'the full and the empty' formula, the two elements are presented as 'the existent and the non-existent' and 'the thing and the nothing'. At least two commentators, David Sedley and Frédéric Nef, have arrived at a conclusion, independently, it seems, that Democritus might have been conceiving of a negation of substance, 'the void is the privation or negation of the full'[10], the void is 'not empty space but the negative substance which occupies empty space'.[11]

One cannot go any further than these conjectures, but one thing is clear, the void of Democritus is complex and it certainly cannot be

[8] Heisenberg, 1990, 47.
[9] *Metaphysics*, Bk I, Ch. 4, 985b.
[10] Nef, 2011, 113.
[11] Sedley, 1982, 179.

reduced to Newtonian empty space. And there is a further point to be made about presenting the Ancient doctrine as having the atom as its central concept, as Heisenberg does – as a consequence the void is pushed into the background. Yet, it was not the idea of the atom that was the most original ingredient of the theory; it was the concept of the void that distinguished it most. This was a contentious proposition from the beginning. It was attacked by the Eleatic philosophers; Aristotle pronounced himself against it. The fact that the Greeks did not have the concept of zero also must have affected their thinking. (What would the Pythagorean system look like if it had to incorporate zero into it?). The mediaeval thought rejected the void with such meaningless utterances as 'nature abhors the vacuum'. Western philosophy's agenda was dominated by the intricate scheme of the Great Chain of Being, which, following the 'principle of plenitude, (that is, that God does not allow any potentiality for being to remain unfulfilled) expressly forbade the void.[12] Interestingly, many of the scientists and philosophers who were inclined to adopt the atomist theory went along with it only as far as the concept of the atom was concerned but did not accept the void. It was argued that the space between the atoms had to be filled with some very subtle substance. Descartes spoke of a 'subtle matter', which was most often referred to as ether, which would function as an omnipresent medium. Newton, too, throughout his life thought, on and off, of incorporating the idea of ether into his thinking. The final scheme of Newtonian mechanics is based on solid bodies moving in empty space; void is no more than a passive container.

The fact that the void (empty, nothing, non-existent) was consistently seen as the opposite of substance (full, thing, existent) already takes us away from the 'space the container' image of the

[12] On this see *The Great Chain of Being* by Arthur Lovejoy (1964). It is, in fact, Lovejoy who coined the very helpful expression 'the principle of plenitude'.

void. The void is an element itself and so it shares some characteristics of the 'full'; it is also capable of locomotion, for example, something that we can imagine when thinking of a moving gap between cars in the traffic or a vacuum in a thermos flask that we carry around.[13] Other testimonies indicate that the void is not passive; it is also the cause of things:

> [...] they [Leucippus and Democritus] say being no more is than non-being [...]; and they make these the material causes of things.[14]

Finally, we must consider the following comment of Aristotle:

> But people really think that there is an empty interval in which there is no sensible body [...] interval, different from the bodies, either separable or actual – an interval which divides the whole body so as to break its continuity, as Democritus and Leucippus held.[15]

The void is a force, a force that 'breaks up' and 'divides', because of the void reality is discontinuous.[16] Discontinuity as a principal feature of reality is perhaps the most lasting legacy of Democritean thought. '[Q]uantum theory dates 24 centuries further back, to Leucippus and Democritus. They invented the first discontinuity – isolated atoms embedded in empty space'.[17] (This may sound very similar to Heisenberg's statement but Schrödinger emphasises the discontinuous nature of reality while Heisenberg speaks of atoms as

[13] The remainder of this section has been largely influenced by an article by David Sedley: "Two Conceptions of the Vacuum" (1982), which gives a most insightful analysis of the problem of the void in the thought of the Greek sages.
[14] *Metaphysics*, 985b5.
[15] *Physics*, 213a27-b1.
[16] This means that in Democritus' thought discontinuity is constituted by the void and not by such concepts as limit or rupture. This would also mean, for example, that atoms could not push out the void into the periphery and huddle together to form a continuous material reality. (This is an idea that we find in the thought of the Stoics.)
[17] Schrödinger, 1996, 158.

basic building blocks of reality, which is quite a different focus.) This sense of discontinuity is further heightened by the realisation that the void might have been conceived as the negation of the atom; discontinuity is a consequence of an act of negation. What exact form this negation took cannot be ascertained, but it seems quite clear that it is a 'no' that introduces a discontinuity.

The preceding remarks allow us to identify three distinct systems based on the concept of atoms and the void; here is a brief resume of some of the features that each one presents us with:

1. The system of Leucippus and Democritus presents a multiplicity, a dialectics of the void and the atom, negation and indeterminacy which is a consequence of discontinuity.
2. The Epicurean system also presents a multiplicity. We do not find in it, however, either the problem of negation or discontinuity; the indeterminate character of reality is the consequence of the workings of the unpredictable *clinamen*; deviation is the source of variety and changeability. The void is empty space (this space, however, should be conceived as vectorial rather than metric).
3. If Laplace is taken to be the crowning of Gassendi's atomism then we are presented with a world of atoms in the empty void that is governed by immutable laws. It may be a godless universe but these laws operate as much as St Augustine's doctrine of predestination, as we could, for example, in principle calculate which species are destined for extinction. Such a calculation is not in reality possible but that is what makes it so alike predestination; the iron grip of inevitability is certain, only that it remains just as unfathomable as the workings of God. This is not surprising, because the very notion of a fixed immutable universal physical law has its origins in theology.

To sum up these observations one can begin by noting that the term 'atomism' is complex and the seemingly simple formula: 'there are only atoms and the void' has given very distinct philosophical outlooks. To end these remarks a comment on the term 'atomism' is necessary. 'Atomism' is a 17th century coinage, the Greeks did not have an equivalent term, they spoke of the philosophy of Democritus (and Leucippus) or the Epicurean system and did not collapse them into some unified overview, as the so often used term 'Greek atomism' obviously does. It does not make much sense to talk of 'Greek atomism' as there is little sense in talking about 'Indian atomism'.[18] The Greeks distinguished clearly between the earlier and later doctrines and this is apparent when we compare the entries 'Democritus' and 'Epicurus' in Diogenes Laertius' *Lives of Eminent Philosophers* (Books IX 34–49 and X, respectively). These are two different men, from different epochs, with different temperaments, who do, of course, also share many convictions (materialism, lack of *telos* or divinity). The term 'atomism' is an obstacle in that it conjures up an initial guiding image of billiard balls bouncing off each other, which shapes the thought; it does not do justice to Epicurus and it is certainly misleading with respect to Democritus.

As for Democritus this is an interesting comment:

> one can imagine a scholar of the young School of Athens paying a holiday visit to Abdera [...], and on being received by the wise, far-travelled and world-famous old gentleman Democritus, asking him questions on the atoms, on the shape of the earth, on moral conduct, God, and the immortality of the soul – without being repudiated on any of these points. Can you easily imagine such a motley conversation

[18] Paul Masson-Oursel recognizes four different atomist systems in India (Jaina, Buddhism, Vaiseisika and Nyaya (Masson-Oursel, 1925). And even this could be extended as the atomist thinking in Buddhist thought, at least, evolved; the somewhat materialistic atomism of the early Buddhism is very different from the temporal atomism advanced by the Dignaga – Dharmottara - Dharmakirti school that was active between the 5th and 7th centuries.

> between a student and his teacher in our days? Yet, in all probability, quite a few young people have a similar – we should say quaint – collection of inquiries on their minds, and would like to discuss all of them with the one person of their confidence.[19]

These are the words of Erwin Schrödinger, a great admirer of the Greek sage. Democritus was a philosopher of great depth, immensely richer than what goes today under the banner of 'Greek atomism'. The same goes for Epicurus. 'Atomism', as it is used, is a term that oversimplifies, it should not be used when referring to the Greeks, the finesse of their thought gets lost and with it precious philosophical insight.

[19] Schrödinger, 1996, 14.

References

Aristotle, *Metaphysics*, tr. W. D. Ross. *The Basic Works of Aristotle*. Ed.R. McKeon, New York, Random House, 1941.

-- *Physics*, tr. R. P. Hardie and R. K. Gaye. *The Basic Works of Aristotle*. Ed.R. McKeon, New York, Random House, 1941.

Diogenes Laertius, *Lives of Eminent Philosophers*, tr. R. D. Hicks, Cambridge, Massachusetts, Harvard University Press, 1980.

Heisenberg, W., 1990, *Physics and Philosophy*, Harmondsworth, Middlesex, Penguin Books.

Koyré, A., 1973, *Études d'histoire de la pensée scientifique*, Paris, Gallimard.

Lovejoy, A., 1964, *The Great Chain of Being*, Cambridge, Massachusetts, Harvard University Press.

Lucretius, *De Rerum Natura*, tr. W. H. D. Rouse and M. F. Smith, Cambridge, Massachusetts, Harvard University Press, 1982.

Masson-Oursel, P., 1925, L'atomisme indien, *Revue philosophique de la France et de l'étranger*, 99, 342-368.

Nef, F., 2011, *La Force du vide*, Paris, Seuil.

Schrödinger, E., 1996, *Nature and the Greeks* and *Science and Humanism*, Cambridge, Cambridge University Press.

Sedley, D., 1982, Two Conceptions of the Vacuum, *Phronesis*, 27, 175-193.

Models and Molecules: Representation in the Work of John Dalton

Rachel Dunn
(Durham University – UK)
r.l.dunn@durham.ac.uk

Abstract This paper briefly addresses representation in early nineteenth-century chemistry with reference to the work of John Dalton (1766-1844). The focus will be on the design and implementation of Dalton's atomic symbols and three-dimensional molecular models. It will first examine the individual symbols presented in *A New System of Chemical Philosophy* before attempting to categorize them according to design. The discussion will then shift to the ways in which he manipulated the symbols alongside the posters and handbills he employed as pedagogic tools. Dalton's wooden balls were useful musing and pedagogic aids that helped him to visualize his two-dimensional construction. In the final section, their design and function will be discussed; particular attention will be given to their spatial arrangements to suggest he was one of the earliest stereochemists. Finally, conclusions will be drawn on his deductive reasoning to show that his visual thinking was apparent in his symbols and models.

Resumo O presente artigo aborda o tema da representação na química dos inícios do século XIX, por referência à obra de John Dalton (1766-1844). A nossa atenção incidirá no *design* e na implementação dos símbolos atómicos e dos modelos moleculares tridimensionais. Primeiro são examinados os símbolos individuais apresentados em *A New System of Chemical Philosophy* antes de os categorizarmos de acordo com o *design*. Seguidamente, discutir-se-ão os modos através dos quais Dalton manipulou os símbolos, bem como os posters e folhetos que utilizou como instrumentos pedagógicos. As bolas de madeira de Dalton foram instrumentos pedagógicos úteis que o ajudaram a visualizar a sua construção bidimensional. Na secção final, será discutido o seu *design* e função; é dada especial atenção à sua distribuição espacial, sugerindo-se assim que foi um dos primeiros estereoquímicos. Finalmente, serão extraídas conclusões quanto ao seu raciocínio dedutivo para mostrar que o seu pensamento visual aparecia já nos seus símbolos e modelos.

Introduction

In this paper I will briefly discuss representation in early nineteenth-century chemistry with reference to the work of John Dalton (1766-1844). The focus will be on the design and implementation of Dalton's atomic symbols and three-dimensional molecular models. I will first examine the individual symbols presented in *A New System of Chemical Philosophy* (1808-1827) and attempt to categorize them according to design.

The discussion will then shift to the two-dimensional representations of compounds Dalton drew using his individual symbols. I will examine the ways in which he manipulated the symbols and in doing this, the posters and handbills he employed as pedagogic tools will be presented.

Dalton's wooden balls were useful musing and pedagogic aids that helped him to visualize his two-dimensional construction. In the final section, their design and function will be discussed. I will pay particular attention to their spatial arrangements to suggest he was one of the earliest stereochemists. Finally, I will draw conclusions on his deductive reasoning and hope to show that his visual thinking was apparent in his symbols and models.

What is a symbol?

For the purpose of this study, it is useful to explain what I mean when I refer to the term 'symbol'. I associate the term with the images and pictograms Dalton (and others) created to represent the chemical elements and subsequent compounds. Defining the term 'symbol' can be problematic. On a basic level, a symbol is a sign or a graphic mark that represents something. In this paper, it refers to a printed or hand-

drawn image representing a chemical substance, so it has a specified meaning. If the image does not represent anything it is merely a meaningless sign.[1] The social anthropologist Sir Edmund Leach wrote that, "a sign or a symbol only acquires meaning when it is discriminated from some other contrary sign or symbol."[2] In this respect, Dalton's symbols only take on a meaning (that is useful to the scientist) when seen as part of a complete system. Taking Leach's idea, I would suggest that the tables and compound symbols Dalton produced enabled the individual images to be compared and thus acquire meaning. Leach believed that symbolic representation is a relation between three elements: the concept in the mind, sense image and an object in the external world. He pointed out that an object (in our discussion, a chemical substance) in the external world is connected with a sense image (e.g. a visual symbol) by cultural convention, and this relationship is always arbitrary and metaphoric.[3]

The historian Maurice Crosland wrote that early chemical symbols were originally derived from pictograms, abbreviations and arbitrary signs.[4] The derivation of symbols changed towards the eighteenth century when chemists preferred to analyze symbols in search of a rational justification for each stroke.[5] Symbols often represented the name of a substance rather than the substance itself. This then followed that symbols often had a literal meaning and, when combined, represented a substance with a compound name.[6]

The meaning of a sign can change over time, with the situation, and be manipulated by the author or user, so consistency is an important factor. Dalton dallied with his symbols over the five-year

[1] Korzybski, 2000, 79.
[2] Leach, 1976, 49 quoted in Burke, 2005, 5.
[3] Leach, 1976, 19 quoted in Mach, 1993, 26-27.
[4] Crosland, 1962, 232.
[5] Crosland, 1962, 233.
[6] Crosland, 1962, 235.

period from their first recording in his notebook in September 1803 (see Figure 1), most notably addressing the representations of hydrogen, oxygen, sulphur and gold. Once he published the first volume of *A New System of Chemical Philosophy* in 1808, the symbols were fixed and their form remained unchanged through subsequent editions.

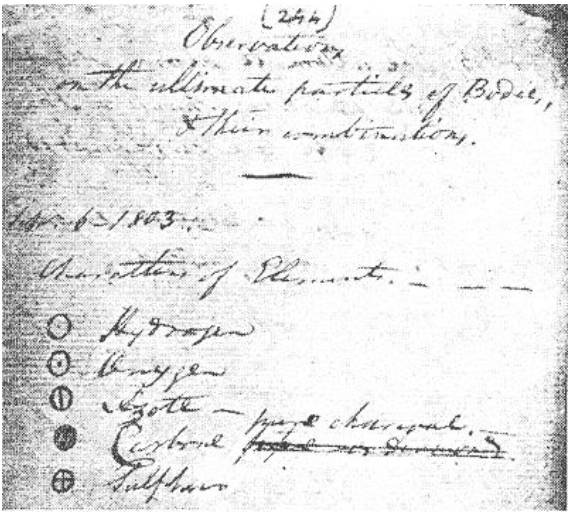


Figure 1: Initial iteration of Dalton's symbols (from his notebook, September 1803, facsimile reproduction from Roscoe and Harden, 1896, 29).

Dalton's symbols

Dalton's atomic symbols represent one atom of each element; he subsequently used the term 'compound atom' to refer to ultimate particles of compounds. Originally drawn by hand, the symbols are pictorial as they are visual tools used to represent given substances in Dalton's notebooks, lecture notes and correspondence (see reproductions in Roscoe and Harden (1896) and archives of The Royal

Society and John Rylands University Library, Manchester). They were later standardized and reproduced in Dalton's books and lectures. Dalton's symbols sometimes suggested pictorially a property or a resemblance.[7] For example, the symbol for Sulphur is similar to the symbol for Sulfur (sic) in alchemical tables with the use of a cross. As Dalton believed that atoms were surrounded by an outer atmosphere of heat, he surmised that every atom would be more or less spherical in shape.[8] The atoms were represented in circles, as a circle is the two-dimensional representation of the sphere.

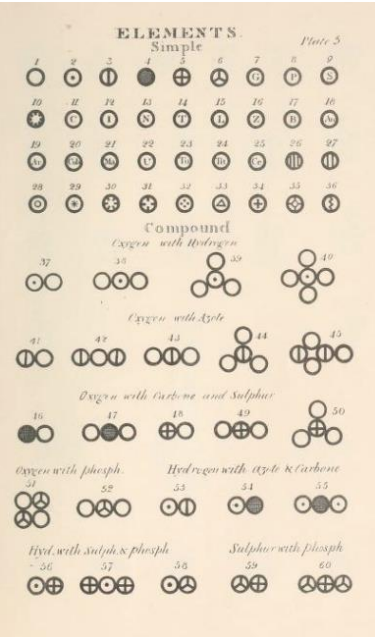


Figure 3: Dalton's symbols from *A New System of Chemical Philosophy* (1808), plate 5, no page number.

[7] Crosland, 1962, 28.
[8] Rouvray, 1995, 54.

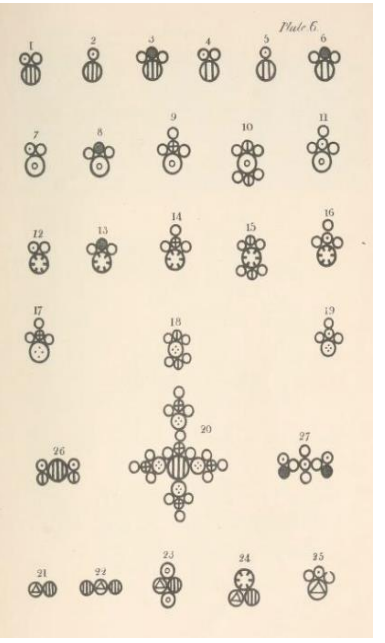


Figure 1: Dalton's symbols from *A New System of Chemical Philosophy* (1808), plate 6, no page number.

Once he had decided on a circular form, he had to compile distinguishing marks to differentiate the individual atoms. For this Dalton used a mixture of letters and patterns enclosed within the circles (see Figures 2 and 3). The letters he used were the initial letters of the English names of the elements. It was surprising that Dalton used symbolic notation incorporating letters given his "great repugnance" to Berzelius' alphabetical symbols.[9]

Aside from those designated by initial letters, where did the designs come from? Dalton invented some of the symbols whilst others show signs of earlier influences. For example, the symbol for copper was the same as Hassenfratz and Adet whereas the full circle

[9] Roscoe and Harden, 1896, 45.

for carbon was original. In a similar way to Geoffroy, some of Dalton's symbols had their origins in alchemy. Dalton borrowed the use of pictures instead of letters from alchemy, but with one important distinction, he meant each individual picture to represent particular quantities of atoms.[10]

Looking at similarities and groupings, in discussion with Dalton's symbols, I would classify them in three categories: (1) abbreviations (e.g. G for Gold), (2) symbolic signs (reminiscent of the substance's properties or form, e.g. a coloured black circle for Carbon) and (3) original signs (e.g. the symbol for Glucine). Dalton's symbols appeared to fall into these three categories, with more original or arbitrary signs present in his earlier symbols and more abbreviations in later lists.

As Dalton's symbols were a scientific tool, they represented the iterative nature of scientific investigation as they changed over time as the system developed from notebook to publication. For example, Hydrogen and Oxygen originally had the opposite symbols. The symbol for Sulphur was similar to the symbol for Sulfur (*sic*) in alchemical tables with the use of a cross. If we examine Dalton's symbols for Hydrogen, Oxygen and Carbon, it is possible that he chose the simplest symbols, (a circle with a central dot, an empty circle and a solid circle) for these most commonly used elements. This would aid their use in handwritten texts and make dies or engraving blocks easier and possibly cheaper to produce. Such simple symbols also suggest a leaning towards the Quaker notion of simplicity, to which Dalton subscribed, or could simply recognise the need for brevity in recording scientific data. Despite the fact pictorial symbols represented a typographical problem, Dalton persisted with them.

[10] A more in-depth discussion of the development of Dalton's symbols is given in Maurice Crosland's 1962 study, *Historical Studies in the Language of Chemistry*. New York, Dover, 256-264.

Manipulating the symbols: making compounds

Alongside his symbols for the atomic elements, Dalton drew two-dimensional representations of compounds using his individual symbols. He placed symbols next to each other in an order believed to be the actual spatial arrangement of the atoms in the compound. In creating these compounds he assumed the simplest possible formulas, e.g. water as a binary compound, HO, rather than, as we now know, H_2O. Essentially, Dalton had to make assumptions as to the numbers of atoms that combined to form each compound. This deductive reasoning has been praised by observers such as Roscoe and Harden, but criticised by others, such as David Knight, who recognised that Dalton's two-dimensional constructions failed to connect with the work of contemporaries such as William Hyde Wollaston who was engaged in work on the composition of crystals.

The creation of a set of symbols to represent each element almost acts as an alphabet, but to make effective use of it, Dalton needed to write words using it, through the creation of compound symbols. Dalton's symbols for compounds were made up of easily recognizable, unaltered symbols that could be deconstructed to their constituent parts. In this respect, I put forward that Dalton's compounds were useful as the reader could recognize the elements involved and draw schematic interactions.

Dalton himself employed his symbols in this capacity. However, their very structure made them cumbersome and led to difficulty in writing formulae for complicated compound atoms, meaning they were impractical and often imposed a spurious structure on some compounds, e.g. alum.[11] It did mean that the reader could visually interpret the chemical formula, e.g. alum is constructed of one atom

[11] Rouvray, 1977, 28.

of potash, four of alumine, five of sulphur and fifteen of oxygen (see Figure 3 from *A New System of Chemical Philosophy*). Composing a new symbol for each compound would have been easier, but Dalton maintained his construction of compound atoms for years to come, as seen in his diary and lecture notes.[12] Using this symbolic representation of invisible atoms, they could be combined on paper, discussed and corrected. This paper modelling enabled atoms to be manipulated, creating interactions that could then be confirmed or denied by experiment; this theoretical chemistry could be considered ahead of its time.

Employing the symbols in teaching materials

Dalton's symbolic notation described above enabled him to easily communicate his ideas. Fellow chemist Thomas Thomson publicized Dalton's symbols in his 1807 work, *System of Chemistry*, the year before Dalton first published them in *A New System of Chemical Philosophy*.[13] He later wrote that representing atoms by Daltonian symbols aided clarity.[14] Thomson subsequently arranged for Dalton to lecture in his classroom in Edinburgh in April 1807 during a tour of Edinburgh and Glasgow. In his Scottish lectures, Dalton used various posters, including one that presented symbols and relative weights of the first twenty substances included in the first volume of *A New System*. The posters were quite large, being 19" by 26 ½". Gee, Coward and Harden described fifty-three lecture diagrams that refer to the atomic theory (for an example, see Figure 4).[15]

[12] As discussed by Roscoe and Harden.
[13] Thomson, 1807, 429-540.
[14] Thomson, 1830, 291.
[15] Gee et al. 1915.

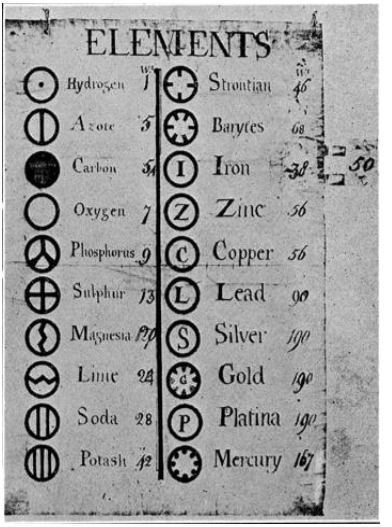


Figure 4: Photograph of Dalton's lecture poster "*Elements*" for his Scottish lectures in 1807 (from Coward, 1927, 29).

The posters were considered lost until just before the First World War when some were discovered in a cupboard at the Manchester Literary and Philosophical Society. There were about one hundred and fifty altogether, which Dalton used in teaching, all apparently prepared by his own hands. Some of them illustrated Dalton's ideas of the chemical constitution of substances, and his exposition of the atomic theory. There can be little doubt that the posters were prepared in 1806 or 1807 for illustrating the lectures delivered in Scotland in April of the latter year.[16] It is with regret that the documents were lost, along with a considerable amount of the Dalton archive, following a fire during the Second World War.

Dalton allegedly made greater use of his symbols in oral presentations than in his published works and laboratory notebooks.[17]

[16] Coward, 1927, 29.
[17] Gee et al., 1915, 41.

This is quite remarkable as he frequently employed his symbols in written work.[18] In an essay read to the Manchester Literary and Philosophical Society in October 1830, Dalton discussed the necessity for the lecture posters in explaining the symbols of atoms and their arrangement in compound atoms:

> All this would have been very difficult to convey to a large and promiscuous audience without suitable diagrams.[19]

He also issued handbills at, among other lectures, a BAAS meeting and an 1835 lecture in Manchester.[20] An example of the handbill given at the 1835 meeting is shown in Figure 5.

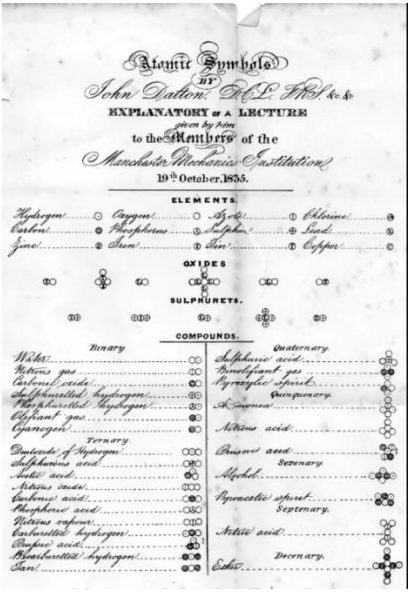

Atomic Symbols
BY
John Dalton, D.C.L. F.R.S. &c. &c.
EXPLANATORY OF A LECTURE
given by him
to the Members of the
Manchester Mechanics' Institution,
19th October, 1835.
ELEMENTS.
Hydrogen Carbon Azote Chlorine
Lime Iron Soda Copper
OXIDES
SULPHURETS.
COMPOUNDS.
Binary
Water
Nitrous gas
Carbonic acid
Olefiant gas
Cyanogen
Ternary
Sulphuric acid
Nitric acid
Nitrous acid
Carbonic acid
Phosphoric acid
Oxalic acid
Quaternary
Sulphuric acid
Pyroxylic spirit
Quinquenary
Nitrous acid
Senenary
Alcohol
Septenary
Nitric acid
Decenary
Ether

Figure 5: Copy of the handbill issued at the Manchester Mechanics Institution in 1835 (from author's personal collection).

[18] Gass, 2007, 371.
[19] Dalton, paper read to the Man Lit & Phil 15 October 1830. Fire damaged manuscript reproduced in Thackray, 1972, 96.
[20] Thackray, 1972, 119; Smyth, 1998, 59.

One of the most important aspects of Dalton's science was that he thought in three-dimensions with respect to atomic dynamics. In 1810, he engaged his friend Peter Ewart (an engineer who worked for Bolton and Watt) to produce some three-dimensional wooden models to assist in his teaching. He described their construction, thus:

> My friend Mr Ewart, at my suggestion, made me a number of equal balls, about an inch in diameter, about 30 years ago; they have been in use ever since, I occasionally showing them to my pupils. One ball had 12 holes in it, equidistant; and 12 pins were stuck in the other balls, so as to arrange the 12 around the one and be in contact with it: they (the 12) were about 1/20 inch asunder. Another ball, with 8 equidistant holes in it, and they (the 8) were about 3/10 of an inch asunder, a regular series of equidistant atoms. The 7 are an awkward number to arrange around 1 atom. The six are an equidistant number of atoms, 90 degrees asunder, 2 at the poles and 4 at the equator. The 4 is a split of the 8, a regular number and equidistant. The 3 are around the equator at 120 degrees asunder; the 2 at opposite poles.[21]

These new tools were the size of gumballs and were drilled with varying numbers of holes that enabled the spheres to be connected by pins. In a number of the balls, the holes were equidistant to reflect the notion of simplicity and most likely association. The use of such three-dimensional, tactile tools meant that Dalton and his students had a better understanding of the actual molecule, demonstrating a stereochemical perspective. This has led to authorities such as Rouvray calling Dalton 'the world's first stereochemist'.[22] Despite this claim, he was not the only early chemical atomist to engage in stereochemical thinking, as William Wollaston also recognized three-dimensions and produced wooden models to illustrate his ideas.

[21] Dalton, 1840-1842, 41-42.
[22] Rouvray, 1995, 52.

Dalton's system of modelling

Having previously discussed Dalton's symbols as a system of representation and a visual, 'paper' tool, it is prudent to address his three-dimensional models in a similar manner. The notion of a scientific tool is a concept that historians of science have discussed for many years; most have an open mind when it comes to what counts as a scientific tool.[23] While 'tool' often infers something tangible, as Dalton's wooden balls were, the word has been redefined in terms of function. Tools assist in investigatory science and scientific education. The most useful tools from the history of science include the balance and the microscope.[24] These can be considered more laboratory equipment as they are fundamental to investigation, whereas I view a tool as something more supportive and descriptive. In this vein, Dalton's wooden balls were useful musing and pedagogic aids that helped him to visualize his two-dimensional construction. In this part of the paper I will examine the design and implementation of Dalton's wooden balls, paying particular attention to their stereochemical properties.

The nature of Dalton's atomic models

Atomic models are useful tools for visualizing the structures and shapes of groups of atoms, which is important for understanding their behaviour. Modelling is an important aspect of modern chemistry. Dalton was one of the first people to use ball and spoke models to represent the structure of molecules.

[23] Gass, 2007, 350.
[24] Gass, 2007, 350.

Some of Dalton's atomic models survive and provide an intriguing insight into pedagogical tools from the early nineteenth century.[25] There are two known sets of Dalton's atomic models: one in the Science Museum in London and the second in the Museum of Science and Industry in Manchester (MOSI).

Looking at the balls in detail, the models from the Science Museum appear to have traces of coloured paint on them, to denote the different atoms. This colouring is common in today's Molymod kits, but could be unusual and innovate for the time. I suggest that the models were originally coloured black, white and brown. The models from the Science Museum seem to correspond with the ball with twelve equidistant holes and the one with two at opposite poles.

The second set, from MOSI (see Figure 6), is more interesting as they represent atoms of different sizes. The models are graduated and could easily be used to represent compound atoms (molecules) such as hydrate of potash, which Dalton depicted as having one larger and two smaller components. Despite having the models created in 1810 and using them in his lectures and for other demonstrations, he failed to publish details of them until 1842, two years before his death. Thus, many of Dalton's conjectures on the spatial positions of atoms in molecules were published after his death by Roscoe and Harden in their work of 1896.

[25] Gass, 2007, 365.

Figure 6: Set of atomic models from MOSI (author's own image). The British 2p coin is given for size comparison.

Dalton and stereochemistry

Stereochemistry refers to the study of the relative spatial arrangement of atoms within molecules. The prefix 'stereo' denotes the three-dimensionality of the molecular models. The emergence of stereochemistry is usually associated with the second half of the nineteenth century and the work of Jacobus Van't Hoff.[26] Until the early 1870s chemists lacked the conceptual tools to picture and model three-dimensional molecules.[27] Even apparently structural formulae indicated only the arrangement of atoms, links in a chain rather than positions in space. Historian of chemistry Trevor Levere noted that it would be difficult to know the real positions of atoms in space.[28] This

[26] For a more detailed analysis of the history of stereochemistry see Peter J. Ramberg *Chemical Structure, Spatial Arrangement: The Early History of Stereochemistry, 1874-1914*.
[27] Levere quoted in Ramberg, 2005, xxi.
[28] Levere quoted in Ramberg, 2005, xxi.

is true, and while many of Dalton's arrangements are incorrect or highly unlikely (such as a dodecahedral molecule) others are related to modern molecular models. The importance of stereochemistry was not fully recognised in the early nineteenth century and as such the majority of Dalton's models are basically topological.

Looking at Dalton's compound atoms, those he termed binary and ternary are linear and as such have predicted bond angles of 180°. These can be easily modelled using Dalton's balls. Molecules such as sulphurous acid (one azote and three oxygen according to Dalton) were drawn with one central atom of azote and three outer atoms of oxygen arranged equidistant as if on the corners of an equilateral triangle. This trigonal arrangement is evident from the existing atomic models and Dalton's own admission of having a ball with three holes "around the equator at 120 degrees". The simplicity of the linear and trigonal arrangements reflects the assumptions he made when combining atoms.

Given Dalton's specifications there was scope to create various molecular models using the balls with twelve and eight equidistant holes. These balls could be used as a central atom and enabled Dalton to vary the number of connections. As they had holes across the entire curved surface and not just around the equator, it suggests arrangements away from the planar (such as the aforementioned linear and trigonal). This move towards the tetrahedral and octahedral is reinforced by a drawing in Dalton's own notebook from March 1834. Figure 6 depicts a three-dimensional tetrahedral arrangement for the oxamide molecule in Dalton's own hand. The formulae for the molecule demonstrate that he believed the compound was composed of one atom of carbon (shaded circle), one atom of hydrogen (circle with a dot), one atom of nitrogen (circle with vertical line), and one atom of oxygen (open circle). Dalton appears to be attempting to

depict stereoisomers of the oxamide molecule.[29] The placement differs from versions of four-atomed compounds in *A New System*, as the central (fourth) atom is drawn overlapping a triangular arrangement of the other three. This layering is unusual in Dalton's work, but there is the possibility that the quaternary elements (compounds) in Figure 2 were in fact two-dimensional representations of tetrahedral molecules. If this is indeed the case, it raises questions about the intended three-dimensional arrangements of the other compounds listed in Figures 2 and 3. Currently this is speculation, but the intricacies of producing a printed plate depicting tetrahedral (or other three-dimensional arrangements) would create problems for the printer and therefore financial considerations for the publisher.

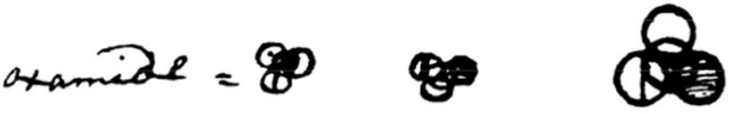


Figure 7: Three-dimensional representation for the oxamide molecule (from Dalton's notebook, No. 11, March 1834, p. 372 reproduced in facsimile by Coward, 1927, p. 34).

An interesting comment from Dalton was that seven atoms were "an awkward number to arrange around one atom" suggesting that he thought in terms of both simplicity and design.[30] Certainly, an aesthetic consideration could point towards his belief that his system of notation was "the only one representing nature" and nature was inherently symmetrical and pleasing to the eye.[31]

[29] Coward, 1927, 34.
[30] Dalton, 1840-42, 42.
[31] Thackray, 1972, 118.

What could a three-dimensional model offer Dalton?

Dalton used his two-dimensional symbols in both his research and teaching. They were reproduced in *A New System of Chemical Philosophy* and also found their way into his correspondence (Figure 8). Thus, they were a useful paper tool and part of a visual language. In terms of his models, the question that we must ask is: can a system of notation such as Dalton's atomic symbols be transformed into a tactile, user-friendly system? It seems evident that this was one of his aims with respect to creating the wooden models as their design correlates to the two-dimensional paper versions of compound atoms. The set of graduated models demonstrated this most effectively as they could have been manipulated to represent a plethora of designs. But the question remains as to whether it was a user-friendly and indeed, useful system.

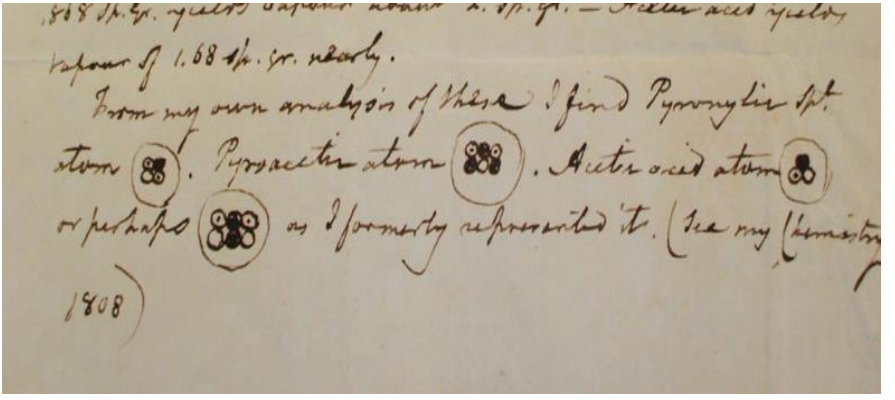


Figure 8: Letter from J. Dalton to J. L. Gay-Lussac (5 August 1833): use of an encompassing circle to (possibly) separate compound atoms (author's own image of MM/1/7 from The Royal Society archives).

Modelling molecules is an important part of modern chemistry teaching, however, the reasons behind the use of models has changed over the past two centuries. Transforming a collection of balls and spokes into a three-dimensional representation of an atom, then combining two or more into a compound demonstrates a chemical process, which can be considered a chemical language. Combining actual chemicals in the laboratory and seeing the reaction is visually rewarding, particularly if an effervescent or explosive reaction is seen, so it is sensible to use a visual language to theoretically represent such a process. Being able to handle and manipulate molecules could also aid retention among his pupils, thus their use as a teaching aid was not only connected with the scientific but also the kinesthetic.

Conclusions

The focus of this paper was Dalton's visualisation of chemical elements and their bonding; Dalton used both pictorial and alphabetical notion to create his symbols within circles, representative of his concept of the atom. Although not conducive to classification, it is interesting to note that when students draw representations of atoms and bonding at school nowadays they initially draw a circle to represent the atom, as Dalton had done two centuries earlier, before attaching the relevant number of electrons.

It has been said that, "Symbolic formulae … would deserve to rank among the chemist's most powerful instruments of research."[32] I would expand this idea to include all chemical symbols and not just limit their use to an instrument of research, but as a useful tool for teaching and learning. Through his use and manipulation of symbols

[32] Hofmann, 1865, 87.

and 3D models it suggests his deductive reasoning and visual thinking.

By creating three-dimensional representations of his atomic symbols, Dalton was able to investigate the composition and structural arrangement of molecules. It is clear from other studies that stereochemistry was very much an emerging discipline in the nineteenth century, so Dalton would not have fully understood the implications of the models. However, the ball and spoke system is simple and by virtue of its application to chemical problems suggests deductive reasoning and visual thinking on behalf of Dalton and his pupils. As a modelling system, it perhaps did not achieve its full potential in the hands of Dalton, but its descendants are seen in the schoolrooms of today. As a pedagogic tool and a physical *aide memoire* the balls could be deemed useful but further investigation is necessary to prove this beyond mere conjecture.

References

Burke, P., 2005, *The Historical Anthropology of Early Modern Italy: Essays on Perception and Communication*, Cambridge, Cambridge University Press.

Coward, H. F., 1927, The early years of the atomic theory as illustrated by Dalton's own note-books and lecture diagrams. *Journal of Chemical Education*, 1927, 4 (1), 23-37.

Crosland, M., 1962, *Historical Studies in the Language of Chemistry*, New York, Dover.

Dalton, J., 1808 – 1810, *A New System of Chemical Philosophy*, Manchester, S. Russell.

Dalton, J., 1840-1842, *On the Phosphates and Arseniates and a New and Easy Method of Analysing Sugar*, Manchester, Harrison.

Gass, G., 2007, Spheres of Influence: Illustration, Notation, and John Dalton's Conceptual Toolbox, 1803-1835. *Annals of Science*, Vol. 64, No. 3, 349-382.

Gee, W. H. et al., 1915, John Dalton's Lectures and Lecture Illustrations. *Manchester Memoirs*, Vol. 59, 1-66.

Hofmann, A. W., 1865, *Introduction to Modern Chemistry*, London, Walton and Maberly.

Korzybski, A., 2000, *Science and sanity: an introduction to non-Aristotelian systems and general semantics*, Brooklyn, Institute of General Semantics.

Knight, D., 1992, *Ideas in Chemistry*, London, Athlone Press Ltd.

Leach, E. R., 1976, *Culture & communication: the logic by which symbols are connected: an introduction to the use of structuralist analysis in social anthropology,* Cambridge, CUP.

Mach, Z., 1993, *Symbols, conflict, and identity: essays in political anthropology*, Albany, State University of New York Press.

Ramberg, P. J., 2005, *Chemical Structure, Spatial Arrangement: The Early History of Stereochemistry, 1874-1914*, Aldershot, Ashgate.

Roscoe, H. E. and Harden, A., 1896, *A New View of the Origin of Dalton's Atomic Theory: A Contribution To Chemical History*, London, Macmillan.

Rouvray, D. H., 1977, The changing role of the symbol in the evolution of chemical notation. *Endeavour*, Vol. 1, (1), 23-31.

Rouvray, D. H., 1995, John Dalton: the world's first stereochemist. *Endeavour*, Vol. 19, (2), 52-57.

Smyth, A. L., 1998, *John Dalton, 1766-1844: A Bibliography of Works by and About Him, With an Annotated List of His Surviving Apparatus and Personal Effects*, London, Ashgate.

Thackray, A., 1972, *John Dalton: Critical Assessments of His Life and Science,* Cambridge, MA, Harvard University Press.

Thomson, T., 1807, *A System of Chemistry, Volume 3*, London, John Murray.
-- 1830, *The History of Chemistry, Volume 1*, London, Colburn and Bentley.

La physique laplacienne dans la seconde moitié du XIX^{e} siècle: Joseph Boussinesq – la pratique et la réflexion autour de l'atomisme en France vers 1875

João Príncipe
(IHC/CEHFCi – Universidade de Évora)
jpps@uevora.pt

Abstract Joseph Boussinesq is one of the most distinguished French scholars that, during the second half of the nineteenth century, develops the Laplacian program of molecular physics. He made important contributions to the theories of aether, elasticity and hydrodynamics and also wrote on the philosophy of science. I analyze some of the different functions of the atomic assumptions in the theories of laplacian molecular physics. Then I show how Boussinesq assigned a central role to the notion of point atoms in his scientific work and in his philosophical reflections, which concern the method and theory of knowledge. This case study is connected with a critical appraisal of some recent literature on empiricist biased philosophy and history of science.

Resumo Joseph Boussinesq é um dos mais distintos académicos franceses que, durante a segunda metade do século XIX, desenvolveu o programa laplaciano da física molecular. Fez importantes contributos para as teorias do éter, da elasticidade e da hidrodinâmica, e também escreveu sobre filosofia da ciência. Analiso algumas das diferentes funções das hipóteses atómicas nas teorias da física molecular laplaciana. Mostro então como Boussinesq atribui um papel central à noção de 'átomos-centros de força' na sua obra científica e nas suas reflexões filosóficas sobre método e teoria do conhecimento. Este caso de estudo relaciona-se com a crítica a alguma literatura recente sobre filosofia e história da ciência de orientação empirista.

Joseph Valentin Boussinesq (1842-1929) est l'un des plus distingués savants français qui au cours de la seconde moitié du XIX^e siècle poursuit sur la voie du programme laplacien de la physique moléculaire ; il a fait d'importantes contributions à la théorie de l'éther, de l'élasticité et de l'hydrodynamique et a aussi écrit sur la philosophie des sciences. L'analyse de ses travaux et réflexions illumine la question générale du statut et fonction des hypothèses atomiques et peut permettre un aperçu plus complet de l'histoire des atomes.

En *Prologue de problématique*, à propos des différentes lectures philosophiques de l'histoire des théories atomiques, je présente quelques considérations qui visent à mettre en évidence le caractère problématique de certaines généralisations historiques. Dans le paragraphe suivant j'analyse des différents travaux développés dans le cadre de la tradition française de physique moléculaire d'origine laplacienne (ou mécanique physique) du point de vue du rôle joué par les hypothèses atomiques. Dans les deux derniers paragraphes, je montre comment Boussinesq a attribué un rôle central à la notion d'atomes-centres de force dans ses travaux scientifiques, mais aussi dans ses réflexions philosophiques, lesquelles s'occupent de la méthode et de la théorie de la connaissance ; je me concentrerai surtout dans ses travaux publiés avant 1880, en particulier sur son ouvrage *Étude sur divers points de la philosophie des sciences*.

1. Prologue de problématique

Le savant néerlandais Andreas van Melsen à la fin de son histoire des atomes (1949), traduite vers l'anglais en 1952, *From Atomos to Atom*, nous a laissé quelques remarques philosophiques et conclusions générales. Pour lui les conceptions atomiques, dans ce

qu'elles portent de philosophique, ont fait partie, à travers les âges, de ce qu'il appela le fond (background) philosophique des scientifiques, couche de croyances inconscientes, qui dans certains cas ont été consciemment réfléchies, en devenant des convictions philosophiques.[1] Les atomes appartiennent à une classe de préjugés d'origine philosophique sans lesquels les sciences physiques seraient incapables de se constituer. Puisque ce fond reste actif dans l'intellect des scientifiques qui ont développé des théories atomiques après 1800, van Melsen juge que la division en deux périodes historiques, le premier philosophique et le second scientifique, avec une période de transition entre les siècles XVII et XIX, ne saisit pas l'essentiel de l'évolution historique.

Par contre, un ouvrage récent (2006) tranche nettement entre les deux versants que van Melsen jugeait constituer un tout évolutif et complexe. Le titre de l'ouvrage de Alan Chalmers met en évidence nettement la distinction et la discontinuité entre deux discours ou attitudes : *The scientist's atom and the philosophers stone: how science succeeded and philosophy failed to gain knowledge of atoms.* Il juge que les hypothèses philosophiques, ne peuvent qu'être 'accommodées' par l'expérience des phénomènes, ce que les distingue des théories scientifiques à part entière lesquelles peuvent être confirmées directement par l'expérience. Un exemple majeur d'une telle hypothèse philosophique c'est l'atome des philosophes dont la conception remonte aux idées de Démocrite. Après un chapitre à allure épistémologique, dans lequel il fournit des critères permettant de trancher entre les deux domaines, il raconte, avec détail, l'histoire des atomes depuis l'antiquité jusqu'au XX^{e} siècle. D'après lui, il a fallu attendre la fin du XIX^{e} siècle pour qu'on puisse trouver une version vraiment scientifique de la théorie atomique. Cette conclusion n'est pas si différente de celle du sens commun, scientifiquement

[1] van Melsen, 1952, 215-216, 218.

renseigné, du début du XX^e siècle, lequel voyait dans les expériences de Jean Perrin sur le mouvement brownien la preuve finale et décisive de la réalité moléculaire. Dans le chapitre sur l'histoire de la physique des atomes pendant le XIX^e siècle, Chalmers ne considère que la théorie cinétique de Clausius, Maxwell et Boltzmann et remarque les anomalies de cette théorie, en particulier le problème des chaleurs spécifiques. Le tout de son argument est biaisé dans le sens de valoriser le contact expérimental avec les atomes, en oubliant d'autres fonctions scientifiques associées à l'hypothèse atomique. Mais en tant qu'ouvrage qui veut arriver à des conclusions philosophiques en partant de l'histoire des sciences, il est difficile à comprendre pourquoi il restreint l'histoire des atomes des physiciens au XIX^e siècle à la théorie cinétique. Van Melsen l'avait fait aussi, mais lui il a écrit son ouvrage à une époque où les historiens des sciences n'avaient pas creusé le riche territoire de la physique moléculaire laplacienne, dont les réussites scientifiques sont indéniables. Après les travaux de Grattan Guiness, Robert Fox, John Heilbron, Bruno Belhoste et Olivier Darrigol, entre autres, il y un matériel assez riche, aussi du point de vue de la philosophie des sciences, lequel, curieusement, avait été cité par le savant-philosophe Émile Meyerson dans son *Réalité et Identité,* ouvrage parue en 1908. La position de Meyerson se trouve aux antipodes de celle de Chalmers ; il valorise la quête de la causalité et ce qu'il nomme le postulat de l'identité dans le temps comme une partie intégrante de notre raison; et il dénonce le biais empiriste, qui concrètement a produit, au début du nouveau siècle, l'effacement d'une partie de l'histoire réelle de la physique:

> De là aussi l'illusion, si fréquente chez les savants et même chez quelques philosophes, qui prétendent distinguer la conception atomique spéciale à une partie de la science comme «vérité expérimentale», de la théorie atomique générale qualifiée d'hypothétique, voire même l'ensemble des théories atomiques modernes de celles des anciens, alors qu'il est facile de s'apercevoir

> qu'en réalité toutes ses conceptions se tiennent, qu'il y a entre elles une véritable communauté de fond.[2]

Dans ce qui suit j'illustre ce qu'on peut en profiter de cette connaissance historique pour une philosophie des sciences consciemment non empiriste.

2. La tradition de physique moléculaire

Vers 1870 les savants français commencent à s'intéresser par la théorie cinétique des gaz dans la version de Robert Clausius ; or, la théorie cinétique, dans ses deux versants de théorie sur la nature de la chaleur et de théorie des gaz, aborde des questions déjà agitées par les physiciens français. La réception de la première théorie cinétique des gaz se fait dans des cadres conceptuels dominés par la tradition de physique moléculaire, issue de Laplace, et de la tradition opticienne, issue de Fresnel et de Cauchy. Ces deux traditions, dans leurs explications ou 'déductions', postulent l'existence d'atomes ou de molécules centres de force.[3]

Le *schéma explicatif général des explications quantitatives de la physique moléculaire* peut être aisément compris en regardant les calculs de Laplace pour la capillarité ou pour l'équation d'état des gaz déduite d'après l'hypothèse du calorique. Dans la sphère d'activité d'une molécule on trouve assez d'autres molécules pour qu'on puisse calculer l'effet des interactions en remplaçant des sommes discrètes par des intégrales. Les forces intermoléculaires deviennent insensibles

[2] Meyerson, 1908, chap. II, 103-104 de la 3ème édition.
[3] Jean Perrin, dans ses réflexions sur les hypothèses atomistiques, a écrit : 'une méthode s'est développée qui consiste à imaginer a priori pour la matière une structure dont la perception directe échappe encore à nos sens imparfaits, et telle que sa connaissance permettrait de prévoir par voie déductive les propriétés sensibles de l'univers'. Perrin, 1903, VII.

à des distances sensibles, ce qui permet de substituer les limites d'intégration finies (correspondant à des distances entre éléments de ligne, de surface ou de volume) par des limites infinies, la valeur des intégrales ayant donc un caractère à peu près universel. En 1821 Navier publia le mémoire dans lequel il donnait les bonnes équations pour le mouvement vibratoire d'un solide élastique isotrope et homogène de type particulier. Ce solide était conçu d'après le modèle laplacien. Il n'avait qu'une seule constante élastique.[4] Pour Navier, dans l'état naturel du corps, l'attraction et la répulsion entre deux molécules quelconques se détruisent réciproquement. Cet état pour lequel le matériau est dénué de toutes contraintes sert de référence pour les déformations. Dans la théorie de l'élasticité linéaire inaugurée par Cauchy en 1827, les composantes du tenseur des contraintes sont des fonctions linéaires des composantes du tenseur des déformations infinitésimales, ce qui correspond à une généralisation de la loi de Hooke. Pour l'état naturel les deux tenseurs sont nuls. Cauchy a été le premier à appliquer cette théorie aux vibrations lumineuses. Dès le mémoire lu le 12 janvier 1829, il admet que l'éther vibre comme un solide élastique continu isotrope et homogène. Son équation contient deux constantes (dans le modèle de Navier, qui admet que l'interaction entre deux molécules ne dépend géométriquement que de leur distance mutuelle, les deux constantes sont reliées). Nonobstant, en 1839, Cauchy réalisa des calculs d'actions moléculaires pour les solides isotropes (ce qui donnait une condition reliant les deux constantes élastiques de Lamé), et les théoriciens français de l'éther élastique actifs pendant la seconde moitié du siècle

[4] La principale prévision qui permettait de distinguer l'approche moléculaire de la théorie de l'élasticité de l'approche basée sur le continuum était la diminution du nombre de constantes élastiques indépendantes, Darrigol, 2001, 319.

(tels que Émile Sarrau, Charles Briot ou Boussinesq) sont restés fidèles à la perspective moléculaire.[5]

Bien que les explications moléculaires proposées par Laplace aient un caractère spéculatif et contiennent plusieurs hypothèses ad hoc (le cas typique étant la 'déduction' par Laplace de la loi des gaz parfaits, dans le cadre du calorique), la conception des atomes centres-de-force suggérait une théorie de la matière en général. Cet aspect se trouve aussi chez André-Marie Ampère. Il est assez connu qu'Avogadro en 1811, et Ampère en 1814, énoncent l'hypothèse sur la proportionnalité entre les volumes des gaz et leur nombre de molécules. La mathématisation mise en œuvre par Ampère pour sa théorie de la combinaison chimique nécessite une claire distinction atome/molécule d'une part et mélange/combinaison d'autre part.[6] Mais André-Marie Ampère ne s'est pas limité à ces spéculations chimiques basées sur des conceptions atomiques. Pour lui l'ontologie atomiste constitue une cosmovision. Ampère, par exemple dans son mémoire de 1835 '*Sur la chaleur et sur la lumière considérées comme résultant de mouvement vibratoires*', fait la distinction entre particule, molécule et atome. Les atomes sont des points matériels centres de forces (attractives ou répulsives). Ils s'assemblent en formant des molécules de forme polyédrique, les atomes occupant, en général, les sommets. Les atomes sont toujours tenus à distance par les forces propres des atomes et peuvent vibrer autour de positions moyennes. Les molécules vibrent elles aussi mais on doit distinguer les deux types de vibration :

[5] Sur l'histoire des théories de l'éther voir : Saint-Venant, 1872; Whittaker, 1951, chap. IV et V; Schaffner, 1972, surtout chap. IV; Buchwald, 1985, Appendix 2. Sur Cauchy voir Belhoste, 1992 ; sur les conceptions atomistiques et finitistes de Cauchy voir ses *Sept Leçons*, Cauchy, 1868 (éditées par L'Abbé Moigno, l'un des grands défenseurs de la physique moléculaire vers 1860 ; voir, par exemple, Moigno 1868).

[6] Voir Scheidecker-Chevalier, 1997, 170-175.

> C'est aux vibrations moléculaires et à leur propagation dans les milieux ambiants que j'attribue tous les phénomènes du son ; c'est aux vibrations atomiques et à leur propagation dans l'éther que j'attribue tous ceux de la chaleur et de la lumière.[7]

Les particules sont des assemblages d'un nombre assez grand de molécules, entourées d'éther, et constituant cependant des portions infiniment petites des corps qui ont encore la même nature qu'eux (ce qui signifie qu'une particule d'un corps solide est solide, celle d'un corps liquide est liquide, etc.). Parmi les savants français actifs vers 1870 la théorie ondulatoire de la chaleur méritait encore considération ; cela n'est pas une spécificité française – dans toute la seconde moitié du XIX^e, plusieurs conceptions cinétiques de la chaleur étaient disponibles et quelqu'un comme Clausius acceptait ce pluralisme ou tout au moins le reconnaissait. Cette situation peut être comparée à la multiplicité des théories mécaniques contemporaines de l'éther optique.[8]

La conception atomique du père Boscovich, qui a été adopté par Ampère, est très attirante pour ceux qui veulent réduire la mécanique des systèmes conservatifs à la mécanique de points matériels centre de force.[9] Cela n'a pas été fait par Ampère qui s'intéressait surtout à la théorie ondulatoire de la chaleur.[10] Parmi les raisons justifiant l'introduction du concept d'énergie au XIX^e siècle, la croyance à la réduction de tous les phénomènes à ce qu'on appelle de nos jours la mécanique des systèmes conservatifs a joué un important rôle dans

[7] Ampère, 1835, 436.
[8] Voir Clausius, 1857, Boussinesq, 1873, et Príncipe 2008, § 4.3.3 et § 5.3.
[9] Sur la conception atomique du père Boscovich voir par exemple Dalmedico, 1989, 81-83.
[10] Cependant, en 1832, il a défini la force vive du mouvement vibratoire, comme la somme de deux termes, lesquels correspondent à notre énergie cinétique et notre énergie potentielle ; Ampère a été le premier à énoncer la constance de cette somme pour le mouvement vibratoire modélisé par l'analogie avec un système de diapasons dans le vide ; voir Darrigol, 2001, 321-322.

les formulations d'un principe de conservation général, faites par Adhémar Barré de Saint-Venant (1797-1886) vers 1830 et par Helmholtz dans son mémoire de 1847. Cette raison a été un peu oublié par les historiens qui ont ignoré le fait que la mécanique de D'Alembert était non conservative, en traitant les mouvements assujettis à des contraintes et considérant des chocs durs : par exemple il considérait comme évident qu'un corps dur se mouvant en direction à un obstacle plan annule soudainement la composante perpendiculaire de sa vitesse. Les traités de D'Alembert et de Lagrange ont été le paradigme de la mécanique rationnelle propagée par les géomètres Laplace et Poisson au moins jusqu'aux années de 1820. Le raisonnement de Saint-Venant peut être mis dans la lignée du principe des forces vifs de Leibniz – la conservation de la force vive dans un système mécanique isolé est justifiée par Leibniz par un argument théologique et par l'impossibilité du mouvement perpétuel. Saint-Venant n'est pas habituellement cité comme étant un de premiers à énoncer le principe général de la conservation de l'énergie. Cela est dû au fait que le mémoire qu'il a présenté à l'Académie des Sciences vers 1828 n'a pas été publié, mais ce mémoire contient explicitement l'idée que la somme de l'énergie cinétique et de l'énergie potentielle est une constante pour l'Univers. Cet Univers est en accord avec la cosmovision laplacienne. Dans les formulations postérieures (1834, par exemple), Saint-Venant abandonne les molécules étendues de Laplace en les substituant par une conception plus simple: celle d'un Univers constitué par des points matériels centre de force. Son argumentation est aussi motivée par un accent théologique favorisant l'idée de permanence.[11]

[11] Sur Saint-Venant, ses convictions atomistiques et la conservation de l'énergie voir Darrigol, 2001, §5 et §6. Puisque dans le cadre de la physique laplacienne, d'autres savants comme Navier et Coriolis, ne pouvaient que trouver naturel le raisonnement mécanique de Saint- Venant, la préférence par l'idée de conservation et permanence, laquelle est équivalente à celle

L'ontologie atomistique était aussi justifiée par des arguments scientifiques à large portée et assez pertinents. Par exemple, Saint-Venant avait démontré en 1844 que l'existence de tensions de cisaillement est incompatible avec un continuum laplacien. Cela se comprend par un raisonnement de symétrie : supposons un solide infini et un plan qui le divise ; dans le cas du continuum laplacien, pour lequel les forces entre ces points sont des forces centrales agissant par paires et ne dépendant géométriquement que des distances, une translation de la moitié supérieure ne produirait pas de contrainte de cisaillement puisque géométriquement la situation resterait telle quelle ; donc un tel matériau serait dépourvu de rigidité. Cette ontologie s'inscrit aussi dans la conviction, très partagée pendant presque tout le XIX^e siècle, de que tous les phénomènes peuvent être expliqués en partant de la mécanique (réductionnisme mécaniste). Saint-Venant, le maître de Boussinesq, a bâti une mécanique dont les postulats affirment que toute la matière est constituée par des points matériels. Darrigol, dans son ouvrage le plus récent, affirme que les théories mécaniques les plus pures sont celle des systèmes à liaisons et celle des systèmes de particules dont l'interaction n'est qu'une fonction de la distance, ce qui souligne la rationalité de l'option de Saint-Venant, Helmholtz et de Boussinesq.[12]

---

d'échange entre les deux formes d'énergie, ne se justifie que par un penchant métaphysique, lequel est introduit par des arguments théologiques, Dieu étant le créateur et conservateur de toutes les choses ; voir Darrigol, 2001, 330-335.

[12] Voir Darrigol, 2001, §4, pp. 125-126. Cette démonstration a été publiée dans le *Bulletin de la Société Philomatique* le 20 janvier 1844, dans le *Mémoire sur la question de savoir s'il existe des masses continues, et sur la nature probable des dernières particules des corps*, voir Saint-Venant, 1876 ; sur ses contributions pour les fondements de la mécanique, dans la lignée du programme laplacien, voir Jammer, 1957, 132, 215-216 et Darrigol 2014, 40-42 et 58. Sur la rationalité de cette mécanique voir Darrigol, 2014, 58-62 et 75. Sur le rôle central de la mécanique voir Harman, 1982, chap. I.

Vers 1870, les traditions laplacienne et ampèrienne avaient porté de nombreux fruits dans les domaines de l'élasticité, de l'hydrodynamique, de la théorie de l'éther élastique, etc. Elles permettaient une cosmovision assurant la cohérence entre les différentes théories, la mécanique céleste ayant un rôle d'archétype ; elles bénéficiaient de l'autorité intellectuelle de grands maîtres comme Newton, Laplace, Fresnel, Ampère, etc. ; et elles avaient une traduction institutionnelle renforcée par le caractère centralisé et hiérarchisé de la communauté savante.[13] Ces traditions étaient contemporaines d'une attitude d'agnosticisme théorique, issue notamment des travaux expérimentaux et des conceptions de Regnault, lequel ne niait cependant pas l'ontologie moléculaire. Le poids de ces traditions est détectable, par exemple, dans le débat sur l'atomisme à l'Académie des Sciences (1876-1877) ; suscité par le résultat des expériences de Kundt et Warburg sur les chaleurs spécifiques de la vapeur de mercure, en accord avec les calculs de la théorie cinétique pour un gaz monoatomique, $\gamma = 5/3 = 1.66$; l'académicien Yvon Villarceau revendique sa priorité sur le résultat théorique et le chimiste Marcellin Berthelot, répondant à Villarceau, doute du résultat expérimental et affirme : 'La notion même d'atome indivisible, et cependant étendu et continu, *aussi bien que celle d'un atome doué de masse et cependant réduit à un point matériel*, semble contradictoire en soi'. Villarceau répond à Berthelot que son attitude positiviste, son refus d'accepter des entités inobservables, ne peut servir qu'à 'retarder indéfiniment les progrès de la science moderne'. Saint-Venant, dans sa réponse, note que la notion d'atome centre-de-force a un long passé qu'il évoque, en citant les noms de Boscovich, Kant, Navier, Poisson et d'autres. L'atome est un être de raison et les répugnances instinctives qu'il peut susciter ont leur origine dans l'imagination humaine,

[13] Sur les caractéristiques générales de la physique française dans cette période voir Príncipe, 2008, 1ère Partie.

laquelle 'sous l'empire variable des sens, réclamera contre ces êtres sans étendue, comme elle a réclamé si longtemps contre les antipodes, le mouvement de la Terre, la pesanteur de l'air, etc.'. Or, à la lumière de la raison, l'idée d'existence n'implique pas nécessairement l'idée d'étendue. Un point matériel obéissant aux lois de la Mécanique est un être concevable et traitable mathématiquement (et non un esprit). Saint-Venant rappelle au lecteur sa démonstration du fait que la conception de la matière comme une étendue continue implique l'absence de solidité.[14]

Il faut signaler, pour être historiquement correct, que si la théorie ondulatoire (ou vibratoire) de la chaleur ait été en faveur pendant les deux derniers tiers du siècle en France cela n'implique pas que beaucoup de physiciens y aient vraiment travaillé. La physique théorique était peu pratiquée (en France comme ailleurs) et la théorie mécanique de la chaleur n'était pas le domaine d'élection pour les théoriciens français, contrairement à la mécanique des milieux continus (incluant l'éther lumineux). Vers 1875, la tradition laplacienne de physique moléculaire, ou mécanique physique à laquelle se rattache la théorie vibratoire de la chaleur, est représentée par des chercheurs illustres comme Saint-Venant ou Boussinesq. Dans l'Académie des Sciences, ils sont surtout cantonnés dans la section de mécanique. Ces savants se sentent menacés et se plaignent du manque de soutien institutionnel. La mécanique physique est sur le déclin dans la seconde moitié du siècle même si elle reste active.[15]

Vers 1885, la version ampérienne du programme de la physique moléculaire était encore vivante. Elle garantissait encore une vision

[14] Voir : Villarceau, 1876, 1128 et 1178 ; Berthelot, 1876, 1129 ; Saint-Venant, 1876, 1224-1225 ; Príncipe, 2008, § 7.1.2.

[15] Ce déclin explique en partie le peu d'intérêt des Français pour la théorie cinétique des gaz. En même temps (et paradoxalement), il est vrai que le maintien de cette tradition sous une forme toujours ampérienne est une autre cause de ce manque d'intérêt, Príncipe, 2008, Conclusions, 365.

unificatrice des phénomènes, une cosmovision. Donnons un exemple assez significatif de cet état des choses. Jules Violle (1841-1923) est le disciple d'Émile Verdet (1824-1866) qui s'est le plus distingué par ses travaux sur la chaleur (Verdet était le professeur de L'École Normale Supérieure qui connaissait mieux la physique britannique et allemande de son temps). Verdet et Violle appartenaient à la tradition opticienne et ils avaient un intérêt réel pour la théorie cinétique. Les deux appartiennent à l'élite de la physique française. Dans la préface du traité de physique moléculaire de Violle, publié en 1884, on trouve l'essence de cet engagement thématique commun à beaucoup de physiciens français dans la seconde moitié du XIX[e] siècle :

> Rien dans l'immensité de l'univers n'est immobile. Déplacement de mondes, frémissement d'atomes, tout est agitation et mouvement. La science de la nature tend vers la mécanique par une évolution nécessaire, le physicien ne pouvant établir de théories solides que sur les lois du mouvement. Ces lois deviennent aussi la préface obligée d'un cours de physique. Les phénomènes de la pesanteur en fournissent immédiatement une application simple et grandiose. A l'examen des propriétés générales des corps succède l'étude des mouvements vibratoires qui constituent à proprement parler le véritable champ de la physique. Par un enchaînement rationnel, l'acoustique où ces mouvements sont nettement perceptibles prépare l'optique qui conduit elle-même à la chaleur. En dernier lieu viendra l'électricité, que les connaissances acquises jusqu'à ce jour ne permettent pas de rattacher directement à ces conceptions, bien qu'il soit possible dès maintenant d'y appliquer sans conteste le grand principe de la conservation de l'énergie.[16]

La mécanique est donc le pilier de la physique ; la majesté de la mécanique céleste démontre sa supériorité ; la physique moléculaire a comme théorie fondamentale et unifiante la théorie vibratoire, laquelle permet de comprendre le son, la lumière et la chaleur. Pour Violle, le

[16] Violle, 1884, V-VI. Briot, 1869 est un des traités, écrit par un des spécialistes français de l'éther mécanique lumineux, qui met l'accent sur la rationalité de l'ontologie des atomes centres de force.

programme laplacien, dans sa version ampèrienne, reste au centre de la structure conceptuelle de la physique.

3. Boussinesq comme théoricien

Boussinesq, né en 1842, a eu un parcours atypique pour un futur académicien. Après des études secondaires au petit séminaire de Montpellier, il obtient sa licence ès sciences (1861) et son doctorat (1867) à la Faculté des sciences de Montpellier. Son premier mémoire, présenté à l'Académie des Sciences en 1865, concerne la théorie de la capillarité et sa thèse doctorale concerne la théorie analytique de la chaleur – *Propagation de la Chaleur (Ellipsoïde des conductibilités linéaires)*. Cette thèse et ses travaux de physique mathématique, publiés avant ses trente ans, ont attiré l'attention de l'académicien Adhémar Barré de Saint-Venant qui est devenu son maître et protecteur. Saint-Venant était un polytechnicien, tardivement devenu membre de la section de Mécanique de l'Institut. Il a joué un rôle fondamental dans l'histoire de la loi de la conservation de l'énergie, dans la théorie de l'élasticité et dans l'hydrodynamique. En 1873, après avoir enseigné les mathématiques dans des collèges en province, Boussinesq est nommé chargé de cours de calcul différentiel et intégral à la faculté des sciences de Lille ; l'année suivante, il sera nommé professeur à cette faculté. Après le décès de Saint-Venant il occupe, en 1886, sa place à la section de Mécanique de l'Académie des Sciences et, cette même année, il devient professeur de mécanique physique et expérimentale à la faculté des sciences de Paris ; dix ans après il y occupera la chaire de physique mathématique et de calcul des probabilités.

Par la suite j'analyserai deux des contributions théoriques de Boussinesq, dans lesquelles les hypothèses atomiques jouent un rôle

central. La première concerne le modèle de l'éther lumineux et est considéré une importante contribution ; la seconde est une nouvelle théorie vibratoire des gaz laquelle est en harmonie avec celle sur l'éther et qui n'arrive pas à dépasser celle de Laplace (basée sur l'hypothèse du calorique), ne permettant donc qu'une déduction de la loi des gaz parfaits.

Pour Saint-Venant, dans une bonne théorie, la simplicité doit aller de pair avec des hypothèses de départ naturelles et une mise en oeuvre analytique qui s'exécute sans longs calculs de manière à se substituer dans l'enseignement 'aux théories incomplètement mathématiques, dont on a été forcé de se contenter jusqu'ici'. Tel est pour lui le cas de la théorie des ondes lumineuses que Boussinesq publie en 1868.[17] Une des originalités de Boussinesq consiste à s'écarter du débat qui divisait en deux partis ceux qui s'occupaient à l'époque des propriétés de l'éther lumineux dans les milieux matériels. Pour un des groupes, les différences dans les propriétés optiques étaient dues à des variations dans l'inertie de l'éther ; l'autre groupe les attribuait à des variations de ces propriétés élastiques. Boussinesq assume que l'éther a partout la même inertie et la même rigidité. Pour lui, c'est à l'interaction entre l'éther et la matière que sont dues les différences dans les propriétés optiques. Boussinesq fait participer les molécules pondérables aux vibrations de l'éther. Il établit une équation aux dérivées partielles qui met en rapport les déplacements locaux de la matière pondérable et de l'éther, dans laquelle figurent les densités locales de la matière et de l'éther et les deux coefficients élastiques. Boussinesq ne déduit pas l'équation en partant de calculs sur les actions moléculaires ; son point de départ 'sont les équations différentielles de Lamé, à deux coefficients μ et λ, sans considérer ce qu'ils peuvent représenter d'après les forces individuelles en jeu'. Cela signifie qu'il n'a pas besoin de prendre parti

[17] Saint-Venant, 1872, 35 ; Boussinesq, 1868.

sur la question controversée du nombre de constantes élastiques indépendantes dans un solide isotrope et homogène. Mais la perspective moléculaire apparaît explicitement pour rendre intelligible l'interaction matière-éther ; pour que la nouvelle équation soit traitable, il faut assumer un rapport entre les deux déplacements ; Boussinesq admet que les molécules pondérables et les molécules de l'éther vibrent :

> Constamment en concordance comme fait un liquide à l'égard de petits corps qui y flottent ; en sorte que, dès que les molécules éthérées reviennent à leurs positions de repos, les molécules pondérables y reviennent sensiblement aussi, leurs actions mutuelles n'ayant été mises en jeu que d'une manière négligeable vu la petitesse relative de leurs déplacements...

Dans cet extrait, on voit l'importance de pouvoir concevoir les mouvements microscopiques d'après des mouvements simples à notre échelle. La simplicité de l'hypothèse d'un éther unique et homogène, l'usage d'approximations linéaires (par exemple dans le rapport entre le déplacement de la matière et celui de l'éther) et le recourt à une image simple de mouvements moléculaires couplés sont des caractères de rationalité de la théorie prisés par Saint-Venant et par Boussinesq. Mais cette théorie restait incomplète puisqu'elle n'expliquait pas les deux constantes élastiques fondamentales et ignorait le problème des ondes longitudinales.[18]

[18] Voir Saint-Venant, 1872, 36-39 ; citation de la page 39. Whittaker note que la théorie de l'éther de Boussinesq est : 'In many respects preferable to its predecessors.... It is Boussinesq's merit to have clearly asserted that all space, both within and without ponderable bodies, is occupied by one identical ether, the same everywhere both in inertia and elasticity; and that all aetherial processes are to be represented by two kinds of equations, of which one kind expresses the invariable equations of motion of the aether, while the other kind expresses the interaction between aether and matter. Many years afterwards these ideas were revived in connection with the electromagnetic theory, in the modern forms of which they are indeed of fundamental importance', Whittaker, 1951, 169. Voir aussi Príncipe, 2008, 136-138.

L'interaction nécessaire entre matière et éther a sûrement favorisé l'incompréhension de Boussinesq vis-à-vis la nouvelle théorie cinétique des gaz, dont la formulation par Clausius (comme Boussinesq le note, l'idée remonte à Daniel Bernoulli) a été la seule connue par lui vers 1870. Dans son mémoire sur la théorie des gaz, de 1873, il propose une théorie alternative. Cette théorie, vouée à l'échec (car en repoussant l'idée d'un mouvement chaotique des molécules, comme dans les théories de Maxwell et de Boltzmann, les phénomènes thermiques ne peuvent pas s'harmoniser avec la mécanique des systèmes macroscopiques, en partant d'un point de vue réductionniste, moléculaire) semble être motivée par le besoin de complétude du cadre conceptuel dans lequel il se meut en tant que théoricien – celui de la mécanique physique, qui a dans la théorie de l'éther élastique et de l'hydrodynamique deux domaines d'élection, des théories où les molécules occupent des positions moyennes fixes. Boussinesq ne s'est jamais intéressé aux formulations sophistiquées de la théorie cinétique de Maxwell et de Boltzmann. Il serait injuste et simpliste d'attribuer cette attitude à une tendance isolationniste puisqu'en hydrodynamique il cite souvent ses collègues anglais. Boussinesq déduit à nouveau la loi de Mariotte, sans mentionner les diverses déductions antérieures, notamment celle de Laplace. Le principe des forces vives joue un rôle central dans le mémoire de Boussinesq ; d'après cette conception le monde serait constitué par un système de points matériels qui échangent des forces conservatives ; cette image atomique, naturelle dans le cadre laplacien, permet l'intelligibilité mécanique de la conservation de l'énergie. Voici sa formulation : Soient *n* points matériels de distances mutuelles, $r_{12}, r_{13}, ..., r_{pq}, ...$; Boussinesq écrit :

$$\frac{1}{2}\Sigma m V^2 = -\Psi\left(r_{12}, r_{13}, ..., r_{pq,...}\right) + C\,,$$

ce qui correspond à la conservation de la somme des énergies cinétique et potentielle ; Boussinesq décompose l'énergie potentielle, Ψ, en deux parties. La première, l'*énergie potentielle des actions newtoniennes*, se réfère à l'attraction newtonienne s'exerçant à des distances perceptibles. La seconde, l'*énergie potentielle interne,* comprend 'les actions beaucoup plus énergiques qui ne s'exercent qu'à des distances insensibles'.[19]

Pour expliquer la fluidité il admet des actions de présence entre molécules pondérables: l'action mutuelle entre deux points matériels dépend non seulement de leur distance mutuelle mais aussi de l'influence d'un certain nombre d'autres points matériels voisins. Ces actions avaient été proposées par des savants britanniques, notamment par Green dans sa théorie de l'éther élastique. Boussinesq décompose l'action entre deux molécules en deux parties. La première explique la solidité et est une fonction de leur distance primitive et de leur écartement (l'état primitif étant d'équilibre stable) ; elle disparaît quand les limites d'élasticité sont dépassées et elle ne peut dépendre que de la seule distance entre les deux molécules, comme dans la théorie de Navier de l'élasticité. La seconde dépend des distances actuelles et aussi des densités actuelles et produit une pression normale à tout élément plan mené dans le corps; cette pression normale devra être une fonction de la densité et de la température.

Boussinesq commence par distinguer les phénomènes lumineux des calorifiques. La propagation de la lumière à travers les corps transparents se fait par ondes régulières – les molécules vibrent à l'unisson avec les atomes qui composent l'éther, les molécules se

[19] Boussinesq, 1873, 314. Cette formulation de la conservation, n'étant pas une nouveauté (bien que son statut de vérité expérimentale n'a pas été établi aussi vite que l'on pense – la thèse doctorale de Jules Violle, de 1870, *Sur l'équivalent mécanique de la chaleur*, concerne encore la détermination expérimentale de cet équivalent) ressemble aux formulations de Saint-Venant (vers 1830) et de Helmholtz (1847), voir Darrigol, 2001, §6.

déplaçant très peu de leurs positions d'équilibre ; les actions réciproques entre les molécules peuvent être négligées et la force vive de l'éther est incomparablement supérieure à la force vive de la matière. Dans ce cas, l'état vibratoire imprimé aux molécules pondérables par les ondulations de l'éther n'est pas concordant avec les états vibratoires que les actions réciproques entre molécules 'tendraient à produire ou du moins à maintenir'.[20] Quand la lumière traverse les corps opaques, il y a concordance entre l'état vibratoire de l'éther et les états vibratoires propres des actions réciproques entre molécules. Alors les molécules exécutent d'assez larges oscillations et leur demi-force vive représente la chaleur sensible d'un corps. Dans ce cas, la force vive de l'éther et ses éventuelles actions de présence sur l'énergie propre de la matière pondérable sont négligeables. Il y a une différence fondamentale entre les vibrations lumineuses et les vibrations calorifiques : le caractère erratique des phases de ces dernières vibrations : 'La phase de ce mouvement change extrêmement vite quand on passe d'un point aux points voisins, et j'admettrai qu'elle reçoit à chaque instant ou presque à chaque instant les valeurs les plus différentes dans l'étendue de la sphère d'activité [contenant toujours un grand nombre de molécules voisines] d'une molécule'. Quand un corps transparent est traversé par des ondes d'éther, ce corps transparent a déjà une certaine température, c'est-à-dire un certain état de vibration calorifique, et Boussinesq admet que 'en vertu du principe de la superposition des petits mouvements, les ondes lumineuses…doivent, à fort peu près, se propager à son intérieur comme si cette agitation n'existait pas'.[21]

Comme le calorique chez Laplace et Fourier, l'éther de Boussinesq est responsable de la thermalisation de la matière:

[20] Boussinesq, 1873, 335.
[21] Boussinesq, 1873, 324, 342.

> Ce milieu ne peut pas, à cause de sa faible masse, en détenir des quantités sensibles [de l'énergie interne] et il ne fait que la transporter rapidement, par petites charges, d'un corps à l'autre du système, de manière à tendre à établir entre eux et avec lui-même une sorte d'équilibre dynamique appelé *équilibre des températures*.[22]

Cependant, Boussinesq reconnaît que 'le détail des mouvements calorifiques et leur mode de propagation d'une molécule à l'autre nous est inconnu'.[23]

Comme Laplace, Boussinesq définit température à partir des propriétés du milieu ténu où se propage le rayonnement calorifique. La différence, c'est qu'il la définit cinétiquement comme dans la théorie d'Ampère : 'On peut appeler *température absolue* d'un petit volume d'éther la demi-force vive qu'il possède sous l'unité de masse, ou une quantité proportionnelle à cette demi-force. Un corps est dit à une certaine température *T*, lorsque ces vibrations calorifiques n'augmentent ni diminuent si l'on vient à le placer dans l'éther à la même température'.[24]

Le mécanisme des chocs de la théorie cinétique de Maxwell-Boltzmann cède la place à la proportionnalité des mouvement de l'éther et des molécules pour le cas d'équilibre thermique, et l'hypothèse du chaos moléculaire a son analogue dans l'hypothèse de chaos dans les phases du mouvement vibratoire calorifique des atomes et molécules. Mais Boussinesq n'arrive pas à introduire un rapport avec le deuxième principe de la thermodynamique, avec l'augmentation de l'entropie dans les processus naturels.

---

[22] Boussinesq, 1873, 308.
[23] Boussinesq, 1873, 339.
[24] Boussinesq, 1873, 337. Sur la théorie de Laplace et le calorique voir : Sebastianini, 1981, Sebastianini, 1982, Arnold, 1983, Heilbron, 1993, Príncipe, 2008, § 4.1. La loi de Stefan-Boltzmann pour le rayonnement noir, dont la déduction (par Boltzmann en 1884) est basée sur la pression du rayonnement et sur des relations thermodynamiques, corrige le supposé rapport linéaire entre la densité d'énergie rayonnante et la température.

4. Les atomes, la méthode et les limitations de notre esprit

Boussinesq est habituellement cité, dans le contexte de l'histoire des idées philosophiques, à cause de ses réflexions sur les conséquences de la 'bifurcation des voies' : dans quelques conditions singulières les équations différentielles qui décrivent le mouvement d'un système mécanique admettent une multiplicité d'intégrales, en partant d'un même état initial ; alors le choix entre ces intégrales pourrait avoir lieu en ne dépensant presque aucun travail ; cela permettrait de concilier le mécanisme avec les phénomènes vitaux et avec la liberté humaine ; ces idées remontent à 1877-78 et Saint-Venant, dans une note à l'Académie des Sciences cite ces travaux de physique mathématique de Boussinesq.[25] Ici je m'en occupe d'autres de ses réflexions contemporaines, moins connues, concernant de plus près la question des atomes. Après avoir publié des mémoires remarquables sur la théorie de l'éther élastique, sur l'hydrodynamique et sur la théorie analytique de la chaleur, Boussinesq publie, en 1879, son ouvrage philosophique *Étude sur divers points de la philosophie des sciences*. Dans cet ouvrage les considérations sur les atomes s'encadrent dans les vues générales de l'auteur sur le problème de la connaissance.

Pour Boussinesq l'intelligence humaine n'est pas parfaite. Notre faculté de représentation, appuyée sur la création d'un ordre géométrique objective et universelle, laisse échapper le fond des choses, puisqu'elle constitue un monde de lois et de phénomènes ayant une nature approchée par rapport au vraie ordre physique ; par ses propres mots, notre esprit subit 'un défaut d'adaptation très léger,

[25] Voir Fagot-Largeault, 2002, 966-982. Le contexte international de ce débat est discuté dans Strien, 2014.

mais peut-être irrémédiable' ; cette pensée est développée dans le passage suivant :

> Les moindres faits observables contiennent tant de complications, il y a peut-être une disproportion si radicale entre leurs plus petits détails et notre intelligence, qu'il nous faut renoncer à saisir les choses absolument tels qu'elles sont. Notre esprit se représente et leur substitue, presque sans s'en douter, des objets abstraits, qui en diffèrent fort peu, et dont les notions sont plus simples, c'est-à-dire plus intelligibles ou mieux adaptées à sa forme propre... notre nature intellectuelle nous... fait remplacer les quantités ou les figures réelles existant hors de nous, et dont l'observation nous fournit seulement une connaissance plus ou moins grossière par les quantités abstraites de l'analyse ou par les figures idéales de la géométrie. Et pourtant, aucun moyen de constatation ne nous donne le droit de regarder celles-ci comme leur étant identiques, bien que les idées que nous en avons présentent seules assez de clarté pour servir de base à nos raisonnements... La réalité est d'ordinaire... si complexe, qu'il nous faut... tolérer dans une première étude des erreurs parfois très appréciables... C'est en procédant ainsi par voie d'approximations successives, dont chacune laissa subsister une erreur plus ou moins petite en comparaison avec des résultats poursuivis, que l'intelligence humaine parvient à démêler, dans les phénomènes naturels, des relations assez simples pour pouvoir les saisir et s'y intéresser.[26]

Nonobstant, c'est la possibilité de modéliser les phénomènes d'après cet ordre idéal qui les rend intéressants : 'Ces phénomènes... ne nous intéressent, ils ne nous sont même intelligibles, qu'à la condition de reproduire assez fidèlement les traits de certains objets simples de l'ordre géométrique idéal, vrai domaine immédiat de notre esprit'. Le rôle du sujet de la connaissance est donc fondamental dans notre connaissance :

> Il est vrai que cette existence du phénomène, en tant que nous lui faisons une place à part, motivée par l'intérêt qu'il nous offre, tient autant à notre manière de voir, à la forme de notre existence, ou du moins à l'analogie qu'il présente avec des réalités d'un ordre supérieur perçues par nous, qu'à son essence même : mais toute science porte inévitablement l'empreinte du *sujet* connaissant, et une loi n'est

[26] Boussineq, 1879, 64, 34-35.

> *naturelle* qu'à la condition de se trouver tout à la fois, dans la mesure du possible, conforme à la *nature* de notre esprit et à celle des choses.

Boussinesq admet que cet ordre géométrique constitue un monde à part 'en dehors de nous' (platonisme), ayant une valeur objective, mais parfois il le formule en tant que hypothèse.[27]

Boussinesq, dans la première étude de l'ouvrage, concernant 'Le rôle et la légitimité de l'intuition géométrique', considère les caractéristiques de cet ordre géométrique. Il y critique ceux qui veulent mettre dans le même plan toutes les géométries en jugeant vague la notion d'intuition géométrique. Cette intuition accompagne ceux qui s'occupent de la géométrie euclidienne, et qui, contrairement aux autres, donne à la notion de similitude une valeur absolue (théorème de Thalès). L'intuition géométrique s'appuie sur des images et donc sur le système visuel :

> Le cerveau *pensant* tout entier paraît n'être, à quelques égards, qu'une extension du système visuel, qui est par excellence l'organe de la représentation et des figures. Nous condensons et précisons toutes nos idées par des formes, des constructions idéales, sans lesquelles nous ne parviendrons pas à les fixer, à les voir nettement; et on dirait, que c'est, précisément, dans la mesure où leur assimilation à des images réussit que nous pouvons en faire l'objet de connaissances positives.

Si pour le cas de la géométrie le rapport aux images semble évident, Boussinesq nous souvient que pour l'analyse on travaille avec la correspondance entre les nombres réels et une ligne droite. Aussi, pour l'analyse combinatoire les métaphores empruntées aux choses matérielles sont présentes : disposition, arrangement, substitution, permutation.[28]

---

[27] Boussineq, 1879, 36-37. Voici un passage où le ton est hypothétique : 'à supposer même que l'ordre géométrique n'eut pas d'existence hors de nous, qu'il se réduisit à être une pure émanation de l'intelligence humaine, ou mieux, une simplification et une généralisation abstraites, appropriés à notre esprit', Boussineq, 1879, 49.

[28] Boussinesq, 1879, 17-18. La position esquissée par Boussinesq diffère de celle de Louis Liard, dont la thèse, *Des définitions géométriques et des*

Cependant, pour Boussinesq, l'intuition géométrique, laquelle nous procure des totalités saisies avec clarté maximale, bien qu'ayant ses racines dans un sentiment instinctif, appartient à un ordre spécial dont l'essence est idéal. La notion d'espace, étant logiquement antérieure à celle des figures est 'la plus essentielle qui nous fournisse le sens géométrique'. Or, l'espace est irréductible aux catégories de la substance (dont le paradigme est la notion de matière) et du mode (attribut d'un être réel, c'est-à-dire matériel) ; ces deux catégories sont 'crées sous la prédominance des sens externes' et surgissent donc associées aux êtres considérés (physiquement) réels ; en plus ces deux catégories émergent des régions obscures de notre esprit et la région claire ne les reconnaît pas. Cauchy avait déjà jugé l'espace comme infini et dépourvu d'existence, dans une argumentation qui recèle des motivations théologiques (les deux sont des croyants).[29]

Les savants construisent donc des analogies entre les deux ordres ou mondes – physique ou réel, géométrique ou idéal. Mais la vraie et absolue ressemblance nous échappe, étant le privilège d'une intelligence supérieure qui aurait une connaissance 'également parfaite, également directe des deux ordres de réalité' ; cette intelligence ne serait pas condamnée 'à ne comprendre les choses physiques qu'à la lumière des conceptions géométriques abstraites' ; ces choses physiques, les 'phénomènes bruts' sont régis par des lois inconnues et sans doute inextricables. Nous sommes condamnés à l'emploi d'hypothèses approximatives, accommodées à la faiblesse de l'esprit humain'. Même pour la notion d'espace, Boussinesq juge 'qu'on doit admettre l'existence de différences très petites entre

---

définitions empiriques, cite les travails de Beltrami, traduits par J. Hoüel en 1869, en faveur de l'idée que 'la géométrie non euclidienne, géométriquement interprétée, est une extension de la géométrie euclidienne', Liard, 1873, 41. Liard (1846-1917), est un disciple du philosophe néo-kantien Jules Lachelier et deviendra très influent comme directeur de l'enseignement supérieur français.

[29] Boussinesq, 1879, 14-15, 19-20 ; voir Cauchy, 1868.

l'espace idéal…et l'espace réel où sont les corps, quoique nous ne puissions fixer ces différences.' Comme Cauchy, il note comme le problème de la divisibilité indéfinie des corps et celui de l'étendue ou de l'inétendue des atomes, montre l'existence d'une "incommensurabilité… subjective ou objective… dans cette question du passage de l'abstrait au concret'.[30]

Boussinesq en décrivant notre faculté de représentation parle des régions clairs de notre esprit et des régions obscures qui semblent être le domaine du sens pratique ou du bon sens, domaine qui nous fournit 'une certaine vue d'une matière continue' dont les plus petites parties doivent être massifs et indivisibles ; ce bon sens qui est commun à tout le monde, y inclus les savants-géomètres, repousse l'idée de la divisibilité à l'infini et donc l'idée d'une matière constituée par 'des petites fractions imaginables…indéfiniment divisibles', comme est le cas des petits atomes étendus des anciens. Ce bon sens se refuse aussi à admettre que la matière soit composée par des atomes ponctuels et massifs, puisque dans ce cas on fait du 'point géométrique qui, pour lui, n'est qu'une abstraction irréalisable ou dépourvue du moins de toute probabilité, la seule réalité matérielle'. Nonobstant, Boussinesq juge que cette conception de l'atome inétendu et centre de force, tel qu'elle a été conçue par le père Boscovich, Ampère, Cauchy ou Saint-Venant, est indispensable aux géomètres, étant la plus simple et la seule qui permet d'expliquer les phénomènes mécaniques – elle est en harmonie avec les régions claires de l'esprit et en 'accord suffisant avec la réalité'. Cette conception suppose que l'espace géométrique ne diffère pas de celui où sont les corps matériels. Puisque notre faculté de représentation est aussi constitué par le sens pratique, Boussinesq croit que l'incompatibilité entre celui-ci et l'hypothèse des atomes inétendus, montre que 'même dans les catégories de la forme et de la quantité, la

[30] Boussinesq, 1879, 37, 39, 42, 22.

nature se dérobe en partie à tous les efforts de notre faculté de représentation.' L'ordre géométrique correspond donc aux 'régions claires de notre esprit'. Notre esprit cherche toujours la simplicité et la perfection absolue en utilisant sa faculté d'intuition idéale ou pouvoir de généralisation et sa capacité de déduction (logique). Boussinesq illustre ces considérations avec l'exemple des méthodes de la mécanique physique.[31]

Dans la mécanique physique, on considère les 'déplacements relatifs éprouvés par des parties de matière très voisines'. Ses déplacements sont dus, en général, à des actions de contact, exercées à des distances imperceptibles. Le savant géomètre ne prétend pas évaluer ce qui se passe avec chaque molécule, ce qui serait trop complexe. Le géomètre admet que les irrégularités complexes des mouvements individuels des molécules se neutralisent mutuellement, amenant à la constitution d'un *état moyen local*, lequel varie avec continuité. Le principe de continuité est indissociable de l'utilisation de l'analyse mathématique. La mécanique physique cherche les lois régissant l'évolution de cet état moyen local. La simplicité des lois de la mécanique physique est liée au fait 'que nous observons des effets moyens où les discordances individuelles disparaissent et se neutralisent en vertu de la *loi des grands nombres'*. Boussinesq admet que, au niveau en dessous de celui des moyennes locales, 's'il nous était donné de voir les détails, nous serions tentés, à cause des bornes actuelles de notre esprit, de ne trouver que désordre et incohérence dans le monde des infiniment petits'.[32]

Les textes philosophiques de Boussinesq nous renseignent donc sur le rôle paradoxal des atomes ou molécules dans les théories de la mécanique physique. Les atomes sont postulés comme des centres de

[31] Boussinesq, 1879, 64, 62.
[32] Boussinesq, 1879, 62, 28, 29 et 33.

force, d'après une tradition qui remonte à Boscovich. À ce propos, en 1889, Boussinesq écrit :

> Le géomètre sent la nécessité de bien se représenter et définir les objets dont il s'occupera. Or il ne parvient à la netteté désirable, dans cette *construction* idéale des phénomènes et dans leur expression analytique, qu'en regardant chaque corps comme un ensemble d'*atomes* sans étendue ni dimensions, dits *points matériels*, dont il voit chacun occuper à tout instant, dans l'espace, une *situation déterminée*… Ce qui oblige à arriver ainsi jusqu'à des éléments sans étendue, jusqu'à des *points,* c'est que, justement les points sont les seuls êtres géométriques dont la situation soit précise… la *distance*, qui est par excellence, dans les phénomènes, la chose mesurable, n'a de sens que lorsqu'elle relie deux points. C'est on le voit, une nécessité logique inhérente à la nature même de l'esprit humain, qui nous fait *idéaliser* de la sorte les éléments matériels… plusieurs géomètres, le P. Boscovich, Ampère, Cauchy, de Saint-Venant, ont-ils regardé cette conception d'atomes sans étendue comme réelle, comme parfaitement conforme à la véritable structure de la matière.

Si Boussinesq évoque des exigences a priori du mode de représentation des phénomènes, position vaguement kantienne, il souligne aussi que la position réaliste est tenable car 'les procédés d'observation n'ont jamais reconnu d'erreur dans les conséquences résultant de l'application de nos idées géométriques aux choses'.[33]

[33] Boussinesq, 1889, 5-6. Il est difficile d'inscrire le croyant Boussinesq dans une courante philosophique contemporaine ou d'en détecter des inspirations dominantes (sauf pour ce qui concerne son maître Saint-Venant), bien que la question de la conciliation du déterminisme et de la liberté, pour laquelle il a donné un important apport, soit un thème majeur du spiritualisme français, fort influencé par Kant. Il est probable qu'il ait eu connaissance des ouvrages de Cournot (l'*Essai sur le fondement de nos connaissances* est de 1851) et de Renouvier (ses *Essais de critique générale* sont parus en 1854 et 1859). Le phénoménisme de Renouvier refuse l'idée de chose en soi et souligne le rôle des abstractions : 'Les sciences physiques font usage de l'abstraction aussi nécessairement que les sciences logiques… [Elles] séparent les faits selon l'expérience, et isolent parmi les phénomènes… ceux dont il faut obtenir une définition plus précise ou rechercher des rapports particuliers… La considération distincte, dans un corps donné, du poids, de la chaleur… nécessite une suite d'abstractions physiques', Renouvier 1859, 505 ; d'ailleurs Renouvier s'intéressait aux sciences physiques et considérait l'hypothèse de l'éther une partie importante 'd'une théorie générale…de la constitution et des

Boussinesq profite de l'occasion pour expliciter comment cette conception générale permet d'expliquer les propriétés chimiques (existence de molécules diverses, énergie dégagée dans les réactions) et diverses propriétés physiques. La différence entre phénomènes chimiques et physiques, s'explique en supposant que les forces intermoléculaires varient plusieurs fois de signe à des distances insensibles. L'impénétrabilité s'explique en postulant que les actions moléculaires doivent être répulsives et indéfiniment croissantes aux plus petites distances. L'existence de molécules chimiques oblige à supposer que, à des distances de l'ordre de grandeur de leurs dimensions, cette action devient attractive. La conservation de l'énergie dans les transformations d'un système clos se comprend en admettant que les actions moléculaires et atomiques dérivent d'un potentiel, car la dynamique moléculaire obéit alors au principe des forces vives ou de conservation de l'énergie. Ce cadre général cohérent est très attractif, mais, dans les théories concrètes de la mécanique physique, le détail du monde des infiniment petits reste caché :

---

mouvements moléculaires des corps' dont le caractère hypothétique était mis en relief, Renouvier, 1859, 514 ; sur Renouvier et les sciences physiques voir Schmaus, 2007. Boussinesq ne partage pas du relativisme absolu de Renouvier. Il y a des résonnances platoniciennes dans quelques passages de Boussinesq ; il est probable que Boussinesq ait eu connaissance des écrits du dernier Maine de Biran (Alfred Fouillée est un autre exemple de l'inspiration platonicienne chez les français), ce qui aurait pu favoriser ce penchant.

Les réflexions de Boussinesq n'ont pas eu, à ma connaissance, d'écho évident sur les philosophes français contemporains s'intéressant pour l'atomisme ; par exemple l'ouvrage d'Arthur Hannequin, parue en 1895, reste encore trop attachée à une distinction entre cinétisme et dynamisme (comme chez Stallo, dont l'ouvrage de 1881, *Concepts and théories on modern physics*, en est une des principales sources d'inspiration pour Hannequin) ; le concept d'atome centre de force est intenable pour lui (critique ancienne et dont Boussinesq parle) et son argumentation reste essentiellement qualitative ; voir Príncipe, 2008, § 8.2. La position conservatrice des secteurs philosophiques de l'époque, et la méconnaissance des détails techniques, se trouve aussi dans la méfiance vis-à-vis les géométries non-euclidiennes avant 1890, voir Torretti, 1978, § 4.2.

> Les actions, dites *de pesanteur*, s'exerçant aux distances de grandeur notable, peuvent être regardées comme infiniment petites en comparaison de celles, appelées le plus souvent *actions moléculaires*, qui existent entre points matériels, à ces distances absolument insensibles auxquelles se produit le *contact physique* des corps, et dont même la plus grande, dite *rayon d'activité des actions moléculaires*, échappe complètement à nos mesures.[34]

5. Épilogue

Ce monde des infiniment petits correspond à l'échelle où la mécanique physique postule la constitution atomique. Pour Boussinesq la conception de la matière constituée par des atomes centres de force est donc une exigence a priori de la partie claire de notre esprit, c'est-à-dire de celle que rejette les catégories de mode et de substance et qui s'appuie sur l'intuition géométrique. Le sens géométrique a un rapport avec les sens, étant dans une certaine mesure une extension du système visuel, mais il appartient à un ordre idéal dont la perfection est la seule perfection à laquelle on peut accéder puisque l'harmonie avec l'ordre physique et le sens pratique n'est que partielle, ce qui est du à un défaut d'adaptation léger mais irrémédiable, qui nous empêcherait de voir l'ordre dans le chaos dans le monde des infiniment petits pour lequel les bonnes images de ces mouvements nous manquent. Ce monde nous a été dévoilé, dans son comportement chaotique, par Maxwell et par Boltzmann lesquels ont su utiliser des raisonnements statistiques mettant en rapport l'échelle moléculaire et l'échelle macroscopique de la thermodynamique. La crainte du désordre, et l'habitus (Bourdieu) théorique de Boussinesq l'ont empêché de s'intéresser aux travaux de Maxwell et de Boltzmann, malgré la reconnaissance de l'insuffisance des méthodes utilisés (évidente pour le cas des gaz), attribués à l'incapacité de notre

[34] Boussinesq, 1889, 28.

esprit. Nonobstant, Boussinesq a laissé une œuvre notable dans la théorie des milieux continus et son adhésion à l'ontologie atomiste lui a permis d'avoir une cosmovision du monde physique, ancré sur la mécanique, exigence de l'esprit encore valorisé dans les écrits de Pierre Duhem, et caractérisant une tradition française. Sa valorisation du rôle de l'intuition, de la possibilité d'établir des images ou de rapprocher les phénomènes à des représentations familières, fait songer à quelques des considérations méthodologiques de Maxwell, mais, contrairement à l'écossais, Boussinesq a restreint son cadre de 'themata' en méprisant l'engagement pluraliste, favorisant la prolifération d'approches, aspect auquel Poincaré adhérera.[35]

[35] Sur l'engagement pluraliste de Maxwell et de Poincaré voir Príncipe, 2010 et Príncipe, 2012.

Bibliographie

Ampère, A.-M., 1835, Note de M. Ampère sur la Chaleur et sur la Lumière considérées comme résultant de mouvements vibratoires. *Annales de chimie et de physique*, 58, 432-444.

Andler, D. et Fagot-Largeault, A. et Saint-Sernin, B., 2002 *Philosophie des sciences* II, Collection Folio/Essais, Paris, Gallimard.

Arnold, D. H., 1983, The Mécanique Physique of Siméon Denis Poisson. The evolution and isolation in France of his approach to physical theory (1800-1840), *Archive for the history of the exact sciences*, part I: 28, 243-367; part II: 29, 37-94; part III: 29, 287-307.

Belhoste, B., 1992, *Augustin-Louis Cauchy: A biography*, New York, Springer Verlag.

Berthelot, M., 1876, Remarques sur l'existence réelle d'une matière monoatomique, à la suite d'une communication de M. Villarceau, Académie des Sciences. *Compte rendus hebdomadaires des séances*, 82, 1129-1130.

Boussinesq, J., 1867, Note sur l'action réciproque de deux molécules, Académie des Sciences. *Compte rendus hebdomadaires des séances*, 65, 44-46.

-- 1868, Théorie nouvelle des ondes lumineuses, Etude sur les vibrations rectilignes et sur la diffraction dans les milieux isotropes et dans l'éther des cristaux, *Journal de mathématiques pures et appliquées* (*Journal de Liouville*), 13 (2), 313, 425.

-- 1873, Recherches sur les principes de la Mécanique, sur la constitution moléculaire des corps sur une nouvelle théorie des gaz parfaits, *Journal de mathématiques pures et appliquées* (*Journal de Liouville*), 18, 305-360. Mémoire présenté à l'Académie des Sciences et des Lettres de Montpellier, le 8 juillet 1872.

-- 1879, *Etude sur divers points de la philosophie des sciences*, Paris, Gauthier-Villars.

-- 1889, *Leçons synthétiques de mécanique générale*, Paris, Gauthier-Villars.

Briot, C., 1869, *Théorie mécanique de la chaleur*, publiée par M. E. Mascart, professeur au Collège de France, Paris, Gauthier-Villars (1883 – 2ème édition).

Buchwald, J. Z., 1985, *From Maxwell to microphysics – Aspects of electromagnetic theory in the last quarter of the nineteenth century*, Chicago, University of Chicago Press.

Chalmers, A., 2006, *The scientist's atom and the philosopher's stone,* Boston studies in the philosophy of science, vol. 279, New York, Springer.

Cauchy, A., 1868, *Sept leçons de physique générale avec appendices sur l'impossibilité du nombre actuellement infini par M. L'Abbé Moigno*, Paris, Les Mondes et chez Gauthier-Villars.

Clausius, R., 1857, Conférences libres de Zurich. Physique. Cours de M. R. Clausius (correspondant de l'Institut). De la nature de la chaleur comparée à la lumière et au son, trad. fr. dans *Revue Scientifique*, 20 Janvier 1866, 121-131.

-- 1869, *Théorie mécanique de la chaleur, deuxième partie*, traduit de l'allemand par F. Folie, docteur ès sciences, professeur à l'Ecole industrielle de Liège, Librairie scientifique industrielle et agricole, Paris, Eugène Lacroix éditeur.

Dalmedico, A., 1989, La notion de pression: de la métaphysique aux diverses mathématisations [Causalité et statut des hypothèses], *Revue d'histoire des sciences*, 42 (1-2), 79-108.

Darrigol, O., 2001, God, waterwheels, and molecules: Saint-Venant's anticipation of energy conservation equation, *Historical studies in the physical sciences*, 31, 285-353.

-- 2014, *Physics and necessity: rationalist pursuits from the Cartesian past to the quantum present*, Oxford UK, Oxford University Press.

Fagot-Largeault, A., 2002, *L'emergence*. In Andler et al. 2002, 939-1048.

Fox, R., 1971, *The caloric theory of gases from Lavoisier to Regnault*, Oxford, Oxford University Press.

Hannequin, A., 1895, Essai critique sur l'hypothèse des atomes, Lyon, Annales de L'Université de Lyon (1ère éd.); 1899 chez Paris, Félix Alcan (2ème édition); 1981, New York, Arno Press.

Harman, P. M., 1982, *Energy, Force and Matter: The Conceptual Development of Nineteenth-Century Physics,* Cambridge, Cambridge University Press.

Heilbron, J., 1993, Weighing imponderables and other quantitative science around 1800*, Historical studies in the physical and biological sciences*, Supplement to vol. 24, Part 1.

Helmholtz, H. von, 1847, Über die Erhaltung der Kraft, Eine physikalische abhandlung, *Vorgetragen in der sitzung der Physikalischen gesellschaft zu Berlin* am 23, juli, 1847, Berlim, Reimer.

Meyerson, É., 1908, *Identité et réalité*, Paris, Félix Alcan (3ème édition 1926).

Liard, L., 1873, *Des définitions géométriques et des définitions empiriques*, Paris, Librairie philosophique de Ladrange.

Jammer, M., 1957, *Concepts of force*, Harvard, Harvard University Press (1962 Harper Torchbooks; 1999 Dover).

Moigno, L'abbé F., 1868, *Physique moléculaire: ses conquêtes, ses phénomènes et ses applications*, Paris, Gauthiers-Villars.

Perrin, J., *Traité de Chimie Physique. Les Principes*, Paris, Gauthiers-Villars.

Príncipe, J., 2008, *La réception française de la mécanique statistique*, thèse présentée pour obtenir le titre de docteur en Epistémologie et Histoire des Sciences et des Techniques, thèse dirigée par Olivier Darrigol, Un. Paris 7.

-- 2010, L'analogie et le pluralisme méthodologique chez James Clerk Maxwell, *Kairos Journal of Philosophy and Science*, 1, 55-74.

-- 2012, Sources et nature de la philosophie de la physique de Henri Poincaré, *Philosophia Scientiae,* 16 (2), 197-222.

Renouvier, C., 1859, *Essais de critique générale. Deuxième essai. L'Homme: La raison, la passion, la liberté. La certitude, la probabilité morale,* Paris, Librairie philosophique de Ladrange.

Saint-Venant, A. B. de, 1872, *Sur les diverses manières de présenter la théorie des ondes lumineuses*, Paris, Gauthier-Villars.

-- 1876, Sur la constitution atomique des corps, Académie des Sciences. *Compte rendus hebdomadaires des séances,* 82, 1223-1226.

Schaffner, K. F., 1972, *Nineteenth-century aether theories*, Oxford, Pergamon Press.

Scheidecker-Chevalier, M., 1997, L'hypothèse d'Avogadro (1811) et d'Ampère (1814): la distinction atome/molécule et la théorie de la combinaison chimique*, Revue d'histoire des sciences*, 50 (1-2), 159-194.

Schmaus, W., 2007, Renouvier and the method of hypothesis, *Studies in history and philosophy of science*, 38, 132–148.

Sebastianini, F.,1981, Le teorie microscopico-caloricistiche dei gas di Laplace, Poisson, Sadi Carnot e la teoria dei fenomeni termici nei gas formulata da Clausius nel 1850, *Physis,* 23, 397-438.

--1982, Le teorie microscopico-caloricistiche dei gas di Laplace, Ampère, Poisson Prevost, *Physis,* 24, 197-236.

Stallo, J., 1881, *The concepts and theories on modern physics*, New York, D. Appleton and Company.

Strien, M. van, 2014, Vital Instability and free will in physics and physiology, 1860-1890, *Annals of science*. http://dx.doi.org/10.1080/00033790.2014.935954. Consulté le 20 Aôut 2014.

Torretti, R., 1978, *Philosophy of geometry from Riemann to Poincaré*, Dordrecht, Holland, D. Reidel publishing Company.

Van Melsen, A. G., 1952, *From atomos to atom*, New York, Harper Torchbooks.

Villarceau, Y., 1876, Note sur la détermination théorique et expérimentale du rapport des deux chaleurs spécifiques, dans les gaz parfaits dont les

molécules seraient monoatomiques, suivi d'une remarque de M. Berthelot, Académie des Sciences. *Compte rendus hebdomadaires des séances,* 82, 1127-1130, 1175-1178.

Violle, J., 1870, *Mémoire de thèse : Sur l'équivalent mécanique de la chaleur*, Paris, Gauthier-Villars.

-- 1884, *Cours de Physique par J. Violle, Professeur à la Faculté des Sciences de Lyon, tome I: Physique moléculaire*, Paris, G. Masson.

Whittaker, E.,1951-1953, *A History of the theories of aether and electricity*, London, Thomas Nelson and sons (Dover edition, 1989).

How to Conceive the Atom: Imagery vs. Formalism

Giora Hon
(Department of Philosophy, University of Haifa, Israel)
hon@research.haifa.ac.il

Bernard R. Goldstein
(Dietrich School of Arts and Sciences, University of Pittsburgh, USA)
brg@pitt.edu

Abstract Given the belief in the universality of Newtonian mechanics, it is hardly surprising that atomic structure was compared to that of a planetary system, taken as a model for it. However, Heisenberg eliminated all pictures and models from his new theory. While Sommerfeld, the theoretical physicist, stressed the didactic importance of the defunct theory, Reichenbach, the philosopher of science, argued that a researcher cannot do without visualization, although this visualization is the "outer clothing" of the theory and does not represent its conceptual "skeleton". The problem underlying Reichenbach's statement may stem from what Born considered the naive assumption that the laws governing the macrocosm and the microcosm are the same. But even Born continued to present the defunct theory as a preliminary step for understanding quantum mechanics, not as a theory of historical interest. The force of the model and its accompanied imagery were apparently too strong to resist.

Resumo Dada a crença na universalidade da mecânica newtoniana, não surpreende que a estrutura atómica tenha sido comparada à de um sistema planetário, tomada como seu modelo. Contudo, Heisenberg eliminou qualquer imagem ou modelo da sua nova teoria. Ao passo que Sommerfeld, o físico teórico, salientou a importância didáctica da defunta teoria, Reichenbach, o filósofo da ciência, argumentou que um investigador não pode trabalhar sem visualização, apesar de esta visualização ser a "roupagem" da teoria e não representar o seu "esqueleto" conceptual. O problema subjacente à posição de Reichenbach pode provir do que Born considerava como a suposição ingénua de que as leis que governam o macrocosmo e o microcosmo são as mesmas. Mas mesmo Born continuou a apresentar a defunta teoria como um passo preliminar para compreender a mecânica quântica, e não como uma teoria com interesse histórico. A força do modelo e a imagética associada eram aparentemente irresistíveis.

Kairos. Revista de Filosofia & Ciência 13: 213-236, 2015
Centro de Filosofia das Ciências da Universidade de Lisboa

> When one contemplates what has really been done [in atomic physics], one sees clearly that [mechanical] model conceptions [*Modellvorstellungen*] have no real meaning. The orbits [*Bahnen*] are not real, neither with respect to frequency nor energy.
>
> W. Heisenberg to A. Sommerfeld, December 8, 1923.[1]

> We should not want to clap the atoms into the chains of our preconceptions [*Vorurteile*] (to which in my opinion belongs the assumption of the existence of electron orbits [*Elektronenbahnen*] in the sense of the usual kinematics), but we must on the contrary adjust our ideas [*Begriffe*] to experience [*Erfahrung*].
>
> W. Pauli to N. Bohr, December 12, 1924.[2]

I. Background

Given the belief in the universality of Newtonian mechanics, it is hardly surprising that the structure of the atom was compared to that of a planetary system. Nevertheless, in the early stages of the development of atomic physics it was clearly perceived and explicitly noted that the atom exhibits a categorically different kind of stability from that of a mechanical scheme. For example, Niels Bohr (1885–1962) spelled out this limitation in his Nobel lecture of 1922. To be sure, the analogy between the planetary system and the structure of the atom "provide[s] us with an explanation," but it has its limitations:

[1] Eckert and Märker, 2004, 157. Translated in Kragh, 2012, 322. By *Modellvorstellungen* Heisenberg means mechanical modeling (as he says earlier in this letter); he refers explicitly to orbits, that is, to the frequency and the energy which the material particle possesses in its motion around a core at the center of its trajectory. We are grateful to Michael Eckert for supplying us with a copy of the original German manuscript of this letter as well as a transcription of it.

[2] Quoted and translated by Serwer, 1977, 243 n. 167; Hermann et al., 1979, 189.

> As soon as we try to trace a more intimate connexion between the properties of the elements and atomic structure, we encounter profound difficulties, in that essential differences between an atom and a planetary system show themselves here in spite of the analogy we have mentioned.
>
> The motions of the bodies in a planetary system, even though they obey the general law of gravitation, will not be completely determined by this law alone, but will depend largely on the previous history of the system. Thus the length of the year is not determined by the masses of the sun and the earth alone, but depends also on the conditions that existed during the formation of the solar system, of which we have very little knowledge.... The definite and unchangeable properties of the elements demand that the state of an atom cannot undergo permanent changes due to external actions.[3]

And Bohr surmised:

> On the basis of our picture of the constitution of the atom it is thus impossible, so long as we restrict ourselves to the ordinary mechanical laws, to account for the characteristic atomic stability which is required for an explanation of the properties of elements.[4]

We may formulate the difference in this way: mechanical stability has "memory", it records mechanical disturbances and retains their influence. Atomic stability, by contrast, can undergo all sorts of interactions, but the atom of some element will always remain the same with the same physical properties, unless it is transmuted into another element. Notice that Bohr spoke of a picture of the constitution of the atom, a picture which, given his analysis, cannot be retained.

To be sure, the planetary system is probably the most stable mechanical system known to man, but as a contingency it cannot – in principle – be related to the inherent necessity of the stability of atoms, the building blocks of matter. However, confidence in the planetary model was greatly enhanced by its success in explaining and predicting phenomena of simple atomic structures. For example, Karl

[3] Bohr, [1922] 1965, 10–11.
[4] Bohr, [1922] 1965, 11.

Schwarzschild (1873–1916) and Paul Sophus Epstein (1883–1966) applied theories and techniques of celestial mechanics to the mechanics of the atom; their theoretical success in explaining the Stark effect consolidated Sommerfeld's approach and persuaded Bohr that the planetary model was worthy of elaboration (see Figure 1).[5]

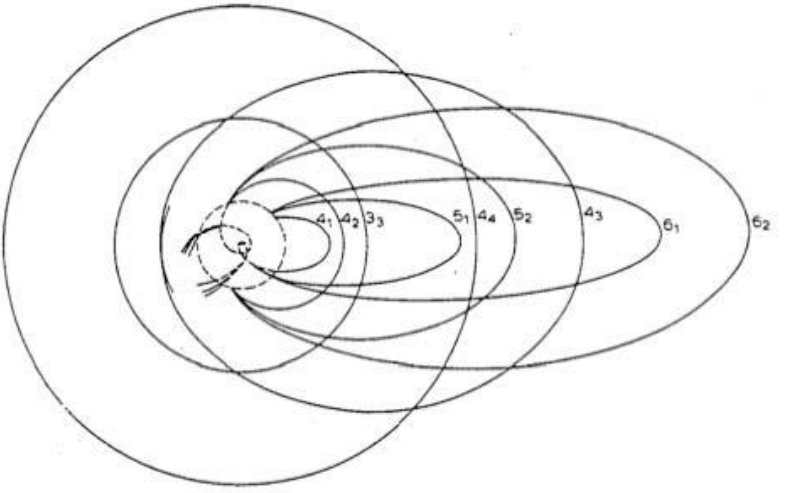


Fig. 1. Bohr, [1922] 1965, p. 32, Fig. 8.

Arnold Sommerfeld (1868–1951) conceived the atom as "a closed mechanical system in which only internal forces act."[6] At stake was the dynamics of this mechanical system.

> The question arises: how can the electrons of the atom maintain themselves in opposition to the attractive action of the nuclear charge? Will this action not cause them to fall into the nucleus? The answer – a

[5] Schwarzschild, 1916; Epstein, 1916. For a discussion, see Jammer, 1966, 103.
[6] Sommerfeld, [1919/1922] 1923, 264.

> possible one which is particularly simple and satisfactory – is furnished by the conditions of the solar system.[7]

Sommerfeld transferred the problem of stability from the atom to the solar system, that is, the stability of the atom is likened to the stability of the planetary system:

> The earth fails to fall into the sun for the reason that it develops centrifugal forces owing to its motion in its own orbit [*Umlauf*], and these forces are in equilibrium with the sun's attraction. If we transpose these ideas to our atomic model we arrive at the following view. **The atom is a planetary system in which the planets are electrons. They circulate about the central body, the nucleus**. The atom of which the atomic number is Z is composed of Z planets each charged with a single negative charge, and of a sun charged with Z positive units. The *gravitational attraction*, as expressed in Newton's law, is represented by the *electrical attraction* as given by Coulomb's law; these laws are alike in form.[8]

The imagery is quite extraordinary (see Fig. 2). Sommerfeld, in so many words, identified the atom as a planetary system: the stability of the atom is due to the same conditions that maintain the stability of the solar system, so his argument went. Moreover, for Sommerfeld it was not merely an analogy, for he stated an identity: "the atom *is* a planetary system." This is, of course, an exaggeration, expressing enthusiasm (surely a sign of his confidence in the model) rather than objective analysis. And Sommerfeld retreated immediately to the analogy: the gravitational force is "represented" by the Coulomb force, the two laws being alike in form. Sommerfeld was aware of the differences too. Although the laws are formally similar,

> There is a difference in that the planets repel one another in our atomic microcosm – likewise according to Coulomb's law – whereas, in the case

[7] Sommerfeld, [1919/1922] 1923, 65.
[8] Sommerfeld, [1919/1922] 1923, 65, boldface and italics in the original; for the German, see Sommerfeld, [1919] 1922, 79.

> of the solar macrocosm they undergo attraction not only from the sun but also from themselves.[9]

Still, the claim that the dynamical laws hold in the "atomic microcosm" exactly as in the "solar macrocosm" runs like a thread in Sommerfeld's study. Confidence in the correctness of the Bohr-Sommerfeld approach had reached its height in 1922. This success made it seem likely that a fully satisfactory theory of this kind would be found in the near future, although later on it became clear that this initial optimism was ill-founded.

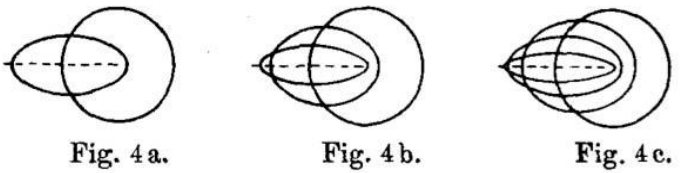


Fig. 2. Sommerfeld, 1916, p. 23.

II. Confidence vs. Confirmation

We distinguish between "confidence" and "confirmation"; the latter is the subject of a vast literature in philosophy of science. It concerns the relation between evidence and theory: in what ways and to what extent does the evidence support the theory (or hypothesis)? One may identify three conceptions: qualitative, comparative, and quantitative. The first simply states that evidence *e* confirms hypothesis *H*; the second claims that *e* confirms *H* more than *H′*, and the third conception calculates the degree to which *e* confirms *H* in probabilistic

[9] Sommerfeld, [1919/1922] 1923, 65.

terms.[10] If a theory has been confirmed, there may no longer be reasons to doubt it, whereas confidence in a theory will vary according to the individual scientist who weighs the evidence idiosyncratically, given certain predispositions which are not always conducive to probabilistic measure. This is akin to scientists ignoring "anomalies", that is, confidence in the theory is not shaken by anomalies.[11] We are not engaged here with the formal analysis of the probabilistic relation between evidence and theory. Rather, we characterize the attitudes which scientists take vis-à-vis a theory under scrutiny. This is the conception of confidence. It is not surprising that close agreement between experimental data and consequences of a theory leads to confidence in the theory. We thus invoke "confidence" as an epistemic disposition which, however, is not related to "degree of belief". Confidence is what keeps a scientist focused on improving a theory so that the theory, as modified, can be confirmed by future experiments. It therefore expresses expectation: Is it worth investing time, effort, and material resources in the theory, or not? Confidence then refers to the expected future state of the theory.

In the case under discussion one was confident that, with some relatively minor adjustments, the quantum theory, complete with the imagery of a planetary system, will work over the entire domain of atomic physics, whereas at its height it had only been shown to work for a nucleus surrounded by one electron. To be sure, the results were excellent and so expectations were running high. This confidence was based on an analogy with Newtonian mechanics; Newton first described a two body problem (and solved it); later it was shown that the same technique with relatively minor modifications yields a very close approximation to the solution of the three-body problem (even

[10] See, e.g., Achinstein, 2001.
[11] This is related to what we have called "snapshot" (as distinct from "baseline"): see Hon and Goldstein, 2009.

though an exact solution could not be found). There was an expectation that the three-body problem could be solved or, at least, one could make progress towards its solution, without casting doubt on any of the fundamental assumptions of Newtonian mechanics. A long series of theoretical physicists and mathematicians made the effort to work towards that solution.[12] This was seen as a useful precedent for confidence in solving the N-body problem at the atomic level, without challenging any of the basic tenets of the theory.

Early in December 1925, in an article that appeared in *Nature*, Bohr offered a brief survey of the development of the quantum theory up to that time, including an acknowledgment that in July 1925 a new phase was ushered by the theoretical work of Werner Heisenberg (1901–1976).[13] In this paper, "Atomic Theory and Mechanics," Bohr noted that "perhaps the greatest successes of mechanics lie in the domain of astronomy".[14] He then elaborated Planck's contribution and the discovery of the quantum of action, which rendered classical theories inadequate when applied to atomic structure.[15] According to Bohr, Rutherford's conception of the atom, that "around the nucleus there move a number of light negative electrons" made "the problem of atomic structure [take] on a great similarity to the problem of celestial mechanics." Still, Bohr pointed out the well-established fact that "there exists a fundamental difference between an atom and a planetary system."[16]

It is, however, important to note that Bohr relaxed the strong contrast between mechanical and atomic stability that he himself identified, and followed a methodology that allows continuous

[12] See, e.g., Barrow-Green, 1997, espec. 7–27. Poincaré, 1893, served as the base for these studies.
[13] Rüdinger and Stolzenburg, 1984.
[14] Bohr, 1925, 845.
[15] Bohr, 1925, 846.
[16] Bohr, 1925, 847.

analysis. This point is critical. In 1925, when quantum mechanics had just arrived on the scene, Bohr – in his reflection on the methodology of his own contribution – remarked:

> Nevertheless, it has been possible to construct mechanical pictures of the stationary states which rest on the concept of the nuclear atom and have been essential in interpreting the specific properties of the elements. In the simplest case of an atom with only one electron, such as the neutral hydrogen atom, the orbit of the electron would be in classical mechanics a closed ellipse, obeying Kepler's laws, according to which the major axis and frequency of revolution are connected in a simple way with the work necessary for a complete separation of the atomic particles.[17]

Bohr's references to "pictures", and "mechanical pictures" at that, should not escape the notice of the reader. Indeed, Bohr added that, as Sommerfeld had shown, the small deviations from Keplerian motion for the electron are consistent with the theory of relativity and offer an explanation for the fine structure of the lines of hydrogen in the spectrum.[18] These considerations made the planetary model attractive and instilled confidence in it; yet Bohr concluded his paper by insisting on the inadequacy of mechanical pictures.[19] In fact, Bohr acknowledged the contribution of Heisenberg who, in a revolutionary move, replaced quantum theory with a new theory, quantum mechanics, in which "the difficulties attached to the use of mechanical pictures may ... be avoided.... In contrast to ordinary mechanics, the new quantum mechanics does not deal with a space-time description of the motion of atomic particles."[20] Indeed, Heisenberg eliminated all pictures and models from his new theory (pictures are static and

[17] Bohr, 1925, 848.
[18] Bohr, 1925, 849.
[19] Bohr, 1925, 850.
[20] Bohr, 1925, 852; see also 845.

models in this case are dynamic).[21] To put it strongly, it is surprising that Bohr did not draw the logical consequences from his own analysis of the physics of the atom, for he continued to seek a mechanical picture of the constitution of the atom.

The three leading physicists who played key roles in developing the quantum theory may be singled out. In brief, Bohr was the originator of the theory with an implicit appeal to the planetary model; Sommerfeld consolidated a full-fledged, indeed Keplerian, planetary scheme, and Max Born (1882–1970) contributed to its demise. In spite of the fact that all three knew very well that the theory could not account for phenomena of complex atomic structures, they continued to appeal to the quantum theory in its planetary presentation. The two Nobel Laureates in physics, Bohr and Born, and the influential teacher of theoretical physics, Sommerfeld, all expressed attitudes which can be characterized as "schizophrenic": the quantum theory failed to capture atomic phenomena and was indeed at some point formally discarded; yet, all three physicists continued to present this theory in one way or another after its demise. What was the attraction of the model? And why was it so popular among powerful thinkers who clearly recognized its limitations? In other words, what was the source of confidence in the quantum theory?

III. The Path of an Electron in Motion

What is the geometrical form of the path of an electron in motion? Or, better, how should one picture the path of a free electron? Obviously, no one asked this question before the discovery of the

[21] In his revolutionary paper Heisenberg (1925) referred neither to a picture nor to a model of the atom. The thrust of the paper is abstract mathematics with which Heisenberg manipulated observable magnitudes.

electron at the end of the nineteenth century. But, even after its discovery and the development of the theory of electron which addressed many sophisticated issues of the physics of the electron, no one recognized any particular difficulty in characterizing the path of an electron in motion. For instance, Max Abraham (1875–1922) in his paper on the dynamics of the electron, and Einstein in his theory of special relativity, talked of a *Bahn*, a path, without saying anything more specific.[22] For another example, Augusto Righi (1850–1920) invoked the expression "la trajectoire d'un électron" in a study of interactions among electrons, ions, atoms, and molecules, again assuming the concept of trajectory without any discussion.[23] The geometrical form of the path of a free electron as it moves in empty space was not considered problematic at that time. The issue, however, had to be addressed once a bound electron, confined to the atom, became the object of study. Given that the atom is a stable and electrically neutral system, and that the electron – a charged component of this system – is not stationary, what is the geometrical form of the path of the electron in motion within the atom such that the atom can radiate and still remain stable and electrically neutral? This is certainly a fundamental problem, one which had no immediate solution. The obvious methodological move, when something new has to be confronted, is to appeal to analogy – "an explanation of the

[22] Abraham, 1902, 27; Einstein, 1905, 921. At Heisenberg's first meeting with Einstein in 1926, Einstein "pointed out to me that in my mathematical description the notion of 'electron path' did not occur at all, but that in a cloud-chamber the track of the electron can of course be observed directly. It seemed to [Einstein] absurd to claim that there was indeed an electron path in the cloud-chamber, but none in the interior of the atom. The notion of a path could not be dependent, after all, on the size of the space in which the electron's movements were occurring" (Heisenberg, [1974] 1989, 113). We may surmise that for Einstein an electron did have a path, a trajectory, both in free space and when bound within the atom. For analyses of observed trajectories in the cloud chamber by Heisenberg and Born, see Figari and Teta, 2013.

[23] Righi, 1908, 602.

unfamiliar by the more or less completely familiar."[24] The mechanical model for the atom, that is, the planetary model, is one such case.

Probably the most stable, well-studied, mechanical system with constituent moving elements is the planetary system. The term *orbit* had been used in astronomy since it was introduced by Johannes Kepler (1571–1630) in the early years of the seventeenth century and, without much discussion, it came to be applied to the path of the electron within the atom.[25] For example, in his paper on magnetism and the theory of electron, the French physicist, Paul Langevin (1872–1946), began his paper of 1905 with the following assumption:

> Let us suppose some electrons in motion in the interior of the molecule, following closed orbits [*suivant des orbites fermées*], that can be considered similar to closed currents circulating along these orbits [*ces orbites*] from the point of view of the magnetic field produced at a distance.[26]

It is of paramount importance to realize that a fundamental assumption is made here. The bound electron, as distinct from the free electron, is assumed to be moving in a closed orbit. No argument is offered to support this claim and, equally, no reference to a planetary system is made, although before the discovery of the electron the term *orbit* in physics had always been associated with a planetary system.[27] In other words, Langevin set up the discussion "silently" within a mechanical context whose paradigm is a planetary system. This theoretical framework seems to have influenced Bohr in the development of the quantum theory.[28] The idea was that the atom resembles a planetary system, that is, a central nucleus corresponding to the Sun with electrons rotating around the nucleus in orbits like

[24] North, 1981, 129.
[25] Goldstein and Hon, 2005.
[26] Langevin, 1905, 73.
[27] Comets, moons, etc. are included in "planetary system"; in particular, comets and moons have "orbits" in classical mechanics.
[28] See Heilbron and Kuhn, 1969, 219–223.

planets around the Sun, such that almost all of the atom is empty space.

Although the development of quantum mechanics, beginning in 1925, rendered the planetary conception obsolete, the model took on a life of its own and continued to flourish. In fact, the planetary model is not even good at representing the theory as it had developed prior to 1925. Yet, this model has continued to be invoked in textbooks of modern physics and now serves as an icon for atomic energy even though it is no longer scientifically meaningful.[29] It is puzzling that the defunct quantum theory continued to play a role in introducing students to the physics of the atom. Why did the planetary model complete with imagery persist?

IV. The Attractiveness of the Planetary Model

The heyday of the quantum theory based on explicit planetary modeling only lasted some seven years, from Sommerfeld's explicit introduction of Keplerian orbits in 1915 to the celebratory presentation of the full-fledged theory in Bohr's Nobel lecture in 1922.[30] Enthusiasm, an expression of great confidence in the theory, was then at its peak. The model offers an attractive metaphysics: the macrocosmos has a complementary microcosmos; physics has a "picture" conducive to understanding and, above all, mathematical techniques of one domain are applicable in the other. However, the decline was rapid: the period of dissatisfaction during 1923 and 1924 along with groping for a new conceptualization of the phenomena came to an abrupt end in the summer of 1925, when Heisenberg

[29] See, e.g., Born, 1933. For the history of the image that turned into the icon of atomic physics, see Goldstein and Hon, 2015.
[30] Sommerfeld, 1915; Bohr, [1922] 1965.

offered a new foundation, replacing quantum theory with quantum mechanics.[31] During these seven years research in quantum theory, such as the impressive achievements of Schwarzschild and Epstein, yielded a theory complete with a visual model which showed great promise.[32] However, after 1925 it was no longer at the cutting edge of research in atomic physics. In fact, the theory and its model had to be discarded; it proved to be wrong. Yet, the planetary model, together with its mechanical imagery, has persisted in a great many contexts (see, e.g., Fig. 3).[33]

Fig. 3. The emblem of the International Atomic Energy Agency (IAEA), adopted in 1960 and still in use.

[31] Heisenberg, 1925.
[32] Schwarzschild, 1916; Epstein, 1916. See n. 5, above.
[33] This figure appears in Wellerstein, 2013.

V. Enter the Philosopher Hans Reichenbach (1891–1953)

A contemporary leading philosopher of science, Reichenbach provided a perspective on this episode by a well-informed non-physicist. Once again, the tension between the defunct quantum theory and quantum mechanics is striking. In his *Atom and Cosmos*, in the section, "The Laws of Atomic Mechanism", Reichenbach described in detail what he called Bohr's model:

> Much more was achieved ... than an explanation of the atom's physical effect; even the chemical properties of the atom could be cleared up by the aid of Bohr's model. The fundamental idea ... is that chemical combination is regarded as the union of planetary systems of the atoms [*Vereinigung der atomaren Planetensysteme*] in more comprehensive systems.[34]

And Reichenbach added figures that illustrate the Keplerian planetary structure (*Ellipsenbahn*) of the various atoms. Chemistry, according to Reichenbach, had thus become a branch of physics.[35]

Reichenbach was well versed in the developments of recent physics. In the next chapter, "The Wave Character of Matter", Reichenbach presented Schrödinger's wave mechanics, informing the reader that Bohr's model finally fell victim to the contradictions inherent in his model of the atom. However, success came at a high price, namely, the new theory lacked a visual model. This was a radical break from the past:

> More clearly than ever does that reciprocal relation here stand out, which exists between mathematical formulae and visual images in modern physics; the kernel of a theory, its conceptual skeleton, is given by the mathematical formulae, whereas the images are only outer clothing [*während die Bilder nur eine dem Wechsel unterworfene*

[34] Reichenbach, 1930, 252; Reichenbach [1930] 1932, 237–238.
[35] Reichenbach, 1930, 257; Reichenbach [1930] 1932, 243. On the teaching of chemistry in light of the Bohr-Sommerfeld theory in the United States, see Martinette, 1940.

> *Umkleidung bedeuten*], subject to change, which have no immediate value for real knowledge. In spite of that, they are of practical worth, for the investigator in search of new paths cannot do without them.[36]

While Sommerfeld, the theoretical physicist, stressed – as we will see – the didactic importance of the defunct Bohr-Sommerfeld theory, Reichenbach, the philosopher of science, took this claim a step further. In his view the researcher cannot do without visualization, although this visualization is the "outer clothing" of the theory and does not represent its conceptual "skeleton".[37] Indeed, Reichenbach went on to explain, as he understood it, the revolutionary approach that Heisenberg had taken:

> Heisenberg's considerations are of a very radical nature; for they rest on a criticism of the very problem of a model. Heisenberg proposed the question whether there is any meaning in representing an atom, in the Rutherford-Bohr manner, by a model. For it is quite impossible for us to conceive of the atom's planetary system, with its tiny dimensions, according to its correct size; what we do is to picture a highly magnified model [*wir malen uns da ein stark vergrößertes Modell aus*], in which the electron is about as large as a pinhead, and circles around a still larger nucleus at an appropriate distance. Is it permissible to think of what really happens as resembling this model? Heisenberg denies it. He objects that we can never really observe the atom in its minute dimensions, and so demands that we remove the model from our thought. We can say only so much about the micromechanism as we can justify by observations.[38]

It is not clear whether Reichenbach fully appreciated that the planetary model has no place in quantum mechanics for, after all, it is not simply a matter of the atomic dimensions that, in principle, makes the simultaneous measurement of an electron's position and momentum impossible. The problem underlying Reichenbach's statement may stem from what Born considered the naive assumption that the laws governing the macrocosm and the microcosm are the

[36] Reichenbach, 1930, 274; Reichenbach, [1930] 1932, 257–258.
[37] This imagery of skeleton and clothing can be traced back to Hertz, 1892, 31.
[38] Reichenbach, 1930, 276; Reichenbach, [1930] 1932, 259.

same.[39] But even Born continued to present the Bohr-Sommerfeld theory as a preliminary step for understanding quantum mechanics, not as a theory of historical interest. The force of the model and its accompanied imagery were apparently too strong to resist.

VI. Born's Appeal to Mechanical Imagery

In 1933, eight years after the introduction of quantum mechanics Born, who was one of the principal architects of quantum mechanics, published a series of lectures on atomic physics, matter, and radiation. Born's *Moderne Physik* is also available in English translation with the title, *Atomic Physics*.[40] When the fourth English edition was published in 1948, Born was 66 years old and quantum mechanics celebrated its 23rd anniversary. One would have expected that the Bohr-Sommerfeld theory was by then a relic of the past – this, however, is not the case.

In fact, Born followed closely Bohr's original argument:

> It seems now a natural suggestion that we should regard the quantization condition for the angular momentum as an essential feature of the new mechanics. We therefore postulate that it is universally valid. At the same time, we must show by means of examples that the postulate leads to reasonable results. Although from the standpoint of Bohr's theory the underlying reason for this quantization rule remains entirely obscure, nevertheless, in the further development of the theory it has justified itself by results.[41]

Born was aware that Bohr's reason for postulating quantization was obscure, and he justifies the appeal to the incomplete theory for its successful results. When it comes to illustrate these results, Born offers the case of the hydrogen atom, whose complete quantization was carried out by Sommerfeld: "By Kepler's laws the orbit of the

[39] Born, 1923, 537–538.
[40] Born, 1933; Born, [1933/1935] 1948.
[41] Born, [1933/1935] 1948, 106–107.

electron round the nucleus is an ellipse; it is therefore simply periodic."[42] And Born directs the reader to the Appendix of his book, where he gives the quantization of "the Kepler ellipse", which in turn leads to the correct energy levels of the Balmer terms.

Again, we note that the year is 1948 and the context is physics, not history of physics. Moreover, Born reproduced Sommerfeld's drawings in the German edition of 1933 and, perhaps not surprisingly, they still appear in the English edition of 1948, except that in the English edition the drawings are no longer acknowledged as Sommerfeld's.[43] We find it puzzling that a defunct theory continued to play a role in introducing students to the physics of the atom. Why did the model persist?

We need not dwell on the many editions of Sommerfeld's textbook, *Atombau und Spektrallinien*;[44] suffice it to note that it was reprinted as late as 1978, and its English translation was reprinted in 1934, nine years after the introduction of quantum mechanics. The fifth German edition came out in 1931; it presented the by now defunct quantum theory. Sommerfeld had already published in 1929 a supplementary volume on wave mechanics. In his Preface to the 1931 German edition Sommerfeld offered an explanation why an invalid theory should be presented, albeit in abbreviated form:

> It has become clear that it is possible to understand the new theory only by building it up from the old theory. For this purpose the present volume necessarily treats not only the basic experimental facts, but also of the orbital ideas [*Bahnvorstellungen*] so far as they are required for introducing the quantum numbers, and for serving as a model [*Vorbild*] for the wave mechanical calculations. The final results are always given in the form in which they are presented by the new theory. Consequently, it has been necessary to refer frequently to the

[42] Born, [1933/1935] 1948, 114; Born, 1933, 81.
[43] Born, 1933, 78–81, and Born, [1933/1935] 1948, 111–114, respectively.
[44] First edition, 1919; cf. van der Waerden, 1969.

> supplementary volume and to leave occasional gaps [*Lücken*] in the proofs.[45]

In a somewhat apologetic tone, Sommerfeld says that a "wrong" theory from the past is beneficial to students in the process of learning and understanding the current theory, namely, wave mechanics (Schrödinger's representation of quantum mechanics). This appears to be the rationale for maintaining the planetary model, but it should be noted that there is no continuous transition from the quantum theory to quantum mechanics.

VII. Conclusion

Did quantum theory, with its underlying planetary model, become a theory of historical interest only? Apparently not. A clear example is Messiah's textbook, *Quantum Mechanics*.[46] This book is based on a series of lectures given in the early 1950s. After a general introduction entitled, "The End of the Classical Period", under the heading "Inadequacy of classical corpuscular theory", Messiah states that "*the evolution in time of a quantized quantity is impossible to describe in strictly classical terms*.... One must give up imagining an exact evolution of the energy as a function of time."[47] Thus the planetary model was dealt a death blow right at the outset of this textbook. Still, Messiah acknowledges the historical and heuristic importance of the Bohr-Sommerfeld theory:

> In spite of the difficulties of principle, and the limitations of [the] Old Quantum Theory, it is useful to know its main features in order to

[45] Sommerfeld, 1931, iv (author's Preface to the fifth German edition), translated by H. L. Brose, in Sommerfeld, [1919/1931] 1935 (slightly modified).
[46] Messiah, [1961] 1972.
[47] Messiah, [1961] 1972, 27–28, italics in the original.

> properly appreciate the later development of the theory. Furthermore, this older theory represents a first example of the application of a heuristic principle which played an essential role in the development of Quantum Mechanics: the correspondence principle.[48]

So, while physicists did see some merit in teaching the defunct theory, Messiah reminds us that:

> The quantization rules are purely formal restrictions imposed upon the solutions of the classical equations of motion; they were determined in an entirely empirical manner. The profound justification of this quantization of classical trajectories is completely absent. In fact, the very notion of trajectory is hard to reconcile with the quantization phenomenon. It implies that the particle possesses at each instant a well-defined position and momentum, and that these quantities vary in a continuous manner in the course of time.... To postulate this quantization amounts to giving up the (classical) idea of a precise trajectory of the electron and, quite logically, the idea of trajectory in general.[49]

For Messiah the old theory "represented a great step forward" and "played a large clarifying role in the history of contemporary physics. But this rather haphazard mixture of classical mechanics and ad hoc prescriptions can in no way be considered as a definitive theory."[50] According to Messiah, the concept of "electronic orbit" exemplifies all that had gone wrong with the quantum theory. It is a concept without experimental foundation. It is perhaps fitting to end on the very note with which we have begun – the orbit:

> Since no experiment allows us to assert that the electron actually describes a precise orbit in the hydrogen atom, nothing prevents us from abandoning the very notion of an orbit; in other words, the fact that the atom is in a well-defined energy state does not necessarily imply that the electron has at each instant a well-defined position and momentum.[51]

[48] Messiah, [1961] 1972, 28.
[49] Messiah, [1961] 1972, 40–41.
[50] Messiah, [1961] 1972, 41.
[51] Messiah, [1961] 1972, 46.

In fact, postulating quantization requires dropping the notion of trajectory. Quantum mechanics begins when the concept of orbit – together with the planetary scheme complete with mechanical imagery – is discarded. What remains is the "skeleton" – the formalism.

References

Abraham, M., 1902, Dynamik des Elekrons. *Nachrichten von der Königl. Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse aus dem Jahre 1902*, 20–42.

Achinstein, P., 2001, *The Book of Evidence*. New York, Oxford University Press.

Barrow-Green, J., 1997, *Poincaré and the Three Body Problem*. History of Mathematics, 11. Providence, RI, American Mathematical Society.

Bohr, N., [1922] 1965, The structure of the atom. In *Nobel Lectures*, *Physics 1922–1941*. Amsterdam, Elsevier, 7–43.

-- 1925, Atomic Theory and Mechanics. *Nature* 116 (December 5, 1925), 845–852.

-- 1984, *Collected works*. Vol. 5: *The Emergence of Quantum Mechanics (mainly 1924–1926).* Edited by K. Stolzenburg. Amsterdam and New York, North-Holland.

Bopp, F. and Kleinpoppen, H. (eds.), 1969, *Physics of the one- and two-electron atoms*. Proceedings of the Arnold Sommerfeld centennial memorial meeting and of the international symposium on the physics of the one- and two-electron atoms (Munich, 10–14 September 1968). Amsterdam, North-Holland.

Born, M., 1923, Quantentheorie und Störungsrechnung. *Die Naturwissenschaften* 11 (27), 537–542.

-- 1933, *Moderne Physik: Sieben Vorträge über Materie und Strahlung*. Berlin, Springer.

-- [1933/1935] 1948, *Atomic Physics*. Translated by J. Dougall. London and Glasgow, Blackie.

Eckert, M. and Märker, K. (eds.), 2004, *Arnold Sommerfeld. Wissenschaftlicher Briefwechsel*, vol 2. Berlin, Verlag für Geschichte der Naturwissenschaften und der Technik.

Einstein, A., 1905, Zur Elektrodynamik bewegter Körper, *Annalen der Physik*, 17, 891–921.

Epstein, P. S., 1916, Zur Theorie des Starkeffektes. *Annalen der Physik*, Series 4, 50: 489–520.

Figari, R. and Teta, A., 2013, Emergence of classical trajectories in quantum system: the cloud chamber problem in the analysis of Mott (1929), *Archive for History of Exact Sciences*, 67, 215–234.

Goldstein, B. R. and Hon, G., 2005, Kepler's Move from *Orbs* to *Orbits*: Documenting a Revolutionary Scientific Concept, *Perspectives on Science*, 13, 74–111.

-- 2013-2014, The Image that Became the Icon for Atomic Energy, *Physis*, 49, 259-272.

Heilbron, J. L. and Kuhn, T. S., 1969, The genesis of the Bohr atom. *Historical Studies in the Physical Sciences*, 1, 211-290.

Heisenberg, W., 1925, Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Zeitschrift für Physik* 33: 879–893. For an English translation, see van der Waerden, 1967, 261–276.

-- [1974] 1989, *Encounters with Einstein and Other Essays on People, Places, and Particles*. Princeton: Princeton University Press.

Hermann, A., Meyenn, K. v., and Weisskopf, V. F. (eds.), 1979, *Wolfgang Pauli: Scientific Correspondence with Bohr, Einstein, Heisenberg A. O*. 1: 1919–1929. New York, Heidelberg, and Berlin: Springer.

Hertz, H., 1892, *Untersuchungen über die Ausbreitung der elektrischen Kraft.* Leipzig, Barth.

Hon, G, and Goldstein, B. R., 2009, Spotlight on: The Nature of Scientific Change; In Pursuit of Conceptual Change: the Case of Legendre and Symmetry, *Centaurus*, 51, 288–293.

Jammer, M., 1966, *The conceptual development of quantum mechanics*. New York, McGraw-Hill.

Kragh, H., 2012, *Niels Bohr and the Quantum Atom: The Bohr Model of Atomic Structure 1913–1925*. Oxford, University Press.

Langevin, P., 1905, Magnétisme et théorie des électrons. *Annales de chimie et de physique*, Series 8, 5, 70–127.

Martinette, M., 1940, The Presentation of Atomic Structure to College Freshman, 1905–1940. *School Science and Mathematics*, 40, 808–814.

Messiah, A., [1961] 1972, *Quantum Mechanics*, Vol. 1. Amsterdam, North-Holland.

Nobel Prize, Physics 1922, Nobelprize.org.
http://www.nobelprize.org/nobel_prizes/physics/laureates/1922/

North, J. D., 1981, Science and analogy. In M. D. Grmek, R. S. Cohen, and G. Cimino (eds.), *On Scientific Discovery.* Dordrecht: Reidel, pp. 115–140.

Poincaré, H., 1893, *Les méthodes nouvelles de la mécanique céleste*, tome II: Méthodes de MM. Newcomb, Gyldén, Lindstadt et Bohlin. Paris, Gauthiers-Villars.

Reichenbach, H., 1930, *Atom und Kosmos: das physikalische Weltbild der Gegenwart*. Berlin, Deutsche Buch-Gemeinschaft.

-- [1930] 1932, *Atom and cosmos: The world of modern physics*. Allen and Unwin.

Righi, A., 1908, Sur quelques phénomènes dus aux rencontres entre électrons, ions, atomes et molécules. *Journal de physique théorique et appliquée*, Series 4, 7, 589–617.

Rüdinger, E. and Stolzenburg, K., 1984, Introduction: Part II. In Bohr, 1984, 219–240.

Schwarzschild, K., 1916, Zur Quantenhypothese. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, Jahrgang 1916 (Januar bis Juni), 548–568.

Serwer, D., 1977, Unmechanischer Zwang: Pauli, Heisenberg, and the rejection of the mechanical atom, 1923–1925. *Historical Studies in the Physical Sciences*, 8, 189–256.

Sommerfeld, A., 1915, Zur Theorie der Balmerschen Serie. *Sitzungsberichte der königl. bayerischen Akademie der Wissenschaften*, 425–458.

-- 1916. Zur Quantentheorie der Spektrallinien. *Annalen der Physik*, Series 4, 51: 1–94, 125–167.

-- [1919] 1922, *Atombau und Spektrallinien*. 3rd ed. Braunschweig, Vieweg & Sohn.

-- [1919/1922] 1923, *Atomic Structure and Spectral Lines*. Translated from the 3rd German ed. by H. L. Brose. New York, Dutton.

-- [1919/1924] 1931, *Atombau und Spektrallinien*. 5th ed. Braunschweig, Vieweg & Sohn.

-- [1919/1931] 1935, *Atomic Structure and Spectral Lines*. Translated from the 5th German ed. by H. L. Brose. New York, Dutton.

van der Waerden, B. L., 1967, *Sources of Quantum Mechanics*. Amsterdam, North-Holland.

-- 1969, The history of quantum theory in the light of the successive editions of Sommerfeld's *Atombau und Spektralinien*. In Bopp and Kleinpoppen, 1969, 21–31.

Wellerstein, A., 2013, The story behind the IAEA's atomic logo, *Restricted Data: The Nuclear Secrecy Blog*, January 11, 2013, accessed August 10, 2014, http://blog.nuclearsecrecy.com/2013/01/11/the-story-behind-the-iaeas-atomic-logo/

Observing stars, representing atoms: images and objectivity in the physical sciences

Eduardo Zubia
(University of the Basque Country UPV/EHU)
eduzubia@gmail.com

Abstract In cultural studies, scientific images are often considered social constructs, not essentially different from works of art. We will argue that imaging plays a key role in physics and that its epistemic realism shall not be systematically dismissed. However, in examining visual practices in science, we will not ignore their relations to artistic practices involved in the representation of nature. We will discuss in some detail two astrophysical theories closely intertwined with the ontological nature of microscopic phenomena: the large scale structure of the universe, in which the statistical analysis of instrumental observations sheds light on the concepts of cosmological inflation and primordial perturbations, and the Hertzsprung-Russell diagram, which relates stellar photometric and spectroscopic data to thermodynamic and atomic descriptions of astronomical bodies. Visual representations as part of good epistemic practice may permit to assess the fundamental constituents of the physical world and to objectively classify the natural phenomena.

Resumo Nos estudos culturais, as imagens científicas são frequentemente consideradas construções sociais, essencialmente não diferentes das obras de arte. Argumentamos que a imagiologia desempenha um papel fundamental na física e que o seu realismo epistémico não deve ser sistematicamente dispensado. Contudo, ao examinar as práticas visuais na ciência, não ignoramos as suas relações com as práticas artísticas envolvidas na representação da natureza. Discutimos com algum detalhe duas teorias astrofísicas fortemente relacionadas com a natureza ontológica dos fenómenos microscópicos: a estrutura de larga escala do universo, na qual a análise estatística de observações instrumentais esclarece os conceitos de inflação cosmológica e de perturbações primordiais, e o diagrama Hertzsprung-Russell, que relaciona os dados fotométricos e espectroscópicos das estrelas com as descrições atómicas e termodinâmicas dos corpos celestes. As representações visuais consideradas enquanto parte de uma boa prática epistémica permitem aceder aos constituintes fundamentais do mundo físico e classificar objetivamente os fenómenos naturais.

Introduction – astronomical insights into fundamental physics

At the core of the question of scientific realism lies the problem of objective representation: how does a product of scientific research, in particular a visual rendition, relate to the ontic nature of a natural phenomenon? The perception of the reality of electrons and atoms offers a canonical approximation to the matter, and astronomy, with its own set of epistemic peculiarities, is historically related to the received image of the subatomic world. In this paper, we will see how a single image of the polarization of the cosmic microwave background in a patch of sky has been recently presented, allegedly, as evidence of the existence of gravitational waves, as confirmation of cosmological inflation, and as a window to quantum gravity, confirming the ever increasing relation between observational cosmology and fundamental physics. We will also discuss the development by visual means of the spectral classification of stars, and how one century ago it not only gave rise to stellar astrophysics, but also offered crucial support to atomic and nuclear theory. Our aim is to assess the nature of visual arguments in science using the tools of epistemology and art theory.

Epistemology – the objective value of visual representations

In recent studies on the relation of science and art, the stress is put on the historicist aspects of scientific imaging. Lorraine Daston reports about the displayed items in eighteenth century cabinets, and points out how the boundary between art and nature was redrawn during this period, even as both categories remained as distinct as ever; nature objects became artificial (considering the playfulness or

"artificialness" of nature) and art objects became natural (as art imitates nature) in the eyes of their collectors.[1] Peter Galison reflects on the concept of objectivity as a historical notion, tracing back its genealogy to the practices of depiction in nineteenth century natural sciences: "truth to nature" idealised pictures made before 1800 were replaced after about 1830 by "objective" representations, i.e. mechanically produced homomorphic images; but a second displacement, brought by a repeated call for judgement in visual practices, saw the emergence of "interpreted" images in early twentieth century.[2] It seems as if visual practice would not only bring out subjectivity in any attempt to systematize the world, but it would also restrict the concept of objectivity to a vanished moment in the history of science, as has been restated by Daston and Galison in a more recent work.[3]

The rich and complex relations between image and nature are certainly present in scientific practice, but they appear also in many artistic representations. From 1821 to 1822, John Constable painted more than one hundred cloud studies and landscapes. In his *Hampstead Heath, Sun setting over Harrow* (figure 1a) he depicts the colours and shapes of clouds using a vivid palette and a dynamical brushstroke, and also notes on the back of the painting how the appearance of the landscape reflects the preceding heavy rains, as the north-west wind clears away the clouds at the end of the evening. In John Thornes' opinion, there is a good agreement between the painters' description and the atmospheric pressure records of 12 September 1821, which show a warm front in the morning and a vigorous cold front in the afternoon.[4] The accurate reporting of weather conditions is not the only meeting point for art and nature.

[1] Daston, 1998.
[2] Galison, 1998.
[3] Daston & Galison, 2007, 307.
[4] Thornes, 1999, 227.

Paul Klee started to collect plants in his childhood (figure 1b), and his interest in natural forms certainly inspired his paintings, but it is also noticeable in his theory of design, even before he taught the discipline at the Bauhaus school. In his journal, Klee states the need to reduce the means of artistic expression, in order to try to say more than nature does.[5] In the two artists, the development of a personal pictorial language allows the mimetic rendering of the light of the countryside, or the creation of abstract designs ultimately inspired in floral shapes.

Figure 1a: Constable, J., 1821, *Hampstead Heath, Sun setting over Harrow*, oil on paper laid on board (private collection). From Thornes, 1999, 227.

[5] Klee, 1908, 274.

Figure 1b: Klee, P., ca. 1890, *Plants on primed paper* (Zentrum Paul Klee). From Baumgartner & Keller, 2008, 55.

In this work, we will examine how visual practices can be used to describe and classify fundamental phenomena. We will analyze the role of astronomical images in two fields of physics from the perspective of philosophy of science, and we will find useful also to understand the aesthetic values implicit in any pictorial representation. Our methodology to study scientific practice borrows from two main sources: Phillip Kitcher states that epistemologists should use the history of inquiry as a laboratory to test their methodological claims, and opposes the current trend of philosophy to search for *a priori* principles.[6] Galison, again, observes scientists as an anthropologist, and contradicts assumptions from positivist and historicist epistemologies about how physicists conduct research; he explains

[6] Kitcher, 2011.

that theorists and experimenters coordinate locally their actions, even when they disagree in their beliefs, without using protocol languages or retreating to their epistemes.[7]

Mimesis – natural phenomena representation in a pictorial language

During the second half of the twentieth century, it became increasingly apparent that there were large structures in the Universe far beyond the size of the clusters of galaxies. The Harvard-Smithsonian Center for Astrophysics Redshift Survey began in 1977, and in 1986 the collaboration presented a map plotting the distribution of 1061 galaxies in a "slice of the Universe". The surprising results showed the galaxies distributed on filamentary surfaces surrounding large voids with typical diameters of 30 megaparsecs, posing serious challenges to the then current models for the formation of large-scale structure.[8] Recent observations have shown however that the largest structures observed are smaller than 200 megaparsecs, which is compatible with the homogeneity required in the cosmological models.[9] With the advances in computational power, a team of astronomers led by the Max Planck Institute for Astrophysics was able to simulate by 2005 the distribution of matter in a volume of space 600 gigaparsecs wide; the formation and evolution of 20 million galaxies was modelled in the study, which showed that large empirical surveys are likely to reflect physics in the early Universe.[10] The leading theoretical framework for the formation and evolution of structure includes, besides cosmological inflation,

[7] Galison, 1997, 802-803.
[8] de Lapparent & al., 1986.
[9] Nadathur, 2013.
[10] Springel & al., 2005.

dark energy and cold dark matter, in fact two place holders for as yet unknown physical entities, accounting for several observations with well established empirical support in astrophysics and cosmology.

The inflationary Universe theory was first introduced by Alan Guth in 1981, as a possible solution to the horizon problem (homogeneous conditions are required in causally disconnected regions of space) and the flatness problem (the initial value of the Hubble constant must be fine-tuned); these problems would disappear if the early Universe experienced a period of exponential growth, as the result of a phase transition (symmetry breaking in unified interaction models).[11] The origin of primordial inhomogeneities, the seeds for the future formation of structure in the Universe, can be explained by the quantum effects due to strong gravitational fields expected in the pre-inflationary stage. The gravitational radiation produced in this early age would also preserve information about the physics of the period, as predicted by Alexei Starobinsky.[12]

The possible detection of such primordial gravitational waves has been recently claimed by a team of astronomers, led again by the Harvard–Smithsonian Center for Astrophysics: a polarimetry experiment, installed at the BICEP2 radio telescope build in Antarctica, detected in a region of sky a pattern in the polarization of the cosmic microwave background radiation compatible with the signal of inflationary gravitational waves.[13] The measure showed a strong significance compared against the simulations, however, the possibility of interstellar dust emission bright enough to explain the signal cannot be excluded at the moment. The long sought polarization modes could be observed directly in the map, and not as a numerical outcome calculated from the statistical analysis of the

[11] Guth, 1981.
[12] Starobinsky, 1979.
[13] Ade & al., 2014.

data, which was one of the key elements in the public communication of the result. An image from the article (Figure 2a) shows the curl only polarization (B-mode), and was presented by science journalists as confirmation of inflation theory, evidence of the existence of gravitational waves, and the first glimpse of the quantum nature of gravity.[14] The same image was widely relayed by the general media as a portrait of our newly born Universe.

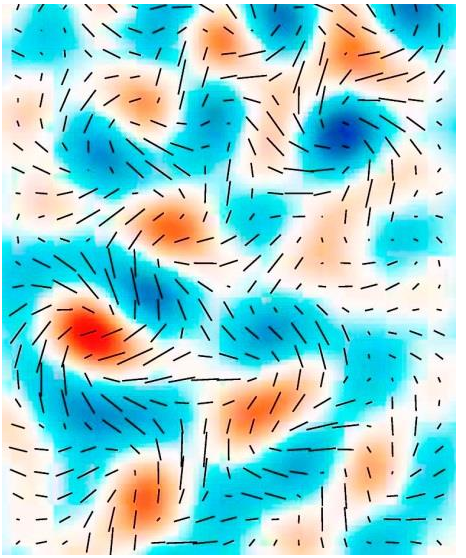


Figure 2a: BICEP2 Collaboration, 2014, B-mode map, figure 3, fragment (*Physical Review Letters*). From Ade & al., 2014.

This particular graph is quite intriguing, as two seemingly opposed factors concur in it. First, the graphic design of the map conveys quantitative information through symbolic language; and second, the image is perceived as the truthful representation of a physical phenomenon. Here we may find guidance in art theory, since artistic representation also confronts the increasing complexity of contemporary abstract languages. We can discuss for instance the

[14] Cowen, 2014.

work of Yves Klein, who refers to the poetics of Gaston Bachelard as an inspiration for his series about the natural elements.[15] In *Cosmogonies* (figure 2b), Klein reflects on what he calls the principles of an explanation of the Universe[16], and tries to capture the signs of the atmospheric phenomena, exposing for example the canvas to the weather. Klein's work in all its diversity can be seen as providing starting points for the *avant-garde* movements of the late twentieth century, and we may presume, for the aesthetic landscape of early twenty-first century. It is interesting in this sense to compare Klein's designs to BICEP2's maps, for their striking visual resemblance, but also for the tension between symbolic abstraction and mimetic immediacy in both of them.

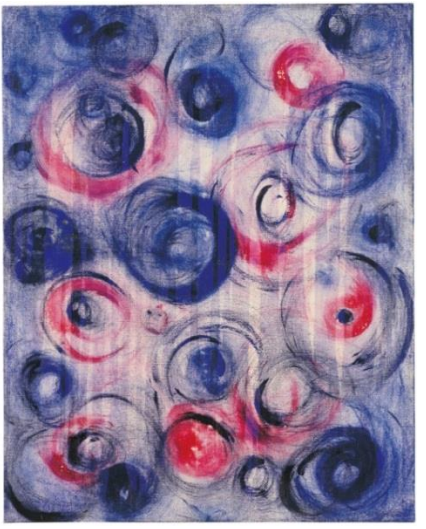


Figure 2b: Klein, Y., 1961, *Cosmogonie Bleue et Rose avec Traces de Vent, Sans Titre (COS 27)*, pure pigment and synthetic resin on paper mounted on canvas (private collection). From http://www.yveskleinarchives.org/. Accessed September 2014

[15] Puthomme, 1999.
[16] Klein, 1961.

The concepts of symbolic representation and mimetic representation may seem to be opposed, but they are not necessarily in contradiction. As Ernst Gombrich explained, Constable developed his personal language working through the exercises from a treatise on drawing by Alexander Cozens, practicing the conventions of the trade, but he still considered landscape painting as a science, part of natural philosophy, and the paintings as experiments.[17] Gombrich described the history of art as the forging of tools for opening our senses; thereby, pictures can be read in terms of natural objects to a degree that depends on our mental set.[18] Any language, with its arbitrary conventions, implies the introduction of some degree of artificiality in the representation of nature, but conventions do not entail its banishment from reality. Conceptual relativity, as Hilary Putnam calls it, is not incompatible with some kind of realism. Even if our concepts are culturally relative, the truth of everything we may say using those concepts does not depend only on the culture.[19] Following Putnam's definition, internal realism would be precisely the kind of realism compatible with conceptual relativity.[20] The precise identification and description of complex natural phenomena is crucial in contemporary science, and pictorial languages offer a useful set of tools that need not be rejected outright. We should keep in mind that the necessity to extract signal from noise, separate fact from artefact, is not exclusive to visual practices, as it encompasses the whole scientific process.

[17] Gombrich, 1960, 150-151.
[18] Gombrich, 1960, 304.
[19] Putnam, 1987, 20.
[20] Putnam, 1987, 17.

Atlas – visual practices and objectivity of natural classifications

The development of astronomical photography and spectroscopy at the end of the nineteenth century presented the occasion for the inclusion of spectral information in the stellar atlases. The Harvard College Observatory began to gather photographic plates in 1886 for the preparation of the Henry Draper Catalogue, published in its final version between 1918 and 1924, with the spectroscopic classification of 225 300 stars. Since the beginning of the project, graduate women were hired as "computers" to process the data, and in view of their expertise, some of them took a leading role in the definition of the spectral classification. In 1901, Annie Cannon published the Harvard classification scheme[21], still in use today. The spectra were divided in seven categories, with ten subcategories each; they were ordered in a continuous sequence, to make progressive changes in the aspect of atomic transition lines as gradual as possible (Figure 3a).

[21] Cannon & Pickering, 1901.

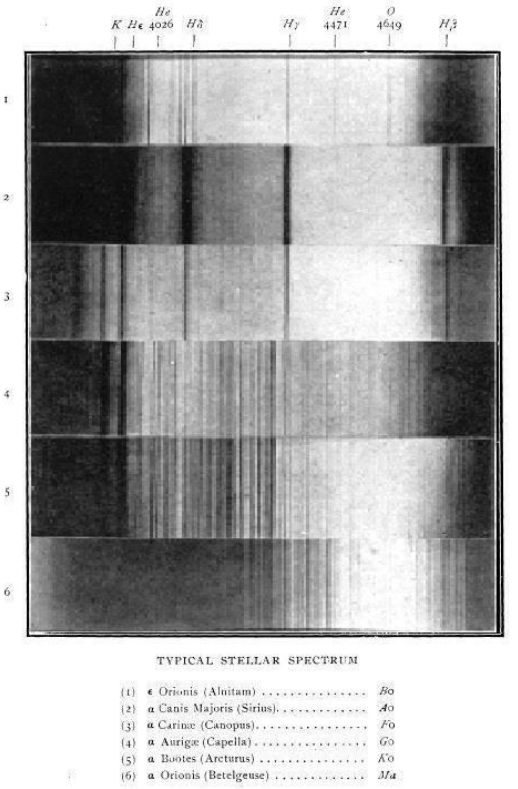


Figure 3a: Cannon, A., 1915, *Typical stellar spectra*, plate XVI (Journal of the Royal Astronomical Society of Canada). From Cannon, 1915.

The Harvard spectral classification was used by Ejnar Hertzsprung, along with his own estimates of absolute magnitude values, to identify two groups of stars, giants and dwarfs[22], which also displayed atomic lines with different widths inside the same spectral class. Working independently, and with much better absolute magnitude data, Henry Russell confirmed this effect. Hertzsprung's and Russell's diagrams of absolute magnitude versus spectral type showed the dwarf stars organized in a single linear sequence, and the giant stars grouped in a distinct branch. In 1914, Russell stated that the only physical

[22] Hertzsprung, 1909.

parameter responsible for the different spectral classes was the effective temperature of stellar atmospheres.[23] He also proposed tentatively a correlation between mass and luminosity along the main sequence. In 1926, Arthur Eddington published his treatise on stellar thermodynamics, including a derivation of the mass-luminosity relation which closely matched observational data.[24] In this influential book, based on general principles from quantum theory and relativity applied to radiation theory and atomic physics, he also postulated nuclear fusion as the probable source of stellar energy.[25] We can see that Eddington builds the foundations of stellar astrophysics using microphysics from early twentieth century as a tool, but we could also interpret that he uses astronomical data to validate fundamental physics.

We can claim that the atlas of stellar spectra published by the Harvard College Observatory lies at the heart of this foundational revolution. But we might question the objectivity of the atlas in the first place. Can we only declare as objective the mechanical production of spectral plates? Do the classification schemes reflect merely the subjective, albeit informed, interpretations of the "computers" and astronomers? Art history may provide again useful analogies. Aby Warburg composed the *Mnemosyne-Atlas* since around 1927, a picture series that combines historical and philosophical approaches, made with photographs arranged to illustrate the genealogy of the ideas behind Renaissance art. Warburg introduced visual language and analogical thinking into art history. For example, in the plates on astronomy (Figure 3b), Warburg traces the adoption of pagan images based in astronomy and astrology (the measurable and the mythological), from Classical Antiquity through the Middle Ages and

[23] Russell, 1914.
[24] Eddington, 1926, 151.
[25] Eddington, 1926, 314.

into the Renaissance.[26] Hence, the plates represent a mental map, the dynamics of the ideas of the author; his subjective, albeit engaging, vision of the material culture of art history. In what sense may Warburg's Atlas and the Harvard Atlas be different? Warburg's program constitutes an impossible archaeology, a project that would entail the completion of a set of particulars that ultimately embodies the universe of ideas in western art, whereas the endeavour of Cannon was to abstract a universal classification from the details of stellar spectra.

Georges Didi-Huberman has recently analysed Warburg's method through the general concept of atlas. As Didi-Huberman beautifully puts it, from the chaos of the world that comes to us in its disparity of prodigies, the sample gives us a chance to perceive a theoretical vision of the universal.[27] We could understand the atlas as a symbolic transition from the bestiary of the Middle Ages to the encyclopaedia of the Enlightenment, in which the discontinuity of free association opposes the continuity of natural classification. Accordingly, the pretence that a collection of mechanically produced pictures could escape subjectivity through lack of authorship does not seem consistent, since the mere fact of selecting the pictures would constitute already an interpretation; objectivity must be found elsewhere. But again, on what grounds can we trust abstractions built on instrumentally mediated observations? Ian Hacking relies on scientific practice, as practical ability breeds conviction to distinguish between instrumental artefacts and real observed structures.[28] Besides, the reliability of the representation is partially independent, as displays are often robust under changes of the instrument's theory: it is the engineering that counts.[29] Around 1900, spectroscopy, born in a

[26] Warburg, 1925.
[27] Didi-Huberman, 2010, 94.
[28] Hacking, 1983, 191.
[29] Hacking, 1983, 199.

lens-making workshop, became a key tool in thermodynamics and astrophysics, and was at the heart of the quantum hypothesis and atomic theory. This network, constructed from formal theories and experimental techniques, provided the framework to interpret the Harvard stellar atlas, and the empirical agreement between the morphological classification of spectra and the quantitative effective temperature of stars was proof of its objectivity.

Figure 3b: Warburg, A., ca. 1924, *Der Bilderatlas Mnemosyne, Planetenkinder*, panel 24 (Warburg Institute). From Warnke, M. (ed.), 2010, 41.

Conclusion – assessing complexity in observational physics

Research in contemporary physics depends on complex experiments, and demands the coordination of large groups of scientists and engineers. The astronomical surveys that are now undertaken require sorting through immense amounts of data, and must undergo exhaustive statistical analysis to pull out the signal from noise. The remoteness of the subject matters makes it difficult to empirically grasp their nature, and still the need to build a mental image of the physical processes remains. Visual practices are essential in this regard.

In conclusion, the artificiality introduced by experimental settings and representational techniques challenges the ability to identify natural phenomena, and the mimetic mapping of their observational features implies the use of pictorial languages, which also demand awareness of the conventions they impose. The systematic description of the observed phenomena, even as a set of images, relies on the judgment of an expert, and always implies interpretation as a first step in the process of abstraction from data; its empirical objectivity follows from the convergence of morphological and quantitative results in a network of theories and techniques, convergence which may eventually lead to a natural classification.

Acknowledgments

The author would like to thank Lídia Queiroz, organiser of the International Colloquium Atomism and Imagery from the 20th Century Onwards, and Olga Pombo, scientific coordinator of the Centro de Filosofia das Ciências da Universidade de Lisboa, for their kindness and support, and the participants in the conference, held the 3rd of November of 2014, for their suggestions and comments. The author is also grateful to an anonymous reviewer, whose comments contributed to improve the final version of this paper.

References

Ade, P. & al., 2014, Detection of B-Mode Polarization at Degree Angular Scales by BICEP2. *Physical Review Letters*, 112 (24), 241101 1-25.

Baumgartner, M. & Keller, M., 2008, *In Paul Klee's Enchanted Garden*, Osfildern, Hatje Cantz.

Cannon, A. 1915, The Henry Draper Memorial. *Journal of the Royal Astronomical Society of Canada*, 9, 203-215.

Cannon, A. & Pickering, E., 1901, Spectra of Bright Southern Stars. *Annals of Harvard College Observatory*, 28, 129-263.

Cowen, R., 2014, Telescope Captures View of Gravitational Waves. *Nature*, 507, 281-283.

Daston, L., 1998, Nature by Design. In: Jones, C. & Galison, P. (eds.), *Picturing Science, Producing Art*, New York, Routledge, 232-253.

Daston, L. & Galison, P., 2007, *Objectivity*, New York, Zone Books.

Didi-Huberman, G., 2010, *Atlas ¿Cómo llevar el mundo a cuestas?*, Madrid, Museo Nacional Centro de Arte Reina Sofía. (Aguilera, M., French to Spanish tr.)

Eddington, A., 1926, *The Internal Constitution of Stars*, Cambridge, Cambridge University Press.

Galison, P., 1997, *Image and Logic*, Chicago, University of Chicago Press
-- 1998, Judgment Against Objectivity. In: Jones, C. & Galison, P. (eds.), *Picturing Science, Producing Art*, New York, Routledge, 327-359.

Gombrich, E., 1960, *Art and Illusion*, London, Phaidon.

Guth, A., 1981, Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems. *Physical Review D*, 23 (2), 347-356.

Hacking, I., 1983, *Representing and intervening*, Cambridge, Cambridge University Press.

Hertzsprung, H., 1909, Über die Sterne der Unterabteilungen c und ac. *Astronomische Nachrichten*, 179, 373-380.

Kitcher, P., 2011. Epistemology Without History is Blind. *Erkenntnis*, 75, 505-524.

Klee, P., 1908, Tagebücher 834 München. In: Kersten, W. (ed.), 1988, *Tagebücher 1898–1918*, Stuttgart, Kunstmuseum Bern.

Klein, Y., 1961, *Manifeste de l'Hôtel Chelsea*.
http://www.yveskleinarchives.org/. Accessed September 2014.

de Lapparent, V. & al., 1986, A Slice of the Universe. *The Astrophysical Journal*, 302, L1-L5.

Nadathur, S., 2013, Seeing Patterns in Noise: Gigaparsec-Scale 'Structures' that Do Not Violate Homogeneity. *Monthly Notices of the Royal Astronomical Society*, 434 (1), 398-406.

Puthomme, B., 1999, Yves Klein : Bachelardien ?. *Philosophique*, 2, 53-60.

Putnam, H., 1987, *The many Faces of Realism*, Illinois, Open Court.

Russell, H., 1914, Relations Between the Spectra and Other Characteristics of the Stars. *Popular Astronomy*, 22, 275-294.

Springel, V. & al., 2005, Simulations of the Formation, Evolution and Clustering of Galaxies and Quasars. *Nature*, 435, 629-636.

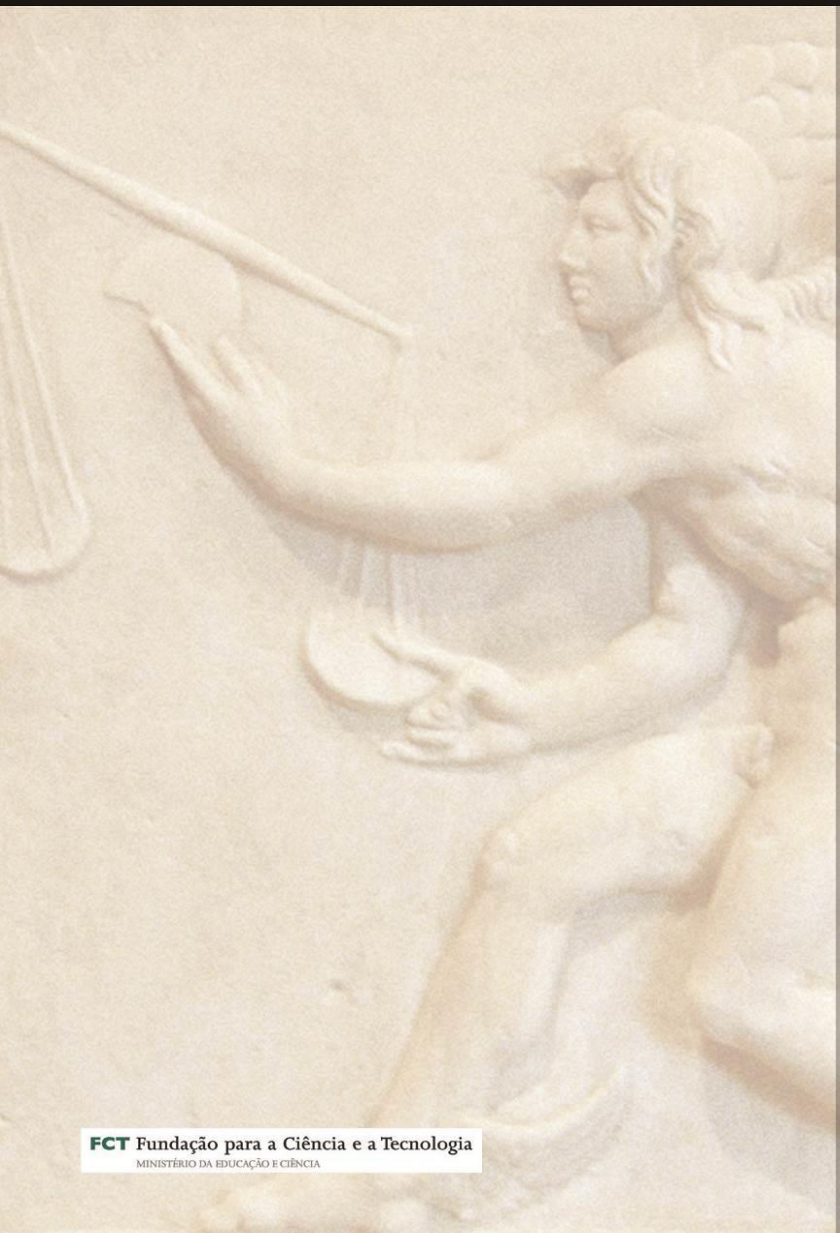
Starobinsky, A., 1979, Spectrum of Relict Gravitational Radiation and the Early State of the Universe. *JETP Letters*, 30 (11), 682-685.

Thornes, J., 1999, *John Constable's Skies*, Birmingham, University of Birmingham.

Warburg, A., 1925, The Franz Boll Lecture. In: Warnke, M. (ed.), 2010, *Atlas Mnemosyne*, Madrid, Akal, 165-172. (Chamorro, J., German to Spanish tr.)

Warnke, M. (ed.), 2010, *Atlas Mnemosyne*, Madrid, Akal.

ISSN (on-line): 1647-659X
ISSN (print): 2182-2824



Kairos
Revista de Filosofia & Ciência
Journal of Philosophy & Science
http://kairos.fc.ul.pt

CFCUL
Centro de Filosofia das Ciências da Universidade de Lisboa
http://cfcul.fc.ul.pt

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA