

Kairos

Revista de Filosofia & Ciência
Journal of Philosophy & Science



Setembro / September 2013

7

Artigos Papers

Em busca da alma no *Banquete*
Giovanni Casertano

Trolley Cases and Autonomy Violation
William Simkulet

John Mikhail on Moral Intuitions
Florian Demont

De Leibniz a Hegel: Força, Lei e Infinitude
na *Fenomenologia do Espírito*
Bernardo Enes Dias

Fabulosas raças de humanóides: monstros e robôs.
A robótica humanóide e a captura da intencionalidade
Porfírio Silva

Genuine Becoming and the Barcan Formula
Emiliano Boccardi

Nota de Abertura
María de Paz

A transversalidade do conhecimento na obra de Poincaré
Isabel Serra

O convencionalismo de Poincaré contextualizado: origem e
significado
María de Paz

Henri Poincaré, ciência e materialismo: o papel das hipóteses
na oscilação entre atomismo e continuísmo
André Carli Philot & Antonio A. P. Videira

As novas concepções da matéria
Henri Poincaré (1854-1912)

Jeremy Gray, *Henri Poincaré. A Scientific Biography*, 2013
María de Paz

Dossier Poincaré. Filosofia, Ciência e Materialismo

Recensão Book Review

Kairos. Revista de Filosofia & Ciência
Kairos. Journal of Philosophy & Science

ISSN: 1647-659X

**Direcção
Editors**

Olga Pombo
(Faculdade de Ciências
da Universidade de Lisboa)

Nuno Melim
(CFCUL)

Comissão Editorial / Editorial Board

Catarina Pombo Nabais
(CFCUL)

Lídia Queiroz
(CFCUL)

María de Paz
(CFCUL)

Nuno Jerónimo
(CFCUL)

Comissão Científica / Scientific Board

Andrea Pinotti
(Università degli Studi di Milano)

Angel Nepomuceno
(Universidad de Sevilla)

Byron Kaldis
(Hellenic Open University)

Danièle Cohn
(Université de Paris X)

Francisco J. Salguero
(Universidad de Sevilla)

John Symons
(University of Texas, El Paso)

José Nunes Ramalho Croca
(Faculdade de Ciências
da Universidade de Lisboa)

Juan Manuel Torres
(Universidad Nacional de Cuyo, Argentina)

Juan Redmond
(Universidad de Valparaíso, Chile)

Marcelo Dascal
(Universidade de Tel-Aviv)

Rudolf Bernet
(Husserl-Archives Leuven:
The International Centre
of Phenomenological Research)

Shahid Rahman
(Université de Lille)

Edição: Centro de Filosofia das Ciências da Universidade de Lisboa

Índice / Index

Resumos / Abstracts	5
Artigos / Papers	
Em busca da alma no <i>Banquete</i> Giovanni Casertano	13
Trolley Cases and Autonomy Violation William Simkulet	35
John Mikhail on Moral Intuitions Florian Demont	49
De Leibniz a Hegel: Força, Lei e Infinitude na <i>Fenomenologia do Espírito</i> Bernardo Enes Dias	61
Fabulosas raças de humanóides: monstros e robôs. A robótica humanóide e a captura da intencionalidade Porfírio Silva	87
Genuine Becoming and the Barcan Formula Emiliano Boccardi	113
Dossier: Poincaré. Filosofia, Ciência e Materialismo	
Nota de Abertura María de Paz	131
A transversalidade do conhecimento na obra de Poincaré Isabel Serra	135
O convencionalismo de Poincaré contextualizado: origem e significado María de Paz	151
Henri Poincaré, ciência e materialismo: o papel das hipóteses na oscilação entre atomismo e continuísmo André Carli Philot & Antonio A. P. Videira	167
As novas concepções da matéria Henri Poincaré (1854-1912)	187
Recensão / Book Review	
Jeremy Gray, <i>Henri Poincaré. A Scientific Biography</i> , Princeton/Oxford, Princeton University Press, 2013 María de Paz	205

Resumos / Abstracts

Em busca da alma no *Banquete*

Giovanni Casertano

Contrariamente ao que acontece noutros diálogos, onde há uma descrição de “o que é a alma”, no *Banquete* nada disso acontece. Naturalmente, a alma não está ausente do diálogo, aparece em todos os discursos, mas como um simples “nome”, que reenvia para algo conhecido de todos, sobre o qual não é preciso demorar-se. Quanto ao tema da alma, comparando o diálogo a uma sinfonia, diríamos que escutamos uma série de cinco tempos em crescendo, nos quais o mesmo tema é tocado em claves diferentes, discursos diferentes num “recitado” *punctum contra punctum*, que representam cada um a negação e a continuação do anterior, numa ampliação do horizonte sinfónico que prepara a sua acme, num largo maestoso, constituído pelo discurso de Sócrates/Diotima. Conclui a sinfonia/banquete, uma retomada presto agitato do tema: o discurso de Alcibíades. Assim, esta sinfonia polifónica abre uma perspectiva fundamental para compreender o problema da alma no pensamento platónico.

Trolley Cases and Autonomy Violation

William Simkulet

In this essay I argue that traditional solutions to the trolley problem, named for a series of cases made infamous by Philippa Foot and Judith Jarvis Thomson, fail and offer a new solution. The problem is that we seem to have conflicting intuitions in cases which are *prima facie* analytically similar, traditionally cases where we have to choose between the life of one or five people. Recently Thomson has argued that our intuitions in some of these cases are wrong; I argue she fails to show this. Then I argue that the most popular attempt to explain the existence of this distinction, the doctrine of double effect, is false. I argue that what makes the difference in these cases is whether or not acting or refraining would be an autonomy violation.

John Mikhail on Moral Intuitions

Florian Demont

John Mikhail's moral psychology is an interesting contribution to philosophical debates surrounding the nature of normativity and moral and legal judgement. The paper initially focuses on Mikhail's metaethical assumptions and how they are combined with the Chomskian framework of his moral theory. Particularly the computational processes which are supposed to generate oughts will be scrutinised. It is then argued that – apart from three other issues – Mikhail does not provide a satisfactory answer to the is-ought problem.

De Leibniz a Hegel: Força, Lei e Infinitude na *Fenomenologia do Espírito*

Bernardo Enes Dias

Apesar do relativo esquecimento a que foi votada por parte da mais recente *Hegelforschung*, a argumentação desenvolvida no terceiro capítulo da *Fenomenologia do Espírito* representa um dos momentos mais interessantes de toda a progressão fenomenológica. O que está em causa na crítica hegeliana às noções de lei e força é nada menos que uma denúncia da defectividade intrínseca de todo o pensamento científico convencional, assente numa estrutura de dedução matemática que não consegue ir além de uma mera remissão à igualdade. E a riqueza dessa denúncia, encoberta pela complexidade algo labiríntica do texto, só pode ser apreendida à luz de um retrato histórico mais vasto, que abarque a revolução dialéctica iniciada por Leibniz, o seu contributo para o projecto crítico kantiano e a sua inclusão no empreendimento fenomenológico hegeliano.

Fabulosas raças de humanóides: monstros e robôs. A robótica humanóide e a captura da intencionalidade

Porfírio Silva

Este texto discute as perspectivas da Robótica Humanóide no quadro historicamente alargado das visões humanas acerca dos humanóides. As crenças ocidentais nas raças fabulosas do Oriente, bem como aspectos da crise dessas crenças por ocasião dos Descobrimentos dos séculos XV e XVI, são usadas para iluminar o significado mais profundo de aspectos recentes desta linha de investigação em Nova Robótica. É proposto que a dinâmica profunda dessa Robótica passa pelo que designamos como captura da postura intencional.

Genuine Becoming and the Barcan Formula

Emiliano Boccardi

Presentism, the doctrine that, necessarily, only presently existing entities are real, is widely thought to be the only ontology compatible with a coherent dynamical picture of the passage of time. This paper contains an argument to the contrary. Presentists, it is argued, are faced with a dilemma: either (1) they acknowledge that their account of change is not consistent (perhaps embracing some form of dialetheism), or (2) they give up the idea of using their conception of change as the foundation for their account of passage.

Artigos / Papers

Em busca da alma no *Banquete**

Giovanni Casertano
(Università degli Studi di Napoli Federico II)
casertan@cds.unina.it

O *Banquete* é um daqueles diálogos que, à primeira vista, parece simples e linear na sua estrutura narrativa (como o *Crátilo*, por exemplo) mas que, pelo contrário, é extremamente complexo e, além disso, extraordinariamente refinado, não só na construção das personagens e dos seus “estilos” discursivos, mas também pela série de reenvios, de antecipações, de menções e alusões entre um discurso e o outro. Isso vê-se logo desde o início: o “prólogo”, com os seus reenvios a contos e a contos de contos, abre espaço para problemáticas sobre o tempo (a grande distância entre o diálogo representado e o diálogo contado por Apolodoro) e sobre o lugar (a estrada que leva do Falero até à cidade, o teatro, a casa de Ágaton, no primeiro e no segundo dias da narração); os vários discursos das personagens (mas não de todas!), que se criticam umas às outras, até ao discurso de Sócrates (que na realidade não é um verdadeiro discurso, mas sim, após uma breve introdução, o resumo de outro diálogo), discursos que acabam por refluir todos, por uma afirmação ou por outra, no discurso socrático; o discurso final de Alcibiades, aparentemente “desafinado”, mas na realidade ligado ao de Sócrates/Diotima; até à última cena, na madrugada do dia seguinte, que antecipa um dia “normal” de Sócrates. De tudo isto, porém, não pretendo falar aqui.

* Comunicação apresentada no XI Simpósio Internacional da Sociedade Brasileira de Platonistas, Simpósio de Platão, VII Seminário Internacional Archai, organizado pela Universidade Federal de Pernambuco, que decorreu em Recife de 17-20 de Maio de 2011. A versão italiana deste ensaio foi publicada em Anastácio Borges de Araújo Jr. & Gabriele Cornelli (a cura di). *Il Simposio di Platone: un banchetto di interpretazioni*. Napoli: Loffredo Editore University Press 2012, 181-93.

Ao invés, pretendo ocupar-me de uma ausência no diálogo: a da alma. Contrariamente ao que acontece noutros diálogos (*Fédon*, *Fedro*, *República*, *Timeu*), onde há uma teorização, ou pelo menos uma descrição, de “o que é a alma”, no *Banquete*, diálogo dedicado ao amor, ao contrário do que esperaríamos, nada disso acontece¹. Naturalmente a alma não está ausente do diálogo, aparece mais ou menos em todos os discursos das várias personagens, mas apenas como um simples “nome”, que reenvia para algo bem conhecido de todos e sobre o qual não é preciso demorar-se mais: em suma, algo adquirido, que basta nomear para que todos percebam do que se trata. De facto, esse algo é nomeado por todos. E ainda assim, se quiséssemos comparar – no que diz respeito ao tema da alma – o diálogo a uma sinfonia, poderíamos dizer que estamos a ouvir uma série de cinco tempos em *crescendo*, nos quais o mesmo tema é tocado em claves diferentes, discursos diferentes num “recitado” *punctum contra punctum*, que representam simultaneamente cada um a negação e a continuação do anterior, numa ampliação do horizonte sinfónico que prepara a sua acme, o sexto tempo, num *largo maestoso*, constituído pelo discurso de Sócrates/Diotima. Conclui a sonata/sinfonia/banquete de discursos uma retomada do tema em *presto agitato*, constituída pelo discurso de Alcibíades. E no final, descobriremos que esta sinfonia polifónica e inteiramente tocada em contraponto abre espaço para uma perspetiva fundamental para a compreensão do *problema alma* no pensamento platónico.

Vamos, pois, à procura da alma no *Banquete*. O discurso de Fedro é, talvez, o mais simples, no que diz respeito ao nosso tema. A sua tese é que Eros se encontra no amante. Os amantes querem morrer na vez dos amados (179b5), porque o deus inspira ardor (179b1: μένος ἔμπνευσαι)² em alguns

¹ É sintomático que no seu livro sobre a psicologia platónica, que releva as características da alma nos diálogos platónicos, Robinson (1995) não fale do nosso diálogo: menciona brevemente dois passos, 208b (127: a referência é à frase «o que é imortal», que concerne mais ao divino do que à alma intelectual) e 212a (125: concorda com Hackforth, 1950, 43), que no passo há um ‘temporary scepticism’ sobre a imortalidade da alma; e 127: a frase «que discerne o belo em si do que o torna visível» é referida à alma intelectual imortal). O *Banquete* também não é tratado em Robinson, 2010.

² A reminiscência é homérica, também no que diz respeito ao termo: *Il. X* 482: Atena inspira força (ἔμπνευσε μένος) em Diomedes, que começa a matar os inimigos, avermelhando a terra; *XV* 262: Apolo ἔμπνευσε μένος em Heitor, que como um leão de grande juba (275) põe em fuga os Danaides, como uma fera que assusta as manadas de bois ou os rebanhos de ovelhas de noite (323-324).

heróis. E isto vale também para as mulheres: Alceste, cuja ação pareceu tão bela aos deuses, que concederam esse dom a pouquíssimos homens, ou seja, que a *alma* regressasse do Hades (179c7). Portanto, também os deuses honram maximamente a dedicação e a virtude (179d1-2: σπουδῆν τε καὶ ἀρετήν)³ relativamente ao amor. Orfeu, pelo contrário, foi mandado para trás de mãos a abanar após terem-lhe mostrado um φάσμα (179d3) da mulher, dado que ele se tinha mostrado fraco e efeminado, sem a coragem de morrer por Eros como Alceste, embora fosse um fabricante de expedientes (179d6: διαμηχανᾶσθαι)⁴ para conseguir entrar vivo no Hades, e fizeram com que a sua morte fosse causada por mulheres⁵. Como se vê, a alma só é nomeada quando se referem mitos e lendas, e nada se diz acerca do seu papel no amor.

Com o discurso de Pausânias, porém, a alma começa, pelo menos, a conquistar o seu papel numa escala de valores. Pausânias fala de dois Eros e, por conseguinte, de duas Afrodites⁶: a *pandémios* é dos homens de pouco valor (181b2: φαῦλοι); eles amam os corpos mais do que as *almas* (181b4), e as menos inteligentes possível (181b4-5: ὡς ἂν δύνωνται ἀνοητοτάτων), por isso, no seu nascimento (181c1) ela participa quer da fêmea, quer do macho. A Afrodite urânica participa só do macho, é mais antiga e não tem devassidão⁷. Os que são inspirados (181c5: ἐπίπνοι) por este amor preferem o que por natureza é mais forte e tem mais inteligência (181c6: νοῦν). No amor pelos rapazes (181c7) pode reconhecer-se os que são estimulados de

³ Começa aqui com Fedro a série de atribuições de virtudes, qualidades, à alma, que em seguida terão uma caracterização mais completa nos outros discursos.

⁴ Esta capacidade, que aqui é considerada uma qualidade negativa, depois será considerada um dos atributos positivos de Eros no discurso de Sócrates (203d).

⁵ Reenvio explícito à *Alceste* de Eurípides (vv. 357-362), onde Admeto diz que, se tivesse a voz de Orfeu, desceria ao Hades para arrancar Alceste às garras da morte. Sobre a morte de Orfeu há várias versões. Aqui, no discurso de Fedro, Orfeu desce ao Hades por soberba para com os deuses, soberba que se vê também noutras versões sobre a sua morte: após a desafortunada tentativa, ele começaria a desprezar todas as mulheres e, ao impedir o culto de Dioniso na Trácia, foi despedaçado pelas Bacantes; ou então, por ordem mesma de Dioniso, porque preferira o culto de Apolo ao seu; ou ainda, foi fulminado por Zeus por ter revelado segredos dos deuses. Sobre a morte de Orfeu cf. Bernabé 2005, fr. 1032-1051.

⁶ Cf. Nucci, 2009, n. 80: sabemos que em Atenas existiam mais cultos a Afrodite (Pausânias I.22.3, VI.25.1, VIII.323.2, IX.16.3) e provavelmente o atributo de *Pándemos* dizia respeito à divindade mais conhecida e comum; a posição de Pausânias adquiriria portanto um sentido aristocrático.

⁷ Cf. Nucci, 2009, notas 82 e 85: as características dos dois Eros são perfeitamente contrárias.

maneira pura (181c7: εἰλικρινῶς), porque amam os rapazes quando eles começam a desenvolver a inteligência. Os que começam a amá-los a partir deste momento, estão dispostos a viver toda a vida junto com eles (181d5: κοινῇ συμβιωσόμενοι), e não a abandoná-los, após os terem enganado. Quem é bom (181e3) impõe a si mesmo esta lei (181e4). Os que amam indiscriminadamente têm desacreditado este amor e alguns têm a coragem de dizer (182a2: τολμᾶν λέγειν) que é feio conceder favores aos amantes (182a2-3)⁸. De facto, qualquer ação que seja realizada de maneira regrada e legal (182a5: κοσμίως καὶ νομίμως), nunca será criticada (cf. 183d)⁹.

O discurso de Pausânias é, com efeito, mais complexo do que o de Fedro. Podemos distingui-lo metodologicamente em duas partes. A primeira parte (180c-182a) abre-se com uma crítica explícita a Fedro¹⁰: «não me parece, ó Fedro, que o discurso tenha sido bem lançado (180c4-5: προσβεβλήσθαι¹¹): é que o amor não é só um, portanto, é mais correto (180c7: ὀρθότερον) decidir qual deles se deve elogiar; por isso, eu vou tentar retificar (180d2: ἐπανορθώσασθαι) este ponto, dizendo primeiro que amor se deve elogiar, e em seguida vou fazer um elogio condigno de um deus». Com Pausânias começa a surgir a questão do método através do qual enfrentar qualquer discussão, questão que terá um maior realce nos discursos seguintes. A segunda parte (182a-185c) é a leitura que Pausânias dá do νόμος ποικίλος περὶ τὸν ἔρωτα (182a7-b1) em Atenas, Esparta, Élide e Beócia, na Jónia e noutros lugares governados pelos bárbaros (182b). É interessante a consideração, embora citada como sendo opinião da massa (183b5-6), que o mais extraordinário (183b5: δεινότατον¹²) é que os juramentos traídos obtenham o perdão dos deuses, porque não existem juras de amor (183b7: ἀφροδίσιον ὄρκον)¹³. Em seguida, Pausânias fala de Atenas e das suas normas mais belas, embora não sejam fáceis de compreender (182d5: οὐ

⁸ Aqui poderíamos ler uma alusão ao discurso de Lísias no *Fedro*.

⁹ A pederastia pode ser distinguida em boa e má.

¹⁰ De facto, o tema sobre o qual entreter-se fora proposto por Erixímaco, mas ele retoma explicitamente uma questão de Fedro (cf. 177a-d), e o tema recebe logo a aprovação de Sócrates, que declara só saber (177d8: ἐπίστασθαι) de ἐρωτικά. Também esta sabedoria socrática revelar-se-á depois superficial e ilusória no diálogo com Diótima, que até chega a pôr em dúvida a capacidade socrática de seguir o seu discurso (210a). Realça justamente a importância da crítica metodológica a Fedro, Nucci, 2009, *ad loc*.

¹¹ De προσβάλλω, que significa lançar, pôr à frente de.

¹² Nucci, 2009, traduz «grave»; prefiro a tradução de Cambiano (1981). Mas a ambiguidade é própria do termo δεινός.

¹³ Reminiscência de Hesíodo fr. 124.

ῥάδιον κατανοῆσαι). Mas as críticas e os louvores, como ele disse (182a), devem ser relativizados: não há nada que em si e por si (183d5: αὐτὸ καθ'αὐτό) seja belo ou feio (183d5-6). Porque feio é quando se concedem favores a alguém malvado de maneira malvada (183d7: πονηρῶ τε καὶ πονηρῶς), belo quando se concedem de maneira bela a alguém excelente (183d7-8: χρηστῶ τε καὶ καλῶς)¹⁴. Malvado é o amante *pandémios* que ama o corpo mais do que a *alma* (183e1), e não é constante, dado que ama algo não constante (183e2: οὐδὲ μόνιμός ἐστιν, ἅτε οὐδὲ μόνιμου ἐρῶν πράγματος): de facto, a flor da juventude murcha, «evola-se e desaparece»¹⁵. Pelo contrário, quem se apaixona pelo *caráter* excelente (183e5: ἦθους χρηστοῦ)¹⁶, assim permanece por toda a vida, pois se funde (183e6: συντακεῖς¹⁷) com algo que é constante. Portanto, é preciso ver a qual das duas espécies pertence o amante, não conceder-se velozmente, fazer passar um pouco tempo, pois é isso que põe à prova (184a7: καλῶς βασανίζειν¹⁸) a maior parte das coisas. Só assim nasce a nobre amizade (184b4-5: γενναίαν φιλίαν): portanto, resta uma só via, com base na nossa lei (184b5: μία λείπεται τῷ ἡμετέρῳ νόμῳ ὁδός¹⁹), que o amado conceda favores ao amante de maneira bela. E para nós também é costume que os amantes se submetam a todo o tipo de escravidão (184c1: ἡγνινοῦν δουλείαν) desejada pelos amados, porque só há um tipo de escravidão que não é passível de crítica, a περὶ τὴν ἀρετὴν (184c3). Assim, tornamo-nos melhores (184c5) em qualquer tipo de sabedoria e em qualquer parte da virtude: é preciso fazer convergir para o mesmo ponto (184c8: συμβαλεῖν εἰς ταυτόν) o amor pelos

¹⁴ Alguns estudiosos, a propósito deste passo, falaram de “formalismo moral”: cf. Centrone, 2009, XVII.

¹⁵ A citação é homérica: *Il.* II 71, onde a expressão se refere ao sonho.

¹⁶ A associação do caráter à alma regressa no discurso de Ágaton, em 195e4.

¹⁷ O verbo συντήκω, fundir, saldar, regressa em 192e8, no discurso de Aristófanes. Como se vê, os reenvios de um discurso a outro pelos participantes no banquete são contínuos; menciono só alguns. Também no discurso de Pausânias, por exemplo, em 183a6, a notação dos amantes que dormem ἐπὶ θύραις regressará no discurso de Sócrates-Diotima em 203b5, d2.

¹⁸ O verbo é importante em Platão. Fundamentalmente, tem dois sentidos: torturar ou usar uma pedra de toque: aqui, provavelmente, é usado com o segundo sentido, como em *Resp.* III 413e1; para o outro sentido cf. *Soph.* 237b2.

¹⁹ Também esta poderia ser uma citação, desta vez retirada de Parménides: μόνος δ' ἔτι μῦθος ὁδοῖο λείπεται, DK28B8.1-2. Um Pausânias “parmenidiano” na ética? Para o νόμος como tradição, costume, mais do que lei escrita, cf. Nucci, 2009 *ad loc.*, e Casertano, 1971, *passim*.

jovens e o amor pela sabedoria (184d1: φιλοσοφίαν²⁰) e por todas as virtudes.

Como se vê, o discurso de Pausânias já nos diz muito mais acerca da alma do que o de Fedro. Antes de mais, no que diz respeito ao estilo, o discurso é muito mais refinado e, teoricamente, muito mais sólido. De facto, insere-se num eixo que vai de Heródoto a Protágoras, a Pródico e a Górgias, eixo que oferece duas perspectivas epistemicamente fortes: a consideração *relativista* dos costumes e das tradições culturais, diversos no tempo e no espaço, consideração por sua vez inserida no quadro das discussões sofisticadas sobre a relação νόμος-φύσις; e a consideração da *neutralidade* das técnicas, que leva a privilegiar, no juízo ético, o uso que se faz de qualquer coisa, ou o modo e a finalidade que levam a ter uma certa atitude²¹. A tese da supremacia da alma sobre o corpo implica também que se atribuam à alma certas características que se podem inscrever no horizonte platónico: de facto, a sua supremacia deve-se não só à posse genérica da “virtude”, mas também ao realçar que a virtude é fruto do empenhamento e do esforço em perseguir as coisas boas que caracterizam o homem valente. Aquilo que o amante deve buscar no amado é principalmente o νοῦς, a inteligência (181b4-5, 181c6), e a inteligência é o que torna constante e caracteriza a constância no amor bom (182e2), é o que representa a excelência do carácter (183e5), quase como se constituísse uma identificação do carácter com a alma. Mas, sobretudo, a alma caracteriza-se e qualifica-se por ser depositária de σοφία (184c5, 184e2), de φιλοσοφία (183a1, 184d1; cf. 182c1), de φρόνησις (184d7): uma clara antecipação da assimilação socrático-platónica de amor com filosofia.

Também o discurso de Erixímaco se inicia com uma crítica ao discurso de Pausânias; isto porque Pausânias, depois de se ter lançado (185e7: ὀρμήσας) de modo belo na exposição do seu discurso, não o concluiu de modo satisfatório (185e7-186a1): «eu [Erixímaco] tentarei dar-lhe uma conclusão (186a1: τέλος)». No que diz respeito à questão da alma, também para Erixímaco o discurso de Pausânias parece limitado: Pausânias distinguiu bem a duplicidade de Eros, mas a sua miopia fê-lo afirmar que Eros existe apenas nas *almas* (186a3) dos homens dirigidos aos belos, quando ele existe também em muitas outras coisas, isto é, nos *corpos* de todos os animais e dos vegetais, e por assim dizer ἐν πᾶσι τοῖς οὔσι (186a6-

²⁰ O termo já tinha aparecido em 183a1, mas fora eliminado por Schleiermacher (porquê?).

²¹ Cf. Casertano, 2010a.

7). Isto, Erixímaco declara ter obtido da sua técnica, a medicina, graças à qual ele pôde ver como Eros é grande e maravilhoso, como abrange e enche tudo (186b1: ἐπὶ πᾶν τείνει²²), as realidades humanas e as divinas (186b1-2). Se por detrás do discurso de Pausânias se encontravam os sofistas, por detrás do discurso de Erixímaco encontra-se Empédocles, com a sua reflexão sobre o valor cósmico da força do Amor, presente na realidade inteira – humana, divina, natural –, que determina a união e a harmonia dos opostos ou, melhor, dos diversos²³.

O discurso de Aristófanes é também claramente polémico face ao do seu predecessor; aliás, com aberta ironia, o “cómico” Aristófanes suplica a Erixímaco, por duas vezes (193b6-7; 193d7-8), que «não transforme em comédia» o seu discurso. Com Aristófanes, a alma volta a assumir o seu papel de protagonista na eterna comédia do amor. Numa espécie de “contracanto” relativamente ao de Erixímaco, que tinha dado todas as explicações fisiológicas e científicas que de Eros se podem dar, o discurso de Aristófanes, muito trabalhado do ponto de vista estilístico e dramático, volta a realçar o aspeto misterioso do amor, desta força que, apesar de tudo, permanece sempre algo de enigmático para o homem. A duas páginas do início encontramos a imagem esplêndida do ἕκαστος ἡμῶν ἀνθρώπου σύμβολον (191d3-4), de cada homem que anda à procura da própria tésseira. A explicação científica da ὁρμή é substituída pelo belíssimo mito da separação do ente originário que infatigavelmente procura reconstituir a unidade perdida²⁴. E esta é a razão pela qual quem ama quer passar toda a

²² O verbo τείνω significa não só tender, expandir, e na forma média expandir-se; mas também – e aqui está na forma ativa – esticar até à tensão, encher penetrando; sentido que alarga aquela ὁρμή (cf. a ὁρμήσας de 185e7) que Pausânias dedicara, retoricamente, só ao seu discurso.

²³ A propósito da expressão «o dissemelhante ama o dissemelhante» de 186b6-7, Centrone, 2009 explica da seguinte maneira: «le cose *tra loro* dissimili desiderano cose *tra loro* diverse» (XIX n. 32). Segundo este estudioso a harmonização dos opostos proposta por Erixímaco é inadequada na perspetiva platónica, no sentido em que vê concórdia e discórdia só numa sucessão temporal. Na perspetiva socrática o desejo erótico é simultaneamente desejo daquilo de que sente falta, logo, do que é dissemelhante, e do que na sua diversidade é profundamente semelhante e afim: desejo do semelhante e simultaneamente desejo do dissemelhante (XX).

²⁴ Como é sabido, Aristófanes oferece também uma classificação, que é simultaneamente uma explicação, mítica naturalmente, da *diversidade* dos amores que comovem os homens: 1. os homens que são o resultado da divisão do ser comum – 191d6: τοῦ κοινοῦ – são amantes das mulheres, enquanto que as mulheres são amantes dos homens; 2. as mulheres que são o resultado da divisão de mulheres são orientadas para as mulheres; 3. os homens que são o resultado da divisão dos

vida junto com a pessoa amada, amantes que *nem sequer saberiam dizer o que desejam receber um do outro* (192c2-4). O que desejam não é só a união sexual (192c5: συνουσία); na verdade, é claro que a *alma* de ambos quer algo mais, que não sabe dizer o que é, por isso, prevê o que deseja e exprime-o *por enigmas* (192c7-d2: ἀλλ' ἄλλο τι βουλομένη ἑκατέρου ἢ ψυχῆ δῆλη ἐστίν, ὃ οὐ δύναται εἰπεῖν, ἀλλὰ μαντεύεται ὃ βούλεται, καὶ αἰνίττεται)»²⁵. O que se deseja é, pois, unir-se, e ao fundir-se com o amado, deseja tornar-se um só de dois que eram (192e10-193a1: συνελθὼν καὶ συντακείς τῶ ἔρωμένῳ ἐκ δυοῖν εἷς γενέσθαι); ora, dá-se nome de amor ao desejo e à busca do inteiro (192e10-193a1: τοῦ ὅλου οὖν τῆ ἐπιθυμίας καὶ διώξει ἔρωσ ὄνομα).

Se Erixímaco usara o Empédocles cientista, podemos dizer que Aristófanes usa o Empédocles poeta. Faz-se referência explícita ao fragmento 63, onde aparece precisamente o termo σύμβολον²⁶. Mas todo o discurso de Aristófanes é um discurso poético, que mistura habilmente elementos cómicos e trágicos²⁷, como se estivesse a anunciar alusivamente

machos amam os homens, e são os melhores dentre os rapazes e os adolescentes, porque preferem o que lhes é semelhante – 192a5: τὸ ὅμοιον – e lhes é congénere – 192b5: τὸ συγγενές.

²⁵ «L'inesprimibilità del sentimento erotico, l'impossibilità di dirne la potenza travolgente e il punto oscuro a cui spinge, è probabilmente uno dei picchi più alti del discorso»: Nucci, 2009, n. 143.

²⁶ DK31B63 (Arist., *de gen. an.* 722b10): em rigor, o termo encontra-se no comentário de Aristóteles, mas muito provavelmente era empedocliano: no macho e na fêmea existe respetivamente como parte de uma marca (σύμβολον). Cf. Nucci, 2009, n. 138. Nucci anota que a referência, no discurso de Aristófanes, às mulheres que amam mulheres é a única referência explícita na literatura grega clássica à homossexualidade feminina (n. 138).

²⁷ A alusão a Hefesto que, no conto de Aristófanes pergunta aos amantes se desejam ser aprisionados juntos e passar assim toda a vida, e porventura passar assim o tempo depois da morte (192d-e), é uma clara citação do aprisionamento de Ares e Afrodite com laços muito resistentes mas invisíveis, e do seu descaramento perante todos os outros deuses (cf. Hom., *Od.* VIII 266-332), uma citação que contém em si elementos simultaneamente cómicos e trágicos. Tal como a pincelada que pinta o abraço mortal das duas metades resultantes da separação do ente originário, aquele abraço inconsciente e interminável que, entrelaçando as duas metades que desejam unir-se, leva-as inevitavelmente à morte (191a-b), é outro exemplo de imagem trágica no discurso do comediógrafo. Cf. Nucci, 2009, n. 145: «l'abbraccio eterno è un'unione priva della tensione propria di eros, un'unione mortale, soffocante, in cui l'eros vivo si spegne insieme al desiderio (cf. Nussbaum, 1986, pp. 171-176). Fare di due uno porta a questo, secondo Platone, che di eros proporrà invece una lettura tutta basata sulla sua funzione dinamica e dialettica». Não concordo, porém, com a última observação: transformar dois num só, e em geral constituir uma unidade da multiplicidade, é uma ideia tipicamente platónica, da política (*República*) à gnosiológia (*Sofista*), segundo a

a síntese de comédia e tragédia que Sócrates defenderá no final do diálogo, perante os últimos dois interlocutores que não perderam as forças apesar de terem bebido demasiado vinho: Aristófanes o comediógrafo, precisamente, e Ágaton o tragediógrafo.

Ágaton oferece uma caracterização diferente da alma de quem é tomado pelo Amor, realçando aspetos diversos dos evidenciados nos discursos anteriores. Mas também o seu discurso começa com uma aberta crítica aos que o precederam: os outros não elogiaram o deus, falaram sobretudo dos benefícios que Eros concede aos homens (194e). Ele, pelo contrário, falará – e isto, segundo Ágaton, é claramente o *τρόπος ὀρθός* (195a1) com o qual se deve fazer um elogio – antes de mais, do que (195a2: οἶος) é Eros e em seguida dos efeitos que causa (195e2). É sintomático ver que esta *metodologia* do elogio é exatamente aquela que Sócrates invocará depois, mas para criticar precisamente Ágaton²⁸. Pois então; Eros: 1) é o mais jovem dos deuses; em todos os mitos antigos narrados por Hesíodo e por Parménides as proezas foram obra de Necessidade e não de Eros, eis a razão das mutilações, dos aprisionamentos e de todo o tipo de violências entre os deuses (195c). Embora com explícita reserva sobre a *veridicidade* daqueles contos²⁹, aqui há o reconhecimento indireto da função cosmológica e cosmogónica de Eros, dos Pitagóricos a Empédocles, além da sua importância para a passagem da geração espontânea dos seres vivos à geração sexuada. Mas agora que Eros reina sobre os deuses há *φιλία καὶ εἰρήνη* (195c5-6; cf. 197b). 2) é delicado (195c7, e1), mas não mora em qualquer um, só nos entes mais ternos (195e3: ἐν τοῖς μαλακωτάτοις τῶν ὄντων³⁰): fixa-se nos feitios e nas *almas* (195e4: ἐν ἧθεσι καὶ ψυχαῖς³¹) dos deuses e dos homens, mas não em todas as *almas*, de facto, se encontra um feitio (195e6: ἦθος) duro, afasta-se. 3) tem uma forma fluida/flexível/gentil (196a2: ὑγρὸς τὸ εἶδος), senão não poderia penetrar e sair das almas sem que elas se apercebesse³². Da sua forma harmoniosa e flexível possui a

qual o ἐν ἐκ πολλῶν não anula a diversidade dos componentes, mas precisamente dialética e dinamicamente constitui-os numa harmonia.

²⁸ Cf. Centrone, 2009, XXIII-XXIV.

²⁹ «Admitindo que eles digam a verdade» (195c3: εἰ ἐκεῖνοι ἀληθῆ ἔλεγον).

³⁰ Estes entes, baseando-nos em tudo o que se dirá depois, só podem ser as almas.

³¹ Quase uma *hendíadis*, que realça que o feitio do homem é a sua alma, como acontecia no discurso de Pausânias, em 183e5.

³² Aqui talvez se faça uma menção ao ἦλθε γάρ, ὡς ἦλθε de Górgias, *Hel. en.* 19 = DK82B11, 19.

formosura/grança (196a5: εὐσχημοσύνη). 4) é justo (196b); 5) é temperante (196c); 6) é corajoso (196d); 7) é sábio (196d-e); 8) é o fabricante das técnicas (197a-c); 9) institui familiaridade (197d1: οικειότης). 10) é ornamento (197e2: κόσμος) de deuses e homens, fascina o seu pensamento (197e4-5: θέλω... νόημα³³). E assim, conclui Ágaton, «ofereci ao deus este λόγος, que participa, em parte, de brincadeira, em parte, de mensurada seriedade (197e7: τὰ μὲν παιδιᾶς, τὰ δὲ σπουδῆς μετρίως... μετέχων)».

Portanto, a alma, no discurso de Ágaton, que também oferece, como no discurso de Aristófanes, uma mistura de brincadeira com seriedade³⁴, mostra-se dotada de qualidades que não apareciam nos discursos anteriores. Parece-me claro que todo o seu discurso, exceto a qualidade mencionada em 197a-c³⁵, faz explícita referência ao *Elogio de Helena* de Górgias, em especial ao facto de o sofista entender o fascinar como obra de sedução e de engano das almas sensíveis e não grosseiras³⁶. E, detalhe estilístico significativo, no discurso de Ágaton aparecem a primeira e a última palavra do *Elogio de Helena*³⁷.

Chegamos assim ao auge do *largo maestoso* recitado pelo duo Sócrates-Diotima. Falei deste diálogo no diálogo no diálogo num estudo escrito há catorze anos³⁸; aqui vou falar dele só no que diz respeito ao tema da alma. A sessão é aberta por Sócrates, com uma clara mas não breve crítica a *todos* os discursos que o precederam (198a-199b), em que Sócrates reconhece que se disseram coisas *belas*, sim, mas não *verdadeiras*³⁹. Todavia, a crítica socrática visa sobretudo o discurso de Ágaton, embora reconheça explicitamente a justeza do método que ele utilizou⁴⁰. Também aqui a crítica

³³ O verbo θέλω significa também seduzir, enganar.

³⁴ Lembro ainda que Aristófanes e Ágaton são os últimos dois interlocutores de Sócrates no diálogo narrado.

³⁵ Que se encontra na *Defesa de Palamedes* de Górgias, DK82B11a.

³⁶ Também aqui o reenvio ao fragmento de Górgias sobre os Tessálios é evidente: cf. 82B23a. O fragmento não se acha em DK, mas foi acrescentado justamente por Untersteiner, 1961, II, 142, com base principalmente em DK82B23.

³⁷ Ágaton usa o termo παιδιᾶ, enquanto que em Górgias temos παίγνιον, e κόσμος: a última e a primeira palavra do *Hel. en.*

³⁸ Casertano, 1997.

³⁹ 198d-e: mas não podemos estabelecer com certeza se este reconhecimento socrático é irónico ou não. Cf. Casertano, 2010b, 4.4. Sobre o argumento, cf. Rosen, 1987, 205 sgg.

⁴⁰ 199c3-5: «Caro Ágaton, pareceu-me que tu te saíste muito bem (καλῶς) no teu discurso, quando disseste que em primeiro lugar se devia mostrar o que era (ὁποῖός τις ἔστιν) Amor e depois os seus efeitos (ἔργα)».

se faz por uma série de perguntas e respostas entre Sócrates e Ágaton, em que se estabelecem outros elementos importantes na determinação semântica de Eros: o facto de ele ser um termo relativo (199d), a perspectiva do desejo, da privação e da posse (200a-e), da beleza e da identidade entre bom e belo (201a-c). Até se declarar, na última troca de frases entre os dois, mais uma vez, que aquilo que agora se dirá será unicamente a verdade⁴¹. Sócrates autoapresentasse como quem dirá finalmente *a verdade* sobre o Amor; como é sabido, o discurso verdadeiro não será pronunciado por Sócrates, mas, em contraponto ao que ele realçara no diálogo com Ágaton, será feito pela sacerdotisa de Mantineia, noutra contraponto com Sócrates. Nele, Sócrates desempenha exatamente o mesmo papel que Ágaton havia desempenhado durante o diálogo com Sócrates, e será refutado precisamente da mesma forma que Sócrates o fora no diálogo com Diotima.

Mas debrucemo-nos sobre o nosso tema, começando pela secção em que se examina de que modo e em que atividade (206b1-2: τίνα τρόπον... καὶ ἐν τίνι πράξει) se pode reconhecer a presença de Amor em quem o busca. Já se tinha dito que os homens desejam o bem, desejam possuí-lo, possuí-lo sempre (206a); ora, se substituirmos o bem pelo belo⁴², podemos dizer que o amor é a *procriação no belo* (206b7-8: τόκος ἐν καλῷ), quer segundo o corpo, quer segundo a alma. De facto, todos os homens estão grávidos e, numa certa idade, a nossa natureza deseja dar à luz (206c3). Isto é que é divino e é isto que é *imortal no vivente mortal* (206c7: ἐν θνητῷ ὄντι τῷ ζῶνι ἀθάνατον ἔνεστιν): engravidar e gerar (206c7), gerar e procriar (206e5) no belo. É precisamente aqui que Platão alarga maximamente o horizonte da alma, ligando-a ao divino e à imortalidade. Mas o que foi pouco realçado é que esta imortalidade é sempre uma imortalidade mortal, é a imortalidade concedida aos seres mortais: e, começando precisamente pela expressão citada em 206c7, durante o discurso os dois termos θνητός e ἀθάνατος estarão sempre conjuntos. Dado que a geração é o que renasce (206e8: ἀιγενές) sempre e é *imortal enquanto mortal* (206e8: ἀθάνατον ὡς θνητῷ), pelo que se concordou obtém-se necessariamente (207a1-2: ἀναγκαῖον... ἐκ τοῦ ὡμολογημένων), 1) que se deseja *a imortalidade* através do *bem* (207a1: μετὰ ἀγαθοῦ); 2) que amor é desejo de posse contínua do bem, e deste discurso obtém-se ainda necessariamente que 3) o amor é amor também de imortalidade (207a3-4).

⁴¹ 201c8-9: «É a verdade, caro Ágaton, que não podes contradizer (ἀντιλέγειν), porque contradizer Sócrates não é nada difícil». Para um conceito análogo, cf. *Phaed.* 91c, *Charm.* 161c.

⁴² Para o parentesco entre ἀγαθόν e καλόν cf. Centrone, 2009, LI-LIII.

Numa primeira fase, a imortalidade é uma característica também da corporeidade e, como tal, é prerrogativa de todos os seres vivos, homens, animais e plantas: através da geração física cada ser vivo ganha uma espécie de imortalidade, que é claramente a imortalidade da espécie e não de cada indivíduo. É *eros* que determina este tipo de imortalidade, porque é *eros* que determina aquela atitude terrível (207a8: δεινῶς διατίθεται) de todos os animais, que quase adoecem quando agem amorosamente (207b1: ἐρωτικῶς διατιθέμενα)⁴³. Platão não deixa de realçar a *relatividade* desta maneira de ser imortal: de facto, só deste modo a natureza mortal (207d1: ἡ θνητὴ φύσις) procura (ζητεῖ), *por quanto é possível* (207d1: κατὰ τὸ δυνατόν), ser sempre e, por conseguinte, ser *imortal* (207d1-2); e pode sê-lo *somente* por esta via (207d2), com a geração, pois deixa sempre outro ser novo no lugar do velho (207d3).

Mas esta imortalidade relativa dos corpos cede logo o lugar a outra imortalidade relativa, precisamente a das almas. E se a alma é a coisa mais importante do ser vivo mortal, e se o conhecimento é a coisa mais importante para a alma humana, temos de reconhecer que além das nossas imortalidades relativas há outra imortalidade não relativa de que nós homens mortais podemos participar: a imortalidade do mundo do conhecimento. Introduce-se esta imortalidade com uma esplêndida analogia entre vida/morte do homem e vida/morte do conhecimento. Tal como de cada indivíduo se diz sempre que é idêntico (207d7: αὐτὸς κελεῖται), mesmo que nunca possua em si as mesmas coisas, aliás nasce/torna-se sempre novo (207d7: ἀλλὰ νέος ἄει γιγνόμενος), não só no corpo mas também na alma, é mais surpreendente ainda (207e5), diz Platão, que isto aconteça também com as ἐπιστήμαι (207e5); não só, mas também com *cada* ἐπιστήμη. É preciso cuidar de cada uma delas, é preciso exercitar-se (208a4: μελετᾶν) continuamente em cada uma delas, pois desaparece: de facto, o esquecimento é a saída de um conhecimento (208a4-5: λήθη γὰρ ἐπιστήμης ἕξοδος), enquanto que a *μελέτη*, ao reproduzir em nós uma nova recordação (208a6: μνήμη), conserva (208a6: σῶζει) o conhecimento de modo a que ele *pareça* o mesmo (208a6-7: ὥστε τὴν αὐτὴν δοκεῖν εἶναι)⁴⁴. A analogia é profunda e muito bela: cada homem é mortal, no corpo e na alma, mas tem a *possibilidade*, na sua vida

⁴³ Note-se também o uso, com o mesmo verbo, dos dois advérbios δεινῶς e ἐρωτικῶς, que realçam a força terrível do amor.

⁴⁴ Para uma análise das páginas 207-208 cf. Rosen, 1987, 253 segs.: a inseparabilidade de alma e corpo, que significa a ausência de uma doutrina de imortalidade pessoal, significa que alma e corpo se encontram numa condição de perene fluxo (254). Sobre os fluxos da alma cf. Sassi, 2007.

mortal, de alcançar uma forma de imortalidade, que consiste precisamente em elevar-se ao mundo imortal dos conhecimentos. Com efeito, eles vivem da vida e da morte dos indivíduos: a sua vida imortal é construída pelas contínuas contribuições de cada indivíduo, que renovam e enriquecem sempre um património que é comum a todos os que são capazes de alcançar esse mundo. Esses homens, se alcançarem esse mundo, atingem também a imortalidade, que dura enquanto, e pelo tempo que, o indivíduo for capaz de habitar naquele mundo. E esta imortalidade mortal – parece sugerir Platão entrelinhas – é superior porventura à dos deuses. Isto pode ser lido nas linhas finais deste passo, onde, ao reafirmar que só deste modo se conserva para sempre tudo o que é mortal (208a8), isto é, deixando outro ser jovem tal qual o anterior (208b1-2), outro corpo, se se é fecundo só no corpo, e uma ideia nova, se se é fecundo também na alma. Esta imortalidade dos conhecimentos é uma imortalidade que é sempre diferente mesmo sendo sempre idêntica a si mesma enquanto conhecimento, que se enriquece e se amplia continuamente; é uma imortalidade diferente da imortalidade divina, que acaba por ser sempre monótona e absolutamente idêntica a si mesma (208a8: τῷ παντάπασιν τὸ αὐτὸ ἀεὶ εἶναι)⁴⁵. Com isto, que é um verdadeiro expediente (208b2: μηχανῆ), *o mortal participa da imortalidade* (208b3: θνητὸν ἀθανασίας μετέχει).

Este tempo da sinfonia é tocado com o fundo de um horizonte amplo e sempre mais abrangente, onde a obtenção da imortalidade é fruto de exercício e de esforço, é o resultado de uma atividade digna. Portanto, a imortalidade é característica fundamental da alma. Mas não pertence a todas as almas, somente às que souberam empenhar-se no cuidado das coisas mais elevadas e belas e souberam elevar-se, na sua vida mortal, à contemplação e ao gozo das coisas imortais, contribuindo assim, cada uma com a sua individualidade e atividade, à amplitude e à vida daquelas coisas imortais. O resto do tempo é apenas uma série de *variações* sobre o tema, onde se exalta a glória imortal de virtudes imortais de personagens cantados pela poesia (208d)⁴⁶, de poetas e também de artesãos inventores, de criadores de cidades, de todos os que puseram em prática sabedoria e justiça (209a8); reafirma-se a superioridade da alma relativamente ao

⁴⁵ O mesmo conceito será reiterado pouco depois, em 208b4.

⁴⁶ Expressão solene do verso lembrado em 208c5-6: κλέος ἐς τὸν αἰὶ χρόνον ἀθάνατον καταθέσθαι, verso de autor desconhecido, mas quem sabe de Platão mesmo, em que se vê claramente uma recordação do κλέος ἀέναν de Heraclito, DK22B29.

corpo⁴⁷; reafirma-se que o que mais convém à alma são a inteligência e todas as outras virtudes (209a3-4: φρόνησίν τε καὶ τὴν ἄλλην ἀρετήν)⁴⁸. Reafirma-se, por fim, que o amor leva a cumprir belas atividades, e não implica uma contemplação estática e afasia: de facto, quando alguém está grávido *na alma* (209b1) deseja dar à luz e gerar, aproxima-se de *uma alma* bela, nobre e bem dotada (209b6: καλῆ καὶ γενναία καὶ εὐφυεῖ), e imediatamente tem facilidade (209b8: εὐπορεῖ) em compor discursos sobre a virtude, criando assim uma comunhão (209c5: κοινωνίαν) muito maior que com os filhos e uma amizade mais firme, dado que as duas almas estão unidas (209c7: κεκοινωνηκότεσ) por filhos mais belos e imortais. Esta é, para Platão, a verdadeira imortalidade: a de almas belas que, graças a Amor, são capazes de gerar os filhos mais belos⁴⁹.

⁴⁷ Cf. também 210b7; no nosso passo há ainda uma notação significativa a propósito dos *erotikoi* fecundos só no corpo (208e2), que obtêm recordação, imortalidade e felicidade, por meio da geração, *assim pensam eles* (208e4: ὡς οἴονται).

⁴⁸ A mesma fórmula no discurso de Pausânias, em 184d7-e1.

⁴⁹ A questão da imortalidade da alma, e os problemas que lhe dizem respeito, continua a ser sempre muito debatida pelos estudiosos. Para me limitar só à questão de uma imortalidade que Platão estaria a atribuir a toda a alma ou apenas à alma racional, e para me limitar aos estudos mais recentes, cf. Vegetti, 1985, que, ao realçar a mudança na reflexão platónica de um horizonte órfico-pitagórico (*Fédon*) para uma teoria que quebra o dualismo alma/corpo (*República*), vê a passagem de uma imortalidade da alma inteira para uma distinção entre partes mortais e parte imortal da alma no *Timeu* e nas *Leis* (210-213). Brisson, 2007, percebe a aparente contradição platónica acerca da alma: imortal na sua totalidade (*Fedro*), ou só no intelecto (*Timeu*)? Brisson defende que, visto «l'anima in quanto tale (si veda *Timeo* 35a-b) è ... una realtà intermedia tra il sensibile e l'intelligibile [e quindi] non può costituire l'oggetto né dell'intelletto (νοῦς), poiché non è una Forma, né di alcuno fra i sensi, perché non è una realtà sensibile particolare», dela pode falar somente um tipo de discurso que não é possível declarar como verdadeiro ou falso: este tipo de discurso é o mito» (26). Poderíamos objetar a Brisson que, em geral, os discursos sobre a alma são mitos (sobre a 'paradoxalidade' do mito nos diálogos platónicos, cf. Ferrari, 2006, 11-64., Ferrari, 2007), mas: 1) o passo do *Fedro* a que se faz referência, na nota 6, 26, é uma autêntica demonstração (o que move a si mesmo); 2) também o sensível é sempre objeto do intelecto, cf. *Teeteto*. Para Brisson é «l'anima nel suo insieme ad essere immortale, non un'anima particolare» (33). «L'immortalità individuale pertanto è legata alla capacità che l'anima possiede di ricordare l'intelligibile: per questo essa è limitata nel tempo, ciò che invece non è vero dell'anima in quanto tale» (34). Para Centrone, 2007, face a posições que reivindicam a Platão uma imortalidade *tout court* da alma e a de, por exemplo, Dorter, 1982, 36-44, 162-175, que exclui a imortalidade da alma entendida como única entidade individualizada e, por conseguinte, lê todas as afirmações platónicas afins em sentido metafórico e mítico, seria necessário assumir uma posição intermédia. Deve-se excluir a imortalidade pessoal, mas pode-se pensar na existência de uma pluralidade de almas individuais imortais (p. 36). O estudioso sugere a possibilidade que também a imortalidade da alma seja uma das “nobres mentiras” de que Platão fala na

Aqui não me demorarei na chamada *scala amoris* que começa em 210a, nem na suposta distinção entre “Grandes Mistérios” e “Pequenos Mistérios”⁵⁰;

República: «Platone ha un fortissimo interesse ad argomentare l'immortalità personale per ragioni etico-politiche; che l'anima di ciascun uomo sia immortale è condizione necessaria perché, in ambito escatologico, una pena eterna o un premio eterno possano esercitare una funzione rassicurante (...) o deterrente, come accade nei miti dell'aldilà del *Gorgia*, del *Fedone* e della *Repubblica*; d'altro canto nella *Repubblica* si teorizza, in determinate condizioni e per determinate persone, la liceità del mentire per fini superiori, il che rende plausibile l'ipotesi che l'immortalità personale possa essere annoverata tra le nobili menzogne. Ed è altrettanto naturale immaginare che tra i soggetti autorizzati a mentire Platone includa se stesso» (36-37; 47-48). No *Banquete* a imortalidade é só do género humano (38-39). Por fim, parece-me muito correto o que Centrone diz da imortalidade ligada à doutrina da metempsicose: «È la stessa forma in cui Platone professa questa dottrina a essere inconciliabile con l'immortalità personale, perché a sopravvivere è la stessa anima con contenuti di coscienza diversi e l'individuo di volta in volta presente è un individuo diverso dal precedente... Non c'è bisogno, del resto, di ricordare il carattere oggettivo e impersonale delle conoscenze rammemorare nel processo dell'anamnesi, così come è descritto nei luoghi classici del *Menone* e del *Fedone*. Qui si ripropone il problema di fondo, se Platone creda realmente nella metempsicosi o se si tratti invece di una dottrina da relegare in ambito mitico. In questo caso la compartimentazione dei dialoghi in sezioni 'filosofiche' e in sezioni 'mitiche' non porta molto avanti, se è vero che la metempsicosi e l'anamnesi, cioè una dottrina con forti implicazioni filosofiche, sono nei dialoghi inscindibilmente connesse» (41). Fronterotta, 2007, defende que as espécies mortais da alma «insorgono nel corpo all'innesto dell'anima immortale» (97), como reações psicofisiológicas corpóreas ao enxerto da alma num corpo mortal (especialmente 94-96). Isto não me parece muito claro: estas reações verificam-se na alma ou no corpo? Mas, em geral, Fronterotta defende que aí «sono fondate ragioni per dubitare che in Platone vi sia davvero, almeno dal punto di vista genetico della sua costituzione, una dottrina della tripartizione dell'anima» (99), e, por conseguinte, uma evolução do *Fédon* à *República*. Bastante ambíguo parece-me o que o estudioso diz sobre a diferença entre princípios imortais e princípios mortais: as recordações anteriores ao nascimento no corpo podem ser conservadas, porque são recordações que pertencem à alma *racional* e *imortal*, e que são esquecidas devido ao corpo, no momento em que a alma 'nasce'; enquanto que as recordações da vida mortal só podem ser dispersas quando o corpo morre, porque «sono ricordi acquisiti e posseduti dal corpo che si dissolve» (102). Os itálicos são meus e servem para realçar a ambiguidade de falar de reações, aqui de recordações, que são *do corpo*: o corpo não pode ter recordações, pois elas são só da alma: cf. *Teeteto*. Todavia, para Fronterotta «Platone non dichiara mai, neanche sotto forma di mito, che l'anima immortale ha un carattere in qualche senso personale» (108). Para Migliori, 2007b, pelo menos, cada alma intelectiva é imortal, mas não há verdadeira sobrevivência do sujeito. E não é um caso que, ao examinar este problema também noutros diálogos, o estudioso não faça referência ao *Banquete*. Também para Robinson, 2010, é significativamente assim, ao tratar o problema, não faz nenhuma referência ao *Banquete*, mas examina principalmente o *Fédon*, para concluir, que nenhum dos argumentos platónicos para imortalidade da alma é convincente.

⁵⁰ Segundo alguns estudiosos aos “Pequenos Mistérios” pertencia a parte inicial do discurso de Diotima, que pretendia libertar Sócrates dos falsos conceitos sobre Amor, aos “Grandes Mistérios” pertencia esta parte que inicia precisamente em 210a, com a

tão-pouco tratarei a ascensão da parte do homem – evidentemente de um homem dotado de alma bela – da contemplação das coisas belas à contemplação da ideia de belo⁵¹. Limitar-me-ei a reafirmar que o que se descreve é uma ascensão e não uma ascese: a aquisição da ideia de belo é exatamente a conquista da *capacidade de ver o belo* nas coisas belas, onde quer que ele se manifeste, não um momento que determina a afasia, mas que, pelo contrário, permite o começo de uma vida vivida num nível mais elevado⁵². E esta vida vivida desta forma implica necessariamente, como bem viu Stenzel, um crescimento coletivo e não um crescimento do indivíduo separado dos outros⁵³.

dúvida, por parte de Diotima, sobre a capacidade de Sócrates em acompanhar o seu raciocínio (cf. Centrone, 2009, XXVII-XXIX e XXXI-XXXIV, e Nucci 2009, notas *ad locc.* Sobre a *passividade* de Sócrates nesta fase do seu diálogo com Diotima, cf. Rosen, 1987, 264). A expressão “mistérios de amor” não me parece encontrar-se no texto platónico, onde encontramos ταῦτα τὰ ἐρωτικά em 209e5. Pessoalmente, não daria ênfase à suposta linguagem misteriosa (solene, sim), desta parte, nem à declarada (por Diotima) incapacidade de Sócrates em seguir o discurso (210a2, 210a4), que me parecem fazer parte da extraordinária capacidade platónica de construir nos seus diálogos discursos diversos quanto ao estilo e à qualidade, além de disseminar o texto de expressões que estilisticamente determinam o conjunto de uma atmosfera dramática.

⁵¹ A este propósito, veja-se Casertano, 1997.

⁵² Muito se discute sobre esta visão instantânea (cf. o ἐξαίφνης de 210e4; o termo é usado duas outras vezes, em 212c6: Alcibiades que toca à porta, e em 213c1: Alcibiades que sobressalta ao ver Sócrates sentado junto dele; sobre o uso deste advérbio em Platão cf. Casertano, 1998) do belo em si, e sobre o tipo de conhecimento, intuitivo ou discursivo, que dele se obtém. Pessoalmente, concordo com o que diz Centrone, 2009, L: «Il raggiungimento della meta... è preludio a una prolungata e incessante attività intellettuale, l’analogo della discesa nell’ultimo segmento della linea della *Repubblica* (511b7-c2). Un’attività, se si vuole, discorsiva, ma diversa dalla διάνοια menzionata precedentemente, perché ora il filosofo si muove esclusivamente nell’ambito delle idee». E parece-me que Centrone não concorde com Nucci, 2009, notas nas páginas 209-210, que insiste na diferença entre Pequenos e Grandes Mistérios, defendendo que nos primeiros o belo é o que permite a procriação de belos discursos e de virtudes, enquanto que nos Grandes o belo é somente objeto de contemplação. Parece-me, pelo contrário, que há uma continuidade num único caminho, no qual alguns podem parar a um certo ponto, e outros podem continuar; nem se pode chegar ao segundo nível sem ter sido educado no primeiro. Nem há contemplação do belo entendida como uma espécie de afasia, porque assim que se consegue ver a ideia do belo, o homem continua a fazer belos discursos, e procura, como faz precisamente Sócrates (212b), convencer os outros da verdade, naturalmente com outros discursos. Na verdade, nem sequer o místico pode eximir-se de fazer discursos, se quiser convencer outras pessoas: a afasia não convence ninguém.

⁵³ Stenzel, 1966: «Proprio l’asceta, due volte descritta attraverso tutti i gradi della aspirazione collettiva, dall’amore per un altro individuo a quello per tutti i membri della comunità, e alle belle imprese, e poi al sapere, e infine al μᾶθημα supremo, mostra

A nossa pesquisa conclui-se com o *presto agitato* de Alcibíades. Poderíamos dizer que o seu discurso, no que diz respeito ao nosso tema e pensando no que foi dito por Sócrates/Diotima, é a demonstração dramática da imortalidade de uma alma, a de Sócrates, e da mortalidade de outra, a do próprio Alcibíades. E isto porque o seu discurso desvela a *alma* de Sócrates e, em contraponto, o que a *alma* de Alcibíades não conseguiu ser. Já na descrição da sua entrada na sala do banquete Platão descreve o seu caráter animoso. De repente (212c6: ἐξαίφνης) ouviu-se um grande barulho à porta e a voz⁵⁴ de uma flautista (212c-d), e em seguida a voz de Alcibíades que gritava alto (212d4), completamente bêbedo, perguntando por Ágaton. Conduziam-no a flautista e “outros”⁵⁵ (212d6-7). E entra cingido por uma espessa coroa de folhas de hera e de violetas, a cabeça⁵⁶ coberta por uma infinidade de fitas, declarando que na noite anterior não pudera vir, mas que vinha agora para coroar o mais sábio e o mais belo⁵⁷ (212e). É embriagado, conhece a sua situação, mas reivindica logo que diz a verdade (213a1⁵⁸). Senta-se entre Ágaton e Sócrates, depois vê Sócrates e sobressalta-se: por que se encontra Sócrates aqui? Para «me preparares uma emboscada (213b9: ἐλλοχῶν⁵⁹) novamente, como costumás fazer, é que apareces de repente (213c1: ἐξαίφνης)», e por que se encontra junto do belo Ágaton?

Sócrates intervém pedindo a Ágaton que o defenda, porque para ele o amor de Alcibíades tornou-se um οὐ φαῦλον πρᾶγμα (213c7). Aqui pode-se ver claramente a costumeira ironia socrática, embora apareça desenhada no horizonte a diferença entre os dois modos de conceber o amor e, por conseguinte, dois modos diversos de fazer agir a alma: a de Sócrates, em que o amor pelo belo o leva a progredir no caminho para a ideia e o

questa crescita della conoscenza entro una comunità di sentimento e d'azione, come di una organica formazione radicata, cresciuta nella vita collettiva» (235).

⁵⁴ Nucci, 2009, *ad loc.*, traduz “o som”.

⁵⁵ Parece-me clara a referência a Heraclito, DK22B117: um embriagado é um homem conduzido por uma criança, porque a sua alma está “húmida”. Pelo contrário, a alma de Sócrates nunca está húmida, porque nunca ninguém viu Sócrates bêbedo, *cf.* 220a.

⁵⁶ Uma “caricatura” de Empédocles? *Cf.* DK31B112.5-6.

⁵⁷ Outra subtil ironia platónica, aliás uma dupla ironia: também Sócrates (174a) havia declarado não ter estado com Ágaton na noite anterior, mas agora ia honrar um belo; e Alcibíades faz a mesma declaração, pensando ir louvar Ágaton, acabando, pelo contrário, por tecer o elogio de Sócrates.

⁵⁸ Reafirmado em 214e6, e em 217e3-4, onde se encontra a metonímia do οἶνος ἀληθής, que é um entrelaçamento irónico de dois provérbios: “vinho e verdade” e “vinho e jovens são verdadeiros”.

⁵⁹ Ἐν-λοχᾶω: estar à coca, à espera.

conhecimento, e a de Alcibiades, que não é capaz de ir além do sentimento imediato, com tudo o que esse comportamento implica. De facto, desde que Alcibiades se encontra apaixonado por ele, Sócrates não pode olhar nem falar com um belo porque Alcibiades, roído de inveja e de ciúmes (213d2), denigre-o e por pouco não bate nele. E se pensarmos que no discurso de Diotima a contemplação do belo é um sentimento absolutamente destituído de inveja (cf. 210d6: ἐν φιλοσοφίᾳ ἀφθόνῳ), teremos já delineada a enorme diferença entre os dois modos de conceber o amor.

Após uma troca de frases, que envolve também Sócrates e Erixímaco, Alcibiades começa o discurso sobre Amor, louvando Sócrates, mas louvando-o δι' εἰκόνων (215a5), porque as suas imagens terão por objetivo a verdade (215a6). Não me demorarei no discurso de Alcibiades, a não ser para realçar aquilo que concerne ao nosso tema. Pode-se dizer que o seu discurso é a descrição dos efeitos dos discursos filosóficos sobre um jovem com uma alma bem-disposta, que se comove e perturba pelo discurso amoroso, facto que é sempre característica da verdadeira filosofia. Embora, no caso particular de Alcibiades, subsistam todas as qualidades para apreender a beleza daquele discurso, a sua alma “timocrática” é tão forte que o impede de «ser senhor de si mesmo». E se pensarmos na explicação que Sócrates oferece desta locução na *República*⁶⁰, podemos dizer que, no final, em Alcibiades prevaleceram os elementos piores.

Mas de que “sofre” (cf. 215d8: πάσχω) a alma de Alcibiades quando escuta o amável discurso filosófico de Sócrates? Permanece perturbado e inspirado⁶¹, o coração palpita-lhe e saltam-lhe as lágrimas (215e2-3)⁶², enquanto que quando escutava Péricles e outros bravos oradores pensava que falavam bem, mas não sentia (215e5: ἔπασχον) nada de semelhante, nem a sua alma se agitava (215e6: ἐτεθορύβητο), nem pensava que *não*

⁶⁰ Em IV 430e-431a: a temperança (430e6: σωφροσύνη) é uma espécie de ordem (430e6: κόσμος) e de domínio (430e7: ἐγκράτεια) sobre os prazeres e os desejos, como se diz quando se usa a expressão «senhor de si mesmo» (430e7: κρείττω αὐτοῦ). A expressão pode parecer ridícula (430e11: γελοῖον), porque quem é senhor de si mesmo é também escravo de si mesmo (430e12: καὶ ἦπτων... αὐτοῦ). Mas este *logos* parece-me querer dizer (431a3: φαίνεται μοι βούλεσθαι λέγειν), esclarece Sócrates, que na alma de cada homem convivem um elemento pior e um melhor, e quando o elemento melhor se impõe sobre o pior, então é-se «senhor de si mesmo», enquanto que quando o elemento melhor é dominado pelo pior, é-se «escravo de si mesmo».

⁶¹ 215d5-6: ἐκπεπληγμένοι ἐσμὲν καὶ κατεχόμεθα. Ἐκπλήσσω = fazer cair atingindo, e por conseguinte, atordoar, perturbar, agitar.

⁶² Mais uma referência ao *Elogio de Helena*, 9, de Górgias.

poderia viver no estado em que se encontrava (216a1). De facto, só os discursos de filosofia mudam a vida do homem, ou pelo menos, fazem-no entrever uma nova vida, mesmo que depois se continue a ser Alcibíades e não se tenha cuidado com a própria alma (cf. 216a). Ainda assim Alcibíades conseguiu ver a alma de Sócrates e percebeu a diferença entre *aparecer* e *ser*. Só a alma de Sócrates lhe fez sentir aquilo que ninguém acreditaria existir nele mesmo, a vergonha (216b2), e isto pelo maravilhoso poder (216c7: δύναντιν) que Sócrates tem. De facto, ele dispõe-se sempre amoravelmente para com os belos⁶³, fica perturbado, ignora tudo e não sabe nada (216d3-4). Mas dentro (como o Sileno⁶⁴) está cheio de σωφροσύνη (216d7). E Alcibíades, que se mantinha escondido na aparência por causa do seu *thymos*, achava que Sócrates considerava a sua beleza como os outros a consideravam e que lhe fazia discursos que os outros amantes costumam fazer ao amado. Portanto, não pôde fugir à doença (217e6: πάθος) de quem foi picado pela víbora. De facto, dizem que quem sofreu essa picada não a confessa a não ser a quem também já foi picado, como se fosse o único capaz de compreender e de se conder (217e8-218a1: γνωσομένοις τε καὶ συγγνωσομένοις). Alcibíades foi picado no ponto mais doloroso (218a3: ἀλγεινότερου) em que se possa ser picado... No *coração*, ou na *alma* (218a4), ou como a queiramos chamar, foi tocado e picado pelos discursos filosóficos, que penetram mais fortemente do que a picada de uma víbora⁶⁵, quando apanham a alma (218a6) de um jovem não destituído de dotes naturais e a obrigam a fazer e a dizer quaisquer coisas (218a6-7: καὶ ποιούσι δῶν τε καὶ λέγειν ὅτιούν)⁶⁶.

Não me parece estar enganado ao ver no desfecho do discurso de Alcibíades – que se dirige aos outros comensais, e os une à loucura e ao furor báquico típicos do filósofo (218b3-4: τῆς φιλοσόφου μανίας τε καὶ βακχείας), e por esse motivo são capazes de lhe prestar atenção (218b4: ἀκούσεσθε), ao contrário dos servos e das almas profanas e grosseiras que

⁶³ 216d2: ἐρωτικῶς διάκειται τῶν καλῶν. Διάκειμαι = estar com determinada disposição; com o genitivo, estar apaixonado. Portanto é mais complexo do que o ἐρωτικῶς διατίθεσθαι de 207a8 e 207b1.

⁶⁴ Cf. Brisson, 1998, 217-218, n. 526.

⁶⁵ Não concordo com Nucci, 2009, n. 331, quando diz que aquilo que Alcibíades compara à picada de uma víbora é a separação entre Pequenos e Grandes Mistérios, e, por conseguinte, a passagem de uma beleza que deve fazer procriar a uma beleza a contemplar.

⁶⁶ Mais uma vez Górgias, *Hel. en.*, 12: um discurso que tenha persuadido uma alma, obriga-a a acreditar nas palavras ditas (λεγόμενοις) e a consentir nos fatos (ποιουμένοις).

devem pôr «portas muito grandes nas próprias orelhas» (218b4: ἀκούσθε) – o remorso sentido de uma alma consciente de ter tido uma oportunidade única, mas que fez a escolha errada – embora possuísse boas qualidades – e que perdeu, por isso, a grande ocasião da sua vida, deitando-a a perder para sempre⁶⁷.

(Tradução de
Maria da Graça Gomes de Pina)

Referências bibliográficas

Brisson, L., 2007, Le parti mortali dell'anima o la morte come oblio del corpo. In: M. Migliori, L. Valditara Napolitano, A. Fermani (Org.), *Interiorità e anima. La psychè in Platone*, Milano, 25-34.

Brisson, L., 1998, *Platon, Le Banquet*, Paris.

Cambiano, 1981, *Platone, Dialoghi filosofici*, vol. II, Torino.

Casertano, G., 1997, Il (in) nome di Eros. Una lettura del discorso di Diotima nel *Simposio* platónico. *Elenchos*, XVIII, 277-310.

--- 1998, L'istante: un tempo fuori del tempo, secondo Platone. In: L. Ruggiu (Ed.), *Filosofia del Tempo*, Milano 1998, 3-11.

--- 2010a, *Sofista*, São Paulo.

--- 2010b, *Paradigmas da verdade em Platão*, São Paulo.

Centrone, B., 2007, L'immortalità personale: un'altra nobile menzogna? In: M. Migliori, L. Valditara Napolitano, A. Fermani (Org.), *Interiorità e anima. La psychè in Platone*, Milano, 35-50.

--- 2009, Platone, *Simposio*, tr. e comm. di M. Nucci; introd. di B. Centrone, Torino.

Dorter, K., 1982, *Plato's 'Phaedo'. An interpretation*, Toronto-Buffalo-London.

Ferrari, F., 2006, Filosofia e mitologia in Platone. In: Ferrari, F., *I miti di Platone*, Milano.

--- 2007, L'*anamnesis* del passato tra storia e ontologia. Il mito platonico come *pharmakon* contro utopismo e scetticismo. In: M. Migliori, L. Valditara Napolitano, A. Fermani (Eds.), *Interiorità e anima. La psychè in Platone*, Milano, 73-88.

⁶⁷ Cf. Rosen, 1987, p. 385: «In the *Symposium* Alcibiades sings the swan song of his genius, delivered at the height of his power but only a step away from disaster».

- Fronterotta, F., 2007, Che effetto fa essere un pipistrello? Il problema mente-corpo nel *Timeo* platónico. In: M. Migliori, L. Valditara Napolitano, A. Fermani (Eds.), *Interiorità e anima. La psychè in Platone*, Milano, 89-108.
- Giombini, S., Marcacci, F., 2010, *Il quinto secolo. Studi in onore di Livio Rossetti*, Città di Castello.
- Hackforth, 1950, *Immortality in Plato's Symposium. Classical Review*, LXIV, 43-45.
- Migliori, M., 2007a, Ma c'è interiorità nei dialoghi di Platone? In: M. Migliori, L. Valditara Napolitano, A. Fermani (Org.), *Interiorità e anima. La psychè in Platone*, Milano, 123-164.
- 2007b, La prova dell'immortalità dell'anima (608c-612c). In: Vegetti, *Platone, La Repubblica, Vol. VII, Libro X*, Napoli, 199-275.
- M. Migliori, L. Valditara Napolitano, A. Fermani (Org.), 2007, *Interiorità e anima. La psychè in Platone*, Milano.
- Nucci, 2009, Platone, *Simposio*, tr. e comm. di M. Nucci; introd. di B. Centrone, Torino.
- Nussbaum, M., 1986, *The fragility of goodness*, Cambridge.
- Robinson, Th.M., 1995, *Plato's psychology*, Toronto [1970].
- 2010, Socrates on soul and immortality. In: Giombini, S., Marcacci, F., 2010, *Il quinto secolo. Studi in onore di Livio Rossetti*, Città di Castello, 389-400.
- Rosen, S., 1987, *Plato's Symposium*, Ann Arbor [1968].
- Sassi, M.M., 2007, Eros come energia psichica. Platone e i flussi dell'anima. In: M. Migliori, L. Valditara Napolitano, A. Fermani (Eds.), *Interiorità e anima. La psychè in Platone*, Milano, 275-292.
- Settis, S. (Ed.), 1996-1998, *I Greci. Storia Cultura Arte Società*, 4 voll., I: *Noi e i Greci*, Torino.
- Trabattoni, F., 2007, Si può parlare di "unità" della psicologia platonica? Esame di un caso significativo (Fedone, 68 b - 69 e). In: M. Migliori, L. Valditara Napolitano, A. Fermani (Eds.), *Interiorità e anima. La psychè in Platone*, Milano, 307-320.
- Untersteiner, M., 1961, *Sofisti. Testimonianze e frammenti*, voll. 4, Firenze.
- Vegetti, M., 1985, *Anima e corpo*, in M. Vegetti, (Ed.), *Il sapere degli antichi*, Milano, 201-228.
- 1996, L'io, l'anima, il soggetto. In: Settis, S. (Ed.), 1996-1998, *I Greci. Storia Cultura Arte Società*, 4 voll., I: *Noi e i Greci*, Torino, 431-467.
- Vegetti, 2007, *Platone, La Repubblica*, Vol. VII, Libro X, Napoli.

Trolley Cases and Autonomy Violation

William Simkulet
Simkuletwm@yahoo.com

I. Introduction

The trolley problem is the problem concerning reconciling our *prima facie* inconsistent intuitions about two kinds of cases, both known as “trolley cases”, with the following basic forms:

1. *The Trolley Case* – A runaway trolley is about to kill five workers unless a bystander flips a switch to divert the trolley onto a different track where only one worker is working.¹

2. *The Fat Man Case* – A runaway trolley is about to kill five workers unless a bystander derails the trolley by pushing a fat man off of a bridge onto the track.²

In the trolley case, our commonsense moral intuition is supposed to be that they bystander should flip the switch, diverting the trolley, killing one person and saving the other five. In the fat man case, though, we’re supposed to have the opposite intuition, that killing one person to save five people would be morally abhorrent. How do we explain these competing moral intuitions?

Perhaps the simplest solution is to conclude that one of these intuitions is wrong. Don Marquis once told me something that has, against all odds, stuck – that we cannot trust our moral intuitions in strange cases. Of the two, the fat man case strikes me as quite strange indeed, if only because it’s hard to imagine a sufficiently fat man perched so precariously on a bridge that a bystander could successfully push him off and derail a trolley. As such, we

¹ This case was originally proposed by Philippa Foot (Foot, 1967/1978). For the sake of this paper, suppose these workers have thick, protective ear-wear that prevents the bystander from getting their attention.

² This version of the case was proposed by Judith Jarvis Thomson (Thomson, 1985).

might be inclined to conclude that our intuitions in this case are less reliable than in the original trolley case, which strikes me as quite plausible. Consider now a third case:

3. *The Transplant Case* – Trapped in a mountain hospital by a storm, a doctor has five patients who will die unless they each get an organ transplant. A healthy (and medically compatible) bike messenger, a former patient, enters the hospital to get out of the storm. The doctor realizes that if she kills the messenger, she can save the lives of the five people in need of organ transplant.

As in the fat man case, we're supposed to have the overwhelming intuition that killing one to save five would be gravely morally wrong. Furthermore, this case strikes me as comparably familiar to the trolley case, and in each of these cases we have strong moral intuitions that appear to contradict each other.

In the next two sections of this article, I argue that two popular approaches to resolve the problem fail. Judith Jarvis Thomson originally attempted to explain our conflicting intuitions in these cases in terms of rights violations, contending that although acting to turn the trolley in the trolley case violated the rights of the single worker, the bystander was justified in doing so.³ Dissatisfied, recently Thomson has argued that our intuitions in the trolley case are wrong; that it is morally wrong to divert the trolley.⁴ In section II, I argue this approach by Thomson is unsatisfactory, and fails to show that turning the trolley is wrong. Section III concerns the doctrine of double effect (DDE). According to DDE, it may be morally acceptable to bring about some unintentional bad consequences (in the trolley case, the death of the single worker) if doing so is not intentional, not a means to a desired end, and the desired end is sufficiently good as to override the wrongness of the bad effect. I show that the DDE rests upon there being a morally relevant distinction between acting and refraining when all else is equal, and this distinction is false. I argue James Rachels has succeeded in showing that, all else being equal, acting and refraining are morally equivalent.

In the final section of this article, I offer an alternative explanation and justification for our intuitions in trolley cases. I contend that in cases like the fat man and transplant case, the agent would be wrong to kill the fat man or messenger without consent because doing so would constitute a grievous

³ Thomson, 1976, 1985. For a good analysis of Thomson's positions here, see Gorr, 1990.

⁴ Thomson, 2008.

autonomy violation, while in the original trolley case there is no such autonomy violation. To violate another's autonomy is to make decisions for that person without her express, tacit, or hypothetical consent when she is, or will be, in the position to make those decisions for herself. In the trolley case, the morally relevant distinction can only be made by the bystander, and although his decision may effect others, he does not violate anyone's autonomy by either turning or refraining from turning the switch. In contrast, in the fat man case, the fat man can decide for himself whether he will jump onto the track, and the messenger is capable of consenting to donate his organs if prompted. Furthermore, because there is at least a *prima facie* strong moral obligation to preserve one's own life, neither the bystander nor the doctor in these last two cases is justified in judging the fat man or messenger to be moral monsters who have lost the right make their own decisions, and thus neither is justified in violating their autonomy. I contend that autonomy violation in such cases is a significant moral harm such that we might be morally obligated to let five people die rather than to violate an innocent person's autonomy.

II. On Thomson

In "Turning the Trolley," Judith Jarvis Thomson has recently argued that our moral intuitions about the trolley case are wrong, and that it would be morally wrong for the bystander to turn the trolley towards the single worker, even if doing so saved five lives. To demonstrate this, Thomson constructs a trilemma variation of the trolley case; here is a version of this case:

1^T. *The Trolley^T Case* – An old-timey villain ties a bystander to track 1 and gives him a remote control track switcher, then releases a runaway trolley onto track 3 towards five workers. The bystander can redirect the trolley to either track 1 or 2, where track 2 has only a single worker.⁵

Thomson contends that it is morally acceptable for the bystander to either turn the trolley towards himself, or refrain from doing so. However, if he refrains from doing so, she says, it would be indecent of him to turn the trolley towards someone else; how could the bystander force the worker to make a sacrifice he was unwilling to make? Thus, for Thomson, because it is acceptable to refrain from turning to trolley towards track 1, and unacceptable

⁵ This is my variation of Thomson's case; for the original see Thomson, 2008. See Shaver, 2011, for another robust criticism of this article.

to turn it to 2, by process of elimination it must be acceptable to let the trolley stay on track 3 and kill five people.

Next, Thomson argues that in the traditional trolley case the bystander has only two of the options that were available to him in the trilemma; one that was not morally acceptable – turning the trolley to kill one worker – and one that was morally acceptable – letting the trolley kill five people. Thus, even if the bystander would have been willing to sacrifice himself in the trilemma, she argues, he is morally obligated to choose to let the five people die, because of the two remaining options, only this is morally acceptable. Before we move on, there is a substantial problem with Thomson's argument; it rests upon the assumption the options in the trolley^T case are morally equivalent to the "remaining" options in the original trolley case.⁶ This is not at all clear that they are morally identical, or even sufficiently analogous to reach Thomson's conclusion; however purposes of this paper let us assume that they are.

Thomson's account, I think, is problematic for three reasons. First, this suggests there is a substantially morally relevant difference between killing and letting die, such that it is wrong to kill the single worker, but acceptable to let five people die! The problem is that on Thomson's view she believes it is morally acceptable for the bystander to kill himself, so killing is, at least in some situations, morally acceptable. On Thomson's view, the bystander is allowed to do so because he is not violating his own rights, but to turn the trolley towards the single worker would violate that person's rights. Surely, though, allowing five workers to be killed when the bystander could easily prevent it would be to violate their rights as well! Suppose that in the trolley^T trilemma, track 2 was empty; surely in such a case it makes sense to say that to not turn the trolley to track 2 would be to violate the rights of the five workers. To the extent the options in the trolley case are comparable to those in the trolley^T case, we are committed to the conclusion that refraining from turning the trolley in either of the trolley or trolley^T cases constitutes a rights violation of those workers. Again, perhaps we can appeal to a distinction between killing and letting die to show why it would be wrong to kill one, but

⁶ In a recent article, Tim Willenken compares similar cases to construct instances of deontic cycling, purportedly showing that our intuitions are governed by inconsistent rules. Stephen C. Makin successfully rebuts Willenken's position by demonstrating that although the options in such cases can be described in a similar manner, they are distinct and non-comparable, demonstrating that Willenken's apparent instances of deontic cycling are nothing of the kind. See Willenken, 2012 and Makin, 2012. In Shaver, 2011, he makes a similar criticism of Thomson's analysis of the two trolley cases.

let five die, but it shall become clear in the next section that this is not viable option.

Second, I think the best way to understand ethics is as an attempt to offer a non-arbitrary answer to the question “What should I do?” While attempts to answer this question in certain cases might lead to the conclusion an agent has multiple possible morally acceptable options, presumably these options are comparable. If our intuition is that it is acceptable to sacrifice the single worker in the trolley case are correct, in the trolley^T case presumably, if we agree with Thomson that the bystander’s sacrificing himself was optional, presumably we’d conclude he could sacrifice either himself or the single worker. These two options are comparable and *prima facie* morally equivalent – either would be a sacrifice of one life to save five. In contrast, the options Thomson concludes are morally acceptable in the trolley^T case are *prima facie* not comparable at all; one is to sacrifice one life to save five, and the other is the exact opposite – to refuse to sacrifice one life to save five. The difference is night and day; they are nothing close to morally equivalent.

A third problem with Thomson’s analysis is that just because one is justified in not sacrificing one’s self doesn’t mean that it is unacceptable to sacrifice others. Although perhaps hypocritical, many normative ethical theories allow moral agents to treat themselves different from others. For example, there is a strong, commonsense moral intuition that it is morally acceptable to kill even innocent people in self-defense. (Think horrible misunderstanding, escaped mental patient, rabid child, werewolf, etc. – take your pick.) If self-sacrifice was supererogatory (perhaps for the same reasons it is acceptable to kill innocents in self-defense), this doesn’t preclude the sacrifice of others from being the next best thing. If self-sacrifice is obligatory, one’s immorally refusing to sacrifice one’s self, coupled with the hypocrisy of being willing to sacrifice someone else wouldn’t necessary make turning the trolley to track 2 morally worse than letting it hit the five people – we would have already concluded the bystander in the trolley^T case is morally blameworthy; the question is which is more blameworthy – hypocrisy and killing an innocent person or letting five people die. If killing and letting die are even remotely morally comparable, the additional hypocrisy of the bystander’s failure doesn’t seem like enough to tip the scale. The only other option is that for Thomson self-sacrifice and letting five people die are both morally acceptable for some other reason than hypocrisy.

III. The Doctrine of Double Effect

Many of our actions have foreseeable bad effects that it would be morally wrong to intentionally try to bring about, yet intuitively many are morally acceptable. The doctrine of double effect is an attempt to explain our intuitions in such cases, and to differentiate between those actions which can be done despite their bad consequences from those that can't. The thrust of the principle is that it is always wrong to intentionally cause some bad effect; however it may be acceptable to cause a bad effect if it meets the following criteria:

- i. the bad effect is not intentional.
- ii. the bad effect is not a means to cause what the agent intends to bring about.
- iii. what the agent intends to bring about is sufficiently good to excuse the bad effect.

DDE offers a *prima facie* plausible explanation of the difference between the trolley case and the fat man and transplant cases. In the trolley case although turning the switch has a bad foreseeable effect (the death of one worker), it also has a substantially good effect (saving the lives of five workers), thus satisfying condition iii. As long as the bystander does not intend to cause the death of the single worker, turning the switch satisfies condition i. Furthermore, although turning the switch leads to the death of the single worker, the death of the single worker itself does not bring about the good effect, meeting condition ii. This is to say that if the single worker somehow realizes there is a trolley coming and saves himself, it has no bearing on whether the five workers are saved.

Note that DDE does not excuse any unintentional bad effect, as demonstrated here:

¹*DDE*. *The Trolley^{DDE} Case* – A runaway trolley is about to run over a nice pair of slippers unless a bystander flips a switch to divert the trolley onto a different track where a single worker is working.

According to both commonsense and DDE it would not be morally acceptable to press the button here. To do so would be to violate condition iii; as saving a nice pair of slippers is not sufficiently good to outweigh the worker's death.

The fat man and transplant cases differ from the trolley case because they are said to violate one or more of the conditions. Note that there are actually two relevant bad effects in each of these cases; the pushing of the fat man

(2a) and the harvesting of the messenger's organs (3a), and the foreseeable, but not intentioned, death of the death of the fat man (2b) and the death of the messenger (3b).⁷ 2a and 3a violate condition ii; while 2b and 3b do not violate any of these conditions.

For the remainder of this section, I argue DDE is false. I will show that the principle rests upon there being a morally relevant distinction between acting and refraining; a distinction I believe James Rachels, among others, has successfully demonstrated the falsity of.⁸ In "Active and Passive Euthanasia," Rachels presents two parallel cases, where one is killing, and one is letting die. In the first, Smith drowns his cousin to receive a large inheritance, while in the second, Jones, intending to drown his cousin to receive a large inheritance, arrives in time to see his cousin hit his head in the tub and sink below the water and then lets him die. (1975, 79) Rachels argues that intuitively Smith and Jones are morally equivalent, despite the fact that Smith acts to kill his cousin, while Jones refrains. Rachels contends that the apparent distinction between acting and refraining is easily explicable; in the majority of cases there are morally relevant differences between acting and refraining that lead us to conclude that the average case of killing is worse than the average case of letting die. In "Killing and Starving to Death", James Rachels presents and defends his equivalence thesis (ET from now on):

If there are the same reasons for or against A as for or against B, then the reasons in favour of A are neither stronger nor weaker than the reasons in favour of B; and so A and B are morally equivalent neither is preferable to the other. (1979, 165)

If ethics is not arbitrary, ET is analytically true. ET explains our intuitions in cases like Smith and Jones, because both Smith and Jones had the same reasons to kill (and not kill) their cousin, and acted or refrained on a subset of these reasons.

To illustrate that DDE relies upon existence of a distinction between acting and refraining, let us consider a version of DDE that is supposedly consistent with the falsity of such a distinction. I will call this theory DDE^{ET}. To construct

⁷ In "Thomson and the Trolley Problem" Michael Gorr rightfully points it would be absurd to understand what is wrong with pushing the fat man in the fat man case without considering its likely consequences, see Gorr, 1990. The relevant detail here, though, is that those likely consequences are unintentional.

⁸ Although a number of contemporary philosophers advocate this position; the two most noteworthy are probably James Rachels and Michael Tooley; see Rachels, 1975, 1979, 2001 and Tooley, 1972, 1980.

this principle, we must replace any instance of language regarding acting or refraining in DDE with a means to denote their moral equivalence. According to DDE^{ET}, it is morally acceptable to act in such a manner that brings about a bad effect if and only if:

i^{ET}. the bad effect is not intentional.

ii^{ET}. the bad effect is not a means to cause/allow what the agent intends to bring-about/allow.

iii^{ET}. what the agent intends to bring-about/allow is sufficiently good to excuse the bad effect.

According to DDE^{ET}, it is morally acceptable to turn the switch (act/refrain) in the trolley case because when the bystander acts/refrains, the bad effect is not intentional (satisfying i^{ET}), the killing/letting-die of the one worker is not a means to saving/letting-live the five workers (satisfying ii^{ET}), and the saving/letting-live of the five workers is a sufficiently good effect to excuse the bad effect of the one worker being killed/let-die.

According to DDE^{ET} it is morally wrong to act/refrain in the fat man case because pushing the fat man over the bridge constitutes a violation of condition ii^{ET}. But to say that it is morally unacceptable for the bystander to push the fat man just is to say that that it is acceptable for him not to do so. Since acting and refraining are morally equivalent according to this principle, this means that the bad effect of refraining to push the fat man, killing/letting die the five workers, is not intentional (which satisfies i^{ET}), and that the killing/letting-die of the five workers is not a means to saving/letting-live the fat man (which satisfies ii^{ET}). The problem is that the saving/letting-live of the fat man is surely not a sufficiently good effect to excuse the killing/letting die of the five-workers! Thus, to refrain from pushing the fat man from the bridge would violate iii^{ET}! Thus, according to DDE^{ET} it would be unacceptable to refrain from pushing the fat man off the bridge! Shockingly, according to DDE^{ET} it is wrong to push the fat man, and wrong to not push the fat man! Of course this is absurd, so either DDE is false, or killing and letting die are not morally equivalent when all else is equal.

A number of philosophers have argued that Rachels has failed to show that there is no morally relevant distinction between killing and letting die. Natalie Abrams has argued that in some cases, active euthanasia is actually better than passive euthanasia; the problem is that in the cases she cites, the preference of active euthanasia is preferable to passive euthanasia because it has better results (and thus there are more reasons to do it than the

alternative).⁹ To show that there is a morally relevant distinction between acting and refraining would require two things – first, a pair of cases where one is intuitively worse than the other but that are identical in every relevant manner except that one is a case of acting and the other refraining; and second either an explanation for why this distinction doesn't hold in cases like Smith and Jones or an analysis of Smith and Jones that explains the additional relevant difference that compensates for the inequivalence of acting and refraining in the cases. In absence of such an argument, we must take Rachels' conclusion seriously; that there is no morally relevant difference between acting and refraining; killing and letting die.

IV. Autonomy Violations

It makes sense to say that the doctor in the transplant case wants to save as many lives as possible, yet intuitively he would not be justified to go around harvesting whatever organs he needs to save lives. Under what conditions could the doctor in that case save those five lives? If we take Thomson's intuitions in the trolley^T trilemma seriously, if it were possible for the doctor to sacrifice his own life to save his patients, he would be morally entitled, but not obligated, to do so. Thomson's bystander can sacrifice his own life because that's his choice! Suppose the messenger in the transplant case was fairly perceptive and learned that by donating his organs, he could save five people, and requested to do so – even though it cost him his life. Shouldn't the messenger in the transplant case have the same opportunity to do so as the bystander in the trolley^T case?

On the analysis thus far, our intuitions in the transplant case are that the doctor is not allowed to kill the messenger to save the five. Suppose the messenger were to volunteer his organs, and the doctor responds "I'm sorry, it would be morally unacceptable to take your organs because doing so would kill you, and I have a strong moral obligation not to kill you." I imagine the messenger, having just made a difficult decision to sacrifice his own life to save others, would be quite perturbed at hearing this! Suppose the doctor were to follow this up with "Also, I believe that killing and letting die are morally equivalent, so because I have a moral obligation not to kill you, I also have a moral obligation not to let you die, should you do something stupid and try to take your own life as a means of allowing me to harvest your organs to

⁹ Abrams, 1978.

save those five people's lives." If the messenger were to attempt to take his own life, the doctor would be put in the position of either letting one person die and saving five, or saving one and letting five die. Thus the doctor's claim that he has a strong moral obligation to save the life of the messenger over that of the other five seems a little bizarre. Now suppose that the messenger casually wandered to the other side of the room, lamenting his predicament, only to pull out a concealed weapon and take his own life before the doctor could prevent him from doing so. Now, with the messenger dead and having given his express consent to donate his organs, it strikes me that the doctor would have a strong moral obligation to harvest those organs and save his five patients.

Just as in Thomson's trilemma, the bystander is morally entitled to sacrifice his own life to save others, so too must the messenger be entitled to do so. But this scenario is deeply unsatisfying – not only for the doctor's apparent inconsistency insofar as he would be unwilling to let the messenger die yet willing to let his five patients die – but because some phantom moral obligations prevent the doctor from assisting the messenger in donating his organs, which would be, no doubt, less horrific for the messenger and improve the chances of successfully transplanting the organs and saving the patients.

In the trolley^T case, Thomson sought to explain the wrongness of killing the single worker in terms of how inappropriate it would be for the bystander to choose the worker over himself, leading her to conclude that it was acceptable to sacrifice the five workers. Consider, though, the following two cases:

^{1T2}. *The Trolley^{T2} Case* – An old-timey villain ties a bystander to track 1 and a passerby to track 2, and released a runaway trolley towards five workers on track 3. He gives both the bystander and the passerby a remote that would turn the trolley to either tracks 1 or 2.

^{1T3}. *The Trolley^{T3} Case* – A cliché horror movie villain ties a bystander to track 1 and a passerby to track 2, and aims a runaway trolley towards five workers tied to track 3. The villain gives the bystander a remote that is capable of redirecting the trolley to tracks 1 or 2. The passerby witnesses this, and begs the bystander to switch the trolley to track 2, so that he can give his life to save the other five.

My intuition about trolley^{T2} mirrors my intuition in the fat man case; that it would be unacceptable for the bystander to turn the trolley towards the passerby, and unacceptable for the passerby to turn the trolley towards the

bystander. However, much as in trolley^T, I think it would be acceptable for either the bystander or passerby to turn the trolley toward himself. To turn the trolley towards the other person is make a decision that the other could have made, thus violating her autonomy.

In trolley^{T3}, even if we agreed with Thomson that it would be morally disdainful for the bystander to be willing to sacrifice the passerby's life but not his own, it strikes me as even more egregious to sacrifice the five people in this scenario rather than to sacrifice the one. There are two key differences between Trolley^{T3} and the fat man and transplant cases; first the bystander and doctor in the latter two cases are incapable of sacrificing themselves to save five lives while the bystander in Trolley^{T3} is, and second, the passerby in Trolley^{T3} expressly consents to turn the trolley towards him (which Thomson contends he is morally entitled to), while there is no consent in either the fat man or transplant cases on the part of the fat man or messenger.

I have a strong moral intuition that in Trolley^{T3} the bystander has a moral obligation to turn the trolley towards the passerby if he is unwilling to turn it towards himself. To do so is to act in accordance with the wishes of the passerby, and to respect his autonomy. Similarly, I have a strong moral intuition that if the messenger were to volunteer his organs in the transplant case, the doctor would have a strong moral obligation to do as he requests. Furthermore, if the fat man were, say, too fat to jump off the bridge on his own and were to ask the bystander for help pushing him over the bridge, I believe the bystander would be obligated to help.

At this point, I think the difference between the trolley case and the fat man and transplant cases becomes clear; while it is *prima facie* wrong to push the fat man or harvest the organs of the messenger without their consent, it would be morally obligatory to do so with their consent. The relevant question, now, is whether it is acceptable for the bystander and doctor to ask them for their consent.

Regardless of any inelegance or impropriety we might attribute to the bystander or doctor if, say, they would be unwilling to sacrifice themselves in Trolley^T, it would be *prima facie* immoral to deny either the fat man or messenger the opportunity to act to save the lives of others for this reason alone. Failing to inform the fat man or the messenger of their ability to save lives would be to deprive them of morally relevant information they need to adequately decide how best to live their lives. To intentionally keep this kind of information from these agents, I think, constitutes an autonomy violation far worse than any moral hypocrisy that might be involved.

What makes it wrong to push the fat man off the bridge in the fat man case is that it denies him the opportunity to choose for himself to either sacrifice his own life or not (either choice would be morally acceptable for him according to Thomson). This is all well and good, but I've argued there is no morally relevant distinction between acting and refraining, so just as it is wrong to violate someone's autonomy, so too it would be wrong to let someone else's autonomy be violated. It is at least *prima facie* tempting to say that letting the five people die on the track lets the trolley (or, at least, whoever aimed it) violate their autonomy. While it is plausible to describe the wrongness of their deaths in terms of all of their missed opportunities to exercise their autonomy – in some kind of bizarre variation of Don Marquis's infamous future of value view – this strikes me as a mistake.¹⁰ Roughly, to violate someone's autonomy is to make decisions for that someone rather than to let them make her own decisions when she is in the position to decide for herself, or to replace another's judgment with your own when it is inappropriate to do so.¹¹ To push the fat man off the bridge is to take away his opportunity to decide to sacrifice his own life. Because acting and refraining are morally equivalent, if someone else was to push the fat man off and you can stop it, you have a moral obligation to do. While letting the trolley hit the five people would cause their death and deny them future opportunities to exercise their autonomy, letting the trolley hit the five people would neither be making any decisions for those people, or be letting anyone else make decisions for those people, and thus refusing to push the fat man does not constitute an autonomy violation of the five.

Suppose the bystander in the fat man case knew the fat man well and believed that he would be willing to sacrifice himself to save others; but also knew that there was not enough time to explain the situation in time for him to make the choice the bystander would have him make. It strikes me that in such a case, the bystander is morally justified in pushing the fat man off the bridge – to do so is to help him act as the bystander genuinely believes he would act if there was sufficient time for him to make such a choice.

¹⁰ Marquis, 1989.

¹¹ For example, suppose you and a friend are dining out. Your friend leaves to use the restroom, and the waiter comes to take the table's order. You are reasonably sure that your friend would, given the opportunity, order the most expensive thing on the menu, but in her place given her financial status, you could order something far less expensive. To order for her as she would order, I think, does not violate her autonomy, but to order for her as you think she should order does.

It is sometimes morally acceptable to violate someone's autonomy; for example we are often justified in restricting and even violating the autonomy of prisoners. This is acceptable, in theory, because they have acted immorally and thus proven themselves incapable of moral self-rule. Note, though, that if you give the fat man the opportunity to jump and he refuses, it is not acceptable to violate his autonomy on the basis that he has acted immorally; his refusal to sacrifice himself is a *prima facie* morally acceptable response, at least according to Thomson.

However, it is morally acceptable to violate the autonomy of innocent persons in some situations, when the outcome is sufficiently good. For example, a crossing guard might be morally justified in sticking their hand in front of you as you attempt to cross the street, as doing so might save your life. This is a relatively minor autonomy violation compared to pushing the fat man off of the bridge. It is outside the scope of this paper to rank autonomy violations, but it strikes me that this violation is substantially less of a harm to someone than pushing the fat man off of the bridge would be to the fat man. However there may be some variations of the fat man and transplant cases where violating someone's autonomy is morally permissible; perhaps it is wrong to violate the fat man's autonomy by pushing him to his doom to save five people, but it might be morally acceptable to do so to save millions of people. However, an alternate explanation for this intuition is to say that it would be unacceptable to refuse to sacrifice one's life if doing so would save a million people. If this is the case, then to not choose to sacrifice yourself is to be a moral monster; the kind of moral monster it is *prima facie* acceptable to violate the autonomy thereof.

I contend that the reason we think it would be morally unacceptable to push the fat man and harvest the organs of the messenger in the fat man and transplant cases is that to do so without their permission would be to violate their autonomy. However, I contend that if they give their permission, the bystander and doctor are morally obligated to respect their wishes and act accordingly. In the original trolley case when the bystander reroutes the trolley, he neither violates anyone's autonomy nor allows anyone's autonomy to be violated, and thus his action is morally acceptable.

References

- Abrams, N., 1978, Active and Passive Euthanasia. *Philosophy*, Vol. 53, No. 204, 257-263.
- Foot, P., 1967, The Problem of Abortion and the Doctrine of the Double Effect. *Oxford Review*, 5, 5-15.
- Gorr, M., 1990, Thomson and the Trolley Problem. *Philosophical Studies*, 59(1), 91-100.
- Makin, S. C., 2012, Action Individuation and Deontic Cycling. *Ethics* 121(1), 545-561.
- Marquis, D., 1989, Why Abortion is Immoral. *The Journal of Philosophy*, 86(4), 183-202.
- Rachels, J., 1975, Active and Passive Euthanasia. *The New England Journal of Medicine*, 292, 78-80.
- 1979, Killing and Starving to Death. *Philosophy*, 54 (208), 159-171.
- 2001, Killing and Letting Die. *Encyclopedia of Ethics*, 2nd edition, vol. 2 ed. Lawrence Becker and Charlotte Becker, New York Routledge, 947-950.
- Shaver, R., 2011, Thomson's Trolley Switch. *Journal of Ethics & Social Philosophy* <http://www.jesp.org/articles/download/RobertShaver.pdf> Accessed 02 April 2013
- Thomson, J.J., 1976, Killing, Letting Die, and the Trolley Problem. *The Monist* 59, 204–217.
- 1985, The Trolley Problem. *The Yale Law Journal* 94, 1395–1415.
- 2008, Turning the Trolley. *Philosophy and Public Affairs* 36, 359–374.
- Tooley, M., 1972, Abortion and Infanticide. *Philosophy and Public Affairs*, 2 (1), 37-65.
- Tooley, M., 1980, An Irrelevant Consideration Killing VS Letting Die. In: *Killing and Letting Die*, ed. Bonnie Steinbock, Prentice-Hall, 56-62.
- Willenken, T., 2012, Deontic Cycling and the Structure of Commonsense Morality. *Ethics* 122(3), 545-561.

John Mikhail on Moral Intuitions

Florian Demont
(University of Zurich)
floriandemont232@gmail.com

John Mikhail's *Elements of Moral Cognition. Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgement* is an ambitious book. It combines themes from Rawls' moral philosophy with cognitive science along Chomskian lines. Instead of bickering about Mikhail's reading of Rawls or the Chomskian framework in general, I start by assuming that Mikhail's exegeses are correct and that the Chomskian framework can in principle be used to study moral and legal intuition.¹ Based on this, I shall draw attention to a specific requirement that Mikhail's theory of moral cognition is supposed to fulfil. The requirement is that the cognitive processes constituting moral and legal intuitions must generate deontic structures (understood to be oughts as manifested in the intuitions). I then go on to argue that Mikhail has not shown that his account of the relevant cognitive processes does generate such structures. The primary upshot of the paper will be that Mikhail must substantially revise his account of how oughts enter moral and legal intuitions. I shall also raise three other objections, which question the philosophical underpinnings of Mikhail's project.

Let us start with an outline of Mikhail's project and its philosophical underpinnings. The primary aim of Mikhail's monograph is to propose and motivate a specific research project which combines elements from moral philosophy, cognitive sciences and legal theory. Mikhail argues at length that the most important philosophical insights underlying the project can already be gathered from the early works of John Rawls. One might want to object to

¹ Mikhail does, as far as I can see, not distinguish between the terms 'judgement', 'intuition' and 'sentiment'. I shall prefer the term 'intuition'. Whatever else one might mean by the term, I shall use 'intuition' here to refer to dispositions to make moral or legal judgements.

this exegetical claim, but such an objection is not what I am after here. There are two reasons why I neglect possible questions concerning Mikhail's Rawls exegesis. First, Mikhail's project would still be interesting if it was not inspired by Rawls. And second, the way Mikhail brings cognitive science to bear on moral theory is a distinct feature of his own and it is exactly this aspect of Mikhail's project that interests me most.

Mikhail's research project is to give an account along Chomskian lines of moral and legal intuitions. Noam Chomsky has argued that a formal grammar can describe and explain the most central aspects of human language. Such a formal grammar is thought to achieve its goals by tracking down the unconscious processes responsible for those linguistic dispositions, which form the kernel of human language.² In analogy to this line of thought, Mikhail seeks to describe and explain moral and legal intuitions by giving a formal account of the unconscious processes, which generate them. He sometimes calls this idea the 'moral grammar hypothesis'.

An important feature of Chomskian linguistics is that it is not only naturalist – i.e. it does not only treat the linguistic properties it examines as, for example, physics treats the properties it examines – Chomskian linguistics is also internalist: linguistic properties are to be explained in terms of internal states of organisms.³ And internal states are computational states as described by a formal grammar.

The moral theory that Mikhail is after is also internalist and he calls it a 'theory of I-morality'. Mikhail hence seeks to explain moral and legal intuitions in terms of internal states of organisms. And again, the relevant states are computational states. But unlike Chomsky, Mikhail does not spend any time defending the conceptual soundness of the Chomskian framework. This is understandable, because defending internalism against philosophical nagging is not an easy task.⁴ But it is also highly problematic, because philosophers have produced different powerful objections against Chomskian approaches to human linguistic and psychological capacities.⁵

² The locus classicus of this claim is Chomsky, 1965. An accessible up-to-date account of the Chomskian perspective is Hauser et al., 2002.

³ Chomsky, 2000, 134.

⁴ It has been argued that it is in fact an impossible task as long as internalism requires ascribing such properties to brains which can only be ascribed to persons (cf. Demont, 2012). But objections along these lines will simply be bracketed for what follows below.

⁵ Some of the objections Mikhail should have discussed are those raised in Hacker, 1990, Smith, 2006 and Wright, 2001.

In what follows below I shall simply assume that the Chomskian framework does provide good descriptions and explanations of some linguistic and mental properties. I also grant to Mikhail that it is conceivable that the framework might provide good descriptions and explanations of at least some aspects of moral and legal intuitions. I make these assumptions for the sake of exposition, as it helps bringing to the fore what I regard as the most problematic aspect of Mikhail's theory of I-morality without getting tangled up in familiar problems he shares with most other Chomskians. By this I do not want to insinuate that Mikhail may neglect these problems, I merely claim that they are somewhat independent of the objections I want to raise here.

There are two constraints on a theory of I-morality, which, if they are met, make the theory empirically adequate. The first constraint is descriptive adequacy and the second constraint is explanatory adequacy.⁶ A theory of I-morality is descriptively adequate if it correctly describes human moral and legal intuitions. It is, however, not enough to provide a listiform account, which merely enlists data about moral and legal intuitions. Mikhail, following Rawls and Chomsky, is after a set of principles, which mechanically enumerate all possible moral and legal intuitions and, at the same time, assigns a structural description to each possible intuition. These principles allegedly transcend what people are aware of in everyday life – and this is thought to constitute a first reason why common sense is of limited import in assessing a theory of I-morality.⁷

A theory of I-morality is explanatorily adequate if it provides a description (in the sense just introduced) of the initial state of a human being's moral and legal sense and if it can explain how actual moral and legal intuitions are extrapolated from that when a child grows up and learns to interact with its environment. Meeting explanatory adequacy requires a considerable amount of idealisation and theory-forging by moral psychologists, because the initial state is simply not something which can be directly observed.⁸ The idealisations of moral psychologists are also thought to move the subject

⁶ Mikhail is aware that empirical adequacy might also require a story about how moral and legal intuitions have evolved in the species and how such intuitions are physically realised in the brain (p. 29 and fn.12 on p. 30). This sort of simplification is quite common within Chomskian circles and it is arguably a good idea to study cognitive processes somewhat independently of biological processes (cf. Demont, 2012).

⁷ Mikhail, 2011, 48-51.

⁸ Mikhail, 2011, 22.

matter of the theory further away from what can be assessed in terms of common sense.

Together, descriptive and explanatory adequacy yield empirical adequacy. Empirical adequacy is, of course, not enough to make Mikhail's theory of I-morality a *moral* theory. A moral theory must at least be able to explain what counts as a justified moral principle. Mikhail calls this requirement 'normative adequacy'. A more powerful moral theory also gives an account of how moral principles can be justified. A moral theory meeting this constraint is metaethically adequate.

There is an important *prima facie* problem concerning metaethical adequacy that arises if both empirical and normative adequacy are to be met. An empirically adequate account explains moral intuitions in terms of what is. Normative adequacy, on the other hand, requires an account of what ought to be, simply because rational creatures ought to do whatever is in accord with justified moral principles. If such oughts are to be derived from an empirically adequate moral theory, then Mikhail must say something about why that is possible. He must, in other words, solve the is-ought problem. The problem is a real one for him, because Mikhail explicitly claims that the 'descriptive takes precedence over the normative'.⁹ After all, it has been a common place for a long time that deriving oughts from what is (or from descriptions) requires a substantial amount of philosophical argument in its support. Applying Moore's take on the issue, deriving oughts from what is (or from descriptions of what is) is not possible, because a computational story about human moral and legal intuitions may settle what our intuitions are, but we can then still question whether these intuitions are good or bad. And Mikhail must show that, at least in some paradigmatic cases, such an open question makes no sense.¹⁰ So, how does Mikhail respond to this?

A theory of I-morality is metaethically adequate if it explains how moral principles can be justified. The problem is that a theory of I-morality must at the same time be descriptively and normatively adequate. In order to solve the problem, Mikhail takes his cues from Rawls:

As I interpret him, Rawls presupposes a complicated answer to the general problem of justifying moral principles, which turns on at least three potentially unrelated ideas: first, that moral principles can be presumptively justified by showing that they are a solution to the problem of explanatory adequacy; second, that descriptively adequate moral principles can be further justified by showing that they are part of a solution to the problem of explanatory adequacy;

⁹ Mikhail, 2011, 30.

¹⁰ Compare Miller, 2004, 13-15 for an introduction to Moore's open-question argument.

and third, that moral principles that meet the demands of descriptive and explanatory adequacy can be justified to an even greater extent by showing that the adoption of such principles can be proven as a formal theorem in the theory of rational choice. As I understand it, Rawls' notion of *reflective equilibrium* is intended to suggest that these three apparently disparate ideas can, in fact, be reconciled. In other words, Rawls assumes as a general matter that the same set of moral principles can be part of a single, comprehensive solution to the problems of descriptive, explanatory, and normative adequacy simultaneously.¹¹

An important idea here is that justifying some moral principles is provisional or presumptive. If open questions about the justification of a moral principle are always possible, we merely need a metaethics, which is good enough for whatever our purposes are. And in Mikhail's scheme of things, (descriptively adequate) moral principles are justified enough if they are 'part of a solution to the problem of explanatory adequacy'. Now, scientific theories like the theory of I-morality are always open to revision, but that does not mean that such theories are not practically applicable, because open questions remain about the justification of the theory's moral principles. The theory of I-morality fulfils its purposes as long as it contains those moral principles, which 'free and equal' persons would regard as a rational choice.¹² Mikhail does not spell out in detail what 'free and equal' means here, but he suggests that the pieces fall into place once we understand the philosophical ideal of a reflective equilibrium.

So, how does Mikhail construe the notion of a reflective equilibrium, how does this notion relate to the theory of I-morality and does it solve the problem of metaethical adequacy? Here are two representative quotes from the monograph:

[R]eflective equilibrium is a *technical* concept in Rawls's framework, which strictly speaking refers to a state of affairs rather than a method or technique: namely, the state of affairs in which moral principles and considered judgements coincide, and the researcher thus understands the principles to which those judgements conform, together with the premises of those principles' derivation (Rawls 1971:20). Moreover, Rawls defines the meaning of reflective equilibrium in the context of a conception of moral theory whose principal aim is to solve the problems of empirical and normative adequacy with respect to I-morality.¹³

[T]he primary function of the concept of a considered judgement in Rawls' framework is to select, from among the moral judgements people *actually* make,

¹¹ Mikhail, 2011, 31.

¹² Mikhail, 2011, 32.

¹³ Mikhail, 2011, 289.

those judgements that the moral theorist believes are truly evidential, insofar as they reflect the properties of an underlying cognitive competence.¹⁴

From this we can gather two claims. First, a reflective equilibrium is the state of affairs in which moral principles and considered judgements coincide. But they only should be said to coincide if the moral psychologist thereby understands the computational processes generating the judgements and the elements from which the judgement is generated. The second claim is that considered judgements are those moral and legal judgements, which people actually make and which a moral psychologist takes to be evidential. If we sum this up, the reflective equilibrium turns out to be the state of affairs in which a theory of I-morality generates exactly those moral principles, which account for the moral and legal intuitions that moral psychologists have regarded as relevant data.

This solution to the problem of metaethical adequacy has three shortcomings. First, consider that the psychologists selecting the data are also the psychologists constructing the theory. The problem with this is that some data count as relevant if the psychologists believe that these data reflect ‘the properties of an underlying cognitive competence’ to recognise what is morally or legally correct. The psychologists’ beliefs about what are relevant data will change as they go along constructing the theory of I-morality, which determines what the properties of the underlying competence are. Mikhail has to make clear that the idealisations governing data selection are appropriately independent of the details of the theory of I-morality that is derived from it. Otherwise, the theory’s predictions will be trivially true, as only those findings count as relevant data which theory predicts. This is a general problem about idealisations in scientific reasoning and it is surprising that Mikhail has not addressed it in the monograph. It is hard to see how one can argue that the theory’s predictions are not trivially true if neither data gathering nor theory construction is sensitive to common sense conceptions of what is relevant regarding moral and legal intuitions.

A second shortcoming of Mikhail’s solution to the problem of metaethical adequacy has to do with the concept of computability. It is a formal truth about computational processes that we cannot mechanically determine in advance whether a specific mechanical derivation will eventually produce a definite solution or not.¹⁵ Applied to I-morality, it might be that a computational account

¹⁴ Mikhail, 2011, 283.

¹⁵ This is, of course, the halting problem (cf. Boolos, Burgess & Jeffrey, 2007, 40).

can be given of how a specific moral or legal intuition is generated, but we will not be able to determine through any (Turing-computable) derivation whether or when the intuition will effectively be generated. Yet a rational person can be expected to always have a definite moral or legal intuition when presented with a case, of which she understands all elements and where she has to make a judgement. If we present a rational person with a formal description of a case, she will be able to tell whether she can make a moral judgement if the details are filled in, but we can make no such prediction about computational derivations of moral and legal judgements. So it appears that a rational person always knows more about her moral and legal intuitions than a theory of I-morality can capture. Now, recall that a theory of I-morality is descriptively adequate only if it correctly describes human moral and legal intuitions. With this requirement in the background it appears that descriptive adequacy can never be achieved by a theory of I-morality, because a rational person's knowledge of her moral and legal intuitions is necessarily richer than what a computational account can capture.

The third shortcoming harks back to the is-ought problem outlined at the beginning. We may ask whether Mikhail's conception of a reflective equilibrium tells us how any oughts can be derived from what is. And the answer to this is: no. In a reflective equilibrium, moral principles and considered judgements simply coincide. Any oughts playing a role must have been in the moral principles and considered judgements before they were compared. According to Mikhail's conception of I-morality, deontic structures – moral and legal oughts as they figure in moral principles and considered judgements – are generated whenever moral and legal intuitions are generated. So it is not the conception of reflective equilibrium which tells us how to solve the problem of metaethical adequacy, but the details of the derivation process. The big question now is whether the derivation process can solve the is-ought problem.

In chapter 6.5, Mikhail gives a very short and abstract account of how moral and legal intuitions are derived. If a rational person perceives a situation calling for a moral or legal judgement, her cognitive processes must first identify relevant descriptions of the action. With these descriptions, the cognitive processes must temporally order particular events. On the basis of that, the cognitive processes must identify causal structures in the temporally ordered events. The next step is a distinct element of a theory of I-morality: moral structures are identified by labelling certain effects as good or bad. So, if somebody dies as an effect of some other event, that effect will be regarded

as bad. If, on the other hand, somebody is saved from death as an effect of some other event, that effect will be regarded as good. Note that this does not yet generate oughts. Simply labelling some effects within a causal structure does not suffice to establish an ought, because it is not made clear by this what should be done. That homicide, for example, is a bad effect in any series of causally related events may still be part of the representation of what is the case.

In order to derive oughts, Mikhail first introduces intentional structure. The derivation of intentional structure is accounted for under a presumption of innocence or good intention. Every bad effect will be labelled as a side effect of some actions and every good effect will be labelled as an end or goal. Such a presumption of innocence is, of course, highly problematic, because pursuing good and avoiding evil will be a matter of 'innate instinct'.¹⁶ Somebody who seeks to do evil will, according to Mikhail's proposal, turn out to have defective instincts. The right way of dealing with such wrongdoers is to give them treatment and not to punish them. This is, however, not at all how matters are handled. We distinguish between moral defects and cognitive defects, because any wrongdoing will be punished less severely if the agent has a cognitive defect. People with moral defects will, however, be punished more severely to keep them from harming themselves and others in the future. Mikhail's proposal does hence play down a perfectly sensible distinction between, on the one hand, what is morally and legally wrong and, on the other hand, what is a cognitive defect. They may be related in some cases, but they must not be run together as a matter of principle.

Deontic structure is now explained based on this problematic notion of intentional structure. The point of the last step of the derivation process is to come up with some sort of explanation of intuitions about what is permissible, forbidden or obligatory. Mikhail helps himself to some basic legal definitions to achieve this:

One key insight of the moral grammar hypothesis is that adequate structural descriptions must also incorporate prima facie legal wrongs, such as battery or homicide.¹⁷

Assuming for a moment that all the structures (including intentional structures) can be had just as Mikhail thinks, how do we incorporate prima facie legal wrongs, such as battery, within the confines of I-morality? We

¹⁶ Mikhail, 2011, 173.

¹⁷ Mikhail, 2011, 173.

presumably have available representations of events such as 'X throws Y off a foot bridge', 'Y prevents a train from hitting 5 innocent people' and 'X kills Y'. Now 'X throws Y off the footbridge' (event 1) came before and caused 'X kills Y' (event 2). Event 2 will be labelled a bad effect of event 2. But event 2 will count as a side effect of event 1, because event 1 also caused the good event 'Y prevents a train from hitting 5 innocent people' (events 3). Event 3 will then count as a good effect of event 1 and it will thus count as an end. Adducing *prima facie* legal wrongs, we can now explain what counts as forbidden and then, in a later step, derive permissions (i.e. that which is not forbidden) and obligations (i.e. that which it is forbidden not to do).

Now, X ought not to throw Y off the footbridge, because he thereby commits battery.¹⁸ The legal definition of battery, which Mikhail adduces to derive the ought, has it that X commits battery if he touches Y without his 'express, implied or hypothetical consent'.¹⁹ But how does a rational person come to know *prima facie* legal wrongs such as battery? After all, the theory of I-morality is internalist – it explains moral and legal intuitions in terms of internal states of organisms – and it is not clear how legal definitions, like the definition of battery, are related to internal (computational) states of organisms. I can see two possible replies that Mikhail can make. First, he could claim that the concept of battery is innate. Second, he could concede that the concept of battery has been internalised in some way.

So, what if the concept of battery is innate? The problem with this first possible reply is that Mikhail uses official legal definitions in his derivations. It is quite hard to imagine what would count as verifying that a child has a tacit grasp of battery as touching somebody without that person's express, implied or hypothetical consent. If one insists that the concept is innate and that the legal definition was merely one way of expressing that concept, we must ask whether any empirical finding could topple that claim. There is no such empirical finding, because any alleged evidence against nativism about legal concepts can be rejected by insisting that the data has not been selected in accord with the sort of idealisations that I-morality requires. The account of deontic structures would then be unfalsifiable and a theory of I-morality cannot be, at the same time, empirically adequate and feature an unfalsifiable

¹⁸ Of course, X may throw Y off the footbridge if committing battery and killing him is a side effect of stopping a train, which would otherwise kill 5 people. This more complicated case does, however, not help elucidating the aspect of the derivation of deontic structures that interests me here.

¹⁹ Mikhail, 2011, figure 6.2g.

account of deontic structures as an alleged *result* of empirical enquiry. So, nativism about basic legal wrongs does not have good prospects.

And what if the concept of battery is not innate, but has been internalised in some way? The problem with this second possible reply is that Mikhail's theory of I-morality is an internalist theory. If the derivation of deontic structures requires adducing definitions of prima facie legal wrongs and if these definitions have to be learned (by reading, for example, the Second Restatement of Torts and, perhaps, by looking at court cases where the definitions are applied), then the derivation of deontic structures cannot be explained in purely internalist terms. This amounts to rejecting unbridled internalism for a viable theory of I-morality.

To sum all of this up, we can make four different objections to Mikhail's proposal. The most important one is that he has not solved the is-ought problem. If the derivation of deontic structures builds on innate legal concepts, the theory of I-morality is not empirically adequate, because the derivation bridging the is-ought gap is part of an unfalsifiable empirical claim. On the other hand, if legal concepts are acquired through some sort of training, a correct moral theory cannot be purely internalist.

The second objection was that the derivation of moral and legal intuitions should not blur the distinction between moral and legal wrongs on the one side and cognitive defects on the other side. An interesting proposal such as Mikhail's must not provide a possible excuse for people who neglect their moral responsibilities and who want to lower their liability by pointing out that they are victims of defective instincts. Whatever moral theory one proposes, one should always make sure that it remains applicable.

The third objection was that a computational derivation of moral intuitions is subject to formal constraints on computability – especially the halting problem – whereas a rational person's knowledge of her moral and legal intuitions is not subject to these constraints.

The fourth objection was that Mikhail has not made clear that the process of dividing available data into relevant and irrelevant data is sufficiently independent of a theory of I-morality so as to ensure that the predictions of the theory are not trivially true. It might turn out that this objection cannot be met unless idealisations are made sensitive to what is acceptable from a common sense point of view.

I do think that Mikhail's monograph contains good general ideas and that the whole project is exemplary in how it brings together insights from

philosophy, psychology, linguistics and legal theory. But I also think that there are substantial problems when it comes to the details of Mikhail's theory of I-morality. It might be possible to explain away some problems and it might also be possible that other problems can be met through further research, but Mikhail does have to rethink the conceptual basis of his project.

References

- Boolos, G.S., Burgess, J.P. & Jeffrey, R.C., 2007, *Computability and Logic. Fifth Edition*, Cambridge, Cambridge University Press.
- Chomsky, N., 1965, *Aspects of the Theory of Syntax*, Cambridge (Mass.), MIT Press.
- Chomsky, N., 2000, *New Horizons in the Study of Language and Mind*, Cambridge (Mass.), MIT Press.
- Chomsky N. & McGilvray J., 2012, *The Science of Language*, Cambridge, Cambridge University Press.
- Demont, F., 2012, Chomsky's Methodological Naturalism and the Mereological Fallacy. In: *Philosophical and Formal Approaches to Linguistic Analysis*, Frankfurt a.M., Ontos.
- Hacker, P.M.S., 1990, Chomsky's Problems. *Language & Communication* (10)2, 127-148.
- Hauser, M.D., Chomsky, N., & Fitch, W.T., 2002, The Faculty of Language: What it is, Who has it, and How did it Evolve? *Science* 298, 1569-79.
- Lepore, E. & Smith, B.C. (eds.), 2006, *The Oxford Handbook of Philosophy of Language*. Oxford, Clarendon Press.
- Mikhail, J., 2011, *Elements of Moral Cognition. Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgement*. New York, Cambridge University Press.
- Miller, A., 2004, *An Introduction to Contemporary Metaethics*. Cambridge, Polity Press.
- Rawls, J., 1971, *A Theory of Justice*. Cambridge (Mass.), Harvard University Press.
- Smith, B.C., 2006, What I Know When I Know a Language. In: Lepore and Smith (eds.) 2006, 941-82.
- Wright, C., 2001, *Rails to Infinity*. Cambridge (Mass.), Harvard University Press

De Leibniz a Hegel: Força, Lei e Infinitude na *Fenomenologia do Espírito*

Bernardo Enes Dias
(LIF - Universidade de Coimbra)
bernardoenesdias@gmail.com

De modo muito genérico, poder-se-ia caracterizar a primeira grande secção da *Fenomenologia do Espírito*, relativa à consciência em geral – e aos modelos sucessivos da certeza sensível, da percepção e do entendimento –, como um conjunto de variações sobre um mesmo tema central: a convicção genérica de que existe um objecto exterior à consciência, isto é, um objecto independente da consciência que se tem dele. Mas se o ponto de vista comum tende a reconhecer sem dificuldade esta independência, um olhar mais atento para a estrutura da objectividade não tarda a pôr a nu o seu carácter problemático. Ao procurar definir o seu objecto, a consciência dá-se conta de que a realidade que julgava imediata e independente não é uma coisa nem outra: em lugar de estar aí, imediatamente, para a consciência que o pensa, o objecto está sempre já implicado numa relação dialéctica com o sujeito que o põe. A sua identidade não é simples, mas mediada, e a sua objectividade está desde o início tingida de subjectividade.

O objecto da certeza sensível é o *puramente imediato* – a realidade sensível em estado bruto, anterior a qualquer distinção categorial¹. Isto significa, pois, que o próprio acto de conhecer não representa senão uma pura apreensão. A consciência, a existir de modo autónomo, constitui apenas um filtro de absoluta transparência, que se limita a registar a riqueza e a concretude infinitas do mundo sensível. Mas esta riqueza e esta concretude, quando consideradas atentamente, depressa se convertem em pobreza e abstracção. O *hic et nunc* imediato vem a revelar-se, afinal, a tradução

¹ Cf. Hegel, 1807, 79.

necessariamente universal de um *Meinen* singular. Ao compreender que o *isto*, o *aqui* e o *agora* correspondem a operadores indiferenciados, susceptíveis de traduzir toda e qualquer realidade, a consciência descobre que a singularidade reside, afinal, em si mesma. O modelo é conservado, mas transferido para um plano subjectivo: é a própria consciência a portadora da singularidade que reconheceu, erroneamente, no objecto; é só *para ela* que um objecto é *este* objecto e não qualquer outro, e que o agora é *este* instante temporal e não qualquer outro. No entanto, esse *eu* que se mantém e assegura, alegadamente, a singularidade do mundo exterior, acaba por revelar-se, também ele, universal e indiferenciado. Cada *eu* é, sempre um *eu que x*, um operador abstracto que se define, a cada momento, pela sua relação com uma infinidade de outros operadores abstractos. A verdade dessa “singularidade visada” (*die gemeinte Einzelheit*) não reside, pois, no sujeito ou no objecto, mas na relação entre os dois. A última tentativa de manutenção do modelo olha ambos os termos como essenciais e procura reconhecer a singularidade visada na relação entre eles: se o sujeito e o objecto, quando considerados isoladamente, vêm a descobrir-se como operadores universais, a fuga da singularidade deve residir no próprio acto de isolá-los e considerá-los separadamente. Para corrigir este desvio, o ponto de vista deve manter-se focado sobre os dois em simultâneo: um eu, *este* eu, considerando um aqui e um agora, *este* aqui e *este* agora, e impedindo que qualquer alteridade venha desfazer a imediatez desse olhar. Mas ainda isto se revela insustentável, já que essa focagem pressupõe, sempre, o acto de indicar (*Aufzeigen*), isto é, de nomear a realidade visada e torná-la presente à consciência. Ora, visto que este processo não é estático, mas dinâmico ou dialéctico, a universalidade está sempre já presente e não pode ser anulada. Não apenas o sujeito ou o objecto, mas toda a relação assenta sobre elementos definidos de modo negativo, isto é, que são sempre já *outros* em relação a si mesmos.

A singularidade visada pela consciência não pertence, pois, à certeza sensível. Mas esse *Meinen* perdura e carece de desformalização. Segundo Hegel, é ele que leva a consciência a abandonar o modelo sensível e a dar-se conta de que a certeza imediata de que se julgava portadora é, afinal, percepção. O regime de identidade simples ou de absoluta auto-coincidência implicado no modelo da certeza sensível cede o lugar a um regime de identidade por *determinação*. Uma vez que o *isto*, o *aqui* e o *agora* são instâncias universais, a consciência não pode conhecê-los imediatamente, como meros αἰσθήματα. O modelo atómico é substituído pelo modelo da

coisa e das suas propriedades – a consciência e o seu objecto volvem-se *res cogitans* e *res extensa*.

A consciência adopta, então, o modelo perceptivo, de acordo com o mesmo esquema triádico: em primeiro lugar, a coisa é o seu objecto. Ela contém em si a universalidade resultante da negação do *hic et nunc* singular, ou seja, algo como uma simplicidade feita de multiplicidade. A *coisa*, que a consciência adopta como objecto, é ao mesmo tempo simples – na medida em que é *esta* por oposição a *outras* coisas – e múltipla – porquanto é composta de propriedades universais e indiferenciadas, susceptíveis de integrar a composição de qualquer outra coisa. Um grão de sal é simples e *ao mesmo tempo* branco, cúbico, salgado, etc. Mas a contradição reside justamente no confronto da *unidade passiva* das diferentes propriedades ou “matérias livres”, expressa por um *Auch* unitivo que as agrupa sem comprometer a sua independência e a *unidade negativa* ou *exclusiva* que faz de cada coisa *uma* coisa – e, mais ainda, *esta* coisa, cujo branco é o *seu* branco e não o branco em geral, e cuja forma é a *sua* forma e não, simplesmente, uma matéria universal, etc. Se a coisa é verdadeiramente una, a universalidade da suas propriedades não pode subsistir, e a sua variedade tende a anular-se numa unidade tautológica e absolutamente exclusiva; mas, por outro lado, se a coisa é múltipla, a sua identidade perde-se, reduzida a um agregado mais ou menos complexo de propriedades.

Ao momento objectivo do desenvolvimento segue-se um momento subjectivo, por meio do qual se tenta salvar a integridade do modelo. No entanto, ao contrário do que sucedera na certeza sensível, a consciência não intervém aqui como portadora da verdade do modelo mas, justamente, como portadora da sua não-verdade: a consciência dá-se conta de que a sua captação do objecto não corresponde a um puro *Wahrnehmen* (ou a um puro *Nehmen des Wahres*), transparente e eficaz, mas envolve um momento de reflexão sobre si mesma, que torna opaca a sua apreensão. A contradição entre a universalidade e a singularidade não é, pois, inerente à estrutura da coisa, mas corresponde a um mero defeito de observação. Assim, embora a verdade do modelo pertença ainda ao objecto, o momento essencial da relação é, agora, a própria consciência – ao reconhecer-se a si mesma como a fonte do desvio perceptivo, ela propõe-se corrigi-lo e restabelecer o equilíbrio do modelo.

Esta correcção comporta dois momentos distintos e igualmente vãos. Em primeiro lugar, a consciência aceita a unidade absolutamente exclusiva da coisa e reconhece-se como portadora da multiplicidade das suas

propriedades: o grão de sal só é branco *aos nossos olhos* e salgado *ao nosso paladar*, etc. É a multiplicidade dos sentidos que decompõe a unidade da coisa e a reconhece, ilusoriamente, como uma multiplicidade. Mas esta hipótese é desmentida mediante a verificação de que essa unidade implica, em si mesma, determinação – e, logo, multiplicidade. As categorias “branco”, “salgado”, “cúbico”, etc. só podem pertencer a uma coisa na medida em que esta se opõe a outras, isto é, na medida em que a sua identidade se opõe à alteridade alheia. Confrontada com esta evidência, a consciência inverte a sua posição inicial, aceita a multiplicidade intrínseca da coisa e reconhece-se como portadora da sua unidade: o grão de sal é branco e salgado e cúbico – simplesmente, *na medida em que é branco não é salgado e na medida em que é salgado não é cúbico*. É a própria consciência que, para evitar a contradição, isola cada uma das propriedades e a faz subsistir de modo independente. Todavia, a ser assim, a coisa volve-se num *também* absolutamente indiscriminado – uma mera “superfície englobante”, no dizer de Hegel –, de tal modo que o trabalho unitivo da consciência não é capaz de assegurar a subsistência da coisa e acaba por descobrir-se, mais uma vez, em face de uma pura universalidade.

A focagem alternada dos pólos opostos da consciência e do seu objecto corresponde a uma tentativa de isolar a contradição e mantê-la à margem do modelo perceptivo. Mas esta tentativa revela-se vã e o próprio modelo vem mostrar que a contradição existe, de raiz, no coração das coisas. O regime de identidade simples, que a percepção procura manter, cai por terra e o que emerge para a consciência é um regime de identidade em que o *si-mesmo* e o *outro* estão absolutamente e irremediavelmente entrosados. Para Hegel, a descoberta do entendimento corresponde à descoberta de que não existem identidades simples, ou de que toda identidade implica, correlativamente, a sua negação. O que está em causa, pois, é uma alteração radical do modo de reconhecimento da identidade. A implosão do modelo perceptivo representa o abandono definitivo da ideia de identidade simples, sobre a qual assentava o modelo atómico. À estrutura estática da coisa vem substituir-se um entendimento dinâmico da identidade, expresso pelas noções de *relação* e de *força*. A identidade não é mais o resultado de uma determinação absoluta, mas de uma relação em permanente actualização, do trânsito permanente que resulta da oposição para-si / para-outro que está como que engastada no coração das coisas.

A singularidade visada no modelo da certeza sensível foi suprimida e preservada sob a forma de uma universalidade, mas esta universalidade era,

ainda, uma universalidade condicionada. Com o advento da percepção, a contradição do modelo sensível não foi verdadeiramente removida, mas apenas transferida para o plano colectivo da coisa e das suas propriedades. A universalidade perceptiva veio opor-se à ideia de singularidade, mas continuara a braços com o problema da determinidade: a coisa é universal e, ao mesmo tempo, determinada; é uma e, ao mesmo tempo, múltipla. A sua universalidade conserva, pois, uma oposição que só pode ser superada se os dois termos em oposição, em lugar de permanecerem lado a lado, se assimilarem e forem reconhecidos enfim como dois modos de expressão do mesmo fenómeno. Para que isso aconteça, porém, é preciso reformular a própria estrutura da apresentação comum: a verdade da percepção não é um de dois extremos – nem tão-pouco os dois em simultâneo, já que a sua convivência, como se viu, é inviável; a verdade da percepção é o próprio trânsito dialéctico entre os dois, isto é, o movimento eterno e instantâneo por meio do qual uma coisa é, ao mesmo tempo e do mesmo modo, *outra* em relação a si mesma. Só a consciência desta união faz emergir a verdadeira universalidade – a *universalidade incondicionada*², livre de qualquer distinção ou determinação.

O entendimento representa, pois, a primeira tentativa de constituição de um olhar verdadeiramente científico, isto é, a primeira proposta concreta de superação da inconsistência inerente ao ponto de vista empírico. Está pela primeira em causa a adopção de um modelo de apresentação artificial, que rejeita o modo normal de olhar a realidade para suprir as contradições de que ele padece³. Mas a universalidade incondicionada, aquando do seu

² *Unbedingte Allgemeinheit*, ou apenas *das Unbedingte*. O alemão joga com a raiz *ding* e com a sua negação – o universal condicionado é a negação do modelo das coisas.

³ Por isso, a noção leibniziana de força –, que serve de ponto de partida ao modelo hegeliano do entendimento – é apresentada pelo próprio Leibniz como uma noção eminentemente *metafísica*. Cf., por ex., 1690-1703, 66: « (...) ayant tâché de d'approfondir les principes mêmes de la Mécanique, pour rendre raison des lois de la nature que l'expérience faisait connaître, je m'aperçus que la seule considération d'une *masse étendue* ne suffisait pas, et qu'il fallait employer encore la notion de *force*, qui est très intelligible, quoiqu'elle soit du ressort de la Métaphysique. » Embora a física cartesiana não consiga explicar de modo exacto a natureza dos fenómenos, o seu olhar corresponde, por assim dizer, a uma complexificação do ponto de vista natural – o seu vocabulário é, ainda, o da substância e da oposição movimento/repouso. A força e o dinamismo leibnizianos, pelo contrário, introduzem um modo de ver que não corresponde já a nenhum desenvolvimento empírico, mas a um olhar que pensa a finalidade antes da eficiência e antepõe um fundamento metafísico à contingência dos fenómenos naturais.

surgimento, representa apenas o avesso do modelo perceptivo. Visto que se trata de um conteúdo objectivo, ela carece, ainda, de determinação. Ora, esta determinação não pode ser feita ao modo das coisas, isto é, através do isolamento daquilo que a universalidade é *em si mesma*. A identidade e a alteridade residem nela *uno tempore* e é o seu entrosamento absoluto que assegura o carácter *incondicional* da universalidade. A sua identidade é uma identidade suprimida, que atira imediatamente para a alteridade absoluta das suas propriedades; e, do mesmo modo, a multiplicidade dessas propriedades resolve-se de imediato na unidade que as contém. O novo objecto da consciência é, pois, o próprio movimento de transição entre a identidade e a alteridade, expresso pela noção de *força*.

I. Força e Expressão

Na *Fenomenologia do Espírito*, a descoberta da noção de força implica a supressão da oposição pensada pela consciência e a posição de algo como um movimento absoluto, isento de pólos ou limites formais. Mas, tal como havia acontecido no modelo perceptivo, essa oposição está como que engastada na noção de força e não pode ser suprimida sem que esta perca o seu sentido. O impasse dialéctico que conduzira ao colapso do modelo perceptivo é retomado, agora, a um outro nível: por um lado, o movimento por meio do qual a força perde a sua unidade e se dispersa no meio inclusivo das suas propriedades é a expressão ou exteriorização da força; por outro, o movimento simétrico por meio do qual essa pluralidade é reconduzida à unidade é a força propriamente dita, ou a força recolhida sobre si mesma. Simplesmente, ao contrário do modelo perceptivo, a relação entre os dois pólos é dinâmica, o que significa que o movimento de posição de um deles implica, dialecticamente, o movimento da sua supressão: a expressão é sempre expressão *de uma força* e, inversamente, a força só existe porquanto pode expressar-se a si mesma. Por isso, a exteriorização absoluta e o recolhimento absoluto conduzem, ambos, à destruição da força – o primeiro corresponde a uma pura alteridade e o segundo a uma tautologia.

No terceiro capítulo da *Fenomenologia*, Hegel parece reportar-se, embora de modo inexplicito, a um momento particular da história do pensamento filosófico e científico. Com efeito, o abandono do modelo perceptivo e a descoberta da noção de força equivalem, no plano histórico, ao abandono do dualismo cartesiano e ao advento das dinâmicas de Newton e, sobretudo, de

Leibniz. Enquanto que a física cartesiana assentava, ainda, sobre um entendimento atômico da realidade, concebendo a matéria como um mero agregado cujas partes se definem tão-só pela sua figura e pela sua extensão, a física leibniziana procurará mostrar que a realidade da matéria não depende da sua extensão e que a unidade e o movimento verdadeiros só podem ser explicados através do recurso à noção de força⁴. Para Leibniz, a extensão e a figura são instâncias puramente relativas, incapazes de assegurar os fenómenos da identidade ou da alteridade. Uma pura *res extensa*, isenta de unidade formal, não seria nada mais do que uma multiplicidade absolutamente indefinida de pontos materiais. Para pensar a realidade há, pois, que pensar uma unidade que transcenda o plano da espacialidade e assegure a identidade *real* da matéria. O que está em causa não é, já, um conjunto de átomos físicos ou geométricos, cujas diferentes configurações espaciais obedeçam a um estado accidental de movimento ou repouso, mas a posição de *átomos formais* ou *substâncias reais*, isto é, de unidades monádicas que possuam em si mesmas a determinação da sua identidade. Neste segundo modo de entender a realidade, o movimento e o repouso não são contingentes mas correlativos de uma identidade formal fundamental, que se *expressa* a si mesma na alteridade do mundo exterior.

A mónada, em Leibniz, não é uma unidade elementar ou atômica, no sentido clássico do termo, mas uma unidade expressiva e imanente, cuja identidade corresponde, justamente, a um determinado ponto de vista sobre a alteridade. Isto significa, pois, que o modelo estático das identidades simples deve ceder o lugar a um modelo globalmente dinâmico, em que todos os movimentos e todas as transformações se descobrem como momentos particulares de um fenómeno de *expression* ou *Äußerung*. As oposições movimento/repouso e ipseidade/alteridade não são oposições fixas, mas momentos arrancados a um gradiente dinâmico e infinitamente matizado. Por isso Leibniz procura vincar o carácter pervasivo da força: enquanto que a consciência comum tende a reconhecer a força como uma realidade meramente regional, que modifica o estado de movimento dos corpos e lhes imprime uma velocidade e uma direcção novas, a Dinâmica vem mostrar que a força é omnipresente e constitui o próprio suporte de realidade dos corpos materiais. A diferença entre a força actuante – *active*, ou *geäußert* – e a força estática – *passive*, ou *in-sich-zurückgedrängte* – é apenas uma diferença de expressão. É uma força que determina o

⁴ Cf., por ex., Leibniz, 1663-89, 229, ou 1690-1703, 66-67.

movimento visível dos corpos, mas é também uma força que assegura a sua coesão interna. O que a Dinâmica leibniziana vem mostrar é que a força não é uma instância segunda, acrescentada a um substrato material estático. A força é, pelo contrário, o termo primeiro de toda a realidade e a configuração adoptada, a cada momento, pelos corpos materiais é a expressão de um jogo de forças sempre já constituído.

Começa a prefigurar-se, pela primeira vez, a falência definitiva do modelo da identidade simples. O abandono da gramática da substância implica – como já acontecia em Leibniz, com os conceitos de *entr'expression* e *entr'empêchement* – a tradução da oposição identidade/alteridade num jogo de forças dinâmico em que não existem, já, extremos fixos, mas apenas graus diferentes de interferência recíproca⁵. As forças não são detentoras de uma identidade estanque, que se alterasse continuamente no curso da relação, mas que permanecesse, a cada momento, igual a si mesma. O que aqui está em causa, pelo contrário, é um modelo de identidade *na alteridade* – ou seja, a descoberta de uma identidade puramente puramente *expressiva*. A identidade de cada força é expressa por uma outra força e a identidade dessa força, por seu turno, é expressa fora de si. O ser de cada força não é, na verdade, algo de essencial, mas um “puro ser-posto através de outro” (*ein reines Gesetzsein durch ein Anderes*).

No entanto, a compreensão deste entrosamento absoluto resulta na perda de realidade da força e na descoberta de um puro desaparecer (*Verschwinden*). Uma vez que as forças não existem, já, enquanto extremos fixos de um movimento de transição, não existe nelas nada que possa manter ou assegurar a sua identidade: «a sua essência consiste, antes, pura e simplesmente [na circunstância de] cada uma *ser* apenas através da outra; e, assim, aquilo que cada uma é deixa de sê-lo imediatamente, uma vez que é o *outro* [de si mesma]»⁶. Face à pura actualização da força, o que resta é apenas o seu *conceito* – não já um conceito por actualizar, mas o conceito que se descobre, enfim, como conceito (*Begriff als Begriff*). Confrontada com o desaparecimento da força, o que a consciência retém é um puro pensamento, o pensamento da universalidade que constituíra inicialmente o seu objecto e que se revela, agora, o seu conceito. Recorrendo à formulação de Hegel, a universalidade incondicionada é, agora, não apenas *em si* mas também *para a consciência*.

⁵ Cf. Leibniz, 1663-89, 224.

⁶ Hegel, 1807, 111.

Mas a descoberta da força enquanto mero conceito não põe fim à dualidade sujeito-objecto. O entendimento, ao conceptualizar a evanescência absoluta da realidade, põe diante de si o oposto desse conceito, isto é, um para-lá correspondente à realidade enquanto pura actualidade. Este para-lá é *formal* ou *negativo* porque serve de suporte ou de base ao entendimento: com efeito, não acontece que a verdadeira essência das coisas – ou o seu *interior*, no dizer de Hegel – surja como uma alternativa ao conceito de força, a que a consciência não consegue chegar. O interior das coisas (*das Innere der Dinge*) é aquilo mesmo que o entendimento, justamente porque não tem qualquer notícia do que isso possa ser, procura traduzir através do conceito de força. Contudo, quando se fala de um interior ou de um *em-si* independente, oposto ao território da aparência, tende a pensar-se o problema como a contraposição de duas realidades qualitativamente diferentes, mas semelhantes quanto ao seu estatuto: a ser assim, a aparência corresponderia ao elemento em que se move o ponto de vista, incapaz de conhecer a realidade tal como é, enquanto que o interior das coisas se referiria, pelo contrário, a uma realidade verdadeira, autónoma e impoluta, subsistindo algures no exterior do território da aparência. Ora, este vício de raciocínio é o mesmo que presidia, já, a uma leitura simplista do projecto kantiano e que o próprio Kant se esforçara, iteradamente, por dissipar. Ao chamar a atenção para a natureza integralmente fenomenal do ponto de vista comum, Kant não se limita a estabelecer uma distinção simétrica entre o reino da aparência, a que o olhar humano está eternamente votado, e um *em-si* transcendente que não consegue alcançar. Se assim fosse, o problema esgotar-se-ia na verificação de um desencontro que, em última instância, nada teria de problemático – pois que, nesse caso, a realidade *em si* seria apenas *outra* realidade que não esta e não chegaria a comprometer o seu estatuto referencial. Mas o que está em causa na análises kantiana e hegeliana da fenomenalidade é algo bastante mais complexo: por um lado, visto que o ponto de vista natural lida apenas com fenómenos, a realidade *em si* não é menos fenoménica do que tudo o resto a que tem acesso e só pode ser concebida a partir do território da fenomenalidade; mas, por outro, esse território não se constitui de tal modo que o ponto de vista que o habita seja indiferente ao estatuto da apresentação a que tem acesso. Não acontece, com efeito, que aquilo que lhe aparece se limite a aparecer-lhe, sem mais, nada dizendo quanto ao seu valor de verdade. Toda a apresentação implica a fixação original de um estatuto de verdade ou de um *An-sich* excessivo em relação aos conteúdos

apresentados, mas que tende a confundir-se, a cada momento, com eles. Por isso, a aparência não é vivida como aparência, ou como uma pura apresentação, mas sempre já como a apresentação *da própria realidade* – ou, antes, de uma versão da realidade reconhecida como a própria realidade.

A crítica de Kant procura chamar a atenção para a simultaneidade destes dois aspectos e para a impossibilidade de suspender um deles em favor do outro: se é certo que o númeno é *apenas* um fenómeno, não é menos certo que se trata de um fenómeno peculiar. A fenomenalidade, ainda que constitua uma totalidade fechada sobre si mesma, está desde o início desequilibrada pela posição formal do seu próprio exterior, ao modo de uma *negative Erweiterung* ou de um *reines Jenseits*. Ora, é este desequilíbrio que permite distinguir entre realidade e aparência e é ele que está na base da diferença entre o olhar empírico e o olhar científico⁷.

Neste ponto, a análise de Hegel assemelha-se à de Kant, com a diferença fundamental de que o modo de ver do entendimento não corresponde, já, à *natürliche Denkart*. O sujeito do entendimento reconheceu o carácter aparente da aparência, mas esta não se tornou menos aparente devido a esse reconhecimento. É este, justamente, o paradoxo do entendimento: a descoberta do interior das coisas não equivale à descoberta de algo de novo ou transcendente, subtraído ao território da aparência, mas antes à descoberta da peculiaridade da prisão em que está encerrado o ponto de vista – uma prisão onde a própria descoberta de que se está encerrado (e, portanto, de que existe um para-lá do ponto de vista de que se é possuidor) não diminui em nada o poder absoluto desse encerramento. Por isso Hegel insiste sobre o carácter total ou universal da *Erscheinung*: é a própria aparência que põe o interior das coisas e permite pensá-lo, negativamente, como um interior; mas, por outro lado, essa posição não é a posição de um conteúdo qualquer, mas uma posição-limite – um alargamento

⁷ Trata-se da identificação do mesmo desequilíbrio constitutivo que está na base da antropologia platónica. Tal como em Kant, a distinção entre $\delta\acute{o}\xi\alpha$ e $\gamma\nu\omega\sigma\iota\varsigma$ tende a ser compreendida distraidamente como a oposição de dois termos independentes e mutuamente exclusivos. Mas o que Platão tem em vista é o reconhecimento do carácter sempre já “híbrido” do ponto de vista – um ponto de vista integralmente endoxal, mas afligido pela posição de um para-lá formal que, mais do que acrescentar-se ao plano da realidade, é a própria condição para que qualquer realidade possa ser vivida como tal. O maniqueísmo que Kant e Hegel se esforçam por combater é o mesmo que aflora, involuntariamente, em toda a exegese platónica que vinca a separação de um mundo “concreto” ou “terreno” e um mundo de Ideias ou Formas puras, reforçada por uma leitura imediata e algo simplista dos mitos escatológicos que povoam o *corpus platonicum*.

negativo que visa algo de *positivo* que não está em condições de representar.

II. O Mundo das Leis

Tanto Kant como Hegel – e, antes deles, toda a tradição platónica – dão por adquirida uma mesma convicção fundamental: se existe um para-lá do ponto de vista comum, ele não pode ser pensado como mais um dos conteúdos a que esse ponto de vista tem acesso. Para que se possa apreender a radicalidade dessa instância-limite, é preciso reconhecer, em primeiro lugar, o carácter globalmente aparente do mundo em que o ponto de vista habitualmente de move. O númeno kantiano só é concebível por oposição ao carácter fenomenal da apresentação comum; a sua concepção emerge do reconhecimento da realidade comum como pura aparência e assinala um limite puramente negativo ou absolutamente heterogéneo. Mas enquanto que a análise kantiana se detém neste ponto e se mantém, por assim dizer, face a um abismo de pura negatividade, a análise hegeliana procurará mostrar que esse abismo corresponde, tão-só, à incapacidade da consciência em reconhecer-se como fatora desse *para-lá* negativo – ou, no dizer do próprio Hegel, a incapacidade de reconhecer que esse *Für-sich-sein* autónomo é, afinal, o seu próprio ser-para-si (*sein eigenes Für-sich-sein*)⁸.

Sigamos atentamente o desenrolar da argumentação hegeliana: o interior das coisas, ou o mundo supra-sensível que emerge da descoberta do carácter integralmente fenomenal da realidade, apresenta-se à consciência como um limite puramente negativo. Um vez que representa apenas o avesso ou a nulidade da aparência (*das Nichts der Erscheinung*), o mundo interior é um mundo vazio. Mas esta vacuidade não resulta de uma limitação estrutural da consciência que a impedisse, devido ao modo como está constituída, de avistar esse *Jenseits* de outro modo que não a pura negatividade. Esta vacuidade resulta, tão-só, do modo como o problema é considerado pelo sujeito do entendimento. É certo que o interior das coisas, na sua imediatez, não pode corresponder a nada mais do que uma pura incógnita, já que a sua condição de “puro além” implica, justamente, a abolição de toda a realidade objectiva. É fácil de ver, além disso, que qualquer tentativa de preencher este território-limite – mesmo ao modo de um *Heiliges* etéreo ou onírico – redundará, necessariamente, numa nova

⁸ Hegel, 1807, 112.

concessão ao mundo da aparência. Mas se o interior das coisas não significar nada mais do que este limite vazio e abstracto, a consciência está condenada a uma contradição insolúvel: por um lado, ela reconhece o mundo da aparência como um mundo irreal mas, por outro, é condenada a viver aquilo que lhe aparece como a própria realidade. Ora, a viragem proposta por Hegel consiste na verificação de que esse além que se opõe à aparência, e que se supõe conter a sua verdade, não é verdadeiramente um além, ou um negativo do movimento fenomenal, mas a própria regra que põe a fenomenalidade enquanto tal.

Enquanto que, inicialmente, o mundo sensível era o mundo real, ainda que evanescente, e o mundo supra-sensível um além vazio e incognoscível, o que agora se descobre é justamente o contrário: o mundo fenomenal não é outra coisa senão o reflexo de uma lei da consciência. É a consciência a autora da lei que permite reconhecer a diferença universal dos fenómenos naturais e, por isso, é ela a autora da fenomenalidade. A vacuidade que antes se atribuíra ao interior das coisas é agora transferida para o próprio território da fenomenalidade: enquanto pura aparência, a fenomenalidade é mais do que um puro desaparecer, insusceptível de ser pensado. Mas enquanto reflexo do entendimento, a fenomenalidade é a lei por meio da qual tudo o que há se converte, de imediato, em algo de diferente, em obediência a um trânsito original que corresponde ao modo de ser *necessário* de toda a realidade.

Esta viragem implica, também, uma alteração fundamental de modalidade. A descoberta da implicação universal de todas as forças comporta a negação definitiva da diferença enquanto algo de ocasional ou de ocasionado. A diferença não é, já, algo que ocorre de modo contingente, mas a própria lei que faz da realidade aquilo que é. Do mesmo modo, não acontece, já, que a fenomenalidade exista por direito próprio, como uma pura contingência ou como um acontecer cego que contém em si a sua própria razão de ser – e que o entendimento procura, com maior ou menor êxito, ordenar e confinar a determinadas regras; acontece, antes, que a realidade é o reflexo ou a expressão de uma lei universal e que essa lei é o fundamento a partir do qual toda a realidade é reconhecida como tal.

De facto, Hegel não se limita a afirmar que a realidade fenomenal, embora pareça por vezes casuística ou incompreensível, obedece na verdade a um manual de instruções global. O que aqui está em causa, tal como em Leibniz, é a verificação de que a realidade fenomenal só é aquilo que é enquanto expressão de uma lei universal. É a lei que fornece

inteligibilidade ao mundo da aparência e é ela que confere aos fenómenos o estatuto de fenómenos. Assim, por exemplo, as leis do movimento dos corpos estabelecem relações entre variáveis que, consideradas isoladamente, representam puras incógnitas. Ao mover-se no tempo e no espaço, um corpo descreve uma trajectória que pode ser fixada por meio de uma relação quantificável entre a distância percorrida e o intervalo de tempo entretanto decorrido – mas enquanto que a relação se mantém a mesma, permitindo agrupar os sucessivos estados de movimento do corpo num único movimento discreto, o tempo e o espaço variam continuamente. Por isso, Leibniz definira o tempo e o espaço como conteúdos *irreais*, isto é, puramente relacionais. A heterogeneidade absoluta que opõe os diferentes momentos do espaço e do tempo só pode ser compreendida através da relação que esses momentos estabelecem uns com os outros. Mas essa relação não é superveniente, de tal modo que esses momentos existissem, primeiro, em si mesmos e só em seguida se relacionassem entre si. Pelo contrario, é a própria lei que, ao estabelecer uma relação inteligível entre os fenómenos, põe a sua existência independente.

Todavia, embora o entendimento descubra, enfim, a ascendência da própria lei sobre os fenómenos legislados, essa descoberta começa por ser apenas parcial. O vazio correspondente ao reconhecimento imediato do interior das coisas foi preenchido pela descoberta de uma lei universal da diferença, que esvaziou por seu turno o território da aparência. Mas este processo de transferência deixa subsistir, ainda, uma porção não-legislada do território da aparência. A vigência da lei não é, ainda universal, e os passos subsequentes do capítulo III da *Fenomenologia* consistirão justamente em denunciar os aspectos de que a aparência reclama, ainda, a autoria, como outros tantos modos de expressão de uma mesma lei global.

A lei da fenomenalidade surge, em primeiro lugar, contraposta ao território da fenomenalidade, que persiste como realidade objectiva. Embora as relações de forças entre os dois territórios se tenham invertido, a lei é, ainda, a imagem ou a cópia inteligível do mundo dos fenómenos. Ela não consegue assimilar completamente o plano da fenomenalidade porque deixa subsistir, nesse plano, um núcleo irreduzível de contingência. Esse núcleo, ou esse defeito de que não consegue libertar-se, prende-se com a ambiguidade da sua definição: a lei é, por um lado, uma lei *universal* – e, por isso, indeterminada: a lei da fenomenalidade enquanto tal; mas, por outro, ela espelha a variedade do mundo fenomenal e converte-se numa pluralidade de leis *determinadas* – as diferentes leis que regem o mundo físico. Tal como

acontecera com os modelos da coisa e da força, a dialéctica que opõe estes dois momentos nasce de uma contradição presente na própria arquitectura da noção de lei. Enquanto que só uma lei universal é capaz de pôr cobro à contingência e assegurar a unidade do mundo fenomenal, essa universalidade é de tal modo abrangente que nada determina em relação à natureza dos fenómenos – é uma lei total, que se limita a assinalar a circunstância de toda a realidade ser conforme com uma lei. Mas, por outro lado, ao querer aproximar-se dos fenómenos e traduzir a sua natureza específica, a lei converte-se numa multiplicidade contingente de leis particulares, cuja sucessão é determinada pelas circunstâncias arbitrárias da própria fenomenalidade.

A contradição nasce, como já acontecera no modelo coisal, da oposição entre amplitude e determinação: quanto mais abrangente é a lei, menos determinado é o seu conteúdo, e vice-versa. A lei que rege a queda de um corpo e a lei que rege o movimento dos astros podem ser reunidas numa única lei – a lei da gravitação universal; mas esta lei única não expressa, na verdade, nem uma nem outra, mas apenas a própria legalidade enquanto tal. A lei da gravitação universal, ao manter-se num plano globalmente abrangente e recusar a caracterização de uma realidade particular ou de um movimento determinado, não faz senão proclamar o conteúdo genérico de todas as leis. Trata-se de uma lei que se aplica, igualmente, a todas as coisas e, justamente por não deixar absolutamente nada de fora do seu âmbito de aplicação, corresponde a um predicado tão universal quanto a essência ou a existência. Como todas as leis universais, a lei da gravitação não tem qualquer poder determinativo e limita-se a afirmar, de modo abstracto, que todas as coisas são conformes a uma lei.

Ora, como observa Hegel, se levarmos até ao limite a unidade desta lei universal, aquilo que se nos depara é a superação da própria lei enquanto tal e o advento de um puro conceito. Atente-se no exemplo da gravidade: enquanto que a lei da gravitação estabelece uma relação entre termos distintos – o espaço e o tempo, o corpo que atrai e o corpo que é atraído, etc. –, a gravitação universal enquanto puro conceito limita-se a afirmar a necessidade dessa relação, transcendendo as diferenças particulares dos termos em relação. A unificação dessas diferenças é a necessidade interior da lei, ou a lei anterior à sua tradução no território da fenomenalidade. A dialéctica que vai estabelecer-se até ao final do capítulo vai centrar-se, pois, na oposição entre a lei no primeiro sentido e a lei enquanto conceito absolutamente unitário – a que Hegel chamará, novamente, *força*, por

representar uma assimilação total da diferença. O desencontro entre o interior das coisas e o mundo da aparência, que o entendimento não conseguiu eliminar, volve-se agora no desencontro entre a força e a lei, cuja superação vai depender não já de um movimento unilateral – que fora tentado, sem êxito, nas etapas objectiva e subjectiva da argumentação – mas de uma reunião dialéctica dos dois termos em oposição e da consequente descoberta da realidade como *conceito absoluto*, ou como *Unendlichkeit*.

Como vimos, a descoberta do carácter integralmente legislado da realidade, expressa por meio de uma lei da força ou de uma lei da fenomenalidade, não conseguiu eliminar completamente a contingência enraizada no mundo fenomenal. Segundo Hegel, este núcleo de contingência manifesta-se, ainda, de dois modos distintos. Em primeiro lugar, através do desencontro entre *força* e *lei*: no caso da gravitação universal, por exemplo, o conceito ou a necessidade interna da lei correspondem, tão-só, à força simples *gravidade*; mas a *lei da gravidade* afirma que, no movimento de qualquer corpo, a distância percorrida é igual ao produto entre o quadrado do intervalo de tempo decorrido e um factor de ponderação constante⁹. De igual modo, enquanto que a *electricidade* corresponde a uma força simples, a *lei da electricidade* estipula a atracção entre um campo eléctrico positivo e um campo eléctrico negativo. Contudo, em ambos os casos, a lei não decorre necessariamente da força, mas representa, antes, um mero atributo desta: por isso se diz, habitualmente, que a lei da gravitação “tem a propriedade de se expressar deste modo”, ou que “o modo de ser” da electricidade é a divisão entre positividade e negatividade. E embora chegue a afirmar-se, por vezes, que a gravidade ou a electricidade são, *necessariamente*, relações desta natureza, a necessidade aqui em causa é um termo vazio, pois não se avista a implicação lógica que vincularia estas forças a este modo de ser e não a outro qualquer.

O problema pode ser reconduzido à oposição leibniziana entre necessidade lógica e necessidade hipotética, ou entre finalidade e eficiência. O primeiro tipo de necessidade é aquele de que se faz uso ao afirmar-se, por exemplo, que a soma dos ângulos internos de um triângulo é *necessariamente* igual a dois ângulos rectos, ou que, num triângulo rectângulo, o quadrado da hipotenusa é *necessariamente* igual à soma do quadrado dos catetos. Nestas duas situações, o vínculo necessitante corresponde, tão-só, à afirmação explícita de uma propriedade implicitamente

⁹ A constante gravitacional *g*. Hegel alude à célebre fórmula enunciada Galileu, relativa ao movimento uniformemente acelerado de queda de um corpo: $d = 1/2gt^2$

contida na definição do triângulo. Trata-se, pois, de uma inferência que pode ser reconduzida, de modo mais ou menos mediato, ao princípio da identidade, ou à tautologia que afirma a igualdade do triângulo consigo mesmo. No segundo caso, porém, a necessidade não é de ordem lógica, mas fáctica. É o vínculo reservado, no pensamento de Leibniz, ao conhecimento dos acontecimentos futuros, ou dos desenvolvimentos da noção completa de cada mónada – um conhecimento que não pode ser deduzido matematicamente, mas apenas assegurado *ex hypothesi*, por um ponto de vista capaz de vislumbrar a totalidade da cadeia causal que rege a evolução do universo.

Na argumentação de Hegel, a lei e a força são indiferentes uma à outra porque a necessidade que as une é uma necessidade de ordem hipotética, no sentido leibniziano do termo: nada no conceito de electricidade permite antecipar a divisão entre electricidade positiva e negativa. Se é necessário que essa divisão se produza, é-o apenas devido à verificação contingente de que isso sempre acontece – o conceito tem este modo de ser *apenas porque tem este modo de ser*.

Mas o problema é ainda mais complexo, pois para além do desencontro entre força e lei, subsiste ainda um desencontro entre as diversas partes de uma mesma lei. Para estabelecer, por exemplo, a regra de um qualquer movimento no espaço e no tempo, é necessário considerar a relação entre a distância percorrida e o intervalo de tempo decorrido. De facto, a lei não é outra coisa senão a relação necessária entre estes dois factores. No entanto, se o movimento é pensado como uma realidade simples e unitária, ou como a relação enquanto tal, a lei é incapaz de expressar esta unidade em si mesma. E isto porque a lei indica a relação de um conjunto de factores que, embora se relacionem necessariamente *de determinado modo*, não se relacionam necessariamente *tout court* – ou, dito de outro modo, não decorrem uns dos outros da mesma forma que o teorema de Pitágoras decorre da noção de triângulo. A cada distância corresponde necessariamente uma determinada velocidade, mas a relação entre estas duas instâncias é acidental. Enquanto que, no caso da electricidade, o problema residia unicamente no desencontro entre força e lei – já que, na lei da electricidade, a positividade e a negatividade decorriam logicamente uma da outra, eram o reverso uma da outra – no caso do movimento dos corpos, o problema é interno ao próprio enunciado da lei. A distância é pensada como algo de autónomo, que existe independentemente do tempo, e o tempo é pensado também como algo em si mesmo, anterior à distância de que é

medida. A lei não consegue, pois, libertar-se do modelo atômico e a necessidade por ela afirmada é, mais uma vez, uma falsa necessidade.

Face a estas dificuldades, a crítica de Hegel incide sobre aquilo a que chama o processo de “explicação” (*das Erklären*), mediante o qual o entendimento procura desfazer a contradição entre a necessidade absoluta da força e a necessidade acidental da lei – ou entre a unidade da consciência e a diversidade do mundo fenomenal. Esta contradição subsiste, como vimos, porque a consciência não se descobriu ainda como a autora integral da fenomenalidade. Embora o mundo fenomenal seja um mundo globalmente legislado, as próprias leis contêm, ainda, núcleos de identidade simples que escapam ao movimento do entendimento e forçam, ainda, a oposição estática entre identidade e alteridade. A necessidade da força, pensada pela consciência, é ainda um movimento que vem como que agitar ou diluir uma estrutura rígida de identidades contingentes – ela vem pôr em relação um conjunto de leis que permanecem diferentes umas das outras, mas cuja diferença não é ainda percebida como *diferença interna* ou como *infinitude*. Ora, face a este desencontro, o entendimento começa por igualar a lei e a força, por meio de uma tautologia:

«Uma lei é enunciada. Com isso, o seu universal em-si, ou o fundamento, é distinguido como *força*; todavia, é afirmado que esta distinção não é distinção nenhuma, mas, pelo contrário, que o fundamento é constituído exactamente do mesmo modo que a lei. A ocorrência isolada [de um] relâmpago, por exemplo, é compreendida como universal e este universal é enunciado como a *lei* da electricidade: a explicação reúne, pois, a *lei* na *força*, enquanto essência da lei. Esta força é então constituída *de tal modo* que, ao exteriorizar-se, surgem electricidades opostas que desaparecem de novo umas nas outras – ou seja, a *força é constituída exactamente do mesmo modo que a lei.*»¹⁰

A solução do entendimento consiste em tomar a necessidade fáctica da lei pela exteriorização necessária de uma força ou de um fundamento essencial. A lei tem *esta* configuração específica porque corresponde, tão-só, à tradução de uma mesma força essencial; mas a força, por seu turno, não é outra coisa senão a necessidade desse modo de manifestação. Assim, por exemplo, a razão pela qual a corrente eléctrica se divide em cargas opostas é nada menos que a circunstância de sofrer a acção de uma força que se caracteriza, justamente, por provocar essa divisão. Por outras palavras: a corrente eléctrica divide-se em cargas opostas porque se divide em cargas opostas. A equiparação da força à lei conduz a uma tautologia que nada

¹⁰ Hegel, 1807, 121. Tradução nossa.

adianta acerca da necessidade de uma ou da outra. Com efeito, a *ruse* do processo de explicação consiste em afirmar a diferença de duas instâncias que não são, verdadeiramente, diferentes, para poder em seguida afirmar a sua igualdade. Trata-se, por assim dizer, de uma reiteração disfarçada do princípio de igualdade.

A crítica de Hegel a este *Erklären*, apesar da sua relativa brevidade, retoma um dos temas mais significativos de todo o pensamento hegeliano. O que aqui parece estar em causa é nada menos que uma denúncia da defectividade intrínseca de todo o pensamento científico convencional, assente numa estrutura de dedução matemática que não consegue ir além de uma mera remissão à igualdade. Recorde-se que, no prefácio à *Fenomenologia*, Hegel denunciaria a imobilidade do conhecimento matemático e a sua incapacidade de incorporar a diferença interna que é apanágio da vida e de todo o pensamento dialéctico¹¹. O carácter estático ou “morto” da matemática e de todas as ciências ditas exactas, que a adoptam como método, decorre justamente de uma entronização ingénua do modelo da identidade simples: os teoremas e as fórmulas matemáticas correspondem a desenvolvimentos internos, de complexidade variável, de uma mesma tautologia fundamental e todos podem ser reconduzidos, idealmente, ao princípio da identidade. O conhecimento matemático aspira a um conhecimento da igualdade que exclui de si mesmo toda a diferença e afirma algo como uma identidade absoluta, ou uma *reine Selbstgleichheit*.

De modo análogo, o processo de explicação consiste apenas no isolamento específico de um dos infinitos desenvolvimentos possíveis do princípio de identidade. Ao procurar explicar os fenómenos físicos, o entendimento põe em confronto duas proposições que decorrem de um mesmo fundamento tautológico, mas que correspondem a diferentes graus de complexificação desse fundamento: em lugar de se afirmar, por exemplo, que $2=2$, afirma-se que $2=1+1$ – ou, para retomar o exemplo acima utilizado, em vez de se afirmar que a electricidade é igual a si mesma, afirma-se que a electricidade é a divisão entre positividade e negatividade, divisão que, por seu turno, não é outra coisa senão electricidade. Ora, o que Hegel procura mostrar é que nenhuma destas afirmações afirma o que quer que seja acerca dos fenómenos em causa e que o seu conteúdo explicativo é exactamente o mesmo da tautologia simples $A=A$. A explicação está condenada a esta

¹¹ Cf. a discussão do Prefácio acerca das diferenças entre conhecimento matemático e filosófico, em particular Hegel, 1807, 40-41.

remissão tautológica porque entende a identidade, de modo estanque, como o lugar de desaparecimento de toda a diferença.

Para Hegel – e para a consciência natural de que a *Fenomenologia* traça o percurso evolutivo –, a contradição entre a universalidade da força e a diversidade fenomenal das diferentes leis só pode ser resolvida se a diferença superficial do processo de explicação se converter numa *diferença interna*, isto é, numa diferença que não se limite, já, a afirmar a heterogeneidade de identidades diversas, mantendo a sua autonomia, mas que se instale no interior da própria identidade individual. Ao invés de um movimento exterior que assegure a necessidade da relação entre as leis (ou entre os seus componentes internos), o que a argumentação procurará revelar é uma diferença necessária interna à própria estrutura da lei, que a leva, a cada momento, a reconhecer-se como o outro de si mesma. Eis o principal ponto de viragem do capítulo, cuja articulação argumentativa procuraremos acompanhar, agora, em maior detalhe.

III. Diferença interna e Infinitude

A *Erklärung*, engendrada pela consciência para salvaguardar a unidade do entendimento face à contingência do mundo fenomenal, não é mais do que uma tautologia. Mas esta tautologia, tal como Hegel põe em evidência, não corresponde já a uma igualdade estática. É certo que, ao equiparar a lei à força, o entendimento toma a seu cargo um movimento que não pertence à lei ou à força, mas que lhes é acrescentado. Simplesmente, uma consideração mais atenta desse movimento vem revelar que a aparente anterioridade das duas instâncias é, afinal, fictícia.

Se o entendimento se limitasse a afirmar, sem mais, a igualdade entre lei e força, as duas noções assimilar-se-iam, ao modo de uma unidades simples, isenta de toda a diferença, e não mais haveria que distingui-las. Mas esta igualdade é afirmada, justamente, para poder conservar a diferença entre elas. Com o fito de mostrar que a relação entre lei e força é uma relação necessária, o entendimento é obrigado a pôr a sua diferença e a negá-la logo de seguida. Todavia, ao dividir para poder conquistar, ele não se dá conta da necessidade destas duas operações. Atarefado em salvar a unidade do mundo supra-sensível, o entendimento não se apercebe daquilo que há de verdadeiramente necessário em todo o processo: a saber, o próprio

movimento por meio do qual a diferença é, ao mesmo tempo, posta e cancelada.

Eis-nos, finalmente, em face do ponto de viragem que levará à descoberta de um novo estádio na progressão fenomenológica. Ao considerar atentamente a estrutura do processo de explicação, a consciência foca pela primeira vez o próprio movimento que converte a identidade em alteridade, e vice-versa. Ao longo do capítulo, o jogo de forças, a oposição entre força solicitada e força solicitante e a diferença entre lei e força correspondiam a realidades fabricadas para tentar manter a identidade e a diferença em campos isolados. É certo que cada uma destas instâncias, porquanto representava, já, uma superação do modelo coisal, se caracterizava por um entrosamento total ou por um movimento incessante de permuta de determinações. Mas este entrosamento era, ainda, o fluxo resultante do entrosamento de dois pólos opostos. A necessidade do movimento residia no objecto e nascia da peculiaridade (crescente) do seu modo de ser. Agora, pelo contrário, o que vem a descobrir-se é o próprio movimento como lei universal, isto é, como condição necessária para a posição de toda e qualquer realidade.

O entendimento reconhece-se a si mesmo, então, como a lei da aparência ou a lei da diferença. A lei por meio da qual todas as diferenças são, na verdade, igualdades, e todas as igualdades são, na verdade, diferenças. Aquilo que é igual a si mesmo, por via dessa mesma igualdade, repele-se a si mesmo e aquilo que é diferente de si mesmo atrai-se irresistivelmente e converte-se em igualdade. Esta lei universal afirma, pois, exactamente o oposto daquilo a que anteriormente se chamara lei. Enquanto que a primeira lei afirmava a constância da diferença, ou uma «permanência da não-permanência» (*eine Beständigkeit der Unbeständigkeit*)¹², e assegurava desse modo a unidade do mundo supra-sensível, a segunda lei afirma a impossibilidade absoluta de qualquer permanência na diferença. Assim, esta segunda lei dá origem a um segundo mundo supra-sensível, diametralmente oposto ao primeiro. Este mundo invertido assimilou completamente a diferença que o primeiro mundo votava ao território da fenomenalidade e eliminou por fim a contingência que esse território insistia em manter para si. Neste segundo mundo, tudo se passa ao contrário do primeiro: o que naquele mundo era igual a si mesmo é, neste mundo, diferente de si mesmo;

¹² Cf. Hegel, 1807, 123.

o que era positivo é negativo; o que era preto é branco; o que era honroso é ignominioso, etc.

Contudo, importa tentar perceber mais exactamente o que está em causa na oposição entre estes dois mundos – tanto mais que a descrição de Hegel, neste passo, nos parece especialmente ambígua¹³. À primeira vista, poderia parecer que o mundo invertido vem opor-se ao mundo das leis da mesma forma que o positivo vem opor-se ao negativo, ou que o preto vem opor-se ao branco. Se assim fosse, os dois mundos corresponderiam a duas substâncias opostas, separadas uma da outra por um dispositivo especular. Ora, a ser assim, teríamos novamente uma diferença entre dois pólos independentes e repetir-se-ia o jogo dialéctico que levava à destruição das figuras engendradas, anteriormente, pelo entendimento – a oposição entre o mundo canónico e o mundo invertido seria apenas uma recapitulação da oposição entre o mundo das leis e o mundo fenomenal, ou entre um interior e um exterior das coisas, ou ainda entre um mundo em-si e um mundo para-a-consciência. Mas a descoberta de uma lei universal da diferença não permite, já, este tipo de leitura, visto que implica, justamente, a destruição definitiva de toda a polaridade. Ao dar-se conta do movimento necessário e ubíquo que leva a igualdade a converter-se em diferença e vice-versa, o que a consciência descobre é nada menos que a anulação da diferença entre a interioridade e a exterioridade, ou entre o mundo pensado e o mundo fenomenal. A fenomenalidade não é nada em si mesma, mas apenas o reflexo de uma lei universal que leva a que todas as coisas, ao serem aquilo que são, sejam simultaneamente o contrário de si mesmas.

À luz desta descoberta, o *Verkehren* de que este novo mundo é imagem não pode reduzir-se a uma inversão objectiva ou conteudal. O essencial da noção de *verkehrte Welt* não é a oposição entre dois conteúdos antitéticos, mas entre dois regimes de sentido radicalmente diferentes. Por isso – e é aqui, sobretudo, que nos parece residir a ambiguidade –, este mundo invertido não pode ser pensado como algo de independente, existindo em oposição ao mundo das leis, mas antes como uma instância omnieglobante que abrange, também, o mundo canónico e o assimila. Embora Hegel isole este novo mundo e o destaque do mundo precedente, é importante frisar que não se trata simplesmente de um *outro* mundo, que se opõe ao primeiro e

¹³ Para uma discussão mais detalhada deste passo, cf. H.-G. Gadamer, *Die verkehrte Welt, Hegel-Tage Royaumont*, 1964, 135-54; J. C. Flay, *Hegel's Inverted World, Review of Metaphysics*, 23, 1970, 662-678; M. Moneti, *Hegel e il Mondo alla Rovescia. Una figura della Fenomenologia dello Spirito*, Florença, 1986.

toma o seu lugar. O que está aqui em jogo não é a substituição de uma identidade simples por outra, mas a substituição do próprio regime da identidade simples por um regime de diferença interna. O mundo invertido não é apenas um mundo novo, mas o mundo antigo *e, ao mesmo tempo, o seu contrário* – a identidade e o seu oposto numa mesma unidade absoluta e absolutamente contraditória:

«Há que pensar a pura mudança, ou a *oposição em si mesma*, ou a *contradição*. (...) O mundo supra-sensível, que é o mundo invertido, assimilou assim, ao mesmo tempo, o outro [mundo] e passou a contê-lo em si mesmo; ele é *para si mesmo* o [mundo] invertido, isto é, o inverso de si mesmo; ele é ele-próprio e o seu oposto numa unidade. Só deste modo a diferença é diferença *interior*, ou diferença *em si mesma*, ou *infinitude*.»¹⁴

Pela primeira vez em toda sucessão de figuras da consciência, a instabilidade intrínseca da oposição identidade-alteridade deixou de poder ser decomposta num objecto determinado. Ao descobrir a diferença como diferença interna, ou como lei, a consciência dá-se enfim conta da falência absoluta do modelo de identidades simples. Enquanto que, nos estádios anteriores, sob múltiplas formas, a diferença era ainda entendida como um resultado – isto é, como o desencontro resultante do confronto entre unidades simples e iguais a si mesmas, que no seu conjunto davam origem a um mosaico de diferenças –, o que agora se descobre é que essas unidades simples pura e simplesmente não existem – ou, dito de outro modo, que a sua diferença não é externa mas interna. É ela a matéria-prima de que é feito o tecido da fenomenalidade e, por isso, para onde quer que se olhe, é impossível não surpreender, já, o trânsito eterno e necessário mediante o qual a identidade se converte em diferença e esta, por seu turno, regressa novamente à identidade.

A descoberta de uma diferença interna representa, pois, uma inversão radical do modelo de reconhecimento da realidade que orientara todas as figuras anteriores da consciência. Nos estádios anteriores, o problema consistira sempre em saber de que modo a diferença pode surgir da unidade, qual é a sua origem ou o seu comutador. A diferença fora sempre encarada como algo de derivado, por oposição à simplicidade primeira da *Sebstgleichheit*. Mas a descoberta de que a diferença corresponde a uma instância original, atrás da qual não é possível recuar, vem mostrar que a própria unidade anteriormente considerada era apenas uma abstracção. De acordo com o mecanismo de desconfinamento do ponto de vista que orienta

¹⁴ Hegel, 1807, 126. Tradução nossa.

(e orientará, até ao seu termo) a ascensão fenomenológica, a descoberta de uma diferença interna corresponde à descoberta do carácter mediado de uma realidade tomada, erroneamente, por imediata. Aquilo que se supunha ser idêntico a si mesmo e afectado *a posteriori* por diferença é, na verdade, um dos extremos de uma oposição. Por isso, a pura unidade, ou o $\epsilon\upsilon\upsilon$ de Parménides, não é outra coisa senão o isolamento distraído de um dos termos de uma oposição sempre já posta e impossível de suspender. Aquilo para que Hegel procura chamar a atenção, no fim de contas, é o facto de a *Selbstgleichheit* constituir sempre o resultado de um processo. Tudo o que é igual a si mesmo é-o porque se tornou igual a si mesmo, através da supressão da cisão que era, desde sempre, parte da sua natureza. De igual modo, onde a igualdade se volve diferença, o que tem lugar é um nova supressão – uma supressão da supressão inicial, mediante a qual a unidade isolada volta a cindir-se e a diferença suprimida se converte em diferença aparente.

Face à descoberta deste movimento infinito de cisões e unificações – ou, para retomar a expressão de Hegel, desta «inquietação absoluta do puro mover-se a si mesmo»¹⁵ –, a consciência descobre pela primeira vez que a diferença que serve de base a toda a realidade não é, já, uma diferença qualquer, mas uma diferença absoluta ou uma pura contradição. Enquanto que o mundo das leis procurava legislar ou, por assim dizer, regularizar o devir do mundo fenomenal, o “mundo invertido” que agora se descobre, ao fundir em si mesmo essa regularidade e o seu oposto, institui uma lei da diferença total. Só ao percorrer a totalidade do espectro da diferença, elevando-a a pura contradição, este mundo logra desfazer enfim as oposições intelectuais entre um *für sich* e um *für anderes*, ou entre um interior e um exterior das coisas. Sob a acção desta lei universal, cada determinação, ao ser posta, precipita-se irresistivelmente rumo à sua própria destruição.

Como nota F. Chiereghin¹⁶, a noção hegeliana de *Unendlichkeit* pode ser aproximada da noção de infinito tal como é considerada no cálculo infinitesimal. Da mesma forma que um infinitesimal não corresponde a um nada absoluto mas a um limite assintótico, também a contradição que opõe

¹⁵ Hegel, 1807, 129.

¹⁶ Chiereghin, 1994, 84. A este respeito, o autor aconselha o trabalho de A. Moretto: *Hegel e la matematica dell'infinito*, Trento, 1984. Cf. também F. X. Miranda: *La interpretación del calculo infinitesimal en el sistema de Hegel*, Pamplona, EUNSA, 2003.

a identidade e a alteridade representa um limite dinâmico e irresolúvel: «a contradição não consiste no facto de o igual ser *também* desigual, de a diferença ser posta e *também* retirada, mas sim no interior do *também*, onde acontece a pura permuta da dissolução dos opostos um no outro.»¹⁷ Ao focar a própria contradição, Hegel procura surpreender o momento exacto em que uma determinação, apesar de ser, ainda, *si mesma*, é já, ao mesmo tempo, um *outro*. O si mesmo e o outro não subsistem lado a lado mas também não se anulam definitivamente, perdendo a sua determinação. A infinitude não é outra coisa senão a subida a um trapézio de instabilidade absoluta, em que as determinações se criam e se destroem sem cessar, alimentando desse modo a vida da consciência. Ora, esta vida é justamente o oposto da “morte” que Hegel imputará, no Prefácio, à matemática e a todas as formas de pensamento pré-dialéctico, presas ao cânone tautológico da igualdade e, por isso, «incapazes de mover-se a si mesmas»¹⁸. A infinitude inaugura um sistema de reconhecimento que abre enfim mão do princípio de identidade e pensa a contradição como princípio primeiro de toda a realidade.

Referências Bibliográficas

- Leibniz, G. W., 1663-1689, *Discours de Métaphysique et autres textes* (Ed. Paris, Flammarion, 2001).
---1690-1703, *Système Nouveau de la Nature et de la communication des substances et autres textes* (Ed. Paris, Flammarion, 1994).
---1703-16, *Principes de la Nature et de la Grâce, Monadologie et autres textes* (Ed. Paris Flammarion, 1996).
---1875-90, *Die philosophischen Schriften von Gottfried Wilhelm Leibniz (in sieben Bänden)*, Berlin, C. I. Gerhardt.
Kant, I., 1781, *Kritik der reinen Vernunft*, Riga, Hartknoch (Ed. Frankfurt am Main, Suhrkamp, 1974).
---1968, *Kants gesammelte Schriften*, hrsg. von der Berlin-Brandenburgischen Akademie der Wissenschaften, Berlin, De Gruyter.
Hegel, G. W. F., 1807, *Phänomenologie des Geistes*, Bamberg/Würzburg, J. Goebhardt (Ed. Reclam, Stuttgart, 1987).

¹⁷ Chiereghin, 1994, 84. A respeito do conceito hegeliano de *Unendlichkeit*, cf. também o capítulo homónimo da *Jenenser Logik* (Hegel, 1968, vol. 6)

¹⁸ Hegel, 1807, 40.

---1968, *Gesammelte Werke*, ed. Nordrhein-Westfälische Akademie der Wissenschaften, Hamburg, Felix Meiner Verlag.

Kojève, A., 1933-39, *Introduction à la lecture de Hegel* (Ed. Paris, Gallimard, 2011).

Hyppolite, J., 1947, *Structure et Génèse de la Phénoménologie de l'Esprit*, Aubier-Montaigne.

Labarrière, P.-J., 1979, *La phénoménologie de l'esprit de Hegel. Introduction à une lecture de la phénoménologie de l'esprit*, Paris, Aubier-Montaigne.

Scheier, C.-A., 1980, *Analytischer Kommentar zu Hegels Phänomenologie des Geistes. Die Architektonik des erscheinenden Wissens*, Freiburg - München, Alber.

Moretto, A., 1984, *Filosofia della matematica e della meccanica nel sistema hegeliano*, Trento, Il Poligrafo, coll. Percorsi.

Belaval, Y., 1993, *Etudes leibniziennes – de Leibniz à Hegel*, Paris, Gallimard.

Chiereghin, F., 1994, *La Fenomenologia dello Spirito. Introduzione alla Lettura*, Roma, Carocci (Trad. port. *A Fenomenologia do Espírito de Hegel*, Lisboa, Edições 70, 1998).

Schlemm, A., 2005, *Wie wirklich sind Naturgesetze? Auf Grundlage einer an Hegel orientierten Wissenschaftsphilosophie*, Münster, Lit.-Verlag.

Vieweg, K. Et al, 2008, *Hegels Phänomenologie des Geistes: ein kooperativer Kommentar zu einem Schlüsselwerk der Moderne*, Frankfurt am Main, Suhrkamp.

Alonso, A. et al., 2009, *Figuraciones contemporáneas de lo absoluto – Bicentenario de la Fenomenología del espíritu de Hegel*, Valencia, PUV.

Fabulosas raças de humanóides: monstros e robôs. A robótica humanóide e a captura da intencionalidade

Porfírio Silva*
(Instituto de Sistemas e Robótica - Instituto Superior Técnico)
porfiriosilva@isr.ist.utl.pt

1. Introdução. Questionar a robótica humanóide

O que andam a fazer tantas equipas de investigação por todo o mundo que trabalham na criação de robôs humanóides? Durante muito tempo, a tentativa de construir máquinas inteligentes centrou-se na mente: o Deep Blue, o computador construído pela IBM para jogar xadrez, que em 1997 envergonhou Kasparov, o campeão mundial da modalidade, não era capaz de ver o tabuleiro nem de mexer as peças – e, mesmo assim, considerou-se que o computador tinha vencido o humano nesse jogo. Hoje, poucos acreditam que a inteligência possa ser assim desligada do corpo que anda pelo mundo. A importância da robótica na procura da inteligência para máquinas resulta, até certo ponto, dessa compreensão, o hardware sendo como “o corpo” da “criatura artificial”. Contudo, isso não explica só por si a proliferação de robôs humanóides. Interessando compreender o fascínio dos humanos pelos robôs humanóides, o que sugerimos aqui é que essas máquinas são parte do nosso exercício colectivo de compreensão da nossa própria humanidade. O humanóide é do campo onde está a fronteira entre o humano e o não humano. O robô humanóide permite experimentar com o que nos parece ser próprio do humano, na forma e na função, sem quebrar interditos éticos (fazer experiências invasivas em pessoas vivas). Essa

*A investigação do autor beneficia do apoio da Fundação para a Ciência e a Tecnologia (SFRH/BPD/35862/2007)

experiência passa por construir máquinas que, até certo ponto, pretendemos à nossa imagem e semelhança. Ora, essa “imagem” e essa “semelhança” é que, precisamente, alimentam a dinâmica da relação entre o humano e o humanóide. Para explorar esta dinâmica vamos recuar no tempo, ao tempo de outros humanóides, também eles fabulosos, também eles a interrogar a nossa humanidade.

2. A aparência de ELIZA

Temos vindo a tentar compreender em que sentido é que as ciências do artificial devem ser entendidas como uma das vias das ciências do humano¹. Consideramos como “ciências do artificial” todas aquelas teorias e práticas científicas que procuram realizar, em máquinas concebidas e construídas pelos humanos, certos comportamentos ou capacidades que tenham sido definidas como objecto de atenção por parecerem típicas dos próprios humanos ou de outros animais que encontramos na natureza. As Ciências do Artificial, que não são uma disciplina científica, mas uma constelação de disciplinas científicas, uma constelação em evolução², podem ser exemplificadas, no século XX, pela Inteligência Artificial (IA) e pela Nova Robótica. Muitos praticantes das Ciências do Artificial entendem as suas experiências como (directa ou indirectamente) relevantes para compreender os próprios humanos. É assim que algumas tendências (como a IA clássica) focam aquilo que consideram específico dos humanos, por exemplo competências deliberativas de tipo simbólico altamente sofisticadas, enquanto outras (como a Nova Robótica) atendem preferencialmente à pertença dos humanos ao mundo animal e, desenvolvendo paralelos com outros animais, pretendem chegar a compreender essa dimensão infra-estruturante da nossa humanidade. Do ponto de vista desta abordagem às Ciências do Artificial, um episódio da respectiva história, datado dos anos 1960, continua a suscitar questões que, se têm sofrido transformações no seu aspecto tecnológico, permanecem essencialmente inalteradas no que toca à pergunta pelo humano que atravessa a investigação sobre inteligência para máquinas. Referimo-nos ao episódio do programa de IA conhecido como ELIZA.

¹ Usando “ciências do artificial” estamos a coincidir na expressão com Herbert Simon, embora a coincidência conceptual seja diminuta, como resulta do capítulo 7 de (Silva, 2011).

² Cf. Schopman, 1987; Gardner, 1985.

O programa ELIZA, desenvolvido por Joseph Weizenbaum no MIT entre 1964 e 1966, estabelece um cenário de conversa em linguagem natural (inglês) entre um computador e um utilizador humano³. O utilizador humano escreve as suas “falas” no teclado e recebe respostas também escritas com tempos de reacção que não desmentem a humanidade do interlocutor. Na versão mais conhecida e usada pelo seu criador para efeitos de demonstração, a máquina desempenha o papel de um psicoterapeuta rogeriano. Um dos elementos de credibilização do sistema consiste precisamente no pressuposto de que um psiquiatra dessa escola incentivará o seu paciente a esclarecer todas as suas afirmações, devolvendo sistematicamente as suas falas com pedidos de melhor esclarecimento sobre os tópicos suscitados. Weizenbaum explica que escolheu o psicoterapeuta como o seu “personagem”, porque a entrevista psiquiátrica lhe pareceu um dos poucos exemplos de comunicação em linguagem natural com dois intervenientes em que parece natural, para uma das partes, que a outra parte exiba uma pose de quase completa ignorância acerca do mundo real. Quando um paciente diz “Fui dar uma grande volta de barco” e o psiquiatra responde “Fale-me de barcos”, não pensamos que o psiquiatra seja ignorante acerca de barcos, mas que ele está a perscrutar a mente do paciente.

São relatadas as mais diversas histórias acerca da forma espantosa como muitas pessoas, interagindo com este programa, se convenciam de que estavam a conversar com um psicoterapeuta. Por exemplo, uma das secretárias do sector onde Weizenbaum trabalhava terá chegado a pedir aos circunstantes que a deixassem a sós com o “psicoterapeuta” para poder falar com a necessária privacidade. Contudo, apesar das aparências, como Weizenbaum sempre explicou com total transparência, o ELIZA não tinha qualquer forma de compreensão. O esquema básico do funcionamento do ELIZA era simples: na frase inserida pelo utilizador era procurada uma palavra-chave; a cada palavra-chave correspondia um conjunto de regras de decomposição e uma delas era aplicada para transformar a frase numa sequência de palavras manipulável pelo programa; sobre essa sequência era aplicada uma das regras de recomposição associadas à mesma palavra-chave, de modo a produzir a sequência de palavras que constitui a resposta do computador. O programa variava as respostas: não usava sempre as mesmas regras de decomposição e de recomposição para ocorrências diferentes da mesma palavra-chave. Como base desta estratégia, o

³Weizenbaum, 1966.

programa dispunha de um dicionário de palavras-chave, que permitia determinar se alguma das palavras contidas numa frase inserida pelo utilizador era uma palavra-chave. O utilizador podia inserir, de uma só vez, mais do que uma frase ou uma frase composta, mas o ELIZA só podia transformar uma frase simples de cada vez. Por isso, quando analisava uma inserção do utilizador e encontrava uma vírgula ou um ponto final, se já encontrara até aí uma palavra-chave ignorava tudo o que aparecia a seguir a esse sinal de pontuação; se ainda não encontrara nenhuma palavra-chave, apagava tudo o que “lera” até aí e concentrava-se no restante. Se numa entrada não encontrava nenhuma palavra-chave, o ELIZA retomava um tópico anterior ou respondia com uma frase do género “Porque é que pensa assim?”, destinada a ter cabimento em qualquer contexto. Weizenbaum sempre foi claro: quem atribui conhecimento e inteligência ao seu interlocutor é o humano; os pressupostos são lá postos pelo humano; além dos truques relativamente simples da operação interna do ELIZA, tudo o resto é fornecido pelo humano utilizador. Neste caso, o autor do programa é completamente transparente: mostra toda a operação interna do ELIZA e explica que, além dos truques relativamente simples que lá colocou, tudo o resto é fornecido pelo humano utilizador.

Anos mais tarde, Weizenbaum, no quadro de uma reflexão sobre a responsabilidade social da investigação em Inteligência Artificial, volta a questionar-se sobre o significado deste episódio⁴. Se se mostra impressionado com o facto de muitas pessoas se envolverem emocionalmente com um programa de computador, como se estivessem mesmo a consultar um psiquiatra, é ainda mais notável que profissionais, psiquiatras no activo, tenham sugerido seriamente que o ELIZA podia ser desenvolvido para ser transformado numa forma automática de psicoterapia. Um dos exemplares mais notáveis é K.M. Colby, que desenvolveu ainda nos anos 1960 programas de análise da neurose, começando com um programa que “tratava” uma mulher que acreditava que o seu pai a tinha abandonado, mas não aceitava conscientemente que o odiava por isso⁵. Este conjunto de reacções ao ELIZA, atribuindo à sua obra maravilhas que ele próprio negava veementemente, levou Weizenbaum a interessar-se pelos problemas suscitados pela facilidade com que as pessoas fazem atribuições extraordinárias a uma tecnologia que não compreendem.

⁴ Weizenbaum, 1976.

⁵ Para uma descrição razoavelmente detalhada do “programa da neurose” de Colby, cf. Margaret Boden, 1977, 21-63.

O episódio ELIZA revela como, de dentro das ciências do artificial, as máquinas podem ser ocasião de aparências fundadoras: como a encenação de certos elementos de aparência das máquinas pode estabelecer as condições de entrada dos humanos em interações com coisas no mundo que, até então, eram exclusivas das relações entre humanos. Os actuais desenvolvimentos da robótica humanóide abrem novas perspectivas a estes cenários, pelo que vamos tentar captar a dinâmica desse processo. Isso será feito no quadro mais vasto da relação do humano com o humanóide, considerando particularmente o caso das raças fabulosas do Oriente no pensar da humanidade, antes e depois dos Descobrimientos dos séculos XV e XVI.

3. O que é ser humanóide?

Dizemos que são “humanóides” os robôs que, de algum modo, têm uma aparência humana. Às vezes têm duas pernas, dois braços, uma cabeça; outras vezes são apenas pedaços, por exemplo um torso e uma cabeça. Enquanto não confundirmos máquinas com humanos, dizer de certos robôs que são humanóides remete apenas para essa aparência exterior. Se pensarmos, no entanto, que a forma da máquina serve, em muitos casos, de veículo a comportamentos que se querem comparáveis aos dos humanos, de veículo a expectativas de convívio entre humanos e máquinas, uma nova delicadeza desce sobre a questão.

Insistamos, então: o que é isso de ser humanóide? Não pode ter só a ver com a similitude do corpo: corpos de pessoas com extensas deficiências físicas continuam a ser corpos humanos. Pode ser também questão de comportamento, mas tão pouco por aí se traça uma fronteira definitiva: há comportamentos que consideramos inumanos, sem deixarmos de reconhecer que os seus autores continuam a ser humanos. Quer dizer: a questão do humanóide passa por dentro da questão do que é ser humano, não é uma questão para domínios claramente exteriores. É uma questão de fronteiras internas: onde está a linha para cá da qual o humanóide é humano?

Como estamos a falar de robôs, podemos pensar que essa questão pode ser rebatida sobre a fronteira entre natural e artificial, mas essa saída é enganadora. Embora esta seja uma questão inabitual para espécies naturais, ela não é completamente nova. Pensemos nos Neandertais⁶. Desapareceram

⁶ Finlayson, 2004; Trinkaus e Shipman, 1993.

do registo fóssil há uns 30.000 anos e antes disso, na vasta área geográfica da Europa ao sul da Sibéria que habitaram, terão convivido com os nossos antepassados. Supõe-se que seriam aproximadamente tão sofisticados como nós, mas parece não haver completa unanimidade quanto a considerá-los como uma subespécie dos humanos (*Homo sapiens neanderthalensis*, ao lado do *Homo sapiens sapiens*) ou como uma espécie humana separada (*Homo neanderthalensis*, *Homo sapiens*). Se hoje vivêssemos essa situação, questões como a da igualdade de direitos entre “nós” e “eles”, entre diferentes humanos, não seriam fáceis de resolver apelando simplesmente à biologia.

Não sabemos muito desse longínquo convívio com outros humanos, diferentes, que poderiam ser considerados humanóides naturais. Não podemos, pois, socorrer-nos desse antecedente para aclarar as consequências do humanóide para o humano. Temos, contudo, outro ponto de observação para as nossas atitudes passadas face a humanóides: as raças fabulosas do Oriente no imaginário ocidental. Fabuloso é o monstro – o qual, nos seus diferentes aspectos, ajuda a pensar o que nós próprios somos.

4. O que é ser monstro?

O que é ser monstruoso? O “monstro” tem vários sentidos. O primeiro é o das palavras em grego (τέρας) e em latim (*monstrum*) para “monstro”, que significavam originalmente “sinal enviado pelos deuses”, referindo-se a casos individuais, a filhos de casais humanos nascidos com malformações congénitas⁷. Esses “monstros” eram sinais do futuro (anúncio de grandes desgraças) ou do passado (castigos de pecados, designadamente da ordem dos interditos sexuais). O segundo sentido para “monstro” é o de animais resultantes do cruzamento de espécies: centauros, unicórnios. (Nesta acepção cabia, por vezes, o demónio disfarçado, por exemplo de cabra). As “raças fabulosas do Oriente” (ou, mais precisamente, da Índia) são o terceiro sentido para monstro, onde estão em causa raças (e não indivíduos), diferentes da humanidade conhecida sem deixarem de ser humanos. É este sentido que vamos explorar.

Os estereótipos das raças fabulosas emergem na Grécia Antiga, talvez já desde o séc. VI a.C., mantendo-se depois razoavelmente estáveis na cultura

⁷ Roux, 2008, 13-15, 80-81.

ocidental até ao século XVI. Plínio, o Velho, com a sua *História Natural*, terá sido o principal responsável pela sua ulterior retransmissão sucessiva. Contudo, um saudável cepticismo em relação aos testemunhos sem fundamentação escrutinada faz com que nem todos os autores antigos aceitem as histórias das raças fabulosas. É o caso de Estrabão e Ptolomeu, que nem as mencionam⁸.

De outro modo, mesmo autores sofisticadamente cépticos quanto ao uso das fontes contribuíram para adensar a problemática. Um bom exemplo é Santo Agostinho (séculos IV-V), que, embora considere provavelmente falsa a existência dessas raças, procura no plano geral da Criação um quadro teológico para a sua eventual existência⁹: “Para concluir esta questão com prudência e cautela: ou o que se conta dessas raças não se verifica; ou, se se verifica, não são homens; ou, se são homens, provêm de Adão.” O que o autor d’A Cidade de Deus pretende (na conclusão do capítulo 8 do Livro XVI, que acabamos de citar) é mostrar como, qualquer que seja o caso quanto à existência efectiva das raças fabulosas no Oriente, nada disso desmente a unidade da espécie humana, a harmonia da Criação ou a sabedoria do plano do Criador – harmonia e sabedoria que não dependem da nossa capacidade para as entender.

O certo é que muitos autores cristãos medievais acolheram as narrativas, particularmente de Plínio, o Velho, integrando o fantástico no imaginário, de tal modo que as raças fabulosas são generalizadamente mencionadas nas grandes enciclopédias dos séculos XII e XIII. Um aspecto interessante na reprodução deste imaginário é o papel da imagem no processo: a imagem é capaz de resistir ao cepticismo, de extravasar o enquadramento discursivo que lhe é explicitamente dado, ganhando uma força própria contra o desmentido racional. A representação pictórica ajudou a manter os monstros vivos, mesmo contra o cepticismo dos que representavam. Por exemplo, Sebastian Münster, na *Cosmographia*, publicada pela primeira vez em alemão em 1544, defendia que as raças monstruosas não existiam, porque, dava como razão, não havia delas nenhum relato que se pudesse considerar fidedigno. Não obstante, acompanhando essa opinião com a inserção de ilustrações das supostas raças, favoreceu a crença na sua existência¹⁰. É como se, espontaneamente, fizesse vencimento o pressuposto de que não

⁸ Wittkower, 1942.

⁹ Santo Agostinho, 1995, 1476.

¹⁰ Priore, 2000, 42.

pode ser pura inexistência o que tão magnificamente pode ser representado pela imagem.

Que as raças fabulosas do Oriente povoaram durante séculos o imaginário ocidental, está bem documentado. Questão outra é compreender o lugar dessa região do fabuloso no pensar da própria humanidade: no pensar do que é próprio de ser humano, do que é essencial ou contingente a esta condição humana particular, do que seja moralmente próprio da humanidade. Georges Canguilhem escreveu que “normal é o grau zero da monstruosidade”¹¹. Nesse sentido, o Outro ajuda a olhar para nós. Nas palavras de José Gil¹²: “Provavelmente, o homem só produz monstros por uma única razão: poder pensar a sua própria humanidade. Seria possível traçar a história das diferentes ideias ou definições que o homem se deu de si próprio através das diversas representações da monstruosidade humana que o acompanharam.” Cada espécie monstruosa, naquilo que nela é especificamente monstruoso, ilumina um aspecto da nossa condição humana conforme à norma.

O que faz de um monstro um monstro é o seu desvio característico. Monstros são sempre desvios do normal, em algum sentido. Certas partes do seu corpo são excessivas: muito grandes (orelhas, boca); em demasiado número (seis braços, quatro olhos). Falta algo importante: um olho, o nariz. Não se conformam à distinção entre masculino e feminino, sendo andróginos. Têm uma organização corporal errada: têm os pés virados para trás; ou, como os monópodes, têm um pé muito grande que usavam para se proteger do sol.



Versão monocromática de imagens de raças monstruosas na obra de Hartmann Schedel, *Chronica mundi*, publicada em Nuremberga, em 1493.

¹¹ Canguilhem, 1952, 160.

¹² Gil, 1994, 56.

O desvio do normal no plano físico, de forma sistemática (raças fabulosas), não é o único plano em que o monstro serve a caracterização da nossa humanidade. Os monstros tinham, também, um sentido moral. Não só os monstros individuais (crianças nascidas com malformações congénitas) eram sinais enviados pelos deuses, sinais de desgraças futuras ou manchas do pecado (ligados a tabus sexuais). Ao nível colectivo, certas raças fabulosas também tinham um significado moral: um povo de pigmeus era um povo humilde; os gigantes eram poços de orgulho; os cinocéfalos (cabeças de cão) eram quezilentos e caluniadores; os acéfalos (não tinham cabeça, sendo representados com olhos no peito), seriam povos sem comando, povos sem organização política, anárquicos (numa identificação entre o corpo, como microcosmos, e a sociedade, como macrocosmos)¹³. Deste modo, o que era distante, na sua diferença, ajuda a compor a norma, à medida que a explicita por contraste, tanto no plano físico como no plano comportamental, no plano do indivíduo como no plano do colectivo.



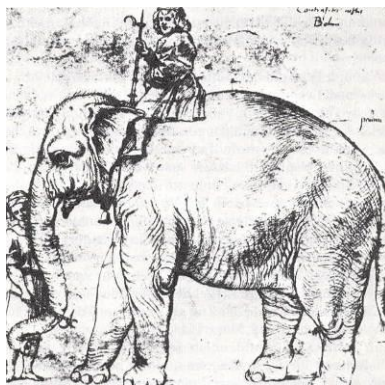
Imagem de um povo acéfalo na América, na obra de Levin Hulsius, *Kurtze Wunderbare Beschreibung, Dess Goldreichen Königsreichs Guianae in America oder neuen Welt*, publicada em Nuremberga, em 1603.

5. A imagem e a viagem: os efeitos contraditórios dos Descobrimentos

Durante séculos, a distância entre Oriente e Ocidente serve de estabilizador da relação entre o conhecido e o fabuloso. A imagem faz a viagem: nós não vamos, nem conhecemos quem vá, ao Oriente, mas a representação pictórica tem uma tremenda força de apresentação. Alguns

¹³ Roux, 2008, 87; Gimenez, 2001, 288.

viajantes famosos, logo no século XIII, fazem relatos que, por muito impressionantes que sejam, não movem substancialmente o imaginário tradicional: os testemunhos contrários, dispersos, não conseguem destruir a ideia da real existência de raças fabulosas no Oriente, de algum modo correspondentes aos relatos e às imagens que circulam¹⁴. Vai ser preciso massificar a viagem para mover o terreno da tradição adubado pela imagem. Poderíamos, assim, pensar que as grandes viagens de descobrimento, no século XVI, quebrariam o encanto e, obrigando ao confronto directo com o real do Oriente, transformariam de forma definitiva o imaginário ocidental do fabuloso da Índia. Ora, se esse efeito existe, de facto, há impulsos contraditórios que tornam o processo mais complexo.

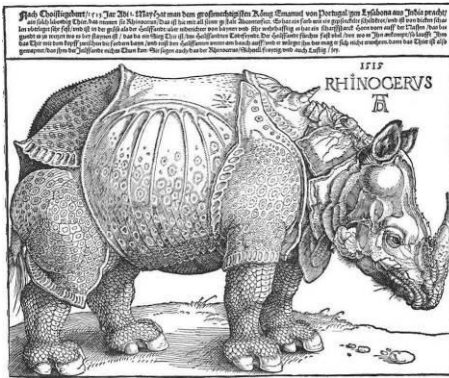


Desenho do elefante Hanno, atribuído a Rafael (ou Giulio Romano, a partir de um desenho de Rafael), c. 1514/1516 (Staatliche Museen, Berlin).

Efectivamente, ninguém encontrou as raças fabulosas, os humanóides clássicos. As populares enciclopédias eram fantasiosas nas suas descrições e as ilustrações enganavam. Contudo, outras maravilhas do Oriente renovam a força do imaginário como capaz de se mostrar traduzido na realidade: os animais raros. Hanno, o elefante indiano enviado por D. Manuel I que chegou ao papa Leão X em 1514, dá a ver a Roma uma espécie que a cidade não via há séculos. Fez sucesso, tornando-se a grande atracção dessas manifestações de poder e glória que eram as procissões católicas. Vivo, em carne e osso, dava força à tradição pictórica: afinal as gravuras do fabuloso mostravam coisas reais.

¹⁴ Woortmann, 2005.

O rinoceronte que o mesmo Rei português enviou, mais tarde, ao mesmo Papa, morreu pelo caminho, mas isso não o impediu de alcançar ainda maior popularidade, graças a uma gravura de Albrecht Dürer, datada de 1515, que, embora baseada numa descrição enviada de Lisboa, deve tanto à imaginação pictórica como ao seu referente no mundo. Este caso ilustra o lento evoluir das relações entre o fabuloso e o real enquanto vivem na imagem: o antigo monoceros (um corno) dividiu-se em dois: rinoceronte (animal real), unicórnio (nunca visto). Embora conhecendo a diferença de estatuto ontológico, o mesmo Dürer da gravura do rinoceronte também representou um unicórnio: na imagem, o nunca visto e o claramente observável podem parecer no mesmo modo¹⁵.



Esquerda: Albrecht Dürer, *Rinoceronte*, 1515, British Museum. Direita: Albrecht Dürer, *O rapto de Proserpina num Unicórnio*, c. 1516, Fine Arts Museum, San Francisco.

A harmonização dos relatos antigos com os novos métodos e dados foi uma questão problemática mesmo para os mais exigentes dos cientistas. Um exemplo do século XVII é o Dr. Tulpp, o médico que Rembrandt pintou na “Lição de Anatomia”. O Dr. Tulpp, que desenhou, com incrível precisão, alguns monstros biológicos que ele tinha dissecado, também desenhou um símio, com a inscrição: “Homo sylvestris – Orangoutang”. (Era, de facto, um chimpanzé.) Num texto, onde estuda a questão de uma das célebres espécies fabulosas do imaginário medieval, intitulado “Satyrus Indicus”, conclui: ou os sátiros não existem, ou, se existem, são, afinal, este animal. A um nível tão exigente da prática científica também era uma tarefa trabalhosa tentar harmonizar tradição textual e observação.

¹⁵ Barbas, 2000.



Nicolaas Tulp, *Homo sylvestris*. Orang-outang, *Observationum Medicarum Libri Tres*, Amsterdam, 1641, p. 275, Figura XIV.

Portanto, no tempo dos Descobrimentos, os animais de maravilha prolongam os efeitos contraditórios da combinação da viagem com a imagem na concepção do real. Não se pense, contudo, que isto quer dizer que tenham desaparecido as questões mais directamente implicadas com a questão do humano e do humanóide. Na verdade, com o avanço das descobertas, a questão das fronteiras do humano torna-se, galgadas as distâncias, uma questão de grande relevância prática. As questões acerca dos limites da humanidade são transpostas para África e para o Brasil – e os debates teológicos acerca de os indígenas serem ou não providos de alma, e da sua natureza em geral, não se circunscrevem já ao domínio da teoria, passando a ser assuntos de administração das possessões, questões políticas decisivas para todos os que nos reinos passaram a pensar em muito mais larga escala¹⁶. Assim, a questão das fronteiras entre o humano e o humanóide, se adquire novos matizes, não se dissolve pelo encontro com a pátria original das raças fabulosas, as Índias Orientais. Ainda aparecem, a par de relatos de novos animais (por exemplo, o “ganso de Magalhães”, o pinguim), sugestões de novos monstros humanóides no Novo Mundo (por exemplo, relatos de homens marinhos no Brasil, nos séculos XVI e XVII, como os deixados por Fernão Cardim ou Gabriel Soares)¹⁷.

¹⁶ Massimi, 2003.

¹⁷ Priore, 2000, 85ss; Gimenez, 2001.

6. Novos humanóides: robôs

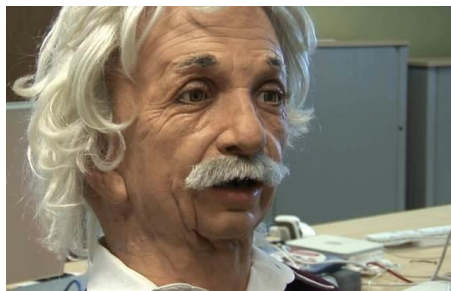
Até que ponto a pergunta “por que construímos robôs humanóides?” pode ser esclarecida pelas modalidades, que visitámos, de outros contactos que tivemos em outros tempos com outros humanóides? As raças fabulosas do Oriente, mesmo não existindo, ajudavam a construir a norma da nossa humanidade. O que nos era dado ver delas (a representação pictórica) entrava na relação com a humanidade conhecida no papel de Outro, tanto física como moral ou comportamentalmente. Questões centrais na antropologia dominante ao tempo (a espécie humana no plano da Criação) tinham de responder a dificuldades que, além de testemunhos vagos, só tinham como suporte as imagens em proliferação. A força das imagens na criação de um mundo fabuloso era enorme, resistindo inclusivamente a enquadramentos discursivos explícitos que contrariavam a credibilidade da espessura ontológica dessas imagens. Em suma: a imaginação, e os seus suportes materiais, ajudavam a criar um mundo fabuloso que dialogava com o mundo conhecido na posição de fronteira da humanidade em exploração. Sugiro que é assim que devemos tentar compreender por que se fazem, hoje, robôs humanóides: mais do que construir máquinas, quer-se compreender os humanos. Quer-se compreender os humanos como máquinas, julga-se que esse é um sinal de cientificidade na busca pela compreensão do humano.

Se nem todos os investigadores em robótica humanóide assumem explicitamente o objectivo de compreender os humanos através dos robôs, encontramos casos onde esse objectivo é apresentado de forma transparente. É essa a orientação, por exemplo, de Hiroshi Ishiguro, cujo laboratório em Osaka (Japão) tem produzido alguns dos robôs humanóides mais impressionantes dos últimos anos, incluindo a série Geminoid, cujo primeiro exemplar é um duplo do próprio Ishiguro em tamanho natural. Lemos, dessa equipa, num artigo com um título programático (“construir humanos artificiais para compreender os humanos”), uma defesa da “ciência andróide”. A “ciência andróide” tem uma meta (“realizar um robô humanóide e encontrar os factores essenciais para a representação do que é ser semelhante ao humano”) e persegue essa meta combinando duas abordagens: (1) construir andróides, robôs muito parecidos com os humanos

(em aparência e em comportamento) e (2) usar esses andróides para explorar, no quadro das ciências cognitivas, a natureza humana¹⁸.

É certo que a investigação em robótica humanóide pode ter uma motivação meramente funcional: um utilizador, face a um dispositivo robótico com a aparência de, digamos, uma mão, conjectura com alguma segurança qual o tipo de interacção que o projectista provavelmente concebeu que tenhamos com esse dispositivo. A expectativa desta identificação é um argumento a favor de recorrer a formas humanóides e, neste sentido, essas formas humanóides têm uma motivação funcional. Aliás, esta motivação funcional pode combinar com a motivação referida anteriormente (explorar a natureza humana). Esta motivação funcional é claramente reconhecida noutro texto produzido pela equipa de Ishiguro¹⁹: "desenvolvemos o Geminoid porque acreditamos que, para facilitar uma interacção humano-robô eficaz, tanto as funções como a aparência do robô deve ser optimizadas para tirar proveito de todas as especializações cognitivas que humanos possam ter para reconhecer outros humanos."

Para alguns, esta compreensão mecanicista é necessária para desfazer as ilusões humanistas, ou para encontrar vias de compreensão que julgam inacessíveis às Humanidades. Fazer humanos será, para alguns, o passo que se segue a fazer humanóides. E, para isso, os robôs tomam formas que pretendem aproximar-se dos humanos, desempenham papéis que costumavam estar reservados aos humanos, misturam-se na forma de vida dos humanos. Vários exemplos de robótica humanóide mostram várias linhas de progressão nessa senda. Vejamos.

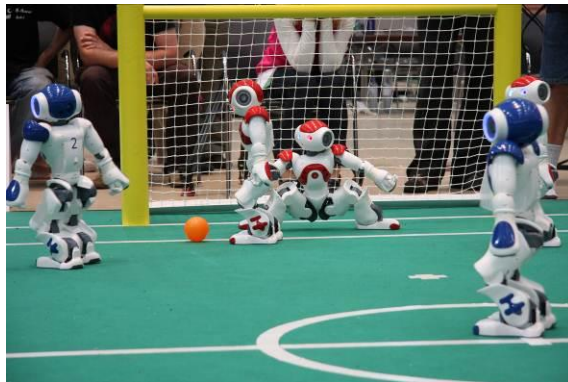


O robô Einstein

¹⁸ Ishiguro e Nishio, 2007, 134-135.

¹⁹ Ogawa et al., 2012, p. 55.

O robô Einstein foi projectado segundo a linha de robôs hiper-realistas da Hanson Robotics, que procura para o robô uma aparência (facial) virtualmente indistinguível de um humano (pele, voz, olhos e contacto visual imitando perfeitamente modelos humanos)²⁰. Desenvolvido cognitivamente pelo Laboratório de Percepção para Máquinas da Universidade da Califórnia em San Diego, o robô Einstein pretende ser uma máquina emocionalmente inteligente. O modelo do famoso físico (ou melhor, da sua cabeça), produz um vasto leque de expressões faciais diferenciadas (graças a 31 motores internos que movem os "músculos" da "face"), segue os movimentos dos olhos dos humanos que se colocam como seus interlocutores, reconhece pistas auditivas e certos gestos (como o abanar da cabeça), identifica e interpreta expressões faciais dos humanos e trata de imitá-las, capta pistas sobre a idade e o género dos humanos que o rodeiam, bem como certas indicações particulares sobre os mesmos (como usarem ou não óculos). Consegue, por esses meios, uma interacção forte com humanos²¹. Um dos usos previstos para este "Einstein" é que ele ajude a entusiasmar adolescentes pelo estudo da física. Podemos, um dia, precisar de um esforço especial para distinguir um destes robôs hiper-realistas de verdadeiros humanos?



O futebol dos robôs. RoboCup 2009, Graz (Foto de Porfírio Silva)

Outro exemplo são os robôs jogadores de futebol. Desde 1997 que se disputa o RoboCup, o Campeonato Mundial de Futebol Robótico, uma

²⁰ Vídeo do Machine Perception Laboratory em <http://www.youtube.com/watch?v=pkpWCu1k0ZI>

²¹ Wu et al., 2009.

iniciativa que combina objectivos educacionais, e de divulgação, com objectivos de investigação ligados aos robôs autónomos com inteligência²². Reúne, anualmente, dezenas de equipas de todo o mundo, mobilizando centenas de participantes em diferentes modalidades, das quais se destacam as que visam contribuir para o objectivo último do RoboCup: que, por volta de 2050, uma equipa de robôs autónomos humanóides vença num jogo de futebol a equipa campeã mundial (dos humanos) segundo as regras da FIFA²³. Nem todas as equipas que participam nas ligas de futebol robótico são compostas de robôs humanóides (em alguns casos, são robôs com rodas), mas todas tentam implementar nas máquinas comportamentos colectivos sofisticados (desempenhar uma acção colectiva no mundo físico, cooperação dentro de uma equipa, competição entre equipas). Nos robôs humanóides que jogam futebol, visando vir a jogar com humanos, há, simultaneamente, uma tentativa de aproximar a forma das máquinas da forma corporal dos humanos e uma tentativa de construir máquinas que entrem numa actividade que faz um sentido específico na forma de vida dos humanos (jogar um certo jogo colectivo).

A “robótica do desenvolvimento” leva-nos a um novo patamar da interacção entre humano e humanóide. Uma diferença importante entre, por um lado, as máquinas resultantes da generalidade dos ramos da robótica e, por outro lado, muitas espécies naturais, das mais sofisticadas, é que as máquinas “nascem adultas”: uma máquina é feita, o melhor que os seus construtores conseguem, para estar dotada de todas as suas capacidades logo que é colocada ao serviço. Diferentemente, os humanos (e muitas outras espécies animais) só chegam a um estado de maturidade após um (mais ou menos) longo processo de desenvolvimento: todo o complexo de processos que, nas espécies que se reproduzem sexualmente, levam da célula única resultante da fecundação ao indivíduo adulto completamente formado. No caso dos humanos, e apenas para o período pós-natal, os juvenis desenvolvem-se, tanto corporal como mentalmente, passo a passo, ao longo de vários anos, graças à interacção física e simbólica com outros espécimes da mesma espécie, sendo que estes, melhor ou pior, adaptam o seu modo de interacção às capacidades que, tipicamente, aquele ser terá naquela fase do seu desenvolvimento. Mesmo sem grande preparação específica, sabemos que não falamos com crianças de dez e de três anos da

²² Informação extensa e actualizada sobre o RoboCup em <http://www.robocup.org>.

²³ Asada e Kitano, 1999.

mesma maneira, tal como sabemos que o tipo de interacção física saudável é muito diferente para cada uma dessas idades ou para um bebé de meses.

Mesmo ramos da Nova Robótica que procuram análogos artificiais de processos naturais descaram a questão do desenvolvimento. É o caso da Robótica Evolutiva²⁴, que opera, por assim dizer, com gerações sucessivas de robôs produzidos “adultos”. Essa falta de atenção ao desenvolvimento (pré-natal ou pós-natal) é o espaço que pretende ser ocupado pela Robótica do Desenvolvimento, ou Robótica Epigenética, como resposta ao diagnóstico de que esse pode ser um entrave crucial às ambições das Ciências do Artificial. Como escrevem Lungarella e os seus colegas²⁵: “A mera observação de que quase todos os sistemas biológicos – em diferentes medidas – passam por processos de amadurecimento e desenvolvimento, comporta a convincente mensagem de que o desenvolvimento é a principal razão pela qual a adaptabilidade e a flexibilidade dos sistemas compostos orgânicos transcende a dos sistemas artificiais”.

A Robótica do Desenvolvimento não constitui ainda um campo de investigação bem delimitado e permanece muito heterogéneo. Autores diferentes concentram-se em momentos e aspectos diferentes da interacção entre organismos e ambiente no desenvolvimento de um organismo. Por exemplo, na esteira dos trabalhos de Teuscher e seus colegas²⁶, vem uma preferência por abordagens centradas na concorrência de três processos (filogenia, ontogenia, epigenia) que, em escalas temporais diferentes, conformam os organismos adultos de uma dada espécie. Já Zlatev e Balkenius²⁷ induzem uma abordagem mais interessada pelos aspectos psicológicos do desenvolvimento pós-natal. De qualquer modo, a “robótica do desenvolvimento” leva muito a sério esta diferença entre criaturas artificiais e criaturas naturais, considerando muito poderoso o processo de desenvolvimento: indivíduos da mesma espécie, agindo informalmente (não interagimos com um bebé para o programar) mas ajustando aproximativamente a interacção ao juvenil (não faço movimento muito bruscos com o bebé, não tento discutir geografia com ele), estimulam uma progressão suave, incremental, que tira partido das competências inatas e das aquisições anteriores para levar o indivíduo ao melhor desdobramento possível das suas potencialidades. Ora, e se avançássemos uma etapa nesta

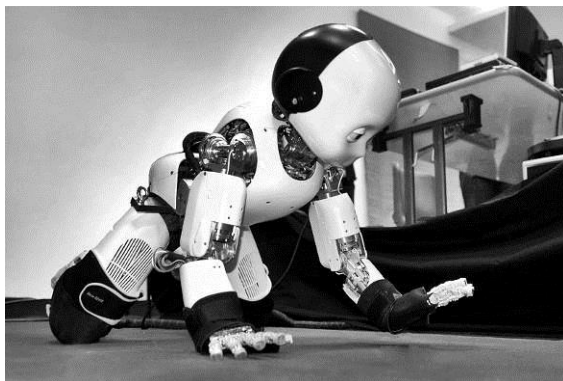
²⁴ Nolfi e Floreano, 2000.

²⁵ Lungarella et al., 2003, 179.

²⁶ Teuscher et al., 2003.

²⁷ Zlatev e Balkenius, 2001.

interacção entre humano e robô humanóide, criando um “filhote-robô” e dando-lhe a oportunidade de um processo de desenvolvimento artificial?



Filhote de robô gatinhando (sítio do iCub em <http://www.icub.org>)

É isso que faz o projecto RobotCub com os seus robôs iCub. Uma breve menção ao projecto RoboCub ilustrará alguns dos aspectos mais interessantes a esperar da Robótica do Desenvolvimento. O “filhote-robô” é um projecto internacional (iniciado em 2004) que construiu uma série de robôs designados como iCub²⁸. Trata-se de um robô humanóide representando as características físicas e cognitivas de uma criança humana de dois a três anos, capaz de gatinhar e de manipular objectos – e de aprender pela interacção com humanos. O seu “corpo”, com 53 graus de liberdade, nove dos quais nas mãos com três dedos independentes e outros dois para estabilidade e suporte, seis dos quais nas pernas que deverão permitir locomoção bípede; as câmaras digitais para a visão, os microfones e outros sensores; no futuro uma pele artificial; e um poder computacional fornecido por máquinas exteriores ligadas por cabos – estão já a permitir experiências de interacção com humanos, dirigidas para perceber melhor como é que as capacidades sensoriomotoras e cognitivas de um espécime jovem resultam dessa interacção com outros membros de uma espécie natural.

²⁸ Para uma primeira apresentação geral do projecto, cf. (Sandini *et al.*, 2004). Toda a informação posterior, também sobre o robô iCub, incluindo as imagens, foi recolhida no site oficial do projecto, em <http://www.robotcub.org/>. O Instituto de Sistemas e Robótica (Instituto Superior Técnico) é um dos intervenientes relevantes neste projecto.

Aqui, o robô não “nasce adulto”, ao “filhote de robô” é dada a oportunidade de entrar num processo onde a plasticidade do seu sistema de controlo é largamente influenciada pela interacção com humanos. Por exemplo, não o programando directamente para reconhecer determinados objectos no seu ambiente, mas dotando-o das capacidades para adquirir para o seu mundo de significado objectos que lhe são dados a conhecer de forma adequada aos seus processos de aprendizagem. Ou, ainda por exemplo, deixando-o descobrir por experiência própria como lidar com certos objectos, dependendo nomeadamente da respectiva forma e peso, em lugar de o instruir explicitamente acerca de como fazer em cada caso. Desse modo, robôs parecidos à partida vão “crescer” de forma diferente na sua relação com o mundo, de acordo com as diferentes experiências que lhe serão proporcionadas: assim se tornando uma espécie de “filhotes” dos seus “parentes” humanos.

Estes exemplos, algumas das mais notáveis promessas da investigação actual em robótica humanóide, já não constituem apenas representações imagéticas de possíveis seres nas fronteiras do humano: prometem interferir no plano das aparências fundadoras reveladas pelo programa ELIZA, referido inicialmente. Vamos terminar propondo, como via para compreender essa dinâmica, que está em causa um processo de captura da postura intencional pelo projecto das máquinas humanóides.

7. A captura da intencionalidade

A nosso ver, os exemplos de robótica humanóide que introduzimos mostram como, de novo, estamos a produzir representações de variantes do humano que nos servem de espelho. Desta vez são máquinas, mas com essas máquinas exploramos as nossas fronteiras – e exploramos as fronteiras das nossas capacidades para nos compreendermos e nos relacionarmos com outros humanos. Pode ser que nos habituemos a esse novo Outro. Nas palavras de José Gil: «Neste fim de século, os monstros proliferam (...). Cessarão, muito em breve, de nos parecer monstruosos e ser-nos-ão até simpáticos (...). Havemos de falar então da “monstruosidade banal”»²⁹. Embora Gil não estivesse a falar propriamente de robôs, deixa-nos uma pista para tentar compreender o que andamos a fazer quando construímos robôs humanóides: estamos, talvez, a criar as condições para

²⁹ Gil, 1994, 9.

um convívio entre humanos e humanóides onde aqueles já nem sempre se distinguem nitidamente destes, onde humanos já nem sempre tratem outros humanos e humanóides de formas em tudo diferentes. Vamos recorrer à noção dennettiana de “postura intencional” para esclarecer este ponto, para podermos depois passar à noção de captura da intencionalidade.

A postura intencional é uma forma de interpretarmos os sistemas com que interagimos. Se, com o fito de compreender um determinado sistema, ou para nos relacionarmos eficazmente com ele, não for praticável prever o seu comportamento a partir do seu estado físico actual e de todas as leis físicas aplicáveis (postura física), nem conhecermos o suficiente sobre a sua arquitectura funcional para sermos capazes de prever o seu comportamento (postura do design), podemos por vezes adoptar a postura intencional: olhamos para esse sistema como tendo certas crenças e desejos e prevemos o seu comportamento nessa base. Seja o exemplo de um computador programado para jogar xadrez: como não é praticável tentar conhecer todas as minúcias do funcionamento dos seus circuitos electrónicos para jogar contra ele, nem sabemos como foi programado, adoptamos a postura intencional: supomos que nos quer ganhar, que tem um certo conhecimento do jogo e que se comportará de modo a satisfazer esse desejo. Nesta postura podemos até pensar que o computador faz certa jogada para nos distrair ou irritar: que isso seja ridículo não importa; o que importa é que essa postura nos orienta na interacção com o computador programado para lidar com o xadrez. Ora, segundo Dennett, também nas relações entre humanos adoptamos a postura intencional. Não se trata de saber se esses sistemas (artificiais ou humanos) têm realmente crenças e desejos – mas da utilidade de os vermos assim: “a definição que dei de sistemas intencionais não diz que os sistemas intencionais realmente têm crenças e desejos, mas que podemos explicar e prever o seu comportamento atribuindo-lhes crenças e desejos”; “A decisão de adoptar a estratégia é pragmática, não é intrinsecamente certa ou errada”³⁰. E exactamente o mesmo se dirá da postura intencional face a máquinas. Esta posição é completamente instrumentalista: não procuramos compreender a realidade do comportamento de outros agentes, limitamo-nos a adoptar uma estratégia para lidar com eles. Apesar de Dennett ter andado durante muitos anos a tentar corrigir este instrumentalismo primário da sua tese³¹, é precisamente

³⁰ Dennett, 1971, 7.

³¹ Cf. Dennett, 1981 e Dennett, 1991.

este instrumentalismo pouco refinado que tem recolhido as preferências de muitos praticantes da chamada Robótica Social³².

Ora, o que dizemos é que a robótica humanóide explora a nossa tendência para a postura intencional. Quando interagimos com uma criatura que nos parece humana, espontaneamente tratamo-la como humana: não começamos por tentar averiguar se é realmente da mesma espécie que nós; assumimos que é esse o caso, o que tem, como implicação notável, que adoptamos face a ela o pressuposto básico de ser dotada do mesmo tipo de intencionalidade que nós. Quem projecta e constrói robôs humanóides, procurando que sejam realistas nessa “humanóide”, está a tentar que os humanos adoptem em face desses robôs o mesmo pressuposto de intencionalidade, que esses humanos se relacionem com esses robôs como se estes fossem basicamente movidos por crenças e desejos ao mesmo título que nós.

Os construtores de robôs humanóides realistas tentam que as suas máquinas despertem em nós esse tipo de comportamento, porque isso denotaria o sucesso dos seus esforços para criar modelos bem-sucedidos de máquinas que parecem e se comportam como humanos. Fazer durar uma interacção entre humanos e robôs onde aqueles tomem estes por humanos e assim se orientem na interacção, como se essa interacção fosse uma relação, será um sinal de sucesso de qualquer projecto de robótica humanóide realista. E esse sucesso da robótica assentará na captura da postura intencional: manipular processos cognitivos humanos de modo a ser capaz de provocar, de forma controlada e sustentada, a adopção da postura intencional face a robôs.

Sem pretender que todas as linhas de investigação em robótica humanóide visam essa meta, poderá ser útil exemplificar como essa captura da intencionalidade é prosseguida em experiências relevantes neste campo. Nesse quadro, é interessante mencionar um conjunto de experiências com a série de robôs Geminoid, já referida. Uma parte das experiências relatadas por Ogawa e seus colegas pretendiam averiguar como é que o robô “gémeo” de um humano particular, conhecido, comparava com meios de comunicação mais tradicionais (por exemplo, vídeo-conferência) em termos de capacidade para transmitir “presença”, naturalidade e humanidade, bem como até que ponto o Geminoid era persuasivo a passar uma mensagem (de tipo comercial), desta vez comparado com o seu “original” humano e com uma

³² Cf., por exemplo, Breazeal, 2002.

gravação vídeo da mesma pessoa a transmitir a mesma mensagem. Dos resultados dessas experiências, concluem os autores que “o Geminoid tem potencialmente um nível de presença similar ao de uma pessoa real e, devido a este nível de presença, o Geminoid pode ter uma influência social nos humanos no contexto de uma comunicação persuasiva ou face a face”³³. Embora sem discutir aqui e agora se as experiências em causa efectivamente permitem as conclusões extraídas, porque isso nos obrigaria a questionar os enviesamentos induzidos por múltiplas simplificações no plano dos pressupostos, o que podemos afirmar é que estas experiências revelam precisamente um projecto de captura da intencionalidade, na medida em que procuram os meios para que robôs consigam espoletar em humanos o tipo de aceitação que, anteriormente, só outros humanos chegavam a provocar.

Outra série de experiências com os Geminoid lida com o conceito de “transferência do corpo próprio”³⁴. Quando falamos de “transferência do corpo próprio” falamos de que, em certas circunstâncias, o humano que está a teleoperar o robô sente como se o “corpo” do robô fosse uma extensão do seu próprio corpo (por exemplo, sente um incómodo físico quando o robô que está a comandar é sujeito a uma acção que, no corpo humano, causaria dor). Ora, o que queremos fazer notar é como esta equipa fala claramente de “induzir a ilusão de corpo próprio” através da experiência de teleoperar o andróide, ao mesmo tempo que se espera que essa ilusão do operador torne mais efectivo o próprio controlo remoto do humanóide, tornando-o mais eficiente a transmitir presença a outros humanos³⁵. Há aqui, claramente, uma tentativa de misturar o corpo próprio com a máquina, com consequências no plano da interacção entre humanos, envolvendo novas possibilidades de captura da intencionalidade.

Um texto anterior desta equipa já tinha traçado a estratégia técnica para esta captura da intencionalidade. Vejamos como aí se apresenta a questão. Estando os humanos afinados pela evolução natural para identificar e processar estímulos particularmente relevantes no seu entorno habitual, aquele em que mais frequentemente deparamos com oportunidades e desafios relevantes, a percepção da aparência humana provoca reacções específicas. Tal é a força dessa preparação natural que acabamos por antropomorfizar muitos objectos: a interpretação humanóide é normalmente tentada, mesmo que venha a ser defraudada em certas ocasiões. O que está

³³ Ogawa et al., 2012, 58.

³⁴ Nishio et al., 2012.

³⁵ Ogawa et al., 2012, 59.

em causa, concretamente, é que vários módulos cerebrais, ao observarmos uma cena, tentam interpretar os estímulos sensoriais em termos de modelos humanos. Mesmo que essa tentativa só obtenha sucesso parcial, ela pode influenciar as nossas reacções. Por exemplo, podemos reconhecer que um andróide é um robô e não um humano e, mesmo assim, reagir-lhe de maneiras que usualmente reservamos aos humanos. O que estes proponentes da “ciência andróide” tentam fazer é explorar esses mecanismos cerebrais e procurar afinar a aparência de certos robôs humanóides para que ela ocasione esse tipo de ocorrências no funcionamento do nosso cérebro³⁶. Nesse sentido, trata-se de uma abordagem que recorre à captura da postura intencional – o que, em termos instrumentalistas, equivale à captura da intencionalidade.

Capturar a postura intencional – tornar porosas as fronteiras entre a relação humana e a relação de humanos com humanóides – será, pois, uma estratégia básica de certas linhas da Nova Robótica. Se estamos ou não a cruzar outra fronteira – a fronteira entre a captura da postura intencional e a captura da intencionalidade genuína dos intervenientes numa relação – depende de um esclarecimento de maior alcance acerca da própria intencionalidade. Para quem tenha uma visão instrumentalista da intencionalidade – o que existe na intencionalidade é apenas o sucesso da postura intencional dos interlocutores –, capturar a postura intencional não será menos nem mais do que capturar plenamente a intencionalidade. Essa questão fica propriamente como questão: se alguma linha de investigação em robôs humanóides capturar consistente e duradouramente a nossa postura intencional, teremos entrado num processo de convergência entre a intencionalidade humana e a intencionalidade de certas máquinas? As realizações mais recentes da robótica humanóide sugerem que esta não é uma questão ociosa.

A imagem, como mostrámos em secções anteriores, esteve historicamente implicada na interrogação pela humanidade. A imagem, enquanto aparência humanóide, continua hoje em jogo em linhas de investigação robótica que usam a construção dessas máquinas para tentar compreender o humano. O nosso ponto é que, sendo a questão do humanóide uma questão de fronteiras internas (passa por dentro da questão do que é ser humano, como dissemos antes), temos de pensar nas experiências com humanóides como experiências com o humano. Daí a

³⁶ Nishio et al., 2007, 345.

proposta do presente texto para pensarmos essa questão em perspectiva histórica, iluminando os novos usos dos humanóides (robôs) com os antigos usos de outros humanóides (monstros).

Referências Bibliográficas

Santo Agostinho, 1995, *A Cidade de Deus*, tradução de J. Dias Pereira, Lisboa: Fundação Calouste Gulbenkian.

Asada, M. e Kitano, H., 1999, The RoboCup Challenge. *Robotics and Autonomous Systems*, 29, 3-12.

Barbas, H., 2000, Monstros: O Rinoceronte e o Elefante. Da Ficção dos Bestiários à Realidade Testemunhal. In: *Actas do V Encontro Luso-Alemão / Akten der V. Deutsch Portugiesischen Arbeitgespräche*, Köln-Lisboa, 103-122.

Boden, M., 1977, *Artificial Intelligence and Natural Man*, Hassocks, The Harvester Press.

Breazeal, C., 2002, *Designing Sociable Robots*, Cambridge, The MIT Press.

Canguilhem, G., 1952, *La connaissance de la vie*, Paris, Vrin (Reed. 1989).

Dennett, D.C., 1971, Intentional Systems. Reimpressão in D.C. Dennett, *Brainstorms*, Londres, Penguin, 1977, 3-22.

Dennett, D.C., 1981, True Believers: The Intentional Strategy and Why It Works. Reimpressão in D.C. Dennett, *The Intentional Stance*, Cambridge, The MIT Press, 1989, 13-35.

Dennett, D.C., 1991, Real Patterns. Reimpressão in D.C. Dennett, *Brainchildren*, Londres, Penguin, 1998, 95-120.

Finlayson, C., 2004, *Neanderthals and Modern Humans: An Ecological and Evolutionary Perspective*, Cambridge, Cambridge University Press.

Gardner, H., 1985, *The Mind's New Science*, New York, Basic Books.

Gil, J., 1994, *Monstros*, Lisboa, Quetzal Editores.

Gimenez, J.C., 2001, A presença do imaginário medieval no Brasil colonial: descrições dos viajantes. *Acta Scientiarum*, 23, 207-213.

Ishiguro, H. e Nishio, S., 2007, Building artificial humans to understand humans. *Journal of Artificial Organs*, 10, 133-142.

Lungarella, M. et al., 2003, Developmental robotics: a survey. *Connection Science*, 15, 151-190.

Massimi, M., 2003, Representações acerca dos índios brasileiros em documentos jesuítas do século XVI. *Memorandum*, 5, 69-85.

- Nishio, S. et al., 2007, Geminoid: Teleoperated Android of an Existing Person. In: *Humanoid Robots: New Developments*, Viena, I-Tech Education and Publishing, 343-352.
- Nishio, S. et al., 2012, Body Ownership Transfer to Teleoperated Android Robot. In: *Social Robotics. Proceedings of the 4th International Conference (ICSR 2012)*, Lecture Notes in Artificial Intelligence 7621, Berlin Heidelberg, Springer-Verlag, 398-407.
- Nolfi, S. e Floreano, D., 2000, *Evolutionary Robotics*, Cambridge, The MIT Press.
- Ogawa, K. et al., 2012, Android Robots as Tele-presence Media. In: *Biomedical Engineering and Cognitive Neuroscience for Healthcare: Interdisciplinary Applications*, Pennsylvania, Medical Information Science Reference, 54-63.
- Priore, M., 2000, *Esquecidos por Deus*, São Paulo, Companhia das Letras.
- Roux, O., 2008, *Monstres*, Paris, CNRS Éditions.
- Sandini, G. et al., 2004, RobotCub, An Open Framework for Research in Embodied Cognition". In: *Proceedings of Humanoids 2004 (IEEE-RAS/RSJ International Conference on Humanoid Robots)*, Los Angeles, Novembro de 2004 (Disponível no sítio do projecto <http://www.robotcub.org>)
- Silva, P., 2011, *Das Sociedades Humanas às Sociedades Artificiais*, Lisboa, Âncora.
- Schopman, J., 1987, Frames of Artificial Intelligence. In: *The Question of Artificial Intelligence: Philosophical and Sociological Perspectives*, Londres: Croom Helm, 165-219.
- Teuscher, C. et al., 2003, Bio-inspired computing tissues: towards machines that evolve, grow, and learn. *BioSystems*, 68, 235-244.
- Trinkaus, E. e Shipman, P., 1993, *The Neanderthals: Changing the Image of Mankind*, New York, Knopf.
- Weizenbaum, J., 1966, ELIZA - A computer Program for the Study of Natural Language Communication between Man and Machine. *Communications of the ACM*, 9, 36-45.
- Weizenbaum, J., 1976, *Computer Power and Human Reason*, San Francisco, W.H. Freeman.
- Wittkower, R., 1942, Marvels of the East. A Study in the History of Monsters. *Journal of the Warburg and Courtauld Institutes*, 5, 159-197.
- Woortmann, K., 2005, O selvagem na 'Gesta Dei': história e alteridade no pensamento medieval. *Revista Brasileira de História*, 25, 259-314.
- Wu, T. et al., 2009, Learning to make facial expressions. In: *Proceedings of the International Conference on Development and Learning (ICDL)*, Shanghai.
- Zlatev, J. e Balkenius, C., 2001, Why 'Epigenetic Robotics'?. In: *Proceedings of the First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, Lund: Lund University Cognitive Studies, 1-4.

Genuine Becoming and the Barcan Formula⁺

Emiliano Boccardi
(Centre in Metaphysics of the University of Geneva)
emiliano.boccardi@gmail.com

1. Introduction

Presentism is the doctrine that, necessarily, only presently existing things exist. Past things and states of affairs, according to this view, ceased to exist. The chief allure of Presentism is its promise to do justice to the idea that the passage of time is an objective (and absolute) feature of reality, and that it involves a genuine change in what exists. Caesar, one would like to say, simply does not exist: not any more, at any rate. We have no chance of establishing any direct epistemic contact with *him*, and it is awkward to think that *he* should nevertheless exist somewhere, enjoying some peculiar kind of non-being, or some very tenuous kind of (non-concrete) being. Much better to simply think that Caesar does not exist at all.

This theory of time appears to foot a double bill. On the one side, it promises to provide us with a clear and coherent explication of qualitative change. On the other side, if it proved to be a viable account of time, it would also provide us with an explication of the elusive notion of temporal passage. Qualitative change, according to this picture, ultimately occurs in virtue of the fact that absolute becoming occurs at all times. The explanation goes roughly as follows. Qualitative change consists in different monadic states of affairs obtaining at different times.¹ According to the presentist, at all times there are states of affairs that come to exist, and others which cease to exist. It is this

⁺ This paper was presented at the Seventh European Congress of Analytic Philosophy, Milan, Italy and at the Space and Time Seminars of the UAB, Barcelona. I'm grateful to the respective audiences for their useful comments. Many thanks to Dr. Federico Perelda for his insightful comments.

¹ Cf. Tooley, 1996, 16.

relentless beginning and ceasing to exist which explains why different states of affairs obtain at different times, hence why things change. In a nutshell: different states of affairs *obtain* at different times because different states of affairs *exist* at different times, and different states of affairs exist at different times because new states of affairs keep springing into existence while others keep falling off the cliff.

This dynamic explanation of change promises to provide us also with a neat explication of the fact of passage. That a time that is previously not present should be afterwards present, or previously present and afterwards past, without experiencing change, is impossible, many think. What accounts for the continuous becoming present by previously non-present times? It is natural to think that absolute becoming could provide us precisely with the correct explanation for this kind of change. If being present is the same as existing, and being past involves failing to exist, then *becoming* present is the same as *coming* to exist, and becoming past the same as *ceasing* to exist. If things are so and if, as the Presentist thinks, at all times there are states of affairs that come to exist, and others which cease to exist, then it is a matter of conceptual necessity to infer that time really passes, for it is always the case that a time becomes present and immediately past.

The good thing about the idea of grounding qualitative change and passage on absolute becoming is that it appears to free the notion of change from the air of contradiction that beset it since the dawn of philosophy. If I was standing and I am now sitting, I changed from standing to sitting. *Prima facie*, this appears to require that someone who is standing could be sitting. After all, in fact, if it were not true that *I* could be sitting, before I sat down, then I could not be sitting now, for what is impossible cannot become true. But if he who is standing could be sitting, and standing and sitting are incompatible qualities, then the mere possibility that a standing person could be sitting involves a contradiction. A chief requirement of genuine change, in fact, is that it be *the same* persisting entity which bears the incompatible qualities. Presentism appears to avoid this difficulty precisely because absolute becoming involves the disappearance from reality of the trouble-making states of affairs. If I have changed from standing to sitting, then my (past) standing simply does not exist (any more), so it is never true to say that I am sitting and standing.

The obvious drawback of Presentism stems from the observation that, of that which does not exist, one cannot truly say anything, not even that *it* existed. That which does not exist any more, in fact, must simply fail to exist:

the expression “any more” that we append to the proposition that Caesar does not exist, does not qualify a particular way in which Caesar does not exist: there is only one sense in which an entity can fail to exist, and that is by failing to be an entity at all. Thus, if Caesar really went out of existence, there can (now) be no truth that is distinguishably about *him*. Yet we do appear to be able to say true things about the past (or about the future); for example, under any account of what it takes to be past, it must be (now) possible to truly say that Caesar did once exist. Likewise, it must be (now) possible to express the (true) proposition that Obama is taller than Caesar was. But how can that be, if Presentism is true and contingent truths are existence entailing (as many think)² Obama is claimed to be taller than... whom? This difficulty, often called the *truthmaker problem*, constitutes the foundation for the master argument against Presentism.

Standard expositions of the truthmaker problem (on the part of Presentists as well as of their foes), take the form of a *reductio ad absurdum*: (i) Suppose Presentism is true. Then (ii) our world doesn't include past or future objects or events. But (iii) if our world doesn't include past or future objects or events, then there is nothing in the world that could ground propositions about the past or future. Therefore (iv) propositions about the past and future lack truthmakers. But (v) at least some propositions about the past and future are true. Therefore, (vi) Presentism must be false.

Here I shall argue that there is a tension between the need to solve the truthmaker problem and the account of change and temporal passage that I have sketched above.

2. Going out of existence

Consider the following two sentences:

- (1) Caesar did exist.
- (2) *He* doesn't exist any more.

The anaphoric occurrence of ‘he’ in (2) inherits its semantic value entirely from the antecedent occurrence of ‘Caesar’ in (1). The proposition expressed by the conjunction of (1) and (2), in fact, is an instance of the following propositional function:

²Cf. Armstrong, 2004.

(3) x did exist & x does not exist any more

What accounts, in a Presentist world, for the referential dependence ($x=x$) that this discourse anaphora appears to require?

It is common to distinguish two kinds of anaphor-antecedent relations: identity of reference and identity of sense. There is identity of reference if 'he' and 'Caesar' (necessarily) refer to the same entity, while there is identity of sense if the two expressions serve to introduce the same description, or the same individual essence. I shall discuss these two options in turn, arguing that identity of reference is incompatible with the strict ontological diet of the Presentist, while identity of sense is incompatible with the Presentist's account of change.

Let us start by considering the hypothesis that (3) is turned into a true proposition by assigning an individual entity as the value for x . The natural assumption is that the value in question is Caesar himself, the man. After all, it is *he* who is claimed to have gone out of existence. But this cannot be, for then we would have:

(a) (Caesar did exist $\rightarrow \exists x: (x \text{ did exist} \ \& \ x=\text{Caesar}) \rightarrow \exists x: (x = \text{Caesar})$)

And:

(b) Caesar does not exist any more \rightarrow Caesar does not exist $\rightarrow \neg \exists x: (x=\text{Caesar})$

So (3) would be contradictory (hence false).

Surely the Presentist can't deny the validity of (b): if the expression 'Caesar' is intended as referring to the past entity (Caesar, the man), then it surely presently fails to refer, for, we are told, there exists no such thing as a past entity. So the presentist must block the inference (a).

2.1. The identity of reference hypothesis

One way of doing so³ is by denying the validity of the second inference contained in (a), while retaining the validity of the first. This entails conceding that the past existence of Caesar entails the present existence of an entity of which it is true to say that it *used to be* Caesar:

(Barcan formula) Caesar did exist $\rightarrow \exists x: (x \text{ did exist} \ \& \ x=\text{Caesar})$

³ See Bigelow, 1996, for an example of this strategy.

What is being denied is that the consequent of the Barcan formula further entails that the entity that used to be identical with Caesar is now (still) identical with Caesar:

$[The\ y: WAS\ (y=Caesar)] \neq Caesar$

This maneuver effectively blocks the unwanted conclusion that there currently exists an entity that *is* Caesar. But is it compatible with the Presentist account of change?

Let $c \stackrel{\text{def}}{=} [The\ y: WAS\ (y=Caesar)]$. According to the claim under consideration, c is not a past entity (there are no such things); it is a *present* entity that has the backward-looking property of having once been identical with Caesar.

Before evaluating the consequences of this manoeuvre for the Presentist's account of change, let me briefly discuss an obvious objection that comes to mind. Many philosophers, including many Presentists, balk at the idea that matters of identity (and distinctness) can be contingent, hence that two things may be identical at a time and different at another. The objection runs roughly as follows. If Caesar and c were identical (sometime before the year 44BC, say), then they surely *can be* identical, for what is impossible cannot have been the case. But if two things *can be* identical then they *are* identical (by the necessity of distinctness). These philosophers, therefore, will claim that accepting the validity of the Barcan formula *forces one* to also accept the second inference in (a). They will then claim that the manoeuvre under consideration is a non-starter. Presentists who subscribe to the necessity of distinctness (the doctrine that if two things *can be* distinct then, necessarily, they *are* distinct) will therefore opt for blocking the Barcan formula (I shall discuss this option later).

Let us turn back to the consequences of this manoeuvre for the Presentist's theory of change. Assume, for the sake of the argument, that the philosophers mentioned above are wrong, i.e. that it is somehow possible that c was identical with Caesar, but that it *isn't* any more. Does the existence of an entity like c solve the problem of the referential dependence that sentences like (1) and (2) appear to require? I argue that it does not.

If it is (now) true that c used to be identical with Caesar (when Caesar existed), then, it is intuitive to assume that whatever truth that could then be uttered about Caesar should have corresponded to an identical truth uttered

about *c*. Thus, for example, back then it was true to say that Caesar would have later ceased to exist:

(4) $WAS(WILL(\neg\exists x:(x=Caesar)))$

Now, if *c* was really identical with Caesar (back then) then the true proposition expressed by (4) must be paralleled by the following (identical) truth:

(5) $WAS(WILL(\neg\exists y:(y=c)))$

But this cannot be, for (5) clearly entails that there currently is no entity that satisfies the description of *c*: $\neg\exists y:(y=c)$. And this amounts to denying that the Barcan formula applies to temporal operators, contrary to the hypothesis under consideration. The Presentist may be tempted to avoid this conclusion by pointing out that, when *c* was identical to Caesar, it did not satisfy the description [*The y: WAS (y=Caesar)*], which serves the purpose of identifying the relevant entity only now, after Caesar's existence has already become past. After all, when Caesar was alive it wasn't necessarily true to say that there (then) existed an entity which had the property of having being Caesar in the still further past (Caesar, in fact, did not exist before he came into existence). Thus, according to this response, when Caesar was alive, *c* did not exist *qua* satisfier of the description [*The y: WAS (y=Caesar)*]. It existed instead *qua* satisfier of a description of Caesar. This blocks the conclusion that it was then true to say that *c* would later (for example now) have failed to exist, for (5) now only entails that *c* would later fail to exist only *qua* Caesar. This response, however, has the unwanted consequence that it implicitly commits the Presentist to an Eternalist ontology.

If, in order to avoid the conclusion that there is no entity which used to be identical with Caesar, we were forced to specify a sense in which it was then true to say that *c* would later cease to exist, nothing would prevent us from saying that also Caesar has not really ceased to exist in the absolute sense, but only in the sense that he has ceased to exist-*qua*-Caesar. After ceasing to exist-*qua*-Caesar, he continued to exist *qua that entity which used to be Caesar*, i.e. he continued to exist-*qua*-*c*. But this is clearly something that also Eternalists (of the dynamic brand) would be ready to concede! Sure it would account for the referential dependence required by the truth of proposition (3), but it would do so at the expense of conceding that entities *always* persist through change, although different descriptions may be true of them at

different times. This is certainly not compatible with the Presentist's ontological diet.

2.2. The identity of sense hypothesis

Many suggested, partly in response to problems such as the ones mentioned above, that we should regiment talk of past (and future) truths by introducing irreducible temporal operators. What makes past truths true, according to this proposal (I shall call it "Cambridge Presentism", for reasons to be discussed), is how things *were*, and not how they *are*: tenses do not tell what the properties of an individual are, but rather *when* an individual satisfying a certain description has its properties.

Formally, Cambridge Presentism blocks the unwanted conclusion derived in (a) by not allowing the Barcan inference, i.e. by not allowing that Caesar's past existence entails the current existence of an entity that was once Caesar. According to this option, the referential dependence involved in the anaphora (3) is certainly not identity of reference. The thesis is precisely that there cannot be identity of reference in cases of genuine absolute becoming. If there are some entities which existed in 44BC which now simply fail to exist⁴ and we are not allowed to apply the Barcan formula to temporal operators, then now it is utterly impossible to express any singular proposition about those entities. For example, no singular proposition can now be expressed that is distinguishably about Caesar, the man. The proposition that *Obama is taller than Caesar* is only partially singular: it is singular in as much as Obama is concerned, but general as far as Caesar is concerned.⁵

Likewise, there are many propositions that we can now express that could not be expressed when Caesar was alive. Consider for example the (singular) proposition that you express when you say "I exist". Before you existed, say during 44BC, there was no individual entity of which it was true to say that *it* will later come to exist. Therefore no one could then express (however falsely) exactly the same proposition that you now express by saying "I exist".

⁴ For convenience, we have chosen Caesar as an example of an entity which has ceased to exist. This is entirely accidental. If one is worried that Caesar is so complex an entity that his departing from reality cannot have clear (non-vague) truth conditions, or if one believes, on independent grounds, that human beings never really cease to exist, one should just pick up her favourite subjects of absolute becoming (atomic instantaneous states of affairs, for example) and apply the arguments proposed in this paper to them rather than to Caesar.

⁵ Cf. Prior, 1968, 13.

However, it must have been possible, even during 44BC, to claim that an entity which satisfies your description would have later come to be. The latter is a completely *general* proposition, one for which issues of individual identity (or distinctness) of reference simply do not arise.

Suppose that in 44BC a prophet prophesized your current existence. He claimed: “in 2013 there will exist a person with such and such characteristics ... etc”. Now, does it make sense, under these circumstances, to ask whether you are exactly the person who the prophet meant, or whether you are instead a *different* person, who happens to satisfy the same description that the person meant by the prophet *would have* satisfied if, counterfactually, the prophecy would have proved to be fulfilled? No! There is no issue of identity in these cases because, *ex hypothesis*, there is no persisting entity about which the question of identity may be significantly raised. The present truth of your past non-existence, like the past truth of your (then) future existence, and unlike the truth of your current existence, are wholly *general* truths.

It may be objected that, in expressing your past non-existence from the standpoint of the present, you *do have* the resources for expressing a singular proposition to the effect that it was true (during 44BC) to claim that *you* would have come to exist so many thousands of years later. You just have to grab yourself as the direct reference of any proposition expressed about *you*, and then you move back in time, to consider reality as it was during 44BC. By confronting yourself (the entity) with the descriptions of the entities that existed in 44BC, you can then express the *singular* proposition that it was *you* who did not exist then. These observations, however, are inconsistent with Presentism.

First, notice that the proposition that you would express using the semantic resources discussed above would not (and could not) be the same proposition as one which could be expressed in 44BC. The proposition to the effect that you don't exist, as expressed (and expressible) in 44BC, in fact, is a wholly *general* proposition; while the one you would express by dragging the direct reference to yourself to the past scenario would be *singular*. Secondly, and most importantly, notice that the truth of your past non-existence is a *past* truth. Accordingly, the Cambridge Presentist must think that the deployment of the semantic resources under consideration here require that one can drag the reference to a presently existing entity (yourself) into the scope of temporal operators, so as to be able to claim that it is *you* who then failed to exist. But this cannot be, for if direct reference to *you* could be made within the scope of a past operator, then, within the scope of that

same operator it would be true to infer that *you* exist. In short, if you allow for the inverse Barcan formula to apply to temporal operators, then all present entities can be argued to exist at all times, contrary to what the Presentist thinks.

Let us go back and check how Cambridge Presentism scores *vis a vis* the challenge posed by propositions like (3). The past existence of Caesar is now rendered by the general proposition that *there was* an entity that satisfied a given description:

$$(6) \text{ WAS}(\exists x: (x = \textit{Caesar}))$$

The general proposition expressed by (6) has no ontological implications as to what is presently the case (the Barcan scheme does not apply). In particular, it is compatible with the (equally general) proposition that Caesar presently doesn't exist:

$$(7) \sim \exists y: (y = \textit{Caesar})$$

The identity of sense hypothesis, therefore, allows the Cambridge Presentist to express the facts of absolute becoming without incurring in the contradictions which beset the identity of reference hypothesis discussed in section 2.1. Propositions of absolute becoming like that expressed by the sentence "Caesar ceased to exist", according to Cambridge Presentism, are to be analyzed as expressing general comparative facts like:

$$(8) \text{ WAS}(\exists x: (x = \textit{Caesar})) \& \sim \exists y: (y = \textit{Caesar})$$

As I shall argue, however, while it is true that blocking the Barcan formula frees the ontology of the Presentist from the inconsistencies discussed above, it fails to cohere with the Presentist's account of change and passage.

3. Change and absolute becoming

As I said in the introduction to this paper, absolute becoming is the secret ingredient which promises to allow the Presentist to formulate at the same time a *coherent* dynamic account of qualitative change and a coherent account temporal passage. As I shall argue, however, given the restricted ontological resources of Presentism, absolute becoming cannot afford to pay this double bill.

3.1. Being past and becoming past

Although the notion of passage is extremely hard to pin down, and different authors have radically different conceptions of it, I think that any dynamic account of the passage of time ought to comply with the following desiderata:

- 1.1. The fact(s) which make it true that a time (T) *became* past, or that it *ceased* being present (fact α), must be distinct from each of the facts which make it true that T *was* present and *is now* past (fact β).
- 1.2. Fact(s) α must fully explain fact(s) β .

As I have argued elsewhere⁶, these desiderata apply to any conception of time which takes tense seriously, and derive from the observation that being past is a product, or an effect, of a process of becoming past (desideratum ii). The distinctness condition (desideratum i) follows from the necessary distinctness of the explanans from the explanandum (no state of affair can productively explain itself).

The current year is 2013. Things haven't always been so, however, as this year has been present only since last January, when it *became* present. Did something happen at New Year's Eve, in virtue of which 2013 became present? What did this becoming present consist of? Should we think that the fact that 2013 became present last January (fact α) is only the earliest part of the fact that 2013 has been present throughout a certain period of time starting from last January (fact β)? Or should we rather think that α and β are two ontologically independent states of affairs?

There are inferential relations between the truths grounded by α and β that suggest that there must be an intimate relation between them. Whatever feature of reality makes it true that the year 2013 *became* present, in fact, must be such as to guarantee that the year 2013 *has been* present for some time: nothing can *become* present without subsequently *being* present for some time (if only for a brief instant). Conversely, if it is true that 2013 has been the current year for some time, then it must be true that it *became* present at some point: if 2013 had never become present, it surely wouldn't be present now. Thus α and β 'entail' each other, or, as they say, they have the same existence set.

⁶ See my "If it ain't Moving it shall not be Moved", forthcoming in *Topoi*, Special Issue on Time and Time Experience, *Topoi*, An International Review of Philosophy.

If you do not 'take tense seriously', i.e. if you believe that there is no mind-independent, absolute fact of the matter as to what time it is, you will not be impressed by the intimate relation that obtains between facts of type α and β . Events, you will probably concede, are temporally ordered. We were all born *after* the French revolution, for example: (nearly) no one wants to deny this comparative fact. The year of our birth is long gone and we are considerably older than we were then. Yet this would not be a *consequence* of the fact that we have been *growing* older at all intervening moments of time, or of the fact that the flow of time has *reached* later and later stages of our life up to the present one. In spite of our intuitions to the contrary, we would be older without having ever been growing older. If you think that reality is 'static' in this sense, you will believe that the becoming present of 2013 last January points to nothing over and above the fact that 2013 has been the current year since then (relative to our temporal perspective). Tense predicates such as 'being present', according to this view, in fact, do not cut nature at its joints, and do not point at irreducibly dynamic features of reality. You may then either opt for an 'error theory' of temporal flow, according to which it is simply false that the year 2013 *became* present last January; or you may prefer a revisionary stance, and claim that the fact that 2013 became present (fact α) is simply the earliest part of the fact that makes it true that this year has been present since last January (fact β). Either way, the static view explains perfectly well the relation between α and β . Under the error variant, fact α simply does not exist, so the problem does not arise; under the revisionary variant, fact α is a mereological part of fact β , which guaranties that they always exist or fail to exist together.

If, instead, you believe that tense properties and temporal passage are objective features of reality, like Presentists typically do, you do not have it that easy. Dynamicists, in fact, believe that events would not be temporally ordered (or things have different properties at different times) if time didn't really flow, viz. if what time it is were not a mind-independent, absolute and *changing* matter of fact. Caesar did exist, but he no longer does. If the Dynamicist is to uphold the explanatory role that passage is supposed to play in her account of change, she should not be (merely) in the business of grounding this comparative truth; she should be further asking for the metaphysically independent grounds of Caesar's *disappearance*, i.e. for the grounds of his *ceasing* to be.

If Caesar's disappearance – whatever this fact consists of – is to explain – in an informative non-tautological way – the comparative truth that he did exist

while he doesn't now (or that he existed only up to the year 44 BC), then such disappearance cannot be grounded on the same facts that ground the comparative truth itself. By the same token, the Dynamicist cannot claim, like the Staticist does, that fact α does not exist, or that it is a part of fact β . Nothing, she should think, can be changed if it is never found in a *state of changing* before the comparative change has been produced. In a motto: if reality is not changing it shall not be changed.

The following are but a few examples of how (i) and (ii) are implicitly assumed by authors with radically different views about time. McTaggart,⁷ for example, argues that if there were to be an end of time, then the last moment, once become present, will remain present indefinitely, for it will never *become* past. On a similar vein, Sanson claims that “one could argue that the first moment of time, if there were one, *was* present but could never *become* present, because it was never future”, and that “if there were only a single moment of time, and if it were present, it would be permanently present”.⁸ These observations appear to presuppose that (if time really passes) a time can *be* past only if it has *become* past first. William James considers the hypothesis that “God, as the orthodox believe, created the space-continuum, with its infinite parts already standing in it, by an instantaneous *fiat*. Past time now stands in infinite perspective, and may conceivably have been created so, as Kant imagined, for our retrospection only, and all at once”.⁹ Clearly the mere metaphysical possibility of James' admittedly paradoxical scenario presupposes that to *be* past is not the same as to *become* past.

The case for desiderata (i) and (ii) can be strengthened by noting that the proposition that a time T *is* past (or that it *was* present) represents a state of affairs that is *post factum*, in that it describes how things are strictly *after* T has already become past. The proposition that T *ceased* being present, on the contrary, represents a state of affairs that is *in medias res*, in that it must have occurred while T was still present: only present times can *cease* being present. In short, the event through which T loses its presentness must occur while T is still present, while the fact that T *is* past can only obtain *after* T's presentness has *already* been lost.

As I shall argue, the Presentist account of passage fails to satisfy these desiderata.

⁷ McTaggart, 1908.

⁸ Sanson 2011, 12, my emphasis.

⁹ James, 1987, 1066.

3.2. Absolute becoming and the passage of time

In order to make her account compatible with her strict ontological diet, the Presentist was forced to claim that truths of absolute becoming (such as *beginning* to exist, *becoming* present, or *ceasing* to exist) are made true by wholly *general* comparative facts: facts that do not involve reference to any *particular* individual. So, for example, my particular becoming non-existent in the future (apparently a singular fact about myself) is conflated with the posthumous general fact that my description will be non-satisfied; while, as prescribed by desiderata (i) and (ii), it should precede it and explain it.

While there may be a general truth about my description being non-satisfied in the future, without *me* being there to substantiate this truth (hence without a correspondent grounding truth that involves *me* going out of existence), there can be no truth about *my* description *becoming* non-satisfied, without *me* being there to become non-existent. Surely, in fact, a description that is (already) not satisfied cannot *become* non-satisfied: it is already! Moreover, my description not being satisfied is not necessarily a fact about me. If it is a singular fact at all, it is a fact about my *description* (or individual essence) and about the world. But *my* going out of existence *must* be a fact about *me*, if any fact is! It must be a fact about *me* because it must serve as a (posthumous) explanation for the general fact that, in the future, *my* description will be unsatisfied (desideratum ii).

We get the illusion of trading with propositions that express genuine change because we are presented with (general) propositions which, upon reflection, change their truth-values depending on what time it is, from which it surely follows that there has been genuine change. The proposition that $\exists x: (x = \textit{Caesar})$, for example, understood as pointing at a general fact, is clearly one that can be expressed at all times, and which may have had different truth-values at different times. If it did (and it did), then surely reality must have genuinely changed! This is why many have been tempted to think that absolute change could be the key to understanding qualitative change and passage. But this cannot be the whole story about the dynamic aspect of reality for, if it were, then *qualitative* change and passage would be ultimately grounded on the fact that some *Cambridge* change has taken place.

Cambridge changes are not *genuine* changes, at least in the sense that they cannot occur without being paralleled and grounded by underlying *intrinsic* changes. True, if my daughter gives birth to a child, I thereby *change* from not being a grandfather to being a grandfather. Likewise, if I will cease to

exist, then my description (or individual essence) will change from being satisfied to being non-satisfied. But such extrinsic change will have to be paralleled and substantiated by an intrinsic change which only *I* can undergo, just like my changing from not being a grandfather to being a grandfather must be paralleled and substantiated by some intrinsic change in my daughter. In short, it must be because *I* will become non-existent that *my* description will be non-satisfied. Likewise, it must be *because* Caesar, the man, *ceased* to exist, that it later became true that he did exist while he doesn't now. If Caesar's disappearance is to *explain* the comparative truth that he did exist while he doesn't now, then that disappearance, I argue, cannot be grounded on the same facts that ground the comparative truth itself.¹⁰

According to Cambridge Presentism, the fact that Caesar ceased to exist is not a fact that is distinguishably about Caesar.¹¹ The fact that Caesar went out of existence, we have seen, is analyzed as the conjunction of two wholly general facts: (1) the fact that there was a *person* called 'Caesar' and about whom it *was* true to say... etc., and (2) the fact that there is now no one called 'Caesar' and about whom it is true to say that etc. Now, fact (1) is compatible with Caesar being still existent, hence it contributes not at all to the truth of his departure from reality. Fact (2), instead, expresses the absolute non-existence of an entity which satisfies the description of Caesar, hence it expresses a circumstance that can only obtain *after* the explanatorily relevant change has already occurred. Now, while it is correct to infer the truth of Caesar's departure from reality from the conjunction of his past existence and current non-existence, still the very fact of his departure has eluded the expressive capacity of Cambridge Presentism.

¹⁰ Notice that the relevant notion of explanation here is not *causal* explanation: we are not in the business of revealing the causes of Caesar's death. We are in the business of explicating the metaphysical grounds of Caesar's disappearance from reality, a fact that ought to explain, in its turn, the former existence and current non existence of Caesar.

¹¹ Notice that not even the proposition that this is not a fact about Caesar can be taken to express a truth about Caesar: it must be paraphrased as expressing the proposition that there is no person of whom it is true to claim that it was called 'Caesar' and that it has now gone out of existence (cf. Prior, 1968, 14).

4. Conclusions

I have argued that, in order to uphold her ontological tenets in a coherent way, the Presentist is forced to analyze facts of absolute becoming as conjunctions of fully general facts which compare how reality was *before* changing with how it is *after* it has already changed. Unfortunately, so construed, absolute becoming cannot ground qualitative change and temporal passage, as it is often advertised to do. Under the proposed explanation, yesterday became past because some states of affairs which existed yesterday went out of existence, and others came to be. But this explanation, once it is unpacked in accordance with the need to solve the truthmaker problem (as discussed in sections 2.1. and 2.2.), amounts to the claim that yesterday *became* past because it *is* past, which is no explanation at all. Similar considerations apply to the proposed explication of qualitative change in terms of absolute becoming.

It is tempting to think that this result was only to be expected. As noted by many authors (e.g. Thomas and Prior)¹², absolute becoming, lest it could be understood in terms of a persisting subject, is not a radical kind of change (from non-being to being, or *viceversa*): it is not a kind of change at all. It would be very surprising if we could analyze genuine change and passage in terms of the obtaining of facts which do not entail that change is happening. The A-theoretic fact that I am sitting now, or that I was standing, by themselves, are just as static as the B-theoretic fact that I am (tenselessly, hence changelessly) standing at 17:00 PM on 26/06/2013 is. To claim that I *was* standing, in fact, is not the same as claiming that my standing was *becoming* past (which would indeed express the fact that non comparative change was taking place); the fact that I was standing simply conveys the thought that my standing *is* past, which is a state of affairs that can obtain only *after* the relevant change (viz. becoming past) has already happened.

Any realist account of the passage of time must explain in virtue of what each present moment *transits* into the next, and in virtue of what the domain of objects which existed in the past has *shifted* into that of the objects which presently exist. The difference between the domain of entities in which our variables presently range and that in which they used to range in the past, is only a posthumous unchanging product of the relevant shifts of which the passage of time consists.

¹² Cf. Prior, 1968, 73.

I have argued that the price for restricting the ontology only to presently existing things is that absolute becoming, contrary to what is often claimed, cannot ground the truth passage.

References

Armstrong, D. M., 2004, *Truth and Truthmakers*. Cambridge: Cambridge University Press.

Bigelow, J., 1996, Presentism and Properties, in J.E. Tomberlin (ed.), *Metaphysics*, Vol. 10 of Philosophical Perspectives (pp. 35–52), Cambridge, Mass.: Blackwell.

Boccardi, E., If it ain't Moving it shall not be Moved, forthcoming in Special Issue on Time and Time Experience, *Topoi, An International Review of Philosophy*.

Crisp, T., 2007, Presentism and The Grounding Objection, *Nous*, 41:1, 90–109.

Prior, A. N., 1970. *The notion of the present*. *Studium Generale*, 23:245-248.

--- 1962, Changes in Events and Changes in Things, Reprinted in Prior 1968, pp. 1-14.

--- 1968, *Papers on Time and Tense*, Oxford: Clarendon Press.

Sanson, D., 2011, Once Present, Now Past, retrieved October 2011, <http://files.davidsanson.com/research/oncepresent.pdf>

Tooley, M., 1997, *Time, Tense, and Causation*. Oxford University Press.

Dossier

Poincaré. Filosofia, Ciência e Materialismo



Nota de Abertura

María de Paz
(CFCUL & UCM)
maria.depaz@hotmail.com

O objectivo do presente Dossier é salientar algumas particularidades da obra de Henri Poincaré (1854-1912) que têm sido objecto de investigação no CFCUL nos últimos tempos. A ideia de fazer um *Dossier* sobre Poincaré surge em dois momentos. O primeiro é o centenário do seu desaparecimento, comemorado o ano passado através de numerosas iniciativas internacionais (conferências, publicações...). Assim, são aqui apresentados alguns trabalhos desenvolvidos e discutidos em conferências e seminários durante 2012. O segundo momento é a tradução de André Philot do texto *Les Conceptions nouvelles de la matière*, que serve de inspiração ao título do *Dossier: Filosofia, Ciência e Materialismo*.

O *Dossier* é composto por três artigos, pela tradução do texto mencionado e por uma recensão crítica da primeira biografia científica de Henri Poincaré, escrita pelo historiador da matemática Jeremy Gray. No primeiro artigo, *A transversalidade do conhecimento na obra de Poincaré*, Isabel Serra explica como a investigação sobre as equações diferenciais se torna fundamental na obra de Poincaré, dado que é precisamente a partir desse trabalho que se podem estabelecer ligações entre todos os domínios científicos trabalhados pelo génio francês. Apesar de a importância das equações diferenciais na obra de Poincaré já ter sido abordada, por exemplo, por Jeremy Gray que a investigou na sua relação com a noção de grupo, o artigo de Isabel Serra vai mais longe porque oferece uma explicação do ‘universalismo’ de Poincaré. A partir desta explicação, compreendemos as conexões entre as diferentes disciplinas nas quais este autor realizou importantes descobertas (desde a geometria até à física). Além disso, relaciona o trabalho científico de Poincaré

com a sua filosofia, de tal forma que evidencia o cruzamento entre ciência e filosofia fazendo-nos perguntar se a sua ciência teria sido a mesma sem a sua concepção filosófica.

O segundo artigo, *O convencionalismo de Poincaré contextualizado: origem e significado*, é ao mesmo tempo o esclarecimento de um dos pontos mais controversos do pensamento de Poincaré e uma proposta de compreensão desse ponto no âmbito da filosofia e dos problemas da ciência do último quartel do século XIX. O artigo considera ainda o problema da distinção entre o convencionalismo aplicado à geometria e o convencionalismo aplicado às ciências da natureza, nomeadamente, a física e a mecânica. Em suma, procura clarificar qual a compreensão do conhecimento científico deste autor, perante as ameaças feitas à ciência por outras posições de então.

O terceiro, *Henri Poincaré, ciência e materialismo: o papel das hipóteses na oscilação entre atomismo e continuísmo*, é um estudo sobre o artigo *Les Conceptions nouvelles de la matière* de Poincaré. Neste texto, André Philot e Augusto Videira oferecem-nos uma visão geral da obra científica e filosófica de Poincaré, os seus principais contributos e as controvérsias nas que esteve envolvido. A descrição do contexto histórico do artigo de Poincaré é extremamente útil na medida em que nos ajuda a compreender o clima político e intelectual da França dos finais do século XIX e primeiros anos do século XX. Por último, o texto discute o significado da hipótese materialista e compara-o com outras posições adoptadas por este autor acerca da ciência, tais como o determinismo.

O texto *As novas concepções da matéria* é traduzido para português pela primeira vez. De facto, é a segunda tradução em língua estrangeira desse texto, sendo a primeira em inglês publicada em 2012, o que de algum modo atesta a sua importância. Este texto é a transcrição da conferência “Le Materialisme à l’heure actuelle em France” pronunciada a 7 de Março de 1912, no ciclo de conferências “Foi et Vie”. Poincaré começa por perguntar se a ciência nos conduz ao materialismo; questão muito importante na altura, mais para a filosofia do que para a ciência propriamente dita. O autor está consciente desta importância, de tal modo que, logo no início do texto, afirma que não é à ciência que cabe dar tal resposta. Em seguida, considera as últimas descobertas científicas do seu tempo, tais como as experiências de Perrin sobre os átomos, e expõe a controvérsia entre materialismo atomista e continuísmo. O artigo suscitou um novo debate que procura averiguar se Poincaré teria ou não mudado de posição no que diz respeito ao lugar dos

átomos na sua compreensão da ciência. Esse debate não está ainda concluído e deixamos ao leitor a possibilidade de chegar às suas próprias conclusões.

Por último, a recensão crítica da obra *Henri Poincaré. A Scientific Biography* tem por objectivo apresentar uma das publicações mais recentes sobre esse autor. Embora a obra não esteja traduzida em português, pensamos ser importante incluir aqui uma referência ao primeiro livro que realiza uma análise de conjunto do trabalho científico de Poincaré numa perspectiva biográfica.

A ideia de fazer este *Dossier* prende-se também com um já considerável histórico de investigação em torno da filosofia e ciência de Henri Poincaré realizado no CFCUL. Não podemos deixar de referir dois projectos de investigação. O primeiro, *Henri Poincaré. Filósofo da Ciência* (PTDC/FIL/64748/2006), da iniciativa da Professora Olga Pombo e coordenado pelo Professor Franco de Oliveira. O segundo, *A física dos princípios e a epistemologia científica: os casos de Mach, Hertz e Poincaré* (PTDC/FIL-FCI/116939/2010), da iniciativa de parte da equipa do projecto anterior, dedicado à epistemologia das ciências naturais, em colaboração com o Professor Augusto Videira e coordenado pela Professora Isabel Serra. Os membros, colaboradores e consultores destes dois projectos contribuíram para a formação do meio intelectual adequado para o desenvolvimento dos trabalhos que aqui são apresentados. Para todos eles, o nosso mais profundo agradecimento.

A transversalidade do conhecimento na obra de Poincaré

Isabel Serra
(CFCUL)
isabelserra@netcabo.pt

1. Introdução: o Universalismo de Poincaré

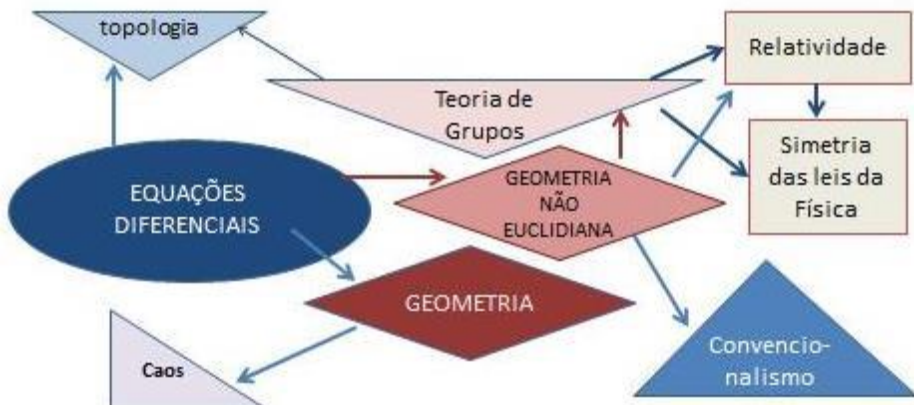
Henri Poincaré iniciou-se na investigação científica com o estudo das equações diferenciais, uma questão central na matemática da época, embora de âmbito restrito no conjunto das numerosas áreas da disciplina. Contudo, a partir desse tema bem específico o matemático francês foi capaz de estabelecer uma rede de interações e cruzamentos de saberes com ramificações em domínios científicos muito diversos. De facto, os seus resultados no estudo de equações diferenciais permitiram transpor barreiras entre áreas tão diferentes como as funções, a teoria de grupos, as geometrias não euclidianas e a topologia. Por outro lado, parecendo relacionar apenas vários domínios matemáticos entre si, essa rede de ligações acabaria por implicar também a física e a filosofia, aproximando Poincaré dessas disciplinas, muito antes de ele as ter estudado e trabalhado¹.

A arte de estabelecer ligações e cruzamentos entre saberes, que se descortina logo nos primeiros trabalhos de Poincaré, torna-se evidente quando se analisa o conjunto da sua obra. Sendo essencial para caracterizar o seu “universalismo”, o relacionamento entre saberes não só marcou todo o trajeto pessoal do cientista como alterou também radicalmente a evolução de alguns ramos da matemática. É difícil, se não impossível, explicar tal

¹ A sua primeira obra filosófica, *La Science et l'Hypothèse*, é publicada em 1902. Os seus inícios na física são em geral associados ao ensino da física matemática que ocorre em 1886.

capacidade de estabelecer ligações, a não ser talvez usando uma justificação do próprio Poincaré em *La Valeur de la Science* que, depois de formular a pergunta, “Quem nos ensinou a conhecer as verdadeiras e profundas analogias, as que os olhos não vêem mas que a razão adivinha?”, responde: “foi o espírito matemático, que ignora a matéria para estudar apenas a forma pura”.²

Foi então o seu “espírito matemático”, a sua ligação à “forma pura”, que lhe permitiu intuir relações que os outros não viam? Não existe uma resposta fácil e definitiva para esta questão. No entanto, compreender o universalismo de Poincaré talvez seja, não tanto encontrar as suas justificações, mas antes analisar a história das ligações interdisciplinares estabelecidas. Percorrendo a rede de interações e cruzamentos de saberes por ele criados, e seguindo os caminhos traçados para desenvolver relações e analogias, talvez seja possível encontrar pistas para entender esse universalismo. É o que se procura fazer nos próximos parágrafos, explorando a teia de interações onde se situa o seu trabalho, sem no entanto aprofundar as matérias abordadas, o que aliás foi já feito por inúmeros autores. Os diferentes temas científicos ou filosóficos serão referidos apenas na perspectiva das suas ligações mútuas. Essas ligações podem ser visualizadas no esquema seguinte, onde se incluíram os principais temas de investigação e também os cruzamentos estabelecidos entre eles. O gráfico pretende dar uma perspectiva panorâmica sobre a rede de interações criada pelo trabalho de Poincaré, difícil de traduzir apenas através de um texto.



² Poincaré, 1905, 106.

2. As Equações Diferenciais e a Noção de Grupo

Nos finais dos anos 1870 Poincaré inicia o estudo das equações diferenciais lineares, sob a orientação de Charles Hermite (1822-1901) tendo em vista o doutoramento em matemática, um trabalho que, de acordo com Jeremy Gray, transformou completamente os conhecimentos sobre as equações diferenciais.³ Um dos resultados relevantes foi a definição das funções automorfas⁴ ou “fuchsianas”⁵, como Poincaré lhes chamou. Mas, na perspectiva aqui proposta – a do estabelecimento de ligações entre domínios matemáticos – o grande momento desse trabalho foi a descoberta da relação entre equações diferenciais e geometrias não euclidianas. Alguns anos mais tarde, ao escrever sobre o processo de invenção matemática, Poincaré dá relevo a esse acontecimento, ocorrido durante uma viagem:

No momento em que pousei o pé no estribo, surgiu-me uma ideia, sem qualquer ligação com os pensamentos anteriores: as transformações que eu tinha usado para definir as funções fuchsianas eram idênticas às da geometria não euclidiana.⁶

Esta descoberta é relatada por Jeremy Gray, com grande pormenor do ponto de vista da matemática e da sua história.⁷ Antes de referir as suas consequências vale a pena determo-nos numa afirmação de Gray:

As publicações de Poincaré denunciam a grande clareza do seu pensamento, associada a uma quase dramática ignorância da matemática contemporânea.⁸

Esta curiosa observação perfeitamente fundamentada ao longo de dezenas de páginas, e onde Gray refere a ignorância “dramática” de Poincaré acerca da matemática do seu tempo, talvez seja interessante para caracterizar a invenção matemática. A afirmação permite levantar a questão do valor epistemológico, na produção científica de um autor, dos conhecimentos anteriormente acumulados por outros. Esse valor é em geral considerado positivamente. Mas não existirão casos em que uma certa “ignorância” tem um papel libertador dos preconceitos e se torna geradora de

³ Gray, 1981, 281.

⁴ Poincaré, 1916-1956, 336-372.

⁵ Poincaré partiu de alguns resultados do matemático alemão Lazarus Fuchs (1833-1902) com quem manteve uma correspondência publicada em *Obras Completas*, Vol. 11, pp 13-25.

⁶ Poincaré, 1908, 361.

⁷ Gray, 1981, 303-332.

⁸ Gray, 1981, 298.

novas ideias? Assim parece ter acontecido no caso presente, o da relação estabelecida por Poincaré entre funções automorfas e geometrias não euclidianas.

Inventadas por János Bolyai (1802-1860) e Nikolai Lobatchevsky (1792-1856) em meados do século XIX, as geometrias não euclidianas desencadearam o aparecimento de outros trabalhos nesse domínio. Bernhard Riemann (1826-1866) e Hermann von Helmholtz (1821-1894) são os primeiros a usar essas geometrias⁹, mas seriam as contribuições de Félix Klein (1849-1925) e de Sophus Lie (1842-1899) a abrir novas perspectivas na área, ao considerarem que a geometria se reduz ao estudo de um “grupo de transformação”.¹⁰ Foi o “grupo de transformação” que permitiu a Poincaré fazer a associação entre as funções fuchsianas e a geometria não euclidiana, dois temas da matemática com estatutos muito diferentes na disciplina. O primeiro inscrevia-se no estudo das equações diferenciais, um ramo clássico da Análise com um enorme desenvolvimento e inúmeras aplicações, já muito trabalhado por matemáticos de prestígio. As geometrias não euclidianas, pelo contrário, estavam nessa época a fazer o difícil caminho que as novas áreas têm muitas vezes de percorrer até conquistarem direito de cidadania no mundo da ciência. Até aí consideradas como uma espécie de curiosidade lógica, as geometrias não euclidianas adquiriram um novo estatuto através da relação, estabelecida por Poincaré, com as funções automorfas e, doravante, “tornaram-se um terreno fértil para os novos métodos da teoria de grupos”.¹¹ No seguimento das publicações de Poincaré surgiu um grande interesse pelas geometrias não euclidianas, que passaram a ser ensinadas nas universidades¹²; o que contribuiu de forma decisiva para o seu desenvolvimento. Este sucesso foi essencial, também no reconhecimento de Poincaré pela comunidade científica.¹³

Embora não se saiba qual a origem das suas concepções geométricas, Poincaré já havia adoptado, no início da sua carreira, o ponto de vista de Félix Klein, moderno e pouco conhecido, o de que a geometria não é mais do que o estudo de um grupo.¹⁴ A ideia de grupo, que foi essencial para fazer a ponte entre os dois domínios matemáticos – as funções e as geometrias não

⁹ Torreti, 1984, 154.

¹⁰ Torreti, 1984, 137-142, 171-188.

¹¹ Gray, 1984, 1.

¹² Walter, 1996, 95-96.

¹³ Gray, 2012, 179.

¹⁴ Gray, & Walter, 1997, 15-16.

euclidianas – e que lhe permitiu progredir rapidamente¹⁵ no seu estudo das soluções das equações diferenciais, tornar-se-ia crucial em futuros trabalhos. No entanto, ele nunca investigou em teoria de grupos, limitando-se a desenvolver o estritamente necessário para obter os seus resultados: “Era a ideia de grupo, e não propriamente a teoria de grupos, que o interessava”.¹⁶

Esta outra observação de Jeremy Gray permite levantar mais uma questão acerca da maneira poincareana de produzir conhecimento em matemática. Não sendo um especialista em geometria não euclidiana e apesar de ignorar – manifestamente – muito do trabalho da escola alemã¹⁷, Poincaré usa uma concepção geométrica arrojada, aplicando-a à Análise Matemática. O que ninguém havia ainda feito. Dado que posteriormente usou com sucesso a noção de grupo também noutros estudos, poder-se-á considerar que ele tinha intuição, talvez mais do que os especialistas da área, acerca da importância e das possibilidades de utilização de tal objecto matemático. De facto, essa importância confirmou-se nas décadas seguintes. O impacto da noção de grupo e as suas aplicações foram crescendo, tanto na matemática como nas ciências da natureza, acabando por originar uma nova e pujante área científica, a álgebra abstrata.

3. Das equações Diferenciais à Topologia, à Ideia de “Caos” e ao Convencionalismo

O estudo das equações diferenciais atravessa toda a obra Poincaré, estando na origem de ligações e influências estabelecidas em diferentes áreas. A ideia de grupo, uma das pontes utilizadas nessas ligações, e que tão frutuosa havia sido no estabelecimento da relação com as geometrias não euclidianas, estará presente também nas suas interrogações topológicas.¹⁸ Da aplicação da noção de grupo aos espaços topológicos resultou a conjectura de Poincaré, enunciada nos inícios do século XX, e demonstrada por Grigori Perelman (1966-) só cerca de cem anos depois. Para além de ter feito o primeiro estudo sistemático da topologia, Poincaré revolucionou a investigação na área com a introdução de novos instrumentos

¹⁵ Gray, & Walter, 1997, 15.

¹⁶ Gray, 2012, 178.

¹⁷ Gray, 1981, 298.

¹⁸ Nabonnand, 2000.

matemáticos. Os seus resultados estão publicados em *Analysis situs*¹⁹, o nome que então se dava à topologia.

Do estudo qualitativo das equações diferenciais surge ainda, nos anos 1880, outra descoberta, não menos importante que a da relação com as geometrias não euclidianas, embora não tão valorizada na altura. Poincaré é o primeiro a identificar o que hoje se designa por “caos”, um fenómeno cuja relevância é reconhecida pela comunidade científica só a partir dos anos 1960, após o trabalho do meteorologista Edward Lorenz (1917-2008) e da generalização do estudo dos fenómenos caóticos.

Tal como acontece com a ideia de grupo, também neste caso Poincaré utilizou uma perspectiva já existente mas pouco usual, ao escolher uma abordagem geométrica para estudar o comportamento das soluções das equações diferenciais. Ele inspirou-se no trabalho de outros matemáticos, mas desenvolveu-o de forma inovadora tal como fez também com as geometrias não euclidianas. De facto, já Augustin Cauchy (1789-1857), Charles Briot (1817-1882), Jean-Claude Bouquet (1819-1885) e Lazarus Fuchs (1833-1902), haviam estudado o comportamento geométrico das soluções das equações diferenciais, mas sempre localmente, enquanto Poincaré vai analisar o seu comportamento global. É essa análise global que lhe permite detectar o fenómeno de “sensibilidade às condições iniciais”, ou “caos” como hoje se chama. A primeira definição de caos, e que conserva toda a sua actualidade, é dada por Poincaré:

Pequenas diferenças nas condições iniciais podem originar grandes diferenças nos fenómenos finais. Um pequeno erro inicial poderia ter então um efeito enorme. Nestas condições, a previsão torna-se impossível.²⁰

Ele próprio se apercebe das potencialidades do seu inovador processo de análise, referindo nos seus artigos o “vasto campo que ele abre aos géometras”.²¹ A mecânica celeste foi uma das áreas onde o aplicou, estudando o “problema dos três corpos”. Os seus resultados, publicados no *Journal des Mathématiques Pures et Appliquées*²² dariam a Poincaré, em 1887, o prémio do Rei Oscar da Suécia.

Tal como no caso das funções fuchsianas, também neste problema Poincaré usou a geometria, mais uma vez num terreno essencialmente ligado

¹⁹ Poincaré, 1895.

²⁰ Poincaré, 1908a, 62.

²¹ Poincaré, 1881, 377.

²² Poincaré, 1881, 375-422; 1882, 251-286; 1885, 167-244; 1886, 151-217.

à Análise Matemática, pelo menos nessa época. A geometria continuaria, nos anos seguintes, a desempenhar um papel importante na sua actividade e no seu pensamento. Aliás, apesar de ter trabalhado também com objectos que faziam parte de outros ramos da matemática, Poincaré considerava-se um géometra: e como tal “deixava-se guiar pela intuição”, conforme escreveu em *L'intuition et la logique en mathématiques*.²³ É certo que palavra intuição, embora pareça adequada para caracterizar a inventividade de Poincaré, tem múltiplos significados²⁴ e presta-se a interpretações diversas. No entanto, é difícil encontrar outro termo para traduzir a sua capacidade de utilização da geometria.

A primeira ligação de Poincaré à filosofia também será estabelecida através da geometria. É aliás pelo “convencionalismo geométrico” que Poincaré é mais conhecido na filosofia das ciências.²⁵ A ligação prolongar-se-ia até à filosofia da física, com o “convencionalismo físico”. Estes dois aspectos da sua filosofia das ciências foram apresentados em vários artigos, reunidos posteriormente em *La Science et l'Hypothèse*. As geometrias não euclidianas podem ser consideradas como o elo de ligação entre as equações diferenciais e o convencionalismo geométrico. De facto, Poincaré fundamentou essa posição filosófica na compatibilidade das várias geometrias. Depois de analisar as geometrias não euclidianas²⁶ conclui que não existe uma geometria melhor que outra, sendo uma questão de convenção a escolha de uma, entre as várias geometrias possíveis.²⁷

4. As Geometrias não Euclidianas e o Espaço Físico

Na física do século XIX a geometria é naturalmente interpretada como a ciência do espaço, sendo o espaço concebido como uma entidade real. Mas, paradoxalmente, as proposições da geometria não podem ser nem corroboradas nem refutadas pelos acontecimentos reais que têm lugar nesse espaço real. De facto, de Platão a Kant, e também para os empiristas, a geometria descreve a realidade, embora resulte de conhecimentos *a priori* e independentes da experiência. A emergência das geometrias não euclidianas veio avivar as questões filosóficas em torno do significado da geometria, e

²³ Poincaré, 1905, 27-40.

²⁴ Pombo, 2012.

²⁵ Gray, 2013, 525.

²⁶ Poincaré, 1902, 63-76.

²⁷ Poincaré, 1902, 75.

também estimular a controvérsia acerca da natureza do espaço físico. Uma das questões levantadas foi a da compatibilidade das várias geometrias com as medidas físicas. Mesmo supondo que, do ponto de vista matemático, a existência de várias geometrias não é problemática, qual é a “verdadeira” geometria das medidas físicas?

Neste debate, são fundamentais as intervenções de Helmholtz, que pensava ser possível estabelecer os fundamentos da geometria a partir da mecânica. Na sua conferência de 1870, « Sobre a origem e o significado dos axiomas da geometria », ele apresentou o conjunto dos trabalhos de Carl Friedrich Gauss (1777-1855), Bolyai, de Lobachevsky, de Riemann – e também o seu próprio – como constituindo o fundamento científico de uma filosofia empirista da geometria, oposta ao apriorismo kantiano.²⁸ As suas teses, que exerceram profunda influência em Poincaré, contêm também as primeiras afirmações de carácter convencionalista na geometria.²⁹ Para as apoiar, Helmholtz dá exemplos de visualização de situações não euclidianas, pondo assim em evidência que os axiomas da geometria não são dados *a priori*.

De acordo com Rougier, os trabalhos de Felix Klein e de Sophus Lie, ao colocarem em evidência que os movimentos dos sólidos indeformáveis definem um grupo de transformações, contribuíram para a fundamentação da teoria de Helmholtz:

São os trabalhos de Helmholtz suscitados pelos resultados de Riemann que estabelecem a relação lógica entre estes últimos e a teoria dos grupos de transformação.³⁰

Em 1872, no “Erlangen Programme”, Felix Klein define, unifica e classifica as diferentes geometrias usando, precisamente, o conceito de grupo de transformação.³¹ Klein e Lie mostram que a geometria, euclidiana ou não, se reduz ao estudo de um grupo e que as geometrias se equivalem entre si.

A equivalência das geometrias que resulta dos trabalhos de Riemann, Helmholtz, Klein e Lie, foram determinantes nas reflexões de Poincaré publicadas em 1887³², em 1891³³, e mais tarde em *La Science et*

²⁸ Torreti, 1984, 163.

²⁹ Torreti, 1984, 163.

³⁰ Rougier, 1920, 56.

³¹ Klein, 1891, 87.

³² Poincaré, 1887, 203-216.

³³ Poincaré, 1891, 769-774.

l'Hypothèse. Fundamentadas também na utilização que ele próprio fez das geometrias não euclidianas, as concepções elaboradas por Poincaré conduziram-no ao “convencionalismo geométrico”, doutrina que tantas questões e interpretações suscitou até aos dias de hoje.³⁴ Os vários estudiosos da obra de Poincaré parecem não estar de acordo, nem sequer acerca da origem do convencionalismo de Poincaré; para Scott Walter, tais origens encontram-se provavelmente nos debates em torno da coerência lógica e significado físico da geometria não euclidiana durante a década de 1870.³⁵ Para Rougier, contudo, há uma relação forte entre a elaboração do convencionalismo e a investigação em funções fuchsianas.³⁶

Foi referido o interesse que os matemáticos demonstraram pela descoberta de Poincaré da relação entre funções fuchsianas³⁷ e geometrias não euclidianas. Já o mesmo não se passou com a sua teoria convencionalista, que não agradava aos géometras nem, naturalmente, aos físicos. De facto, o convencionalismo geométrico tornava inútil a discussão sobre a natureza do espaço físico e condenava a possibilidade de se estabelecer empiricamente a estrutura geométrica do espaço.³⁸

Então, a geometria não euclidiana, que se tinha revelado uma das ligações fortes entre áreas de conhecimento, parece agora desempenhar o papel oposto, que é o de afastar Poincaré de um problema importante, o da natureza do espaço. No entanto, o convencionalismo geométrico, responsável pelo divórcio entre Poincaré e essa questão fundamental, vai permitir-lhe outra ligação com a física, ao abrir caminho para o convencionalismo físico. Apesar da controvérsia que se gerou em torno do convencionalismo de Poincaré, numerosos autores são unânimes ao considerar que o convencionalismo físico, contendo embora características específicas³⁹ foi inspirado no convencionalismo geométrico.⁴⁰

As ligações das geometrias não euclidianas com a física não se limitam, contudo, aos aspectos filosóficos e ao convencionalismo. De facto, a noção de grupo vai ser essencial também no estudo de um problema de física, o das equações de Maxwell, que conduziria Poincaré a novos resultados.

³⁴ Giedymin, 1992, 423-443.

³⁵ Walter, 1996, 89.

³⁶ Rougier, 1920, 147.

³⁷ Walter, 1996, 95.

³⁸ Walter, 1996, 90.

³⁹ Neste mesmo número de *Kairos*, o artigo de Maria de Paz, “O convencionalismo de Poincaré contextualizado: origem e significado”, discute a questão com pormenor.

⁴⁰ Giedymin, 1977, 271.

Os exemplos citados permitem afirmar que, embora não tivesse propriamente feito investigação em geometrias não euclidianas e teoria de grupos, esses temas constituem, na obra de Poincaré, o ponto de ligação entre diversos trabalhos onde se cruzam a matemática, a física e a filosofia.

5. A Noção de Grupo e a Física

Nos finais do século XIX o estudo do electromagnetismo conduziu a aplicações de grande importância científica e social. A par desse sucesso, e na falta de uma teoria unificada, a interpretação física dos fenómenos levantava inúmeros problemas e contradições. Poincaré apercebeu-se dessa grande crise conceptual e enfrentou-a começando por estudar e avaliar as várias formulações da electrodinâmica. Os resultados desse trabalho foram publicados em vários artigos, mas o essencial do seu pensamento em relação à crise da física e ao estudo da electrodinâmica estão reunidos em duas obras, *La Science et l'Hypothèse* e *Electricité et Optique*.

Professor na Sorbonne desde 1886, Poincaré ensinou nessa escola várias matérias, entre as quais a física matemática. Embora até essa época ele fosse essencialmente um matemático, os seus temas de investigação haviam-no levado a interessar-se pela física matemática⁴¹, uma área onde as equações diferenciais desempenhavam um importante papel. É o ensino de física matemática, a partir de 1888, que está na origem do seu envolvimento com o electromagnetismo.⁴² Nas suas lições, longe de se contentar em reproduzir os conhecimentos já bem estabelecidos, Poincaré apresentava e discutia os resultados e teorias mais recentes.⁴³ Assim, introduziu em França a chamada teoria de James Maxwell (1831-1879) e também os resultados de Helmholtz, de Heinrich Hertz (1857-1894), de Hendrik Lorentz (1853-1928) e de Joseph Larmor (1857-1942), fazendo comparações e estabelecendo relações entre eles. *Electricité et Optique* foi publicada precisamente para apoiar as suas aulas, mas tornou-se uma obra de referência, ainda hoje útil para quem queira conhecer, no quadro da história da ciência, as diferentes formulações do electromagnetismo dos finais do século XIX. É curioso observar que para estudar a teoria de Maxwell a escola alemã privilegiava a apresentação de Poincaré face aos textos originais do próprio Maxwell.

⁴¹ Houzel e Paty, 1999, 7.

⁴² Darrigol, 2000, 352.

⁴³ Houzel e Paty, 1999, 2.

Poincaré considerava que, de entre todas as teorias, a de Lorentz, apesar de tudo, era a mais coerente com os factos observados e a que melhor se apresentava como uma construção definitiva. Apesar dos seus defeitos, a teoria de Lorentz “dava uma explicação simples de alguns fenómenos”⁴⁴, que até então tinham escapado à teoria de Maxwell. Foi da interacção de Poincaré com o trabalho de Lorentz que surgiu mais outra utilização das geometrias não euclidianas e da noção de grupo.

A crise conceptual da física nos finais do século XIX teve origem, entre outras razões, nas dificuldades sentidas com as leis do electromagnetismo que, estranhamente, ao contrário das leis da Mecânica Clássica, não permaneciam invariantes na mudança de um referencial de inércia para outro.⁴⁵ Lorentz, que trabalhava nesta questão, definiu transformações de coordenadas que conservavam invariantes as equações do electromagnetismo em todos os referenciais de inércia. Em 1900, na conferência de Leyden, Poincaré chamou-lhes “transformações de Lorentz”⁴⁶ e mostrou que definem um grupo⁴⁷, uma propriedade “que constitui a essência do Princípio da Relatividade”.⁴⁸ De facto, ao perceber que o Princípio da Relatividade se pode deduzir das propriedades de simetria das leis da física – ou seja, da invariância das equações da física por acção de um grupo – Poincaré obteve uma versão da teoria da Relatividade baseada no grupo de transformações de Lorentz. É uma versão matemática dessa teoria mas que assenta num aspecto essencial para a física: o da simetria, ou invariância, das suas leis. Poincaré foi o primeiro a reparar nas propriedades de simetria das leis da física⁴⁹ e a caracterizar a sua invariância usando, mais uma vez, a noção de grupo. Os seus resultados, publicados no artigo *Sur la*

⁴⁴ Poincaré, 1902, 242.

⁴⁵ Referencial de inércia é um referencial para o qual, se uma partícula não está sujeita a forças, ou está parada ou se move em linha recta e com velocidade constante. As leis da mecânica são invariantes (têm a mesma expressão) em todos os referenciais de inércia.

⁴⁶ Poincaré, 1900, 464-488.

⁴⁷ Poincaré, 1905-a, pp. 1504-1508. Poincaré chamou-lhe Grupo de Lorentz. Actualmente o termo «transformação de Lorentz» designa diferentes transformações, conforme a teoria física em que se trabalha. Em todos os casos o conjunto das transformações define um subgrupo do grupo de Poincaré.

http://fr.wikipedia.org/wiki/Transformations_de_Lorentz consultado em Julho de 2012.

⁴⁸ Reignier, 2004, 9.

⁴⁹ Feynmann, 1989, 121.

*dynamique de l'électron*⁵⁰, estão na origem de outros trabalhos em Relatividade, em particular de Minkowski.⁵¹

O “grupo de simetria de Poincaré”, assim chamado por Wigner⁵², define, na física actual, o conjunto de transformações que conservam a estrutura do espaço-tempo em relatividade restrita. Os grupos continuaram a desempenhar um papel muito importante na Relatividade, mas também na física em geral, pois as transformações que descrevem a simetria normalmente definem grupos e as propriedades de simetria estão intimamente relacionadas com as leis de conservação. A teoria de grupos veio, aliás, a tornar-se uma área da matemática muito importante na física⁵³, em particular na mecânica quântica.

A noção de grupo, aplicada anteriormente ao estudo das equações diferenciais e à topologia, revelou-se mais uma vez essencial no percurso de Poincaré, agora no tratamento das equações de Maxwell. Os resultados que obteve foram cruciais também na emergência de uma nova área, a Relatividade, mesmo que o papel desempenhado pelas suas descobertas tenha sido objecto de grande controvérsia.⁵⁴ Stephen Hawking sintetiza as conclusões dessa controvérsia afirmando que se deve a Lorentz e Poincaré as transformações matemáticas e a Einstein a interpretação física⁵⁵. Não se pretende, no entanto, discutir aqui a prioridade da descoberta da relatividade, mas antes usar este exemplo para sublinhar, mais uma vez, a capacidade de Poincaré em cruzar conhecimentos de diferentes áreas. Também na física ele utilizou um objecto matemático de forma totalmente inovadora – o grupo – tal como havia feito noutros campos. E uma vez mais a sua ideia estaria na origem de importantes desenvolvimentos, num futuro que ele mesmo já não pôde apreciar.

6. Conclusão

Em resumo, poderíamos dizer que os resultados de Poincaré, embora obtidos no quadro da investigação em matemática, foram aplicados com sucesso a outros domínios do conhecimento. Esta síntese, talvez útil para

⁵⁰ Poincaré, 1905a, 494-550.

⁵¹ Walter, 1999.

⁵² Wigner, 1967, 15-19.

⁵³ http://en.wikipedia.org/wiki/Symmetry_group

⁵⁴ Darrigol, 2004, 6-7.

⁵⁵ http://fr.wikipedia.org/wiki/Controverse_sur_la_paternité_de_la_relativité

descrever o trajecto de alguns cientistas, é contudo insuficiente para caracterizar a travessia de Poincaré, fortemente marcada pela capacidade de transpor barreiras entre saberes. A explicação dessa capacidade é um desafio para a psicologia, para as ciências cognitivas ou ainda para quem investiga sobre a criação matemática.⁵⁶

Será também possível valorizar esse aspecto da obra de Poincaré – as ligações entre conhecimentos – a partir de algumas observações de Laurent Rollet, alguém que conhece bem certos momentos do seu trabalho. Afirma Rollet que as reflexões sobre a ciência de Poincaré, olhadas frequentemente como “filosofia de cientista”, não têm tido a mesma atenção que a sua obra científica.⁵⁷ No entanto, “a filosofia de Poincaré está ancorada na filosofia tradicional, ou seja, numa esfera intelectual que não tem, obrigatoriamente, a física e a matemática como objecto de estudo”.⁵⁸ Partindo da afirmação implícita em Rollet – a de que o lugar da filosofia não tem sido devidamente valorizado no pensamento de Poincaré – será possível construir uma nova perspectiva acerca da interacção ciência-filosofia no desenvolvimento do seu trabalho. Não ousando ir ao ponto de inverter os termos dessa interacção, é no entanto legítimo perguntar até que ponto a filosofia influenciou a sua evolução científica. Concretizando, será que os seus conhecimentos de filosofia e a sua capacidade filosófica não determinaram as suas escolhas científicas? E que o seu “espírito filosófico” não o guiou na investigação? Se assim foi, então o pensamento filosófico terá servido de apoio ao seu trabalho científico que, por sua vez, constituiu uma fonte de ideias para a sua filosofia. E mais, a dualidade pensamento científico – pensamento filosófico teria sido porventura, para Poincaré, uma forte inspiração na arte de cruzar conhecimentos, que é própria da unidade, da transversalidade e da universalidade da filosofia.

⁵⁶ Van-Quynh, 2013.

⁵⁷ Rollet, 2007, 7.

⁵⁸ Rollet, 2007, 7.

Referências Bibliográficas

- Darrigol, O., 2000, *Electrodynamics, from Ampère to Einstein*, Oxford Univ. Press.
- 2004, Faut-il réviser l'histoire de la relativité?, *La Lettre de l'Académie des Sciences*, 14, 6-7.
- Feynmann, R., 1989, *O que é uma lei física*, Gradiva, Lisboa (Edição original: *The Character of Physical Law*, Cambridge: MIT Press, 1967).
- Giedymin, J., 1977, On the Origin and Significance of Poincaré's Conventionalism, *Studies in History and Philosophy of Science*, 8, 271-301.
- 1992, Conventionalism, the Pluralist Conception of Theories and the Nature of Interpretation, *Studies in History and Philosophy of Science*, 23, 423-443.
- Gray, J.J., 1981, *Linear differential equations and group theory from Riemann to Poincaré*, PhD Thesis, Warwick.
- 1984, Fuchs and the Theory of Differential Equations, *Bulletin (New Series) of the American Mathematical Society*, 10, 1, 1-26.
- & Walter, S. A. (Eds.), 1997, *Three Supplements on Fuchsian Functions* by Henri Poincaré, Berlin, Akademie-Verlag.
- (2012), "Poincaré and the idea of a group", *Naw* 5/13, 178-186.
- (2013), *Henri Poincaré, A Scientific Biography*, Princeton Univ. Press, Princeton.
- Houzel, C. & Paty, M., 1999, Henri Poincaré (1854-1912), *Encyclopædia Universalis, Dictionnaire de l'Astronomie*, Albin Michel, Paris, 1999, 696-706. (reproduzido em http://www.scientiaestudia.org.br/associac/paty/pdf/Paty,M_1997g-PoincareEU.pdf)
- Klein, F., 1891, *Ann. Ec. Norm. Sup.*, 87-102; 173-199. Ed. original: *Vergleichen.de Betrachtungen über neuere geometrische Forschungen, Programm zum Eintritt in die philosophische Fakultät*, Erlangen (1872).
- Nabonnand, P., 2000, Les recherches sur l'œuvre de Poincaré, *SMF - Gazette*, 85, 33-48.
- Poincaré, H., 1881, Sur les courbes définies par une équation différentielle I, *J. Math. Pures Appl.*, S. 3, 7, 375-422.
- 1882, Sur les courbes définies par une équation différentielle II, *J. Math. Pures Appl.*, S. 3, 8, 251-286.
- 1885, Sur les courbes définies par une équation différentielle III, *J. Math. Pures Appl.*, S. 4, 1, 167-244.
- 1886, Sur les courbes définies par une équation différentielle IV, *J. Math. Pures Appl.*, S. 4, 2, 151-217.
- 1887, Sur les hypothèses fondamentales de la géométrie, *Bulletin de la Société Mathématique de France* 15, 203-216.
- 1891, Les Géométries non-euclidiennes, *Revue générale des Sciences Pures et Appliquées*, Vol. 2, 769-774.
- 1895, Analysis situs, *Journal de l'École Polytechnique*. (2) 1: 1-123.
- 1900, La théorie de Lorentz et le principe de réaction, *Œuvres*, Tome IX, 464-488.
- 1901, *Electricité et Optique*, Paris, Gauthier-Villars.

- 1902, *La Science et l'Hypothèse*, Paris, Flammarion, (Reed. 1968).
- 1905, *La Valeur de la Science*, Paris, Flammarion (Reed. 1970).
- 1905a, Sur la dynamique de l'électron, Académie des Sciences, *Comptes-rendus*, 140, 1504-1508.
- 1908, L'Invention Mathématique, *L'enseignement mathématique*, 10.
- 1908a, *Science et Méthode*, Paris, Editions Kimé, 1999.
- 1916-1956, *Œuvres*, 11 Volumes, Paris, Gauthier-Villars.
- Pombo, O., 2012, Conceptions of Intuition in Poincaré's Philosophy of Mathematics, *Philosophy Study*, Vol. 2, 6, 384-397.
- Reignier, J., 2004, Poincaré synchronization: From the local time to the Lorentz group, *Proceedings of the Symposium Henri Poincaré* (Brussels, 8-9 October 2004) www.ulb.ac.be/sciences/ptm/pmif/ProceedingsHP/Reignier.pdf, consultado em Julho, 2012.
- Rollet, L., 2007, *Henri Poincaré : Des Mathématiques à la Philosophie, Étude du parcours intellectuel, social et politique d'un mathématicien au début du siècle*, Thèse, Archives - Centre d'Études et de Recherche Henri-Poincaré.
- Rougier, L., 1920, *La philosophie géométrique de Henri Poincaré*, Paris, Librairie Félix Algan.
- Torreti, R., 1984,, *Philosophy of Geometry from Riemann to Poincaré*, 2e éd. (Dordrecht : Reidel, 1e éd. 1978).
- Van-Quynh, A. (Coord.), 2013, Introspection and Intuition in Mathematics, Dossier *Kairos*, *Journal of Philosophy & Science*, 6, 159-237.
- Walter, S., 1996, *Hermann Minkowski et la Mathématisation de la Théorie de la Relativité Restreinte (1905-1915)*, Thèse, Université de Paris.
- 1999, Minkowski, Mathematicians, and the Mathematical Theory of Relativity. In: Goenner, Renn, Ritter & Sauer (eds.), *The Expanding Worlds of General Relativity* (Einstein Studies, volume 7), Boston/Basel, Birkhauser, 45–86.
- Wigner, E., 1967, *Symmetries and reflections*. Indiana, Indiana University Press.

O convencionalismo de Poincaré contextualizado: origem e significado

María de Paz
(CFCUL & UCM)
maria.depaz@hotmail.com

O objetivo deste artigo é definir o contexto filosófico em que surge o convencionalismo de Poincaré, distinguir as suas teses de outras posições filosóficas que, sendo também convencionalistas, se afastam da perspectiva deste autor. Pretende-se assim clarificar o pensamento filosófico do génio francês no que diz respeito à ciência, ainda hoje origem de grande controvérsia. Além disso, procura-se também apresentar, em traços gerais, os problemas científicos que levaram Poincaré a desenvolver as suas posições filosóficas. Procura-se ainda proporcionar uma interpretação sobre o estatuto que o conhecimento científico tem para Poincaré.

Para esclarecer em que consiste o convencionalismo, seja o de Poincaré, seja o de qualquer outro, é preciso enunciar várias questões. Nomeadamente: (1) como é que surgem as convenções, isto é, qual é a sua origem? (2) Como é que se sustentam estas convenções, ou seja, que é que faz que estas convenções possam ser parte integrante das teorias? (3) Como se escolhe entre convenções alternativas, isto é, que critérios se usam para selecionar determinadas convenções? (4) Porque é que se deve adequar as teorias a estas convenções; ou, dito de outro modo, porque é que as convenções servem de guia no proceder científico?¹

Para dar resposta a todas estas questões, seria preciso expor a filosofia de Poincaré no seu conjunto, o que não é objetivo deste artigo. Tentar-se-á definir o marco filosófico que assinala a presença do convencionalismo no seu pensamento, para o distinguir de outras posições também

¹ Estas perguntas são expostas com algumas variantes em Rescorla, 2011.

convencionalistas mas que, na história da filosofia, não surgem no mesmo momento que a de Poincaré.

Durante a primeira metade do século XIX o Idealismo Alemão, que domina o panorama filosófico, é uma filosofia que toma o “eu” como ponto de partida e se centra no sujeito, deixando de lado o mundo natural. No curso desse mesmo século surge a filosofia positivista de Auguste Comte, que destaca o primado dos factos e se afasta da metafísica, procurando o seu fundamento na ideia de que a ciência constitui a forma privilegiada de conhecimento. A hegemonia destas duas posições permanece, no âmbito académico, durante a segunda metade deste século. Assim, em 1870, podemos dizer que existem duas correntes filosóficas que predominam no panorama filosófico: o idealismo e o positivismo². Note-se no entanto que as duas filosofias, sendo contemporâneas recebem influências uma da outra, dando lugar a novas correntes de pensamento que tomam elementos tanto de uma como de outra.

O positivismo, dada a sua particular interpretação da ciência, atraiu o interesse não só de filósofos mas também de cientistas, cuja reflexão epistemológica sobre os resultados da sua própria actividade profissional ocasionou uma aproximação entre ciência e filosofia. No último quarto do século XIX, quando a filosofia das ciências não era ainda uma disciplina academicamente constituída³, emergiram e foram defendidas, tanto por filósofos como por cientistas, posições epistemológicas de grande relevo, de carácter neo-kantiano, positivista crítico, instrumentalista, indutivista, convencionalista, etc. Esta última posição é a que foi sustentada por Henri

² É claro que existem outras posições filosóficas no marco histórico ao qual nos referimos, tais como o materialismo, o espiritualismo e algumas formas de existencialismo, mas todas elas surgem como “reação contra” ou como “matização de” estes dois movimentos principais. Esta mesma ideia é defendida por Mandelbaum, 1971, 4-5.

³ Normalmente considera-se a cátedra de Moritz Schlick de Filosofia das Ciências Indutivas, fundada em 1922 na Universidade de Viena, como a instauração académica da filosofia da ciência e Schlick como o primeiro filósofo da ciência profissional (cf. Friedman, 2001, 12). No entanto, a cátedra de Mach, História e Teoria das Ciências Indutivas (nessa mesma universidade), de 1895, também ocupada por Boltzmann, é sem dúvida nenhuma um precedente. Do mesmo modo, a cátedra de História Geral das Ciências no Collège de France, criada em 1892, por influência de Comte e a cátedra de História da Filosofia na sua relação com as Ciências, criada em 1909 na Sorbonne por Gaston Milhaud. Porém, a do Collège de France acabou em 1922 (cf. Brenner, 2003, 5-7). Outras semelhantes na Europa são a cátedra de Filosofia das Ciências Indutivas na Universidade de Zurique que ocupou Friedrich Lange desde 1870 até 1875 (cf. Schnädelbach, 1983, 110).

Poincaré, tornando-se bem conhecida através da sua obra e também pela controvérsia que gerou.

Na história da filosofia as posições convencionalistas aparecem, por exemplo, no funcionamento de uma sociedade, traduzindo-se neste caso pela existência de um conjunto de normas que, estabelecidas por acordo entre as pessoas de um determinado grupo, definem o que é válido ou aceitável. A ideia de reger a sociedade com base nos acordos entre indivíduos procede da distinção entre aquilo que é dado pela natureza e aquilo que não o é. Ou seja, o que tem a sua origem na natureza é diferente do que é criado pelo ser humano. Esta distinção é atribuída a Hípias nos diálogos *Hípias* e *Protágoras* de Platão. Os sofistas consideravam que as leis são convenções humanas, ou seja, normas que os homens adotam para não viverem do mesmo modo que os animais. Desta forma, para o convencionalismo social a sociedade é um produto humano e, em consequência, não natural. Com a implicação consciente do ser humano nas normas sociais, chega-se à necessidade de um acordo, isto é, de um convénio entre os cidadãos que desejam viver em sociedade. Este tipo de posições fundamenta-se na noção de “contrato social”. O contratualismo considera a sociedade como se, num momento histórico determinado, nela se tivesse realizado um pacto entre os membros do grupo que a compõem. No entanto, estas ideias não resultam quando tentamos aplicá-las à ciência. Dado que a ciência é entendida como um conhecimento da natureza, como é possível que este conhecimento seja o resultado de um acordo entre indivíduos? Ou ainda, como é possível que o conhecimento da natureza seja algo que não tem a sua origem na própria natureza mas que, em vez disso, seja criado pelo homem?

A consideração destas questões põe em evidência que é indispensável esclarecer em que consiste uma visão convencionalista na ciência. As posições convencionalistas na ciência podem entender-se, em geral, de duas maneiras: de um lado, aquelas que consideram que a verdade é algo sempre convencional; de outro lado, aquelas que consideram que algumas coisas comumente tidas por verdadeiras, não o são, mas são, na realidade, convenções.

A primeira destas posições, ou seja, aquela que considera que a verdade é sempre convencional, está mais próxima do contratualismo social, no sentido em que a verdade resultaria de um acordo entre um grupo de indivíduos. No que diz respeito à ciência, esse grupo seria constituído pelos indivíduos que controlam ou produzem a ciência, nomeadamente, os

cientistas. Desta forma, a verdade das teorias ou dos elementos que compõem a ciência é decidida em função dos interesses ou critérios comuns deste grupo de indivíduos. Neste caso, o grupo não é estritamente necessário, pois as convenções também podem ser escolhidas por um único indivíduo, ou seja, pode ser um cientista particular a decidir quais as afirmações científicas consideradas como verdadeiras e que, portanto, governam o seu sistema. Esta consideração responde à pergunta sobre como surgem as convenções em ciência, cuja resposta mais exacta, tendo em conta o que foi dito, seria a de que a convenção surge a partir da decisão de um cientista ou de vários. Quanto à segunda e à terceira perguntas, ou seja, como é que são mantidas estas convenções e como é que podemos escolher entre várias alternativas, a resposta seria que as escolhas podem ser completamente arbitrárias. Mas esta resposta conduz a posições nominalistas extremas, segundo as quais a ciência já não traduz o que se passa na natureza nem tem conexão alguma com ela, isto é, todas as decisões respondem exclusivamente às necessidades do cientista. E são precisamente estas necessidades que dão resposta à última questão posta mais acima, a de porque é que se devem adequar as teorias às convenções. Neste caso, a resposta seria simplesmente: na medida em que elas são necessárias para o cientista, a ciência deve adequar-se a elas.

Este tipo de concepção denomina-se nominalista porque, em geral, não está comprometida com nenhum tipo de ontologia e apenas admite definições dos termos científicos estipuladas arbitrariamente como convenções linguísticas. Deste modo, nega-se que uma definição possa declarar a natureza de um objecto e ser algo mais do que a explicação de um símbolo. Sendo assim, todas as verdades científicas se deduzem de acordos linguísticos relativamente ao significado das palavras que são elas próprias convenções.

Mas a posição que nos interessa caracterizar aqui, e que surge no período histórico que demarcámos no início, nomeadamente o final do século XIX e a transição para o século XX, é a que pode ser identificada com a visão filosófica de Poincaré. Esta posição inclui-se no segundo modo de compreender o convencionalismo, que já referimos. Isto significa que representa uma perspectiva filosófica segundo a qual nem todos os elementos da ciência considerados verdadeiros o são de facto. Ou seja, alguns destes elementos não são verdadeiros mas sim resultado de convenções. De acordo com esta posição, os nossos sistemas científicos encontram-se, pois, regidos por este tipo de elementos convencionais.

Esta visão surge no mesmo contexto de um “retorno a Kant”. O slogan de Liebmann é de 1865, um pouco anterior à época em que Poincaré desenvolve a sua filosofia, mas o movimento neo-kantiano está muito vivo no momento em que Poincaré escreve sobre filosofia da ciência. Embora surja inicialmente no campo da filosofia, o neo-kantismo afecta também a ciência e é adotado por vários cientistas, que tentam uma fundamentação crítica das ciências e, assim, consideram a epistemologia como o aspecto maior da filosofia. A ideia de repensar Kant é, em parte, produto de uma reacção ao idealismo, o qual, de certo modo, se tinha afastado da posição original deste filósofo. No entanto, o neo-kantismo, dado que não tem em conta os aspectos do pensamento de Kant contidos nos *Primeiros princípios metafísicos da ciência da natureza*, não pode ser visto com um regresso literal à interpretação kantiana do mundo, mas sim como uma reacção ao desprestígio da metafísica. Por outro lado, este movimento foi também uma consequência da mudança de posicionamento estratégico da ciência que se foi tornando, na cultura, uma actividade predominante. Essa transformação teve efeitos na própria ideia de ciência, que deixou de ser um sistema de conhecimentos *a priori*, para passar a incluir grande número de conhecimentos empíricos e a ter, portanto, uma composição variável⁴. Da tentativa de reconsiderar as posições kantianas surge, além do neo-kantismo, uma filosofia que põe em questão alguns aspectos do pensamento de Kant. Essa filosofia questiona, nomeadamente, se serão de facto sintéticos *a priori* os conhecimentos que Kant considerava como tal, ou se não terão um outro estatuto. Se, em lugar de verdades sintéticas *a priori*, não serão antes convenções⁵.

Por outro lado, a filosofia convencionalista, reagindo ao positivismo comtiano, a outra corrente filosófica dominante na altura, põe em causa a ideia de que, na ciência, tudo é originado na natureza. Ou seja, questiona se tudo é descoberto a partir de observações empíricas e, conseqüentemente, se o papel do cientista se limita ao registo de dados, ou se, pelo contrário, em lugar de ser um mero colecionador ou descritor de acontecimentos, tem algum papel na constituição da ciência. O convencionalismo procura dar

⁴ Cf. Schnädelbach, 1983, 109.

⁵ É preciso ter em conta que ignoramos até que ponto Poincaré conhecia a filosofia de Kant, ou seja, não sabemos se realmente tinha lido este autor ou se se limita a citar as ideias kantianas a partir da sua recepção e compreensão na época, dado que nunca o cita, embora utilize frequentemente a sua terminologia (juízos sintéticos *a priori*, principalmente). O desconhecimento de Kant por Poincaré é assinalado em Brenner, 2003, 43.

relevo ao papel criador do cientista e mostrar que algumas proposições científicas reflectem a posição de livre escolha do investigador. Um dos problemas da ciência positiva estava no facto de tomar a experiência como único critério científico e de a considerar como um conhecimento em constante mudança. Daí resultou a rejeição da ideia de verdade e, conseqüentemente, a sua inclusão dentro dos mitos da metafísica⁶, o que, do ponto de vista do ser humano, significa olhar para a ciência de uma forma exclusivamente funcional, atribuindo-lhe como único valor a capacidade de produzir acção⁷. A crítica de Poincaré à posição positivista é bem clara no seguinte passo:

A marca que o catolicismo imprimiu na alma ocidental foi tão profunda que muitas mentes pouco emancipadas tiveram a nostalgia da servidão e esforçaram-se por reconstituir as igrejas; foi assim que certas escolas positivistas se tornaram um catolicismo sem Deus. O próprio Auguste Comte sonhava em disciplinar as almas e alguns dos seus discípulos, exagerando o pensamento do mestre, ter-se iam tornado rapidamente inimigos da ciência se fossem mais fortes⁸.

Em resumo, a interpretação convencionalista da ciência de Poincaré caracteriza-se principalmente por: primeiro, questionar o estatuto de algumas “verdades” científicas, com a finalidade de mostrar que essas verdades são apenas convenções; segundo, esclarecer o papel do cientista na constituição da ciência e dessas convenções. Isto significa que a posição de Poincaré não pretende em momento algum eliminar o conceito de verdade do âmbito da ciência, mas sim pôr em causa a maneira como algumas proposições científicas têm, habitualmente, sido tomadas por verdadeiras. Ou seja, sendo o conhecimento considerado um produto humano, podemos dizer que o convencionalismo de Poincaré se opõe a um empirismo *stricto sensu*.

Na perspectiva do empirismo, o conhecimento humano é um mero reflexo da natureza. A mente regista de uma maneira passiva as impressões sensoriais como se fossem pegadas, e todas as ideias que temos são derivadas das impressões primeiras que consistem em dados dos sentidos. A característica principal destas impressões é que são imediatas, ou seja, são apresentadas directamente à nossa mente, unicamente com intervenção dos órgãos dos sentidos. Assim, para a teoria do conhecimento empirista, tanto o

⁶ Cf. Schnädelbach, 1983, 116.

⁷ Tal é o caso do convencionalismo radical de Édouard Le Roy, ao qual Poincaré se opõe em numerosas ocasiões. Cf. Poincaré, 1905, 151-153.

⁸ Poincaré, 1913, 205.

conhecimento como a ciência se baseiam na observação e na experiência, sendo o princípio de não-contradição o único princípio racional que lhes serve de guia.

O convencionalismo, pelo contrário, defende que o conhecimento é um produto de fabricação humana, o que não significa que tudo o que conhecemos seja 'fabricado'. Para Poincaré, o conhecimento também deriva de dados empíricos, a partir dos quais o cientista organiza e estrutura este material. Portanto, a origem do conhecimento encontra-se no facto bruto (empírico), mas a partir deste facto constroem-se hipóteses e, depois, leis e teorias. Ou seja, ao assinalar o papel ativo do cientista no processo de conhecimento, o convencionalismo demarca-se tanto de posições positivistas, nas quais o cientista apenas *descreve* registos tomados da natureza e do empirismo na sua versão mais básica. O conhecimento da forma como surge o convencionalismo de Poincaré dá-lhe, até certo ponto, um novo significado e complementa a sua caracterização.

A filosofia convencionalista de Poincaré tem origem nas suas concepções sobre a geometria. O aparecimento de geometrias não euclidianas tinha posto em questão o *dictum* kantiano segundo o qual os axiomas da geometria euclidiana são juízos sintéticos *a priori*, e que constitui, precisamente a principal objecção de Poincaré à concepção kantiana do espaço e da geometria. Ou seja, Poincaré afirma que se os axiomas de Euclides fossem verdades universais e necessárias não se poderia conceber nenhuma proposição que lhes fosse contrária, e menos ainda basear nessa proposição a construção de todo um edifício teórico, pois isso significaria que não existe geometria não euclidiana. Como de facto existem estas geometrias, e são logicamente consistentes, Poincaré interroga-se, então, se os axiomas da geometria poderão ser verdades experimentais. A resposta é também negativa porque nunca podemos fazer experiências com os objectos da geometria, que são linhas rectas ou esferas perfeitas. Assim, para dar conta do estatuto dos axiomas da geometria, Poincaré cria uma nova categoria epistemológica: a convenção. Deste modo, ao afirmar que os axiomas da geometria são convenções, Poincaré opõe-se tanto a uma interpretação racionalista da geometria quanto a uma interpretação empirista.

A geometria é apenas um exemplo da maneira como Poincaré procede na construção da sua filosofia convencionalista e que serve para ilustrar, para além da problemática filosófica, a questão científica que lhe está subjacente: o surgimento das geometrias não euclidianas. O ponto de partida das suas concepções geométricas é o artigo de 1887, "Sur les hypothèses

fondamentales de la géométrie”⁹, após o qual Poincaré consolida a sua posição em diversos artigos e também em quatro famosas obras filosóficas¹⁰.

Como já foi dito, o convencionalismo geométrico é uma reacção tanto à afirmação de que os axiomas geométricos são juízos sintéticos *a priori*, como à de que estes axiomas resultam de dados empíricos. Assim, esta filosofia define uma posição bem clara relativamente à natureza dos axiomas geométricos e ao estatuto da geometria como ciência. De facto, ao considerar os axiomas geométricos como convenções, Poincaré defende a ideia de que a geometria é uma ciência exacta que não está submetida a revisão, ao contrário do que acontece com as ciências empíricas¹¹; acrescenta-se que se trata de uma disciplina exacta, no sentido de que é matematicamente consistente, o que significa ser isenta de contradição.

Porém, nas experiências, ao tratar com os corpos sólidos, é necessário escolher uma determinada geometria para definir em que tipo de espaço estão inseridos os ditos corpos. É aqui que, segundo Poincaré, se revela mais profundamente o carácter da geometria. De acordo com as suas concepções, a escolha de uma dada geometria em vez de outra é convencional, mas, no entanto, não arbitrária. De facto, mesmo sendo livres para escolher entre uma ou outra geometria para modelar o nosso espaço físico, somos guiados nesta escolha por considerações empíricas relevantes, tais como o tipo de deslocações a que estão submetidos os corpos que nos rodeiam. Ou seja, a escolha de uma determinada geometria é feita por razões de comodidade e, portanto, não descreve as propriedades desses corpos. A geometria baseia-se em idealizações dos corpos circundantes e é por esta razão que não pode ser considerada nem verdadeira, nem falsa,

⁹ Brenner, 2003, 13, afirma: «Poincaré propõe uma teoria original da convenção desde 1891». É verdade que Poincaré publica nesse ano um artigo intitulado “Les géométries non euclidiennes”, no qual estabelece os pontos fundamentais da tese da convencionalidade da geometria de uma maneira mais profunda. Mas as observações finais do artigo de 1887 constituem a primeira expressão do convencionalismo geométrico neste autor.

¹⁰ *La Science et l’Hypothèse* de 1902, *La valeur de la science* de 1905, *La science et la méthode* de 1908 y *Dernières Pensées*, publicado postumamente em 1913. É preciso ter em conta que em 2002 publicou-se uma quinta obra filosófica, a partir de um projeto original de Rougier de 1919, que contava com o apoio de Le Bon (editor dos quatro volumes anteriores). Este projecto não foi apoiado pelos herdeiros de Poincaré e por isso não foi publicado na altura. O editor final do projecto é Rollet. Cf. Poincaré, 2002, X-XXVI.

¹¹ Cf. Poincaré, 1902, 75.

pois não fazendo afirmações sobre propriedades empíricas, a experiência revela-se incapaz de a confirmar ou refutar.

Apesar de tudo, a experiência joga um papel na geometria ou, melhor, um duplo papel. Sem dúvida que os elementos empíricos (os corpos sólidos) fornecem a 'ocasião' de criar uma determinada geometria¹². Mas, simultaneamente, faz-se abstração desses mesmos corpos, com o objetivo de constituir como ciência exacta a geometria criada. Assim, a partir da relação do corpo humano com os corpos circundantes, e por meio da idealização destes corpos reais, obtêm-se as propriedades dos corpos geométricos. Por outro lado, como já foi assinalado, a experiência serve de guia na escolha da geometria mais cómoda, que não é arbitrária mas sim adaptada aos corpos reais¹³. Porém, nem o papel desempenhado pela experiência na génese da geometria, nem a sua aplicação aos corpos reais, fazem que esta seja susceptível de verificação empírica e, por conseguinte, se possa considerar verdadeira ou falsa.

Sendo assim, a geometria não nos dá um conhecimento empírico sobre a natureza, mas sim sobre um grupo particular de sólidos ideais, cujo conceito procede do 'nosso espírito', porque o conceito geral de grupo é uma forma do nosso entendimento; por conseguinte, a sua definição não é originada pela experiência e também não pode ser comprovada por ela¹⁴. Ou seja, os conceitos da geometria são aplicados unicamente por causa da sua comodidade e não representam nenhum conhecimento do mundo natural.

Posto isto, já é possível dar uma resposta às questões acerca do convencionalismo acima colocadas. As convenções geométricas (os axiomas) resultam da interação dum conceito do nosso entendimento (o de grupo), com os dados dos sentidos. Sustentam-se pela sua coerência (compreendida como ausência de contradição) dentro de um sistema (uma geometria determinada, seja euclidiana, lobatchevskiana, etc.) e são escolhidas com base em considerações de comodidade, ou seja pela facilidade de adaptação ao meio físico em que nos encontramos. Isto significa que se as propriedades do mundo natural fossem diferentes, a

¹² A ideia de que a experiência proporciona à geometria a ocasião da sua criação é a causa de que Heinzmann denomine 'ocasionalismo' à filosofia de Poincaré. Cf. Heinzmann, 2001, 458.

¹³ Cf. Poincaré, 1902, 94.

¹⁴ Cf. Poincaré, 1902, 93. Não se pretende discutir aqui as implicações mais profundas deste tipo de conhecimento geométrico baseado no conceito geral de grupo, dado que o objetivo deste artigo é caracterizar o convencionalismo de Poincaré no contexto filosófico e dos problemas científicos do último quarto do século XIX.

geometria mais conveniente, ou a mais cómoda, não seria provavelmente a euclidiana. E é precisamente com base nesta comodidade sugerida pela experiência que tentamos adequar as nossas teorias com as convenções.

Poincaré estende o seu convencionalismo à mecânica e à física. No entanto, o pensamento convencionalista relativamente às ciências naturais não surge apenas como uma mera extensão do convencionalismo geométrico, mas também como uma reflexão sobre o uso de determinados conceitos fundamentais para a mecânica. Desta reflexão resultam também posições epistemológicas sobre questões específicas das ciências da natureza e que não jogam papel algum na geometria, tal como é o caso do conceito do tempo ou da evolução das leis da mecânica. Como resultado desse trabalho, Poincaré centra-se num programa científico que destaca a importância da “física dos princípios”.

A física dos princípios consiste na consideração de um grupo de princípios tomados como fundamentais para o desenvolvimento da mecânica e da física, princípios que servem de guia para a construção das teorias. Praticamente todos estes princípios são derivados de uma concepção científica prévia que tem a sua origem na física de Newton e que é denominada “física das forças centrais”. Alguns deles, tais como o princípio da conservação da energia, o princípio de acção e reacção ou o princípio de relatividade foram postos em causa pelo desenvolvimento do electromagnetismo ocorrido nos finais do século XIX e pela descoberta de novos fenómenos radiativos.

No contexto da física dos finais do século XIX, a formulação da teoria de Maxwell do electromagnetismo é, em alguns aspectos, incompatível com a mecânica newtoniana. Em particular, as equações do Maxwell não são invariantes para as transformações galileanas. Essa violação da invariância leva à elaboração de novas transformações, chamadas por Poincaré “transformações de Lorentz”. Mas a teoria electrodinâmica de Lorentz é, por sua vez, incompatível com princípios fundamentais da mecânica, tais como o princípio de acção e reacção ou o princípio de relatividade. A impossibilidade de fazer concordar entre si os princípios clássicos, nas várias formulações da electrodinâmica, leva a Poincaré a refletir sobre o seu estatuto.

Tanto a mecânica como a física estudam corpos que existem na natureza ou corpos análogos a estes, o que não acontecia com a geometria, pois como mostrámos, o estudo desta última é sempre uma abstracção em condições ideais. Assim, podemos dizer que a geometria não é uma ciência natural, dado que não se ocupa de corpos nem de fenómenos que

acontecem na natureza. Já as ciências naturais constroem as suas hipóteses a partir da observação. Os resultados das observações são posteriormente transformados em leis que devem ser submetidas a verificação experimental, definindo-se assim um segundo momento do processo científico. Por último, quando uma lei é confirmada num grande número de casos, é elevada ao nível de princípio, tornando-se então, por decisão do investigador, imune à comprovação experimental.

Os diferentes processos de constituição da geometria e da ciência natural supõem uma divergência nos seus estatutos epistémicos. Poincaré considera que o estatuto da geometria é o de uma ciência dedutiva, permanente (estabelecida definitivamente e não revisível empiricamente) e que não é nem verdadeira nem falsa, mas em que os métodos são regidos por critérios de “comodidade”¹⁵, no sentido em que são úteis quando aplicados ao mundo. Dado que se trata de uma ciência puramente matemática, não proporciona informação sobre o mundo pois os objectos com que trata são idealizações daqueles que encontramos na natureza¹⁶. Contudo, se a geometria e a ciência natural partilhassem o mesmo estatuto epistémico, quer dizer, se o convencionalismo de Poincaré se aplicasse tanto à geometria como às ciências físico-mecânicas, então estas últimas, limitando-se a tratar com idealizações dos objectos de experiência, também não dariam qualquer informação sobre o mundo natural. O questionamento do tipo de conhecimento que a ciência natural permite ter sobre o mundo é a chave para distinguir o convencionalismo físico do convencionalismo geométrico, tomando como origem desta distinção a diferença de estatuto epistémico existente entre as duas disciplinas. Se considerarmos a ciência natural do mesmo modo que a geometria, isto é, simplesmente cómoda ou útil para agir, então essa ciência não proporciona conhecimento sobre o mundo, mas fornece apenas instrumentos capazes de funcionar dentro dele. Nessas condições, relegaríamos a filosofia de Poincaré para uma posição nominalista ou instrumentalista.

No entanto, há lugar para a verdade na concepção da ciência natural de Poincaré. Ele considera que o conteúdo das teorias científicas corresponde a uma certa realidade, de tal forma que entre aquela e o mundo existe algum tipo de relação, pelo que a sua posição não pode, de maneira nenhuma, ser identificada com posições nominalistas ou convencionalistas extremas.

¹⁵ Cf. Poincaré, 1902, 76.

¹⁶ Cf. Poincaré, 1902, 152.

Segundo o testemunho do próprio Poincaré, o objetivo da ciência da natureza deve ser não só a acção mas também a procura da verdade¹⁷, uma verdade que, embora não seja definitiva, nos aproxime do conhecimento do mundo. Pelo contrário, a geometria, dado que não é nem verdadeira nem falsa, não tem como objetivo a procura da verdade, mas apenas a funcionalidade.

É precisamente na questão epistemológica que Giedymin estabelece a diferença de estatuto entre a física e a geometria nas concepções de Poincaré, ao afirmar que:

A diferença principal entre a geometria e a física do ponto de vista epistemológico é: enquanto as assunções da geometria são todas convencionais, as da física só o são em alguns casos. De outro modo a física não seria uma ciência empírica¹⁸.

Isto significa que nas ciências físico-mecânicas algumas afirmações são completamente empíricas, tais como as leis que descrevem o funcionamento dos fenómenos, que prevêm novos fenómenos e que devem ser verificadas à custa de factos do mundo natural, ou seja, de experiências. A existência de dados empíricos e de factos brutos da experiência como base da constituição das nossas teorias permite diferenciar, de modo definitivo, os estatutos do convencionalismo geométrico e do convencionalismo físico¹⁹.

Se, de facto, as duas disciplinas partilhassem o mesmo estatuto, então os novos desenvolvimentos na física, tais como o electromagnetismo, não teriam sido para Poincaré um problema tão grave, dado que numa concepção nominalista, é o cientista que cria os factos a ter em conta na ciência. Para o nominalismo, a questão das incompatibilidades entre factos e teoria poderia, pois, ser resolvida escolhendo a teoria de Maxwell ou a teoria de Newton conforme fosse mais cómodo, sem que tal procedimento causasse qualquer problema. Mas não foi o que se passou com Poincaré, que se preocupou em traçar a história da física a fim de, a partir da sua evolução, expor e justificar as suas concepções filosóficas convencionalistas.

Ao fazer a história da física, Poincaré começa com a mecânica que, baseando-se na concepção das forças centrais, dá conta dos fenómenos em termos de massas e forças de atracção entre elas. Das leis do movimento destes fenómenos aparecem como consequências alguns princípios que se

¹⁷ Cf. Poincaré, 1905, 19.

¹⁸ Giedymin, 1977, 293.

¹⁹ Deste modo, a nossa posição contrasta com a de autores como Adolf Grünbaum, que considerava o convencionalismo físico como uma simples extensão do geométrico. Cf. Grünbaum, 1963.

apresentam sob a forma específica de teoremas, isto é, de verdades logicamente demonstráveis a partir de certas hipóteses. Na física, estes princípios mecânicos servem de guia na tentativa de dar uma explicação dos novos fenómenos do electromagnetismo. Este é o programa da física dos princípios, de que a teoria electromagnética de Maxwell é, para Poincaré, o exemplo mais notável²⁰. Na física dos princípios renuncia-se ao estudo minucioso de todas e cada uma das partes do 'mecanismo universal', como se tentava fazer na física das forças centrais, e «contentamo-nos em tomar por guia certos princípios gerais»²¹. Os princípios são o resultado de experiências generalizadas e elevadas pelo cientista a um grau máximo de validade, dada a sua generalidade e comodidade. Consequentemente, na mecânica, os princípios não aparecem no momento inicial e, embora derivem da experiência, não constituem propriamente verdades experimentais, mas são elevados ao estatuto de convenções. Na física, ao serem considerados como a base experimental de novas teorias, estes princípios passam a desempenhar um novo papel, Ou seja, na nova física, como por exemplo o electromagnetismo, procura-se a compatibilidade das teorias com estes princípios considerados bem estabelecidos. É assim que Poincaré pode interpretar a história da ciência a partir da continuidade entre teorias, estabelecida através dos princípios.

Em resumo, na ciência natural as convenções surgem principalmente a partir da experiência, ou seja dos dados obtidos pelos cientistas que, com base neles, elaboram hipóteses e depois as transformam em leis e princípios, estes últimos com carácter convencional. Os princípios fazem parte da nova teoria porque têm provado a sua utilidade prática em numerosos casos, mas revelam também virtualidades teóricas, na simplificação das teorias. Por exemplo, ao agrupar um conjunto numeroso de leis sob um único princípio, escolhe-se aquele que tem maior poder unificador e que se provou ser mais preditivo. Por último, é importante dizer que estando em presença de uma nova teoria, e em ausência de novos princípios, tenta-se adequar a teoria aos princípios já existentes, porque eles constituem a melhor ferramenta que se tem, o que não significa que estejam definitivamente estabelecidos. Por vezes, quando novos factos obrigam a rever os princípios estabelecidos, basta uma pequena modificação para estabelecer um novo princípio, a partir do antigo. Foi, por exemplo, o que aconteceu com o princípio de conservação

²⁰ Cf. Poincaré, 1905, 127.

²¹ Poincaré, 1905, 126.

da energia que passou a ser, em algumas situações, princípio de conservação da massa-energia.

O desenvolvimento da filosofia científica de Poincaré foi originado pelos problemas científicos do último quartel do século XIX. Primeiro a emergência das geometrias não euclidianas, que o levou a afirmar a impossibilidade de encontrar uma geometria mais verdadeira que outra para descrever o espaço físico. Depois vem o desenvolvimento da física, especialmente os fenómenos electromagnéticos, mas também a física do calor, a ideia de átomo e os fenómenos de radiação. A preocupação de Poincaré, no entanto, não foi só a de criar uma filosofia que desse conta de determinados problemas científicos, mas também defender a ciência como um meio de conhecimento. Esse papel da ciência tinha, de facto, sido posto em causa pelas concepções nominalistas ostentadas por alguns católicos radicais como Édouard Le Roy, que tentavam minimizar o papel da ciência com o objetivo de destacar o da religião. A sua defesa da ciência não constitui apenas um objetivo do cientista preocupado em justificar a proliferação de teorias no âmbito da sua disciplina. Os seus escritos traduzem também a posição de um intelectual que, tendo em conta o carácter provisório das teorias científicas, pretende manifestar que apesar disso, a ciência constitui uma forma de conhecimento com valor de verdade e não se reduz a um mero jogo entre a comunidade dos cientistas. Muitas das respostas de Poincaré aos ataques de posições anti-intelectualistas e anticientíficas guardam ainda a sua actualidade. Em particular, podem ser vistas como uma antecipação de intervenções que teriam todo o cabimento em muitos dos debates ocorridos na segunda metade do século XX no quadro da sociologia da ciência post-kuhniana, e que hoje tem importantes repercussões nas concepções de muitos estudiosos da ciência. Neste sentido, mesmo se a origem do convencionalismo de Poincaré são os problemas da ciência da sua época, podemos afirmar que o seu significado vai muito para além desses problemas, é muito mais profundo, pois a sua filosofia valoriza o conhecimento científico precisamente como tal, como um conhecimento, e não simplesmente como um 'tipo de discurso'.

Com a apresentação do contexto filosófico e científico em que se desenvolve o pensamento de Poincaré queremos destacar a ideia de que a filosofia da ciência de Poincaré é uma filosofia que se desenvolve através da prática científica. Isto significa que não estamos a tratar com um teórico da filosofia, que ao estudar a ciência, vai gerar uma reflexão de segunda ordem sobre ela. Estamos, em vez disso, perante um cientista que proporciona

soluções filosóficas a problemas científicos. Neste sentido, trata-se de um pensamento desenvolvido especificamente para a ciência e não de uma concepção geral epistémica que possa aplicar-se a problemas particulares. É por esta razão que a perspectiva de Poincaré é heterogênea e difícil de explicitar. De facto, as soluções surgem como resposta aos problemas, como soluções particulares a problemas particulares e, ao mesmo tempo, são soluções que mudam, na medida em que os problemas também mudam. O convencionalismo de Poincaré é uma filosofia que se reescreve continuamente no curso do desenvolvimento da ciência e não propõe soluções definitivas nem permanentes. Porém, a sua heterogeneidade não significa uma falta de coerência pois, como dissemos, o significado mais profundo do convencionalismo revela que as preocupações de fundo do génio francês, nomeadamente o valor da ciência, estão sempre presentes na sua obra.

Referências Bibliográficas

- Brenner, A., 2003, *Les origines françaises de la philosophie des sciences*, Paris, PUF.
- Friedman, M., 2001, *Dynamics of Reason: The 1999 Kant Lectures at Stanford University*, Chicago, University of Chicago Press.
- Giedymin, J., 1977, "On the origin and significance of Poincaré's conventionalism", *Studies in History and Philosophy of Science*, 8 (4), 271-301.
- Grünbaum, A., 1963, *Philosophical Problems of Space and Time*, New York, Alfred A. Knopf. (Reed. 1973, Reidel, Dordrecht).
- Heinzmann, G., 2001, "The foundations of geometry and the concept of motion: Helmholtz and Poincaré", *Science in Context*, 14, pp. 457-470.
- Mandelbaum, M., 1971, *History, Man, and Reason: A study in Nineteenth Century Thought*, Baltimore- London, John Hopkins Press.
- Poincaré, H., 1902, *La Science et l'Hypothèse*, Paris, Flammarion. (Reed. 1968).
- 1905, *La valeur de la Science*, Paris, Flammarion. (Reed. 1970).
- 1913, *Dernières pensées*, Paris, Flammarion. (Reed. 1920).
- 2002, *L'opportunisme scientifique*, Rougier, L. (comp.), Rollet, L. (ed.), Birkhäuser Basel.

Rescorla, M., 2011, "Convention", *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), E. N. Zalta (ed.),
URL = <<http://plato.stanford.edu/archives/spr2011/entries/convention/>>.

Schnädelbach, H., 1983, *Philosophie in Deutschland 1831-1933*, Frankfurt am Main, Suhrkamp Taschenbuch Wissenschaft. (Trad. esp. *Filosofía en Alemania, 1831-1933*, por P. Linares, Madrid, Cátedra, 1991).

Henri Poincaré, ciência e materialismo: o papel das hipóteses na oscilação entre atomismo e continuísmo

André Carli Philot* & Antonio A. P. Videira**¹
(*PPGFIL/UERJ ** UERJ, CNPq)
andrephilot@gmail.com
guto@cbpf.br

Introdução

Nosso objeto de estudo é um artigo produzido por Jules Henri Poincaré no mesmo ano de sua morte, ocorrida em 17 de julho de 1912. Em 2012 completam-se cem anos deste acontecimento que oferece uma oportunidade para que revisitemos um dos seus últimos escritos, intitulado “Les conceptions nouvelles de la matière”. Esse texto nunca havia sido traduzido para outra língua, até que em fevereiro de 2012 uma versão inglesa foi publicada.² A tradução que apresentaremos ao final dessa introdução é a primeira em língua portuguesa.

Neste texto Poincaré apresenta um panorama dos principais modelos teóricos disponíveis à sua época para explicar a natureza da matéria. Além disso, o texto é uma interessante fonte de estudo para a compreensão do pensamento filosófico do *savant* francês, responsável por contribuições frutíferas e inspiradoras para o meio científico e filosófico até hoje.

¹ AAPV agradece a bolsa de pesquisa concedida pelo CNPq, à UERJ pela bolsa Prociência no período 2008-2011 e à FCT (Portugal) pelo apoio financeiro através do projeto PTDC/FIL-FCI/116939/2010 – “The Physics of Principles and the Scientific Epistemology; the cases of Mach, Hertz, and Poincaré”. Somos gratos aos comentários e sugestões de María de Paz e Isabel Serra ao nosso artigo.

² Demopoulos et al., 2012, 221-225.

O objetivo desta introdução é apresentar um breve resumo de sua produção científica e filosófica, abordar as circunstâncias nas quais o texto foi originalmente exposto e, também, fazer uma análise da relação que Poincaré estabelece – ou não – entre ciência e materialismo.

A produção científica e filosófica de Poincaré engloba aproximadamente quinhentos artigos e foi a partir do aprimoramento de alguns deles que quatro livros foram publicados: *La Science et l'Hypothèse (A ciência e a hipótese)* de 1902, *La valeur de la science (O valor da ciência)* de 1905, *Science et méthode (Ciência e método)* de 1908 e *Dernières pensées (Últimos pensamentos)* lançado postumamente em 1913.

Poincaré nunca demonstrou possuir o pensamento enraizado numa única posição, tornando difícil reconhecer suas concepções filosóficas. Notamos que ele sempre evitou as dicotomias, por acreditar que posições polarizadas tendem a perder de vista os problemas reais. Na introdução de *A Ciência e a Hipótese*, ele escreveu uma frase elucidativa sobre esta questão: “Duvidar de tudo ou acreditar em tudo são duas soluções igualmente cômodas: uma e outra nos dispensam de refletir”.³ Perceberemos que suas posições sempre seguirão este tom ponderado, como na polêmica entre ceticismo e dogmatismo, na qual adotará a posição de que nem tudo é vão, assim como nem tudo é certo.

Desta forma, demonstrava não ser um cético, pois acreditava no progresso científico e na sua utilidade social, nem ser um dogmático – dogmatismo aqui entendido em oposição ao ceticismo como sendo a posição filosófica que defende a existência de verdades incontestáveis –, pois reconhecia as dificuldades existentes na construção do conhecimento científico. Tais dificuldades, em sua visão, seriam advindas principalmente da dualidade inevitável que se estabelece entre sujeito e objeto na busca do conhecimento, tal como se pode perceber, por exemplo, no artigo que acompanha esta introdução.

O principal objetivo de sua concepção filosófica era mostrar que, apesar de mudanças no meio científico provocarem a aclamação e o ostracismo de modelos científico-teóricos com grande rapidez, era possível identificar um progresso contínuo do conhecimento científico. Como exemplo, vejamos as palavras que ele emprega para descrever o processo de crise da antiga física matemática newtoniana:

³ Cf. Poincaré, 1984, 15.

Poderemos nós dizer que a primeira (física matemática) foi inútil, que durante cinquenta anos a ciência tomou o caminho errado, e que só resta esquecer tantos esforços acumulados que uma concepção viciosa condenava de antemão ao insucesso? Absolutamente. Achar que a segunda fase poderia ter existido sem a primeira?⁴

A discussão acerca de quais autores e correntes influenciaram seu pensamento filosófico envolve algumas dificuldades, sendo a principal delas a escassez de referências diretas nos textos poincareanos de obras de caráter filosófico ou mesmo de menções a autores que desenvolveram reflexões sobre tais assuntos. Porém, sabe-se que em meados dos anos 1870 ele e os intelectuais franceses Émile Boutroux, Jules Tannery, Paul Tannery e Benjamin Baillaud criaram um grupo de discussão interdisciplinar que mais tarde veio a ser conhecido como o “círculo de Boutroux”. As ideias defendidas pelos membros desse grupo são influências inequívocas no pensamento de Poincaré. Na sua filosofia convencionalista, por exemplo, há “paralelos e influências diretas da epistemologia de Boutroux e Jules Tannery, junto com os estudos históricos de Paul Tannery”.⁵

A importância das convenções no conhecimento científico foi sua reflexão filosófica mais proeminente⁶, ainda hoje sendo minuciosamente analisada por filósofos da ciência, além de torná-lo conhecido como uma das mais importantes e representativas figuras do convencionalismo na ciência. Poincaré não chegou a escrever uma exposição sistemática do seu convencionalismo, tratando sobre o assunto em diferentes artigos e sob diversos focos. Cabe ressaltar que até hoje não existe uma interpretação unificada de sua concepção, pois não há consenso entre os especialistas sobre em que grau Poincaré seria adepto do convencionalismo.

Poincaré considera que a convenção exerce um papel central na geometria, seja euclidiana ou não euclidiana, e também na mecânica clássica. Os axiomas da geometria seriam convenções ou, o que seria equivalente, definições disfarçadas (*définitions déguisées*), porque “a experiência não pode nos ensinar nada sobre a verdadeira estrutura do espaço; conseqüentemente, a escolha de uma geometria para a descrição dos fenômenos físicos é uma questão puramente convencional”.⁷ O mesmo ocorreria com os postulados da mecânica, como, por exemplo, o princípio de

⁴ Cf. Poincaré, 1905a, 113. Parênteses nossos.

⁵ Cf. Nye, 1979, 118.

⁶ Sobre o convencionalismo de Poincaré veja: Giedymin, 1982.

⁷ Torreti, 1982, 325.

igualdade da ação e da reação. Portanto, os fundamentos destas ciências não seriam deduzidos a partir de verdades da experiência ou de juízos sintéticos *a priori*, mas seriam convenções.⁸

Este convencionalismo de Poincaré deu ensejo a uma interpretação do filósofo Édouard Le Roy, discípulo de Henri Bergson, que passou a defender que toda a ciência se baseia em convenções. Em sua filosofia, chamada de nominalista, o cientista não só criaria as convenções, como criaria também o próprio fato científico. Esta livre criação seria responsável pela aparente certeza da ciência que, desta forma, se tornaria uma mera linguagem, cuja utilidade residiria unicamente em servir como guia para a vida prática, mas que fracassaria ao tentar dizer algo de verdadeiro sobre a realidade.

Esta apropriação de discurso foi recusada por Poincaré no seu livro “O valor da ciência” que teve dois de seus capítulos dedicados a esta refutação. Nele, Poincaré defende que a certeza da ciência advém de sua capacidade de prever fenômenos e que o cientista pode criar a linguagem em que o fato científico é descrito, mas não o próprio fato. É provável que a reação forte e decidida, mas polida, de Poincaré também se deva à polêmica que atingiu a ciência no final dos anos de 1890 quando Ferdinand de la Brunetière declarou, em artigo que se tornaria célebre, a falência da ciência. Nas palavras de Ferdinand: “Incapazes de nos fornecer um começo de resposta às únicas questões que nos interessam, nem a ciência em geral, nem as ciências particulares – físicas ou naturais, filológicas ou históricas –, podem reivindicar, como faziam há cem anos atrás, o governo da vida presente”.⁹ Essa posição de Brunetière encontrou enorme aprovação nos meios católicos e conservadores da intelectualidade francesa, os quais eram também, em geral, contra o crescimento do pensamento científico e das filosofias a ele associadas.

Conferências da *Foi et Vie*

O texto, que apresentamos mais abaixo, fez parte de uma série de conferências organizadas pela *Société Foi et Vie* (Sociedade Fé e Vida). Esta sociedade de cunho religioso era dirigida por Paul Doumergue, um dos principais líderes protestantes da época. Doumergue iniciou em 1898 a

⁸ Cf. Poincaré, 1902, 54, e Poincaré, 1908, 115.

⁹ Cf. Brunetière, 1895, 104.

publicação de uma revista de evangelização “no sentido amplo do termo”¹⁰, homônima à sociedade, com tiragens quinzenais. Segundo suas palavras, a revista surgiu “da constatação de que as revistas religiosas eram revistas eclesiásticas e teológicas, sem utilidade para o mundo laico”.¹¹

Em resposta a uma “propaganda a favor do ateísmo, iniciada no *quartier latin* de Paris”¹², a sociedade passou a organizar, em 1907, conferências abordando assuntos amplos, como moral, literatura e sociedade. O próprio Doumergue costumava abrir diversas destas conferências, proferindo discursos em que apregoava a revitalização da moral religiosa. Em 1911, ele largou a atividade paroquial “para se dedicar às atividades de relações públicas, buscando melhor integrar o protestantismo francês com o movimento intelectual”¹³ e, como parte deste trabalho, convidou figuras proeminentes do meio intelectual para as conferências. O matemático Henri Poincaré, os filósofos Henri Bergson e Émile Boutroux e os teólogos Gaston Frommel e Henri Bois foram alguns dos palestrantes.

A sociedade *Foi et Vie* era uma organização religiosa protestante e, em seu tempo, este ramo do cristianismo na França não era numeroso. O censo de 1872 indica que os protestantes representavam 1,6% da população francesa.¹⁴ Como dito anteriormente, Doumergue, diretor da revista *Foi et Vie*, foi afastado das atividades paroquiais para se dedicar exclusivamente à tarefa de criar e manter uma relação entre a intelectualidade francesa e o protestantismo. Os protestantes acreditavam que sua religião tinha mais chances de se fortalecer e crescer entre os segmentos sociais que possuíam alta escolaridade e prestígio social; portanto, a tática de convidar grandes nomes do movimento intelectual para discutir temas que tinham repercussão tanto na vida quanto na fé poderia garantir a presença do público que eles desejavam atrair. Não parece constituir um exagero dizer que as ideias protestantes eram recebidas de modo mais favorável pelas pessoas que acreditavam na manutenção e ampliação dos ideais da Terceira República Francesa do que as de qualquer outro tipo.

Doumergue procurou Émile Boutroux na esperança de conquistar um forte aliado em sua missão. Afinal, Boutroux era um filósofo espiritualista que defendia a união entre a ciência e a religião, além de forte crítico da filosofia

¹⁰ Cf. Encrevé, 1993, 179.

¹¹ Encrevé, 1993, 179.

¹² Encrevé, 1993, 179.

¹³ Encrevé, 1993, 179.

¹⁴ Cf. Encrevé, 1986, 64.

materialista. Doumergue, por ocasião da morte do filósofo no ano de 1921, em sua homenagem, descreve a ocasião na qual se conheceram:

Lembro-me da época já distante onde foram fundadas as conferências *Foi et Vie*. Era o tempo em que militantes do livre-pensamento colocavam cartazes nos muros da *Boulevard St. Germain* e da *Boulevard St. Michel* divulgando as conferências direcionadas a estudantes e espíritos cultos, às organizações intelectuais (...) e, entre os cartazes da conferência que discutia os “Problemas de nosso tempo presente” vieram se intercalar, como em um canto, cartazes anticristãos. (...)

Tomei coragem e, desconhecido, mas certo de que a causa era digna, solicitei ao Sr. Boutroux uma conversa. Eu expliquei minhas preocupações e pontos de vista: deve-se manter o espírito público e defender os fundamentos da vida moral e religiosa, fortalecer a base onde podem ser construídos os abrigos espirituais dos homens. Sr. Boutroux escutou e imediatamente disse: Sim. Devemos.¹⁵

Doumergue e Boutroux aproximaram-se por comungar pensamentos semelhantes em relação à associação entre fé e vida e, desta forma, Boutroux tornou-se um frequente colaborador da *Société Foi et Vie*, tanto através dos seus periódicos como através da sua presença nas conferências. E foi através de Boutroux, seu cunhado, que Poincaré teve contato com Doumergue. Em 28 de outubro de 1909, Boutroux enviou uma carta a Poincaré¹⁶ pedindo que ele comparecesse a uma conferência organizada por Doumergue para falar sobre o tema “Ciência e moral”. Não temos a resposta dada a esta carta, porém, no dia 17 de março de 1910, Poincaré compareceu a tal conferência e proferiu o texto que intitulou de “A moral e a ciência”, mais tarde publicado em seu livro póstumo *Últimos pensamentos*.

Poincaré e Boutroux compartilhavam algumas ideias que estavam em sintonia com a proposta das conferências *Foi et Vie* como, por exemplo, a de que a ciência e a moral devem se unir ao invés de se anularem. Poincaré demonstrou este pensamento em algumas ocasiões, como no texto supracitado “A moral e a ciência”:

A ciência, largamente compreendida, ensinada por mestres que a compreendem e amam, pode desempenhar um papel muito útil e muito importante na educação moral.¹⁷

¹⁵ Cf. Doumergue, 1921, 887-888.

¹⁶ Carta digitalizada pelo Arquivo Henri Poincaré da Universidade de Nancy em: <<http://www.univ-nancy2.fr/poincare/chp/image/boutroux-emile-1909-10-28a.jpg>>.

¹⁷ Cf. Poincaré, 1913, 197-198.

Dois anos mais tarde, no dia 7 de março, Poincaré foi convidado para mais outra conferência, cujo assunto central era o materialismo. Sua exposição logo foi publicada na revista *Foi et Vie*¹⁸ e, pouco depois, ela e outras proferidas no mesmo ano receberam uma versão organizada por Gustave Le Bon e publicada por Ernest Flammarion com o título “O materialismo atual”.

As novas concepções da matéria

“O materialismo atual” contou com artigos que tratam sobre o tema em seus diversos aspectos: econômico, literário, teatral e, também, científico, como no caso da palestra de Poincaré, intitulada “As novas concepções da matéria”. Este texto carrega muito do estilo próprio do autor, acostumado a abordar temas complexos de forma simples, mas sem aprofundar-se em detalhes.

Laurent Rollet afirma o seguinte: “Poincaré não procura alcançar essas entidades ideais do “leitor médio” ou do “grande público”, mas se dirige a um público cultivado e dotado de cultura científica sólida (...)”.¹⁹ Em outras palavras, seus textos de vulgarização eram destinados a um público de “alto nível”, ou seja, pessoas não leigas; em geral, cientistas de outras áreas que não fossem diretamente familiarizadas com o tema central da publicação. Porém, o público das conferências era composto basicamente de curiosos.²⁰

À época em que a conferência foi pronunciada, Poincaré já era um dos cientistas mais conhecidos da França e mesmo no mundo, sendo membro da Academia Francesa de Ciências desde 1887 e da Academia Francesa em 1909. Na conferência, podemos notar que Poincaré adota a postura de um cientista dialogando com a sociedade, como podemos deduzir através da seguinte passagem: “(...) alguns de vocês poderiam esperar que eu respondesse a uma questão que as pessoas do mundo todo costumam direcionar aos cientistas (...)”. É vital a identificação desta postura para a compreensão do tom do texto; Poincaré não está ali procurando emitir uma opinião, que fosse entendida como sendo estritamente pessoal; de certo modo, ele deseja transmitir à sua assistência uma opinião que mereça ser considerada como exprimindo um juízo impessoal dos cientistas. Ele fala a

¹⁸ *Foi et Vie*. n. 7., 1912. Todas as edições disponíveis em: <<http://gallica.bnf.fr>>.

¹⁹ Cf. Rollet, 2000, 233.

²⁰ Cf. Encrevé, 1993, 178-179.

todo o momento como um cientista preocupado com assuntos próprios da ciência, criando em volta de si uma espécie de muralha de proteção que o divide do restante da sociedade. Qual a razão desta postura? É impossível dizer ao certo, principalmente porque o autor não se justificou explicitamente neste ponto.

Afinal, qual seria esta questão que as pessoas do mundo todo desejam que os cientistas respondam? A questão é: a ciência nos conduz ao materialismo? Para compreendermos melhor a relevância e a complexidade desta pergunta temos que entender melhor a concepção materialista e o que ela representava à época. De forma alguma, pretendemos esgotar o assunto, desejando apenas descrever basicamente o que é o materialismo filosófico e os supostos desafios morais criados a partir da adoção desta posição. Para a nossa exposição, recorreremos ao clássico “História do materialismo” de Albert Lange, que abarca o tema tanto de um ponto de vista histórico, apresentando estudos sobre o assunto desde a Antiguidade até o século XVIII, assim como as diversas influências desta concepção para a filosofia e ciências modernas. O livro de Lange, originalmente escrito em alemão, foi traduzido para o francês e era muito conhecido nos meios cultos da sua época, na segunda metade do século XIX.

O materialismo filosófico é uma explicação ontológica, ou seja, diz algo a respeito da natureza das coisas. É considerada uma ontologia monista, pois elege somente um princípio que permeia toda a realidade: a matéria. Em outras palavras, ela considera que não há no mundo nada além da matéria.

Quando esse materialismo começou a ser adotado por alguns cientistas no século XVIII, ele foi visto com alarde, principalmente entre os religiosos, pois esta visão, se não elimina a possibilidade da existência da alma e de Deus, no mínimo obriga uma redefinição da natureza destas coisas. Os religiosos, no geral, resistiram a esta redefinição e desde então há uma disputa entre religião e ciência neste tópico.

Na França do século XVIII, como nos descreve Lange, o materialismo surgiu aliado ao ceticismo e, principalmente, contra as ideias religiosas: “Tão poderosa continuou a ser a influência da tendência cética na França que, entre os materialistas do século dezoito, até aqueles conhecidos como os mais radicais e decididos permaneceram longe da finalidade sistemática de Hobbes e pareceram empregar seu materialismo somente como um meio de manter as crenças religiosas em xeque”.²¹

²¹ Cf. Lange, 1925, livro I, seção IV, p. 9.

Do outro lado, como ilustração do combate que as religiões travavam contra o materialismo filosófico, podemos citar um trecho do “Livro dos espíritos”, escrito por Allan Kardec em 1857. À questão “De que maneira pode o Espiritismo contribuir para o progresso?” Kardec responde “Destruindo o materialismo, que é uma das chagas da sociedade (...)”.²² Esta posição não era exclusiva dos espíritas. Pierre Jean Corneille Debreyne, padre e médico formado pela universidade de Paris, escreveu a obra “Pensamentos de um católico ou considerações filosóficas, morais e religiosas sobre o materialismo moderno” em 1839. Segundo Debreyne, o materialismo era a causa de diversos problemas enfrentados por seu país: “Esta triste, esta profunda degradação intelectual e moral é um dos frutos mortíferos que a filosofia materialista produziu no seio da bela terra da França (...)”.²³

Foi justamente no século XIX que esta disputa se aprofundou. Neste período, a revolução industrial aliada aos avanços científicos produziu invenções como a locomotiva, a fotografia, o telefone, etc. Estas conquistas causaram um impacto forte na população e a ciência viveu anos de muito prestígio e confiança, apesar da declaração, já mencionada, de Ferdinand de la Brunetière sobre a falência da ciência. O método científico era exaltado e novos ramos de estudo foram criados, como a sociologia e a psicanálise, com o intuito de explicar a realidade em toda sua riqueza.

A ciência sofria um processo de institucionalização, enquanto as religiões, especificamente a igreja católica, foram aos poucos perdendo espaço na sociedade, o que acabou por acarretar perda de prestígio junto ao Estado. Foi na França que este processo ocorreu com maior intensidade, começando com a aprovação das chamadas “leis de Jules Ferry” de 1881 e 1882, as quais tornavam o ensino público laico. No auge deste processo, ocorreu a aprovação da lei de 1905 que separava a igreja do Estado.

A visibilidade e relevância da ciência cresciam e as religiões temiam que com ela crescesse o número de “adeptos” do materialismo. O que estava em jogo não era somente a perda de fiéis, mas, sobretudo, a perda dos valores morais; se a matéria fosse tudo o que existisse, os princípios morais perderiam sua natureza eterna e seriam meras regras determinadas por uma sociedade em um tempo qualquer, isto é, o relativismo se tornaria dominante. Se a alma sucumbe junto com o corpo, os atos feitos em vida, supostamente,

²² Cf. Kardec, 1995, 373.

²³ Cf. Debreyne, 1844, 12.

seriam esquecidos com a morte. O materialismo tornou-se sinônimo de fraqueza moral, egoísmo e caos social. É neste quadro de conflito entre a ciência e a religião que Poincaré, numa conferência organizada por religiosos, tem que se confrontar a respeito da questão sobre se a ciência nos conduz ao materialismo.

Em sua palestra, ele aponta dois problemas: a não existência de uma resposta satisfatória e também a má compreensão do significado da pergunta. A razão pela qual ele acredita não haver uma resposta satisfatória não é exposta imediatamente no texto, porém, ao final do segundo parágrafo há uma tese que pode esclarece esta questão e que abordaremos mais a frente.

Sobre a má compreensão do significado da pergunta, Poincaré justifica-se dizendo não saber muito bem o sentido da palavra “materialista”. De onde vem esta incompreensão da palavra materialista? Será que é uma dificuldade particular de Poincaré? Ou será que esta incompreensão também é compartilhada com seus contemporâneos, ou, quem sabe, com os seus pares de prática científica? Afinal, já que a definição de “materialismo” é tão incerta, porque os cientistas a adotam? Todas estas questões – e muitas outras que vamos por em evidência – poderiam ser esclarecidas por Poincaré, mas ele preferiu não dar muita atenção a elas e focou seu texto na divulgação dos últimos estudos científicos sobre a matéria. Por quais razões ele escolheu desviar-se da discussão sobre a origem, justificação e definição do que é o materialismo na ciência? Vejamos se é possível formular respostas para estas perguntas.

Das poucas vezes em que Poincaré aborda diretamente a questão do materialismo no texto, ele o faz de maneira a não deixar espaço à reflexão. Somente três afirmações referentes ao materialismo são feitas pelo autor: 1) a ciência é materialista se considerarmos que a palavra “materialista” denota dar “à matéria um papel preponderante”; 2) nem todos os cientistas são materialistas e 3) a questão do materialismo não pode ser resolvida pela ciência. Diante desta abordagem curta e sem abertura para uma reflexão crítica, incomum ao autor, o texto chega até a soar mal-humorado. Evidentemente, diante da ausência de discussão no texto, o auxílio a outras obras do mesmo autor torna-se inevitável para tentar justificar seu posicionamento quase silencioso diante da discussão sobre o materialismo.

Uma das teses que Poincaré defende em seu livro *A Ciência e a hipótese* é sobre a importância das hipóteses para a física. Segundo essa tese, o físico, ao teorizar sobre a natureza dos fenômenos, necessita utilizar

hipóteses e, entre elas, Poincaré estabeleceu uma classificação segundo a função que cada uma exerce dentro do corpo teórico. Por exemplo, as hipóteses que ele chama de *fecundas* são aquelas que representam as verdadeiras generalizações dos fenômenos naturais e, caso sejam confirmadas ou negadas pela experiência, sempre geram um resultado positivo à ciência.

Outra categoria de hipótese explicitada por Poincaré são aquelas que ele dá o nome de indiferentes. As hipóteses indiferentes “não serão nunca perigosas, desde que compreendamos seu caráter. Podem nos ser úteis, seja como artifícios de cálculo, seja para apoiar nosso entendimento mediante imagens concretas, para clarear nossas ideias, como se diz. Portanto, não há porque bani-las”.²⁴ As hipóteses indiferentes aparecem implicitamente e são bastante discutidas no texto que estamos apresentando. Elas, segundo Poincaré, “se abandonadas, isto é, excluídas da Física, em nada modificaríamos as conclusões alcançadas pelos físicos visto que estas últimas seriam igualmente obtidas caso o procedimento utilizado fosse justamente o oposto”.²⁵

Há também as hipóteses naturais. De acordo com o próprio autor, estas hipóteses são aquelas das quais “não podemos fugir”. Os exemplos que ele apresenta são “a suposição de que a influência de corpos muito afastados é inteiramente negligenciável”, a de que “os pequenos movimentos obedecem a uma lei linear” e, por fim, a “de que o efeito é uma função contínua de sua causa”. A razão pela qual “não podemos fugir” destas hipóteses é que sem elas se tornaria extremamente complexa a dedução de leis e, conseqüentemente, a criação de teorias físicas. Portanto, estas hipóteses, para Poincaré, seriam “as últimas que devemos abandonar”.²⁶

Neste momento da discussão, se voltarmos ao texto “As novas concepções da matéria”, mais especificamente na parte na qual ele discute a relação entre determinismo e ciência, ficaremos tentados em classificar o determinismo como uma hipótese natural. Porém, se analisarmos atentamente o que é dito por Poincaré, notaremos que o determinismo possui um valor ainda maior do que este tipo de hipótese. Afinal, as hipóteses naturais seriam “as últimas que devemos abandonar”, devido à dificuldade que enfrentaríamos sem o auxílio delas; porém, “um mundo onde o determinismo não reina seria inacessível aos cientistas”, portanto, o

²⁴ Cf. Poincaré, 1984, 122.

²⁵ Cf. Videira, 1997, 4.

²⁶ Todas as citações deste parágrafo cf. Poincaré, 1984, 121-122.

determinismo é, para Poincaré, imprescindível. Desta forma o determinismo adquire o *status* de princípio fundante da ciência. Em favor da nossa interpretação, citaremos a passagem de outro texto dele: “A ciência é determinista (...) *a priori*, e se (a ciência) admite o determinismo, é porque sem ele não poderia existir”.²⁷

Outra importante conclusão que podemos deduzir da breve descrição que fizemos sobre o papel das hipóteses para o pensamento de Poincaré é que ele considerava o materialismo como uma hipótese indiferente. Segundo sua visão, o papel do materialismo era o de fornecer imagens concretas ao entendimento a fim de “simplificar tudo, a esclarecer tudo e a remover tanto quanto possível qualquer mistério”. É devido a este caráter hipotético das teorias que visam explicar a constituição da natureza que a ciência estaria “condenada a oscilar constantemente do atomismo ao continuísmo, do mecanicismo ao dinamismo e vice-versa”. A partir desta discussão torna-se claro porque Poincaré não defende nem refuta o materialismo em sua conferência. Mas, afinal, os experimentos que confirmaram, como ele mesmo descreve em certo ponto do texto, a existência do átomo não refutam o caráter hipotético do materialismo?

Voltando ao texto da conferência no ponto em que paramos, Poincaré dedica um parágrafo para explicar sua filosofia da ciência. Ela consiste em afirmar a partir da dualidade filosófica – o espírito²⁸ que deseja conhecer o objeto que é externo a ele – que a ciência sempre será imperfeita e que, por esta razão, haverá espaço para a liberdade dos cientistas. Aqui, caberia uma importante questão: não seria ingenuidade de Poincaré conceber como certo a existência do espírito e sua conseqüente diferenciação substancial em relação à matéria antes mesmo de considerar a possibilidade materialista? No mínimo este apelo ao espírito soa como uma petição de princípio. A importante conclusão que Poincaré chega ao final deste parágrafo é obtida através de uma comparação entre o determinismo e o materialismo: “A questão do materialismo, não mais do que a do determinismo, pois não faço separação, não pode ser resolvida, em última instância, pela ciência”.

O que Poincaré quis obter com esta comparação? À primeira vista poderíamos concluir que ele estaria elevando o materialismo ao mesmo nível do determinismo, ou seja, a princípio da ciência. Porém, esta posição não

²⁷ Cf. Poincaré, 1924, 205. Parênteses nossos.

²⁸ Utilizamos o termo “espírito” nesse parágrafo tal qual ele é usado nas obras de Poincaré, ou seja, o *esprit* francês é semelhante ao *mind* inglês, não havendo nenhuma relação com outros significados.

conseguiria se sustentar ao longo do texto. Ela estaria, inclusive, em contradição direta com a tese central defendida por Poincaré de que a ciência oscila entre atomismo e continuísmo. O atomismo é uma teoria que depende “particularmente do materialismo” enquanto as teorias continuístas não possuem necessariamente esta dependência. Portanto, se o materialismo fosse considerado por Poincaré um princípio científico, não seria razoável a defesa da oscilação teórico-científica.

Pensamos que a correta interpretação da conclusão deste parágrafo é a de que, seguindo sua filosofia da ciência, ele acredita que a ciência seria incapaz de decidir litígios envolvendo a essência das coisas. O funcionamento ou a formação interna da natureza seriam para sempre inacessíveis ao conhecimento científico. Desta forma, questões como “a natureza é essencialmente ou não determinista?” ou “a natureza é essencialmente ou não materialista?” seriam vãs e, neste sentido, tanto o determinismo quanto o materialismo não poderia ser resolvido pela ciência.

Já que o homem “não pode conhecer perfeitamente”, isto é, através de essências, ele é impelido a criar “pontes” que ultrapassem estes “vazios” para tornar o conhecimento possível. Mas o homem é livre para escolher quais serão estas “pontes” que o auxiliarão a construir seu conhecimento imperfeito. O materialismo é um dos artifícios criados pelo homem para conseguir alcançar este objetivo.

Um dos pontos que poderiam ter sido mais bem esclarecidos por Poincaré se refere ao por que da inexistência de uma resposta satisfatória à pergunta “a ciência nos conduz ao materialismo?”. Não há resposta satisfatória porque nenhuma resposta satisfará a todos os contendores ou porque não existe resposta que resolva a questão? Talvez ele acreditasse haver uma resposta, mas ela não solucionaria todos os problemas que a questão levanta ou a resposta não seria satisfatória devido à má definição dos termos “ciência” e “materialismo” por parte de quem está fazendo a pergunta. Enfim, apesar de todas as especulações que se possa fazer, acreditamos ser impossível precisar as razões que o levaram a considerar que não existiria uma resposta satisfatória para a questão.

Apesar da falta de esclarecimentos sobre este ponto, acreditamos que após a leitura do texto o leitor é capaz de identificar, com alguma dificuldade, que Poincaré defende a não dependência exclusiva da ciência em relação ao materialismo. É só no final do texto que ele discute que o éter poderia ser uma tese alternativa que retiraria o papel ativo da matéria e daria, assim, outra explicação válida para o fenômeno aparente da massa. Para fortalecer

esta tese Poincaré poderia também ter citado como exemplo os trabalhos de Wilhelm Ostwald na formulação da teoria energética – ou o energetismo – que pretendia explicar os fenômenos naturais a partir da interação da energia.

Todas estas questões foram preteridas no texto e sem dúvida elas possuem seu mérito, afinal são questões fundamentais relacionadas à ciência. Portanto é legítimo questionarmos as razões que o levaram a adotar esta postura. Aliás, uma postura que não é incomum; diversas vezes ele tomou a mesma atitude em relação a questões semelhantes. Analisemos estas questões para ver se é possível traçarmos um paralelo que explique sua atitude resignada diante da discussão do materialismo.

Na filosofia da ciência de Poincaré, a ciência só consegue alcançar o produto da relação entre objetos; qualquer tentativa de descobrir uma realidade além desta, como a essência ou substância das coisas, ou seja, o que elas são – sua ontologia – é considerado como uma postura vã e inútil. Tudo o que se pode saber sobre o mundo físico deve ser deduzido das relações entre os objetos; a tentativa de investigar a realidade última das coisas é caracterizada como metafísica, não como ciência.

Citamos um exemplo deste tipo de resposta, dada por Poincaré no livro “A Ciência e a Hipótese” ao final do capítulo XII: “Um dia virá, talvez, em que os físicos se desinteressarão dessas questões, inacessíveis aos métodos positivos e as deixarão para os metafísicos. Esse dia ainda não chegou, o homem não se resigna, tão facilmente, a ignorar, eternamente, a realidade última das coisas”.²⁹ Portanto, podemos concluir que Poincaré pensa que há certas questões que, por se referirem a uma realidade inacessível ao homem, sempre permanecerão insolúveis; como diz a expressão em latim, *Ignoramus et ignorabimus* (Ignoramos e ignoraremos).

Podemos traçar um paralelo entre as posturas de Poincaré: de um lado quando fica diante de questões de natureza metafísica e de outro quando deve expor as razões que levam as ciências naturais a considerarem o materialismo como explicação para os fenômenos naturais. Este paralelo não nos permite desenhar conclusões definitivas, mas, no mínimo, põe em evidência a tensão existente entre Poincaré, metafísica e a justificação do materialismo. Queremos dizer que Poincaré sabe não haver provas que possam ser deduzidas a partir dos fenômenos da natureza que nos confirmem ou contradigam o materialismo, por isto, qualquer tentativa de

²⁹ Cf. Poincaré, 1984, 164.

justificação destes princípios deverá fazer uso de um ou outro critério de cunho pessoal.

Por um lado, se a relação entre ciência e materialismo recebeu pouca atenção de Poincaré nesta conferência, por outro os últimos experimentos envolvendo a matéria e os problemas que surgem da tentativa de acomodá-los num corpo teórico consistente são bastante discutidos. O atomismo recebia sua primeira confirmação: os experimentos de Crookes e Perrin, que partiam de caminhos diferentes, chegaram a conclusões semelhantes: os átomos existem e podemos conta-los. Mas estes átomos não são indivisíveis, como pensou Demócrito e seus seguidores. Como Poincaré diz no texto “mal notamos os átomos e já se impõe sobre eles o mesmo problema que existe com os corpos brutos que nossos sentidos são capazes de perceber”. Os cientistas já discutiam sua formação e a hipótese mais estudada era a de que havia um núcleo formado de elétrons carregados positivamente e, orbitando ao seu redor, elétrons carregados negativamente, como em um sistema solar. O próton e o nêutron só vieram a receber confirmação anos depois.

Outra questão difícil de determinar era a verdadeira massa do elétron carregado negativamente. Devido sua grande velocidade, uma parte dessa massa total seria resultado do seu rápido deslocamento, denominada massa aparente, enquanto a outra parte seria a massa real, sua massa mecânica. Porém, os cientistas Abraham e Kaufmann ao tentarem determinar a parcela de cada uma dessas massas descobriram que a massa real seria nula. A partir disso, Poincaré levanta a hipótese de que a massa seria uma “aparência” e que o papel ativo que os elétrons possuíam na teoria atomista poderia ser substituído por buracos no éter que ao serem deslocados em grandes velocidades causariam esta ilusão de possuir massa.

Dificuldades não faltavam para a teoria atomista conseguir triunfar. Outro empecilho que se entevia em seu caminho eram as novas descobertas advindas dos estudos de Planck sobre as trocas feitas por radiação. Estas trocas não seriam feitas por graus contínuos, mas por saltos. Agora os atomistas deveriam se preocupar em abrigar e explicar a ocorrência dos fenômenos quânticos em sua teoria.

As experiências eram tão variadas quanto suas descobertas e evidenciavam fenômenos onde os corpos tinham comportamentos tão estranhos que não era possível explicá-los através das teorias já consagradas pela física, causando um momento de instabilidade na ciência. Será que ela seria capaz de dar uma explicação aos novos fenômenos? Será que para isso ela teria que demolir as “velhas” teorias? Para estas questões,

Poincaré dá a seguinte resposta: “Como lhes disse, já passamos por uma crise semelhante. Mostrei-lhes que, na segunda física matemática, a dos princípios, reencontramos os vestígios da primeira, a das forças centrais; o mesmo acontecerá, se tivermos que conhecer uma terceira. Tal como o animal que se transmuta, quebrando sua carapaça demasiado apertada e envergando outra mais jovem; sob o novo envoltório reconheceremos facilmente os traços essenciais do organismo que subsistiram”.³⁰

Portanto, apesar do atomismo receber uma parcela de confirmação – conquistada com a demonstração da existência dos átomos – ainda havia muitas questões que a teoria atomista ainda não explicava e, portanto, ela ainda era considerada por Poincaré uma hipótese.

Conclusão

Defender que a ciência pode oscilar entre posições teóricas tão diferentes quanto o materialismo e o energetismo soa como uma contradição. Como pode a ciência, que busca a verdade, defender modelos teóricos que dizem coisas antagônicas sobre a constituição da natureza? Será que Poincaré não estava percebendo o que defendia? Pensamos o contrário e acreditamos que podemos nos justificar utilizando as próprias palavras do autor quando ele diz o seguinte:

Quando um físico constata uma contradição entre duas teorias que lhe são igualmente caras, diz, por vezes: Não nos preocupemos com isso, mas seguremos firmemente as duas extremidades da cadeia, ainda que os anéis intermediários nos estejam ocultos. Esse argumento de teólogo embaraçado seria ridículo se devêssemos atribuir às teorias físicas o sentido que lhes dão as pessoas em geral. Em caso de contradição, pelo menos uma delas deveria ser considerada falsa. O mesmo não acontece se buscamos aí só o que deve ser procurado. Pode acontecer que uma e outra exprimam relações verdadeiras e que só haja contradição nas imagens com que revestimos a realidade.

Àqueles que acham que restringimos em demasia o domínio acessível ao cientista, responderei: essas questões que lhes proibimos e que vocês lamentam não são só insolúveis, mas também ilusórias e desprovidas de sentido.³¹

Esta citação longa, porém instrutiva, nos responde de uma só vez diversas perguntas levantadas neste trabalho. A partir dela se torna claro a

³⁰ Cf. Poincaré, 1995, 132.

³¹ Cf. Poincaré, 1984, 128-129.

razão pela qual Poincaré não condena nem exalta nenhuma teoria específica acerca da constituição da natureza, afinal, o que importa para ele são as relações verdadeiras que estas diversas teorias podem exprimir. Esta postura, inclusive, serve de sustentação para a presença do pluralismo teórico dentro da filosofia da ciência de Poincaré.

Essa citação também nos auxilia a entender a razão pela qual Poincaré foge de certas discussões que ele considera infrutíferas. No texto que traduzimos fica claro que ele mutila a concepção materialista, desconsiderando seu aspecto ontológico para se concentrar na discussão sobre as relações epistemológicas que essa teoria evidencia. Essa estratégia visa evitar o enfrentamento direto com problemas tradicionais da filosofia que emergem da adoção do materialismo ontológico, ao mesmo tempo em que não abre mão do valioso auxílio do materialismo enquanto hipótese científica.

Mas a discussão sobre o materialismo poderia ter outra chance. Ela tinha a possibilidade de ser abordada pela perspectiva pessoal de Poincaré, como uma questão de fé ou de moral. O ambiente da conferência permitia esta aproximação. Mas ela foi excluída a partir do momento que ele adota uma postura de falar ao público somente como um cientista discutindo assuntos científicos, omitindo suas convicções. Essa atitude indica que a polêmica acerca do materialismo estava viva na sociedade e que ainda era um assunto delicado de ser tratado, principalmente por uma figura pública como Poincaré.

Essa separação entre vida científica e vida moral também se manifesta quando ele declara: “Mas, isso não significa que todos os cientistas são materialistas, já que a ciência não deve ser confundida com as suas vidas”. Será que Poincaré não considera que a prática da ciência poderia influenciar a vida e vice-versa? Na conferência *Foi et Vie* em que ele participou com o texto denominado “Ciência e moral” ele chega a admitir que “toda forma de atividade humana reage no próprio homem e lhe dá uma alma nova. Há uma psicologia profissional para cada ofício; os sentimentos do lavrador não são os do financeiro; o sábio tem, pois, também a sua psicologia particular, isto é, a sua psicologia afetiva, e daí ressalta alguma coisa que só ocasionalmente concerne à ciência”³². Desta forma, não seria possível que o cientista, acostumado a dar “à matéria um papel preponderante” em seu ofício pudesse transferir, mesmo que de forma indireta, esta importância para sua vida pessoal, para suas convicções morais? Não sabemos o posicionamento

³² Cf. Poincaré, 1924, 191-192.

de Poincaré nestas questões; elas foram negligenciadas em detrimento de um enfoque mais restrito que não dava espaço a estes questionamentos. Neste quesito Poincaré separou fé e vida.

Por fim, cabe a nós fazermos a derradeira pergunta: Poincaré consegue manter o discurso metafísico fora da fundamentação científica? Parece correto afirmarmos que seus esforços são suficientes para limitar o poder da metafísica dentro da ciência, mas não para eliminá-los.

O mesmo acontece com os motivos que nos levam a adotar hipóteses acerca da constituição da natureza. O materialismo, a teoria do éter ou o energetismo não nos são impostos pela natureza. Nós até podemos criar estes quadros conceituais, mas eles se mostram úteis ao homem por revelarem relações que ele não conhecia antes e, a partir delas, é capaz de fazer previsões. Se Poincaré, mesmo demonstrando grande aversão a especulações metafísicas, ainda mantém o determinismo e as hipóteses indiferentes em sua concepção científica, é porque, sem elas, a ciência estaria pior.

Referências Bibliográficas

Barrow-Green, J., 1997, *Poincaré and the Three Body Problem*, Estados Unidos, American Mathematical Society.

Bell, E. T., 1953, *Men of mathematics*, Melbourne, London e Baltimore, Penguin Books.

Boutroux, A., 2012, *Vingt ans de ma vie, simple vérité : La jeunesse d'Henri Poincaré racontée par sa sœur (1854-1878)*. In: Rollet, L. (Ed.), Paris, Hermann.

Boyer, C. B. e Merzbach, U.C., 2011, *A History of Mathematics*. 3. ed. New Jersey, John Wiley & Sons, Inc..

Brunetière, F., 1895, Après une visite au Vatican. *Revue des deux mondes*, 127, 97–118.

Collins, G.P., 2004, The Shapes of Space. *Scientific American*, 94-103.

Darrigol, O., 2004, The mystery of the Einstein – Poincaré connection. *Isis*, 95(4), 614-626.

Debreyne, P. J. C., 1844, *Pensées d'un croyant catholique, ou considérations philosophiques, morales et religieuses sur le matérialisme moderne et divers autres sujets, tels que l'âme des bêtes, la phrénologie, le suicide, le duel et le magnétisme animal. Ouvrage généralement destiné à la jeunesse lettrée, et surtout aux jeunes*

gens qui se livrent à l'étude de la médecine, du droit, et à ceux qui se consacrent à l'état ecclésiastique. 3. ed. Paris, Librairie de Poussielgue-Rusand.

Demopoulos, W. et al., 2012, "Poincaré's « Les conceptions nouvelles de la matière »". *Studies in History and Philosophy of Modern Physics*, 43, 221-225.

Doumergue, P., 1921, Emile Boutroux. *Foi et Vie*, 23A, 887-891.

Encrevé, A., 1993, *Les protestants: Dictionnaire du monde religieux dans la France contemporaine*, Paris, Beauchesne Éditeur.

--- 1986, *Protestants français au milieu du XIX^e siècle – Les réformés de 1848 à 1870.* Genève, Labor et Fides.

Galison, P., 2005, *Os relógios de Einstein, os mapas de Poincaré. Os impérios do tempo.* Lisboa, Gradiva.

Giedymin, J., 1982, *Science and convention. Essays on Henri Poincaré's philosophy of science and the conventionalist tradition*, Oxford, Pergamon Press.

Kardec, A., 1995, *O Livro dos espíritos – Princípios da doutrina espírita.* 76. ed. Tradução de Guillon Ribeiro. Rio de Janeiro, Federação Espírita Brasileira.

Lange, A., 1925, *Geschichte des Materialismus un Kritik seiner Bedeutung in der Gegenwart* (Eng. Trans. *The History of Materialism and Criticism of its Present Importance*, por E. C. Thomas, London, Kegan Paul, Trench, Trubner & Co., Ltd.).

Le Bon, G. (Ed.), 1920, *Le matérialisme actuel*, Paris, Ernest Flammarion.

Mawhin, J., 2005, *Henri Poincaré. A life at the service of science. Notices of the American Mathematical Society*, 25(9), 1036-1044.

Nye, M. J., 1979, The Boutroux circle and Poincaré's conventionalism. *Journal of the History of Ideas*, 40(1), 107-120.

Poincaré, H., 1879, *Sur les propriétés des fonctions définies par les équations aux différentielles partielles*, Paris, Gauthier-Villars.

--- 1895a, Analysis Situs. *Journal de l'École Polytechnique*, (2), 1, 1-123.

--- 1895b, À propos de la théorie de M. Larmor. *L'Éclairage électrique*, 5, 5-14.

--- 1898, La mesure de temps. *Revue de Métaphysique et de morale*, 6, 1-13.

--- 1902, *La Science et l'Hypothèse*, Paris, Flammarion (Trad. Port. *A ciência e a hipótese*, por M. A. Kneipp. 2. ed. Brasília, Editora Universidade de Brasília, 1984).

--- 1904, L'état actuel et l'avenir de la physique mathématique. *Bulletin des sciences mathématiques*, 28 (2), 302-324.

--- 1905a, *La Valeur de la Science* (Trad. Port. *O valor da ciência*, por M. H. Franco Martins. 1. ed. Rio de Janeiro, Contraponto, 1995).

--- 1905b, Sur la dynamique de l'électron. *Comptes Rendus de l'Académie des Science*, 140, 1504-1508.

--- 1908, *La Science et la Méthode* (Eng. Trans. *Science and method* by F. Maitland, New York, Dover Publications, 1952).

--- 1913, *Dernières Pensées* (Trad. Port. *Últimos Pensamentos* por G. L. B. Rio de Janeiro, Livraria Garnier, 1924).

Rollet, L. 2000, *Henri Poincaré: Des Mathématiques à la Philosophie – Étude du parcours intellectuel, social et politique d'un mathématicien au début du siècle*, Nancy, Éditions du Septentrion.

Torretti, R., 1978, *Philosophy of geometry from Riemann to Poincaré*, Dordrecht, Reidel.

Videira, A. A. P., 1997, Poincaré e as hipóteses indiferentes. *Revista da SBHC*, 17, 3-10.

Walter, S. 1996, Henri Poincaré's student notebooks, 1870-1878. *Philosophia Scientiæ*, 1, 1-17.

As novas concepções da matéria¹

Henri Poincaré (1854-1912)

Como esta conferência faz parte de uma série cujo assunto é o materialismo, alguns de vocês poderiam esperar que eu respondesse a uma questão que pessoas do mundo todo costumam direcionar aos cientistas: a ciência nos conduz ao materialismo? Bem, essa pergunta não tem uma resposta satisfatória e confesso que não entendo seu significado. Eu não sei muito bem qual é o significado da palavra “materialista”; se se é materialista todas as vezes que damos à matéria um papel preponderante, a ciência é, claramente, materialista, já que as ciências da natureza, em particular a física e a química, têm a matéria como seu objeto próprio. Mas, isso não significa que todos os cientistas são materialistas, já que a ciência não deve ser confundida com as suas vidas. Eu compreendo um pouco melhor o significado da palavra “determinismo”, embora, ao olhar mais de perto, não tenha tanta certeza se a compreendo. Mas, de uma vez por todas, sim, a ciência é determinista; ela o é por definição. Uma ciência que não é determinista já não é uma ciência; um mundo onde o determinismo não reina seria inacessível aos cientistas; quando perguntamos quais são os limites do determinismo, isso é o mesmo que perguntar até que ponto pode-se estender o domínio da ciência, onde estão os limites que ela não pode ultrapassar?

Sendo assim, todo novo progresso da ciência é uma vitória para o determinismo; como as conquistas dos cientistas nunca devem parar, temos que concluir que haverá espaço para a liberdade e, conseqüentemente, para o espírito. Esta última conclusão foi um pouco apressada; como a ciência é imperfeita, a liberdade conservará um pequeno espaço. E, se este espaço não for restringido, isto será o bastante para ela poder executar tudo; porém,

¹ Tradução de André Philot. Revisão de Antonio Augusto Passos Videira.

a ciência sempre será imperfeita, e não só porque nossas faculdades são deficientes; mas ela será imperfeita por definição. A ciência é a dualidade entre o espírito que conhece e o objeto que é conhecido. Enquanto esta dualidade subsiste, enquanto o espírito se distingue do objeto, ele não pode conhecer perfeitamente, uma vez que tudo o que for visto, o será do exterior. A questão do materialismo, não mais do que a do determinismo, pois não faço separação, não pode ser resolvida, em última instância, pela ciência.

Feitas estas ressalvas, permanece o fato de que entre as teorias físicas há algumas que dependem particularmente do materialismo, se assim posso dizer; e estas são as mais caras aos físicos, já que elas tendem a simplificar tudo, a esclarecer tudo e a remover tanto quanto possível qualquer mistério. Estas teorias são aquelas relacionadas ao atomismo e ao mecanicismo. O atomismo, desde Demócrito², sempre teve adeptos, e temos que reconhecer que é uma teoria atraente. A mente não gosta de prolongar indefinidamente a análise sem qualquer esperança de chegar ao fim. Ela prefere pensar que poderá um dia descobrir os elementos últimos e depois descansar. Existem apenas duas maneiras de compreender o atomismo; ou os átomos são elementos no sentido absoluto da palavra, perfeitamente indivisíveis, como exigido pelo sentido etimológico da palavra “átomo”. Neste caso, chegando ao átomo, poderíamos descansar, pois teríamos atingido a quietude metafísica completa; infelizmente, esta quietude não duraria. A necessidade fundamental do nosso entendimento, a de descobrir as unidades, foi suprida; mas temos outras necessidades. Compreender não nos é suficiente, nós queremos ver; contar os átomos não nos é suficiente, nós queremos representá-los; damos-lhe uma forma e isto é o bastante para não mais considerá-los indivisíveis, mesmo através de todos os meios ao nosso dispor, ou até dos mais poderosos que podemos imaginar. Irresistivelmente, somos levados a nos perguntar se não há elementos nos átomos, os átomos dos átomos, por assim dizer.

Acontece o mesmo com o mecanicismo. Compreendemos melhor o choque do que a ação à distância; esta última possui algo misterioso, pois naturalmente evoca a ideia de uma intervenção externa ao mundo, e é por isso que eu disse anteriormente que o mecanicismo depende do materialismo. Mas os cientistas são feitos para afastarem mistérios que

² Demócrito de Abdera (cerca de 460 a.C. – 370 a.C.) foi um filósofo grego considerado o maior expoente da teoria dos átomos. (N.T.)

inevitavelmente reencontram um pouco mais à frente, e nunca estão tão afastados quanto eles gostariam; e é isto que fez com que quase todos os cientistas, mesmo quando suas convicções filosóficas pessoais estavam longe do materialismo, sempre tivessem uma queda por explicações mecanicistas. Quando há uma ação à distância, apressamo-nos a imaginar um meio intermediário, cuja finalidade é transmitir a ação à frente; apenas um meio não parece ser suficiente; se esse meio for contínuo, não satisfará nosso amor à simplicidade, ou seja, nossa necessidade de compreender. Se o meio for formado por átomos, eles não devem se tocar; devem estar muito perto um do outro, talvez um bilionésimo de milímetro. Mas um bilionésimo de milímetro continua a ser uma distância, do mesmo modo que um quilômetro é uma distância. Para o filósofo, é a mesma coisa; ele iria exigir que a ação passasse de um átomo para outro e isto se tornaria novamente uma ação à distância. Não muito tarde, teria que imaginar, entre os átomos do primeiro meio, um segundo meio mais sutil para servir como veículo da ação.

Estas razões explicam porque a ciência é condenada a oscilar constantemente do atomismo ao contínuismo, do mecanicismo ao dinamismo e vice-versa, e por que essas oscilações nunca acabarão. Mas isto não nos deve impedir de analisar o atual estado das coisas e de nos perguntar em que fase da oscilação nos encontramos, apesar da certeza de que um dia estaremos na fase oposta.

Bem, eu não hesito: neste momento, estamos numa fase atomista; o mecanicismo se transformou, tornou-se mais exato, mais ligado ao corpo; nós veremos agora o quanto. Trinta anos atrás, minhas conclusões eram diferentes. Naquela época, parecia ter voltado o entusiasmo do período anterior, o que nos pareceu um pouco ingênuo. As razões, que nos fizeram concluir a descontinuidade da matéria, conservaram seu valor, no sentido de que elas forneceram um conjunto de hipóteses convenientes, mas nós não mais lhe atribuímos força probatória - já tentamos dispensá-las. Estávamos dispostos a seguir M. Duhem³, que queria fundar uma termodinâmica livre de hipóteses e baseada apenas na experiência: *hypotheses non fingo*⁴. Uma

³ Pierre Duhem (1861-1916) foi um físico francês que deu grandes contribuições científicas no campo da termodinâmica. Dentro da filosofia da ciência discutiu a relação entre teoria e experimento, além de ser pioneiro nos estudos sobre ciência na Idade Média. (N.T.)

⁴ Referência à expressão utilizada por Isaac Newton na segunda edição de *Principia* (1713) que significa “sem criar hipóteses”. (N.T.)

termodinâmica cheia de integrais, não de átomos. O que aconteceu desde então?

A grande fortaleza do mecanicismo é a teoria cinética dos gases. O que é um gás? Alguns dizem: eu não sei. Esta é, obviamente, a resposta mais prudente, mas ela não leva a nada. Ela nos preserva do erro, mas não nos deixa chance para descobrir a verdade; não se mover, sob o pretexto de que nós poderíamos tomar o caminho errado, não é o caminho para a solução. É cada vez menor o número daqueles que ainda tentam responder a pergunta, enquanto outros dizem sempre a mesma coisa: um gás é um conjunto de moléculas em grande número que circulam em todas as direções, com grande velocidade, se chocando com as paredes e entre elas próprias. Como um enxame de mosquitos fechados num quarto, voando ao acaso até colidirem com as paredes, teto ou janelas. Ao se chocar com as paredes, essas moléculas empurrarão e as paredes cederão a esta pressão se não forem firmes; quando a densidade aumenta, o número de colisões também aumenta, pois há mais mosquitos para bater nas paredes, e a pressão aumenta: esta é a lei de Mariotte⁵. Quando o gás se aquece, a velocidade das moléculas cresce e os choques se tornam mais violentos, aumentando a pressão desde que as paredes não cedam e não permitam que o gás se expanda: esta é a lei de Gay-Lussac.

Em resumo, as propriedades gerais dos gases foram facilmente explicadas, mas, nos detalhes, restaram muitas dificuldades, o que fez algumas mentes se perguntarem se a explicação não foi um pouco simplista. O estudo das soluções, por exemplo, o da água salgada, conduziu a uma aproximação inesperada. Vimos que as moléculas de sal dissolvidas dentro de um copo com água se comportam como moléculas de gás dentro de um vaso, ou seja, como mosquitos em uma sala; algumas correspondências numéricas podem ser atribuídas ao acaso; mas isto já foi uma confirmação, apesar das moléculas de gás serem muito menores que as moléculas de sal.

Há muito tempo um naturalista⁶ examinou líquidos orgânicos no microscópio: viu partículas se deslocando em movimentos aleatórios e

⁵ Mais conhecida como *Lei de Boyle-Mariotte*, faz parte das quatro leis que regem os gases perfeitos: *Lei de Boyle-Mariotte*, *Lei de Charles*, *Lei de Gay-Lussac* e *Lei de Avogadro*. (N.T.)

⁶ Este naturalista foi o escocês Robert Brown (1773-1858) que em 1827, ao examinar grãos de pólen em um microscópio, observou partículas em um movimento agitado que acabou levando seu nome. (N.T.)

rápidos; isto é chamado de movimento browniano. Para ele isto era a vida, mas logo percebeu que as partículas inertes, os grãos de carmim, por exemplo, não se moviam sob qualquer circunstância. Os naturalistas abandonaram o problema, pensando que era uma tarefa para os físicos; por outro lado, os físicos não se interessaram. Então, estes naturalistas sem que se saiba a razão - iluminaram fortemente sua preparação microscópica; a iluminação aqueceu a preparação e o calor determinou correntes irregulares no líquido. Finalmente, o Sr. Gouy⁷ decidiu pesquisar: não era nada disso, era um fenômeno novo. As partículas visíveis se movimentam e, à primeira vista, nos fazem acreditar que não obedecem a nenhuma força motriz, sendo isto o movimento perpétuo. Na verdade, contudo, o que as coloca em movimento são os choques de outras moléculas dissolvidas e invisíveis. Assim, voltando aos nossos mosquitos, se não temos olhos suficientemente bons para ver, e se entre eles há algumas grandes moscas, ao observar seus movimentos podemos concluir que as moscas não desviam de sua rota por capricho, mas para evitar ou perseguir insetos menores que não vemos.

Dessa vez nós vimos. Quero explicar como se faz a contagem média das moléculas. A teoria nos diz que, como resultado de choques incessantes, as moléculas trocam suas velocidades até chegarem a uma distribuição média destas velocidades que são, então, mantidas indefinidamente. Nesta distribuição, grandes moléculas ficam lentas e as pequenas moléculas velozes, assim, a força viva⁸ das moléculas grandes é, em média, a mesma que a das pequenas. Nossas partículas visíveis submetidas ao movimento browniano, nossas grandes moscas citadas anteriormente, são realmente moléculas muito grandes. Sabemos sua velocidade, já que observamos seus movimentos, conhecemos suas dimensões porque as vemos. Por outro lado, a teoria nos fez conhecer as velocidades das moléculas pequenas; como a força viva de cada uma deve ser a mesma que a das outras, uma simples regra de três nos dará a massa das moléculas pequenas, e a das moléculas propriamente ditas.

⁷ Louis Gouy (1854-1926) foi um físico francês dedicado a diversos estudos relacionados ao magnetismo, eletrocapilaridade, movimento browniano, etc. É conhecido pela criação da *Balança de Gouy*, um instrumento que mede a suscetibilidade magnética. (N.T.)

⁸ Conceito antigo considerado a primeira formulação sobre conservação de energia. (N.T.)

Isto não foi alcançado até o Sr. Perrin⁹. Vamos imaginar a atmosfera da Terra. À medida que ela se eleva, a pressão e a densidade do ar diminuem, assim como a temperatura. Mas em todos os raciocínios que se seguirem, assumiremos que, por qualquer método de aquecimento, a atmosfera foi mantida a uma temperatura uniforme e constante. Vocês compreendem que ao utilizar as leis básicas da física, torna-se fácil calcular como se comportaria nossa atmosfera se a temperatura fosse mantida constante, mas nossa verdadeira atmosfera não se comporta assim. Se nossa atmosfera fosse formada de hidrogênio e tivesse sempre a mesma temperatura, a densidade diminuiria vagarosamente, porque as moléculas de hidrogênio são bem menores que as de oxigênio ou nitrogênio; suas dimensões aumentariam em uma proporção previsível. Por outro lado, se elas fossem constituídas de moléculas maiores, a dimensão de nossa atmosfera diminuiria. Então, temos as partículas visíveis, as grandes moscas e as partículas brownianas em suspensão na água; temos uma atmosfera em miniatura que podemos estudar, e que está a uma temperatura constante, uma vez que está imersa na água. Comparando-a com o que seria uma atmosfera de hidrogênio na mesma temperatura, veremos o quanto sua proporção está reduzida, ou seja, o quanto nossas partículas são maiores que as moléculas de hidrogênio.

Foi assim que o Sr. Perrin foi capaz de nos dizer quantos átomos há em um grama de hidrogênio; há muito menos do que acreditamos, existem somente 683 mil trilhões. Mas não vamos dizer: vemos os átomos porque os contamos; ao iniciar um cálculo, sabemos de antemão que vamos encontrar um número, um resultado qualquer, e isto não é nenhuma maravilha. Não há nisto nenhuma prova de que os átomos existem.

Eis que surge algo sério. Outra maneira de ver os átomos, chamada espintariscópio¹⁰: alguns traços de rádio e um pouco de fósforo a qualquer distância, o sulfeto de zinco, por exemplo, e o instrumento está feito; se olharmos veremos ocasionalmente um brilho, uma espécie de faísca, e essas

⁹ Jean Perrin (1870-1942) foi um físico francês com trabalhos importantes sobre a composição dos raios catódicos, o número de Avogadro, a origem termonuclear da energia solar e experimentos sobre o movimento browniano. Foi ganhador do Nobel de Física em 1926 por, dentre outras contribuições, seu estudo acerca da estrutura descontínua da matéria. (N.T.)

¹⁰ O espintariscópio é um instrumento científico inventado por William Crookes em 1903 destinado a observar as desintegrações nucleares causadas pela interação de radiação ionizante com um cintilador. (N.T.)

faíscas, sendo discernidas, podem ser contadas. W. Crookes¹¹ afirmou que cada faísca é uma molécula de hélio que se desprende do rádio e bate no sulfureto, mas nós permanecemos céticos: será que estas faíscas não são uma propriedade natural do sulfureto? Será que elas não surgiriam normalmente quando o sulfureto acumulasse energia lentamente e romperiam assim que aquecido o suficiente? Será que as faíscas que observamos não são apenas resultado de uma variação descontínua que significa que o calor foi recebido todo de uma vez?

Porém, veja: uma vez que temos uma segunda maneira de contar as moléculas, absolutamente independente daquela do Sr. Perrin, vamos comparar os resultados. Desta vez, encontramos 650 mil trilhões. Isto é uma coincidência surpreendente, totalmente inesperada. Como bem podem compreender, uns poucos milhares de trilhões não fazem muita diferença.

Temos que nos surpreender; mais de uma dezena de procedimentos, completamente independentes e que não consigo enumerar sem me cansar, nos conduziram ao mesmo resultado. Se houvesse mais ou menos moléculas por grama, o brilho do céu azul seria diferente; os corpos incandescentes brilhariam mais ou brilhariam menos, etc. Não há como negar, nós vemos os átomos.

Aqui faço uma pausa para refletir. Imagino um gigante com um enorme telescópio. Ele alcança o fundo dos abismos obscuros do céu e mira em direção a uma espécie de nuvem que possui um brilho leitoso: é a nossa Via Láctea. Sabemos o que é porque estamos dentro e também sabemos que ela é formada por um bilhão de mundos semelhantes ao nosso. Mas nosso gigante para e se concentra em um problema; ele se pergunta, depois de muito raciocinar, se a nuvem é feita de um material contínuo, ou se é formada por átomos. No entanto, ele se aproxima e um belo dia seu telescópio lhe mostra nessa nuvem milhares de pontos luminosos. “Ah! Aqui estão os átomos.” O infeliz mal sabe que esses átomos são sóis, que cada um destes é o centro de um sistema de planetas e que em cada planeta há milhões de pessoas que discutem interminavelmente se eles próprios são feitos de átomos.

¹¹ William Crookes (1832-1919) foi um químico e físico inglês conhecido por suas investigações sobre os raios catódicos que o levaram a ser pioneiro na construção de tubos de vácuo, sendo um deles apelidado de *tubos de Crookes*. (N.T.)

Bem, aqui estamos nós; mal notamos os átomos e já se impõe sobre eles o mesmo problema que existe com os corpos brutos que nossos sentidos são capazes de perceber. Cada um deles não é um mundo? De quais elementos são feitos cada um desses mundos? Ou melhor, já estamos mais avançados que nosso gigante, e conseguimos distinguir em cada átomo uma rica diversidade. Começamos a ver os detalhes e os cientistas respondem com um encolher de ombros àquele que queria fazer-lhes crer que os átomos do químico, os que acabamos de contar, são pontos matemáticos, os seres indivisíveis, como exigia o grego.

Antes de tudo, vemos nossos antigos átomos se desintegrarem diante de nossos olhos; as substâncias radioativas, por meio de suas próprias atividades, se transformam constantemente; se partirmos o urânio, veremos que ele perde hélio constantemente e são estas emissões contínuas que lhe conferem propriedades radioativas; então ele se transforma em rádio e este último, por sua vez, perde hélio e depois de várias etapas chega ao polônio. Sem dúvida, eles não param por aí e, finalmente, se transformam em um corpo simples e comum, sem radioatividade. Mas isso ainda é só uma decomposição química ordinária, diferente de outras devido à sua rápida velocidade, o enorme calor que emite e os estranhos fenômenos que o acompanham, mas que podem ser expressos por uma equação, como todas as reações químicas, uma vez que os produtos da decomposição são corpos tangíveis, conhecidos e catalogados. Alguns corpos que pensávamos serem simples são compostos, ou seja, a velha teoria atômica permanece intacta.

Mas se olharmos um pouco mais de perto, veremos o átomo ser decomposto em pedaços bem menores, chamados elétrons. Vocês conhecem todos os tubos usados por físicos e médicos para produzir os raios X e bater radiografias. São grandes ampolas de vidro a vácuo onde os eletrodos se ligam a uma fonte de eletricidade; quando a corrente passa, o vidro se torna luminoso, emitindo um brilho esverdeado: o eletrodo negativo, o cátodo, emitiu uma radiação específica chamada raios catódicos. Estes últimos são os raios que, ao baterem no vidro, produzem a luz; são eles que, ao baterem no anticátodo, isto é, no eletrodo oposto ao cátodo, produzem os raios-x, nos quais não vou me ater no momento. O que é exatamente um raio catódico? É um jato de partículas extremamente tênues, carregadas de eletricidade negativa que podem ser coletadas. Essas partículas são chamadas de elétrons. Ao estudar a ação do magnetismo e da eletricidade nos raios catódicos, podemos medir a velocidade dessas partículas, que é

enorme, assim como a razão entre sua carga e sua massa. Não há razão para acreditar que esta carga é a mesma que transporta um átomo na decomposição de soluções salinas por correntes elétricas; deve-se concluir que a massa de um elétron é mil vezes menor que um átomo de hidrogênio. Assim, somos levados a imaginar o átomo como uma espécie de sistema solar; no centro, um corpo relativamente grande com carga positiva e, orbitando em torno deste astro central, planetas muito menores carregados negativamente, que são os elétrons. O sol central atrai os planetas, já que é carregado positivamente e que a eletricidade positiva atrai a negativa; temos a imagem da gravitação newtoniana¹² que rege o nosso sistema solar. E também para nós, que vemos o átomo por fora, ele não parece eletrizado, precisamente porque há tanta eletricidade positiva sobre o sol quanto eletricidade negativa sobre os planetas.

Este novo passo à frente é uma vitória do atomismo. Não só a matéria, mas a eletricidade também deixa de ser infinitamente divisível, o que resulta em elementos irreduzíveis. Não temos como cortar um elétron em dois e levar a metade da sua carga para outro lugar; o elétron é o verdadeiro átomo de eletricidade.

Mas não podemos parar nesta etapa, na qual os elementos últimos são pequenos corpúsculos com pouca massa e carga elétrica invariável. Há pessoas que tiveram a curiosidade de procurar a origem desta massa e demonstraram que ela não existe, que era apenas uma ilusão devido unicamente a fenômenos eletromagnéticos causados pelo movimento do éter¹³ que envolve a carga elétrica. Eu não posso dar a vocês uma ideia dos raciocínios deste estudo, pois eu não conheço o seu resultado. Mas, se há um atributo da matéria que parece pertencer à própria, este atributo é a massa, a tal ponto que as palavras “massa” e “matéria” parecem quase sinônimas. Lavoisier¹⁴, com a balança em mãos, demonstrou a indestrutibilidade da matéria, comprovando a invariabilidade da massa.

¹² A gravitação é considerada uma força fundamental de atração que age em todos os corpos. Ela foi formulada pelo físico inglês Isaac Newton (1643-1727) em sua obra *Philosophiae Naturalis Principia Mathematica* de 1687. (N.T.)

¹³ O éter, várias vezes citado no texto, não é o composto químico, mas sim o fluido hipotético criado pelos antigos para preencher o vazio. Este fluido, elástico, sem massa ou volume, seria o meio natural de propagação da luz e da eletricidade. (N.T.)

¹⁴ Lavoisier (1743-1794) foi considerado pai da química moderna e sua principal contribuição científica foi a enunciação do princípio da conservação da matéria. Além disso, descobriu e nomeou os elementos *oxigênio* e *hidrogênio* e contribuiu para a reforma da nomenclatura química. (N.T.)

Bem, eis que a massa não é mais do que uma aparência, um acaso, e que a velocidade pode variar. Então, o papel ativo retirado da matéria é transferido para o éter, o verdadeiro palco dos fenômenos que atribuíamos à massa. Não há mais matéria, só há buracos no éter; como esses buracos não podem se mover sem perturbar o éter que o envolve, é preciso fazer um esforço para movê-los, mas eles parecem dotados de inércia, enquanto que esta inércia pertence, na verdade, ao éter.

Isto nos lembra do éter que havíamos esquecido. Porém, o éter nos parece um meio contínuo; é possível que seja formado de átomos; mas isto é apenas uma hipótese jogada no ar, pois não vemos esses átomos como agora vemos os dos químicos, eles são muitos e só podemos imaginá-los; e tendo a continuidade instalada, pelo menos temporariamente, no meio etéreo, o único verdadeiramente ativo.

E para terminar, devo dizer uma palavra sobre o último episódio da luta entre os atomistas e os defensores do continuísmo, e este episódio foi sem dúvida o mais inesperado, o mais surpreendente de toda esta história. Max Planck¹⁵ acredita ter razões para concluir que as trocas de calor entre corpos próximos, trocas feitas por radiação, só acontecem por saltos, por graus descontínuos. Ele chama isto de a teoria dos quanta. Eu não sei se vocês perceberam o quão estranha é esta hipótese. Para vocês a compreenderem, vou explorar até as últimas consequências onde acho que ela deve fatalmente nos conduzir. O mundo não mais varia de uma maneira contínua, em graus imperceptíveis; ele varia por saltos; esses saltos seriam muito pequenos aos olhos de seres tão míopes como o homem, e isto é o que nos daria a ilusão da continuidade. É notório que os míopes quando veem uma página impressa de longe não distinguem o preto do branco e só enxergam uma superfície uniformemente cinza. Já não podemos mais dizer: *Natura non facit saltus*¹⁶, mas o contrário. Não é somente a matéria que seria reduzida a átomos, mas também a história do mundo. Ou melhor: seria o próprio tempo, pois dois instantes, inclusive no mesmo intervalo entre dois saltos, não seriam mais perceptíveis, uma vez que correspondem ao mesmo estado de coisas.

¹⁵ Max Planck (1858-1947) foi um físico alemão que ganhou o Nobel de Física de 1918 pela formulação da teoria dos quanta, sua principal contribuição científica. (N.T.)

¹⁶ Um princípio da filosofia natural desde os tempos de Aristóteles: "a natureza não dá saltos". (N.T.)

Não vamos nos apressar. Vocês veem que não estamos perto de assistir o fim da luta entre as duas formas de pensar; tanto aquela dos atomistas, que acreditam na existência de elementos últimos, cuja combinação em um número finito, mas muito grande, é suficiente para explicar vários aspectos do universo, quanto à daqueles partidários do contínuo e do infinito. Esta luta vai durar enquanto a ciência existir ou a humanidade pensar, porque ela surge da oposição entre duas necessidades irreconciliáveis do espírito humano, e esse espírito não pode abandonar essas necessidades sem antes deixar de existir; a necessidade de compreender, quando só podemos compreender o finito, e a necessidade de ver, quando só podemos ver o espaço que é infinito.

Se essa guerra não levar à vitória final nenhum dos combatentes, não significa que ela é fútil. A cada novo combate o campo de batalha se move. Assim sendo, cada passo em frente é uma conquista não de um dos dois beligerantes, mas da humanidade.

Recensão / Book Review

Henri Poincaré. A Scientific Biography

por Jeremy Gray, Princeton e Oxford: Princeton University Press, 2013, 592 pp. ISBN: 978-0-691-15271-4

Henri Poincaré (1854-1912) é um dos mais eminentes cientistas na passagem do século XIX ao XX. Embora a sua formação e o seu pensamento tenham fortes raízes nas concepções científicas do século XIX, muitas das suas ideias se revelaram em avanço relativamente a esse século, e antecipatórias de muitos dos desenvolvimentos científicos do século posterior, e mesmo até do actual. Dada a grande variedade de temas em que trabalhou, sendo por isso qualificado como o “último universalista”, torna-se difícil realizar uma biografia científica de Poincaré, sem incorrer no excessivamente geral ou deixar de lado pontos fundamentais.

No entanto, a obra de Jeremy Gray conseguiu escapar a estes dois perigos: foi capaz de mostrar de maneira exaustiva a diversidade de temas sobre os quais se debruçou o génio francês. As biografias correm sempre o risco de ser o elogio exagerado dos seus protagonistas ou, justamente, o contrário, a crítica mais exacerbada dos mesmos. Porém, o autor deste monumental trabalho consegue situar-se no ponto de vista neutral de quem examina criticamente mas, na maior parte dos casos, sem tomar partido nas controvérsias desencadeadas em torno de um personagem que nem sempre é fácil de perceber. Por exemplo, relativamente a saber se merecia ou não ter ganhado o prémio Nobel, Gray da conta com admirável objectividade histórica da campanha levada a cabo em favor de Poincaré, assim como das razões perfeitamente legítimas, de acordo com os critérios do comité, que levaram a que o prémio acabasse nas mãos de outros investigadores. Acontecia que, nessa época, a maior parte dos científicos que o receberam, com a excepção de Lorentz, trabalhavam em física experimental, campo onde os contributos de Poincaré são inexistentes. Além do mais, os seus resultados mais importantes na física são de natureza essencialmente matemática.

O livro está dividido em doze capítulos, sendo os onze primeiros dedicados aos principais assuntos nos quais Poincaré trabalhou. O último é

composto pelos apêndices que integram explicações técnicas acerca das funções elípticas e abelianas, as equações de Maxwell e um breve glossário de termos técnicos de grande utilidade, principalmente para quem não é um especialista em matemática ou física. A distribuição temática dos capítulos evita uma leitura cronológica que se tornaria confusa no caso de um autor como Poincaré. Dado o número de temas em que ele trabalhou simultaneamente ao longo da vida, proporcionar uma visão histórica linear do seu trabalho só iria complicar a já difícil narração de uma obra científica tão complexa. Além disso, este tipo de distribuição oferece a vantagem de dar a possibilidade de ler os capítulos como ensaios independentes. Os leitores mais interessados no problema dos três corpos ou na topologia podem, sem dificuldade, começar pelos capítulos respectivos. Todos os temas tratados foram contextualizados relativamente à ciência da época o que, para além de revelar o grande conhecimento da história da ciência do autor, proporciona ao leitor um marco de localização científica no tempo.

O capítulo sobre o trabalho de Poincaré na física (o sexto) merece uma atenção especial pela talentosa combinação entre exposição temática e história da ciência que o caracteriza. Sendo Gray um historiador da matemática, não foi certamente uma tarefa fácil expor tão clara e sistematicamente os conceitos científicos envolvidos na investigação desenvolvida em física por Poincaré e, simultaneamente, ir descrevendo os seus resultados. O capítulo começa com uma brilhante apresentação das teorias da electricidade no momento em que Poincaré inicia o seu trabalho neste campo, vai explicando os contributos do matemático francês através das suas lições de ótica e electricidade na Sorbonne, para finalizar com a análise do seu trabalho sobre a “dinâmica do electrão”, sobre a qual tanto se tem escrito a propósito da descoberta da relatividade. Neste percurso, Gray faz-nos passar pelas necessárias teorias de Larmor, Lorentz, pelas concepções do electrão e pelo tão polémico éter, sem excluir, evidentemente, a referência à famosa conferência internacional de Arte y Ciência de St. Louis de 1904 (“L'état actuel et l'avenir de la physique mathématique”). Foi então que, juntamente com a sua valoração da teoria de Lorentz, Poincaré propôs uma visão da história da física, baseada na ocorrência de mudanças conceptuais na disciplina desde a *física das forças centrais* até a *nova mecânica*, destacando o papel essencial da *física dos princípios*, a sua concepção predilecta. O capítulo termina com uma menção aos inícios da teoria quântica e conta como, na sua conferência do Congresso Solvay de 1911, Poincaré realizou uma demonstração do processo de intercâmbio de

energia entre ressonadores hertzianos, proporcionando assim mais uma confirmação da lei de radiação de Planck.

O ponto mais crítico do livro é o capítulo sobre filosofia da ciência (o décimo primeiro), sendo que Gray dedica menos de vinte páginas a este tema. Ao contrário do que acontece com os outros, é impossível ler o que aí está escrito sem ter presente o conteúdo da Introdução e do primeiro capítulo, “The essayist”, onde é apresentado o essencial do pensamento de Poincaré. No capítulo décimo primeiro, Gray limita-se a enumerar uma rapsódia de questões que figuram nos textos de Poincaré, sem as enquadrar ou discutir, apresentando apenas pontos de vista muito actuais (o de Zahar e o de Heinzmann) para abordar as ditas questões. Ou seja, neste capítulo dedicado à filosofia, Gray não estuda a filosofia de Poincaré, não analisa os seus textos, nem os interpreta, apenas apresenta as posições de outros filósofos da ciência que considera mais experientes no tema, inclinando-se, sem grande discussão, pela opinião de Heinzmann. Obviamente, para Gray, as posições de Heinzmann e Zahar não se excluem mutuamente, o que mais uma vez mostra a superficialidade do texto deste capítulo. Trata, contudo, o pensamento actual relativamente aos temas discutidos, referindo principalmente o realismo estrutural e citando os trabalhos de French e Ladyman sobre o assunto.

No primeiro capítulo que, como foi dito, é onde se discutem pormenorizadamente os contributos filosóficos de Poincaré. Gray, ao descrever com precisão as circunstâncias nas quais ele escreveu os seus artigos, contextualiza o seu pensamento na França de 1900. Analisa, então, a importância da cultura da época, a relevância das exposições internacionais como acontecimentos particularmente representativos da mesma (especialmente a de Paris) e os congressos internacionais de matemática, física, filosofia e outras áreas associados a estes eventos, os quais são o lugar de encontro da intelectualidade ocidental, principalmente europeia e norte-americana. Isto faz da biografia de Gray uma obra pouco comum, porque antes de nos contar quem é o homem, onde nasceu e qual a sua formação, como a maior parte de trabalhos deste género, descreve-nos a figura dum intelectual preocupado praticamente por todos os temas que dizem respeito à filosofia da ciência do seu tempo, desde as geometrias não euclidianas até os problemas que as teorias electromagnéticas põem à metafísica, sem esquecer a sua defesa da ciência na controvérsia com Édouard Le Roy sobre o valor epistemológico da mesma.

Já assinalamos que a complexidade da obra de Poincaré obrigou o autor da sua biografia intelectual a realizar escolhas, como por exemplo, a de fazer uma exposição temática em vez de cronológica. Tudo se passa como se esta biografia reflectisse a filosofia convencionalista de Poincaré: tal como nela, Gray foi forçado a fazer escolhas, a tomar decisões. Provavelmente é por esta razão que inicia a biografia com “O Ensaísta”, apresentando assim o génio de Poincaré nos seus traços gerais e nos seus aspectos mais populares, como divulgador da ciência e como *savant* envolvido em controvérsias científico-filosóficas. A escolha de Gray é sem dúvida nenhuma legítima e totalmente justificada no plano geral da obra, sem que por isso, deixe de ser surpreendente.

O plano e a intenção da obra estão claramente apresentados na introdução, onde Gray anuncia o seu livro como a primeira obra do género: “The first full-length study covering all the main areas of his contributions to mathematics, physics and philosophy”. Demorou mais de cem anos a chegar até nós um estudo deste tipo sobre Poincaré, o que de facto prova a dificuldade do trabalho. É estabelecido nessa parte também o enfoque do estudo: “This is a scientific biography of Henri Poincaré. It is confined entirely to his public life: his contributions to mathematics, to many branches of physics and technology, to philosophy, and to public life. It presents him as a public figure in his intellectual and social world; it leaves the private man alone apart from a deliberately brief account of his childhood and education”. Ou seja, Gray quer manter-se fiel ao subtítulo do seu livro: *a scientific biography*, excluindo assim todo tipo de argumentos sociológicos ou psicológicos sobre a ciência de Poincaré, assim como considerações pessoais sobre o homem ou sobre o génio, que em geral não são úteis para perceber uma obra científica, embora possam alimentar a curiosidade do grande público. Como o próprio Gray refere, encontra-se em fase de escrita uma “biografia completa” por parte da equipa dos *Archives Poincaré*, que verá a luz em 2015 e que, possivelmente, tratará outros aspectos da vida do cientista.

De qualquer maneira, as lacunas que a obra possa ter ficam perfeitamente cobertas com a exaustiva lista de referências bibliográficas que ocupa mais de trinta páginas e que abrange todos os temas tratados no livro, estando também aí incluída a filosofia da ciência e a vida pessoal do autor.

Para finalizar, não é fácil dizer a que público a obra se destina. Os primeiros capítulos (“The essayist” e “Poincaré’s Career”) podem, sem dúvida, ser apreciados por um público geral, como era o público-alvo dos

livros de popularização científica de Poincaré, ou seja culto, mas não especialista na matéria. Já os restantes capítulos, sobre o seu trabalho científico, são muito provavelmente de maior proveito para matemáticos, físicos, historiadores e filósofos da ciência com interesses específicos na obra de Poincaré. São textos técnicos, onde a quantidade de fórmulas matemáticas, os tecnicismos e a dificuldade de alguns conceitos torna árdua a leitura, não porque Gray não se preocupe em explicá-los, mas pela dificuldade dos mesmos. Em qualquer caso, é, sem dúvida alguma, uma obra de grande relevância para redescobrir a figura de um dos maiores *savants* da história da nossa cultura, mais de cem anos após o seu desaparecimento.

(Recensão feita por Maria de Paz)

ISSN (on-line): 1647-659X
ISSN (print): 2182-2824



FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA



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>