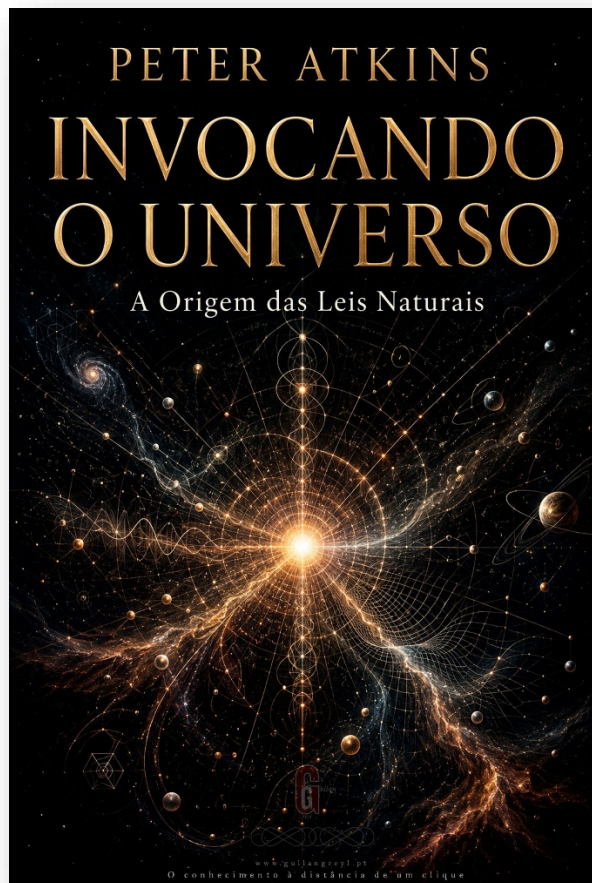


INVOCANDO O UNIVERSO

A Origem das Leis Naturais



Autor: Peter Atkins

Edição e Preparação Digital: Gullan Grey!

Edição Original: *Conjuring the Universe: The Origins of the Laws of Nature*, 2018.

Edição portuguesa concluída em 13-02-2026



Livroteca Gullan Grey!

Sinopse

Em *Invocando o Universo*, Peter Atkins leva o leitor a uma pergunta ainda mais radical do que a origem da matéria: de onde vêm as próprias leis que tornam o mundo compreensível? A maravilhosa complexidade do Universo — estrelas, partículas, calor, luz, movimento, campos, química, vida e consciência — parece repousar sobre um pequeno conjunto de leis e constantes fundamentais. Mas por que existem essas leis? Quem, ou o quê, as “escreveu”? Atkins responde com a sua provocação habitual: talvez elas não tenham sido impostas; talvez tenham emergido da simplicidade extrema, da simetria, da anarquia e, no limite, do próprio nada.

O livro atravessa algumas das regiões centrais da física — conservação da energia, momento linear e angular, mecânica clássica, mecânica quântica, termodinâmica, eletromagnetismo e constantes fundamentais — mostrando como muitas leis podem nascer não de uma autoridade cósmica, mas de condições mínimas: a uniformidade do tempo, a uniformidade do espaço, a simetria, o acaso, a indolência primordial e a liberdade operativa da matéria. Emmy Noether, Newton, Boltzmann, Maxwell, Heisenberg e outros surgem como marcos de uma investigação que tenta mostrar que a ordem pode emergir sem legislador.

No centro da obra está uma inversão poderosa: a ausência de lei não conduz necessariamente ao caos incompreensível; pode, pelo contrário, gerar regularidade. A luz escolhe caminhos, as partículas obedecem à menor ação, a energia conserva-se, a entropia aumenta, os campos eletromagnéticos estruturam-se e as constantes fundamentais delimitam a escala da realidade. Tudo isto aparece como se o Universo tivesse recebido instruções profundas — mas Atkins procura mostrar que essas instruções talvez sejam consequências inevitáveis de algo ainda mais simples.

Invocando o Universo é uma viagem ao subsolo da ciência, onde as leis deixam de ser decretos misteriosos e passam a ser vistas como padrões emergentes da própria estrutura do real. Um livro breve, ousado e conceptualmente intenso, que conduz o leitor até à pergunta final: se a matemática descreve tão bem a Natureza, será porque é apenas uma linguagem humana eficaz — ou porque toca a própria arquitetura profunda do Universo?

Esta tradução foi depurada segundo o princípio:
“Afinar o Instrumento sem Alterar a Melodia”
com apoio de Inteligência Artificial enquanto sistema de amplificação,
organização e clarificação cognitiva.
O objetivo não foi apenas traduzir, mas tornar o texto mais claro, coerente,
imersivo e diretamente experienciável, sem alterar a sua essência.

Gullan Greyl

*Não se trata de saber mais,
mas de ver melhor*

Índice

Prefácio	1
CAPÍTULO 1	2
De Volta à Eternidade — A Natureza das Leis	2
CAPÍTULO 2	21
Muito Barulho por Nada — Como As Leis Podem Emergir Do Nada	21
CAPÍTULO 3	48
A Anarquia Reina — Como Algumas Leis Emergem da Ausência de Leis	48
CAPÍTULO 4	70
O Calor Do Momento — Leis Relacionadas com a Temperatura	70
CAPÍTULO 5	92
Para Além da Anarquia — Por Que é Que Algo Acontece	92
CAPÍTULO 6	118
O Poder Criativo da Ignorância — Como a Matéria Responde à Mudança	118
CAPÍTULO 7	136
A Carga da Brigada Ligeira — As Leis da Eletricidade e do Magnetismo	136
CAPÍTULO 8	157
Medida por Medida — A Origem das Constantes Fundamentais	157
CAPÍTULO 9	185
O Grito das Profundezas — Por Que Funciona a Matemática?	185
Núcleo Ontológico Operacional	202
Glossário	208
Leituras Recomendadas	216
Notas	232

INVOCANDO O UNIVERSO

A Origem das Leis Naturais

Peter Atkins

Prefácio

O funcionamento do mundo tem sido atribuído, por alguns, a um Criador extraordinariamente interventivo, mas incorpóreo, que guia ativamente cada elétron, quark e fóton para o seu destino. O meu instinto recua perante esta visão perdulária do funcionamento do mundo, e a minha razão segue a direção apontada por esse instinto. Nas páginas seguintes, exploro se existe uma visão mais simples do que está a acontecer. Afinal, os cientistas são talhadores de simplicidade a partir da complexidade e, tipicamente, preferem o menos elaborado ao mais elaborado. Exploro os recantos mais longínquos expostos por este talhar e defendo que as leis da natureza — os nossos resumos do funcionamento do mundo — surgiram da forma mais simples possível. Defendo que elas brotam de nada mais do que indolência e anarquia, temperadas aqui e ali com uma pitada de ignorância.

Limitei o meu campo de discurso ao que é comum. Assim, conduzo-o pela mecânica, tanto clássica como quântica, pela termodinâmica e pelo eletromagnetismo. Com confiança decrescente, mas espero que de forma intelectualmente estimulante, levo-o à origem das constantes fundamentais e concluo interrogando-me sobre a eficácia da matemática na formulação das leis naturais e a sua possível revelação da estrutura profunda da realidade. Parte do que escrevo é especulação, pois a compreensão plena continua além do alcance da ciência, apesar do progresso notável realizado em três séculos de esforço sério. Se procurar uma justificação mais aprofundada para algumas das minhas observações, encontrará essa justificação nas Notas no final do livro, que tratei como um espaço seguro para abrigar equações.

Espero, nas páginas seguintes, responder a questões que possam ter-lhe ocorrido, mas que ficaram sem resposta, e que as minhas palavras revelem alguns aspetos da surpreendente simplicidade deste mundo gloriosamente complexo.

PETER ATKINS
Oxford, 2017

CAPÍTULO 1

De Volta à Eternidade

— A Natureza das Leis

Pois todo o conhecimento e maravilhamento (que é a semente do conhecimento) é uma impressão de prazer em si mesmo.

— Francisco Bacon, O Avanço do Conhecimento (1605)

Preciso de preparar a sua mente para um pensamento tão absurdo que até pode ser verdadeiro. A ciência tem progredido através de revoluções, por vezes designadas «mudanças de paradigma», em que aquilo que era tomado como *senso comum* ou simplesmente como uma *atitude dominante* é substituído por algo que parece estar mais próximo da verdade. Aristóteles fornece um exemplo; Copérnico, outro. Assim, Aristóteles, sentado na sua poltrona de mármore, refletia sobre o voo das flechas e inferia que estas eram empurradas por vórtices de ar atrás delas. Tomava também uma visão mais geral, ao observar como os carros eram mantidos em movimento pelo exercício constante do esforço dos bois, concluindo que o movimento tinha de ser sustentado por esforço. Galileu e depois Newton perceberam os efeitos da atmosfera e da lama e substituíram a visão de Aristóteles pela sua oposta, segundo a qual o movimento prossegue sem esforço, a menos que seja travado por um esforço. O ar inibe o voo das flechas e, embora Aristóteles não

o pudesse saber, as flechas voam ainda melhor no vácuo, onde não podem existir vórtices sustentadores. Ele deveria tê-lo notado, mas não teve oportunidade: carros sobre gelo, em vez de se atolarem na lama, não precisam da tração dos bois para sustentar o seu movimento. Copérnico, como é bem conhecido, efetuou uma revolução cósmica e uma profunda simplificação do entendimento (um marcador significativo de aproximação à verdade), ao rejeitar a revolução diária do Sol em torno da Terra, ditada pelo senso comum, e ao adotar um Sol central e estacionário e uma Terra em órbita e em rotação.

Revisões mais subtis, mas não menos profundas, estavam ainda por vir, com as revoluções intelectuais do início do século XX a eclipsarem as convulsões políticas quase contemporâneas e as de um século ou mais antes. A visão de senso comum de que os acontecimentos podem ser considerados simultâneos teve de ser abandonada depois de 1905, quando Albert Einstein transformou a nossa percepção do espaço e do tempo, fundindo-os num espaço-tempo. Essa fusão misturou o tempo no espaço e o espaço no tempo numa extensão que dependia da velocidade do observador. Com o tempo e o espaço assim entrelaçados, dois observadores em movimento relativo não podiam concordar sobre se dois acontecimentos eram simultâneos. Esta revisão fundamental da arena das nossas ações e percepções pode parecer um preço elevado a pagar para nos aproximarmos de um maior entendimento, mas também ela acabou por produzir uma simplificação da descrição matemática do mundo físico: as explicações dos fenómenos deixaram de ter de ser remendadas a partir da bricolage dos

conceitos da física newtoniana e passaram a emergir naturalmente da fusão do espaço com o tempo.

Mais ou menos na mesma época, a recém-nascida classe dos teóricos quânticos transformou o pensamento noutra direção, ao mostrar que Newton tinha sido enganado de mais uma forma e que até a migração da física newtoniana para a nova arena do espaço-tempo, operada por Einstein, era fundamentalmente falsa. Nesse caso, embora tivesse raspado o efeito da lama do movimento, Newton ficara preso à visão de senso comum, inspirada no quintal da quinta, segundo a qual, para especificar uma trajetória, era necessário considerar tanto a posição como a velocidade. Quão chocados ficaram os físicos clássicos, mesmo aqueles que tinham aprendido a viver confortavelmente no espaço-tempo, quando se revelou que esta noção tinha de ser abandonada. No imaginário popular, o ícone desse abandono é o princípio da incerteza, formulado por Werner Heisenberg em 1927. O princípio afirma que a posição e a velocidade não podem ser conhecidas simultaneamente e, assim, parecia minar toda a esperança de compreensão ao eliminar aquilo que era tomado como o alicerce da natureza, ou pelo menos da descrição da natureza. Mais adiante no livro, argumentarei contra a ideia de que o princípio de Heisenberg compromete a perspectiva de entendimento e de descrição completa.

Aparentemente, ainda haveria pior (mas, como acontece com muitas perturbações conceptuais, esse pior era na verdade melhor disfarçado de pior). O senso comum distinguia sem hesitação entre partículas e ondas. As partículas eram pequenas coisas nodosas; as

ondas ondulavam. Mas, numa revolução que abalou a matéria até ao seu âmago, essa distinção revelou-se falsa. Um exemplo inicial foi a descoberta, em 1897, do eletrão pelo físico J. J. Thomson, com todos os atributos de uma partícula; mas depois, em 1911, a demonstração pelo seu filho G. P. Thomson, entre outros, de que, pelo contrário, o eletrão possuía todos os atributos de uma onda. Gosto de imaginar o pai e o filho sentados, gélida e silenciosamente, à mesa do pequeno-almoço.

Mais provas se acumularam. Um feixe de luz, sem dúvida uma onda de radiação eletromagnética, revelou partilhar os atributos de um fluxo de partículas. As partículas ondulavam em ondas de acordo com o tipo de observação que delas se fazia; as ondas, por sua vez, corpusculavam-se em partículas. Quando a mecânica quântica se estabeleceu (em 1927), em grande parte graças a Heisenberg, monasticamente isolado numa ilha, e a Erwin Schrödinger, que, segundo relatou, a concebeu num derrame de paixão erótica enquanto estava numa montanha com uma amante, já não era possível preservar esta distinção fundamental: de um modo totalmente contrário ao senso comum, todas as entidades — dos eletrões para cima — tinham uma natureza dupla e fundida. A dualidade tinha usurpado a identidade.¹

Poderia continuar. Quanto mais a estrutura profunda do mundo é exposta, menos parece que o senso comum — pelo qual entendo a intuição baseada numa experiência local, não controlada e casual do ambiente quotidiano, essencialmente o alimento não digerido do intestino, em vez do alimento intelectual da avaliação de cérebros coletivos, isto é, a inspeção controlada e detalhada de fragmentos

isolados do mundo (em suma, experiências) — seja uma fonte fiável de informação. Cada vez mais parece que o entendimento profundo advém de ir descascando cada camada de senso comum (retendo, naturalmente, a racionalidade). Com isto em mente, e esperando que a sua mente esteja preparada para abdicar do senso comum como investimento na obtenção de uma compreensão futura, gostaria de deslocar ainda mais um aspeto do senso comum.

Gostaria de afirmar que não aconteceu grande coisa na Criação. Tenho consciência, claro, das descrições arrebatadoras que dotaram esse momento de um drama impressionante, pois certamente o nascimento de tudo deveria ser cosmicamente dramático? Um gigantesco cataclismo cósmico. Uma explosão espetacular, à escala do universo, de atividade primordial espantosa. Uma explosão tremenda a abalar o espaço-tempo até aos seus alicerces. Uma bola de fogo prene de intensidade abrasadora, a queimar o espaço. Enormíssima. O «Big Bang», o próprio nome, evoca drama à escala cósmica. De facto, Fred Hoyle introduziu o termo de forma depreciativa e sarcástica, em 1949, favorecendo a sua própria teoria de criação contínua e serena, uma cosmogenia eterna, um mundo sem princípio que implicava um mundo sem fim. O Bang é visto como uma explosão maciça que preenche todo o espaço, criando de facto todo o espaço e todo o tempo, e em cujo tumulto de calor o universo inteiro se expande a partir de um mero ponto de temperatura e densidade inimagináveis até à vasta e muito mais fria extensão ainda em expansão que hoje encaramos como o nosso lar cósmico. A isto acresce a moda atual de considerar uma «Era inflacionária», em que o universo duplicava de tamanho a cada fração minúscula de segundo, antes de atingir,

após menos de um pestanejar, a expansão relativamente contida da meia-idade, com temperaturas de apenas alguns milhões de graus, que deu início à Era que hoje conhecemos.

Não aconteceu grande coisa? Sim, é um grande passo pensar em toda essa hiperatividade, energia e emergência de matéria fundamental em geral como representando que não aconteceu grande coisa. Mas acompanhe-me. Gostaria de explorar o pensamento contraintuitivo de que nada de muito significativo aconteceu quando o universo passou a existir. Não estou a negar que o Big Bang tenha acontecido de forma dramática: há tanta evidência a seu favor, e evidência considerável a favor da Era inflacionária, que seria absurdo rejeitá-lo como descrição do universo primordial, há pouco menos de 14 mil milhões de anos. O que estou a sugerir é uma reinterpretação.

A motivação para esta perspetiva é um passo no sentido de enfrentar um dos grandes enigmas da existência: como algo pode surgir do nada sem intervenção. Um dos papéis da ciência é simplificar a nossa compreensão da Natureza, removendo atributos enganadores. O assombro da complexidade quotidiana é substituído pelo assombro da interligação de uma simplicidade interior. O encanto pelas maravilhas do mundo subsiste, mas é ampliado pela alegria de descobrir a simplicidade subjacente e a sua potência. Assim, é muito mais fácil compreender a Natureza à luz da seleção natural darwiniana do que simplesmente recostar-se e maravilhar-se com a riqueza e a complexidade da biosfera: a sua ideia simples fornece um quadro de entendimento, mesmo que a complexidade que daí emerge seja profunda. O maravilhamento

permanece e é talvez intensificado pelo facto de uma ideia tão simples poder explicar tanto. Einstein simplificou a nossa percepção da gravitação através da generalização da sua teoria especial da relatividade: essa generalização interpretou a gravidade como consequência da curvatura do espaço-tempo pela presença de corpos massivos. A sua «teoria geral» é uma simplificação conceptual, embora as suas equações sejam extraordinariamente difíceis de resolver. Ao eliminar o supérfluo e concentrar-se no essencial, a ciência move-se para uma posição a partir da qual é mais capaz de fornecer respostas. Para o dizer de forma mais direta: ao mostrar que não aconteceu grande coisa na Criação, torna-se mais provável que a ciência consiga resolver o que efetivamente aconteceu.

As palavras evasivas no objetivo que declarei são, evidentemente, o «não muito». Francamente, gostaria de substituir o «não muito» por «absolutamente nada». Ou seja, gostaria que fosse o caso de não ter acontecido absolutamente nada na Criação e de poder justificar essa afirmação. Eliminada a atividade, elimina-se o agente. Se absolutamente nada aconteceu, então a ciência não teria nada para explicar, o que certamente simplificaria a sua tarefa. Poderia até reivindicar retrospectivamente que já tinha sido bem-sucedida! A ciência avançou por vezes ao demonstrar que uma pergunta é destituída de sentido, como quando se pergunta se observadores em movimento podem concordar sobre a simultaneidade de acontecimentos, conduzindo à relatividade especial. Embora não esteja no âmbito da ciência, a questão de quantos anjos podem dançar na cabeça de um alfinete é eliminada se se puder mostrar, de uma forma ou de outra, que os anjos não

existem, ou que, pelo menos, devido a algum defeito fisiológico ou anatómico, são incapazes de dançar. Assim, a eliminação de uma pergunta pode ser uma forma legítima de fornecer uma resposta. Isso poderá ser um passo longe demais — e ser percecionado como uma deserção do dever acadêmico, uma batota, uma típica fuga científica; chame-lhe o que quiser — para que o aceite nesta fase, pelo que me confinarei a defender que «não aconteceu grande coisa» quando o universo passou a existir e, a seu tempo, explicarei quão grande foi esse «não muito».



O objetivo de todo este preâmbulo é mostrar que a evidência que apresentarei a favor de não ter acontecido grande coisa serve para demonstrar que as leis da natureza emergem precisamente disso. Defenderei que pelo menos uma classe de leis naturais tem origem nesse «não muito» ocorrido no primeiro instante. Parece-me que isto constitui uma evidência poderosa a favor da minha posição, pois se a maquinaria do mundo — as leis que governam o comportamento — emerge desta perspectiva, então evita-se a necessidade de postular a elaborada hipótese de um agente legislador, vulgarmente conhecido como um deus. As leis que não emergem da indolência defenderei que nascem da anarquia, da imposição de nenhuma lei. No decurso da elaboração desta exposição, apercebi-me de que a anarquia pode, em alguns casos, ser demasiado restritiva; salvei então o conceito em toda a sua potência, permitindo que a anarquia forme uma aliança com a ignorância. Verá o que quero dizer. Em determinada altura,

chegarei mesmo a invocar a ignorância como uma ferramenta poderosa para alcançar conhecimento.

Devo sublinhar que, nesta exposição, tenho em mente apenas leis físicas, as leis que governam entidades tangíveis — bolas, planetas, coisas em geral, matéria —, bem como radiação intangível, partículas fundamentais, e assim por diante. Deixo de lado a lei moral, que alguns ainda atribuem a um deus promovido a Deus, a suposta fonte inesgotável, incompreensível e livremente fluente de toda a bondade, o árbitro do bem e do mal, o recompensador das ovelhas e o perdoador do bode. A minha posição, para maior clareza, é que os fenómenos biológicos e sociais emergem integralmente da lei física; assim, se escavasse fundo nas minhas convicções, encontraria a ideia de que o apelo à indolência e à anarquia abrange todos os aspetos do comportamento humano. Mas não desenvolverei aqui esse pensamento.

Então, o que são leis da natureza? O que estou a tentar explicar ao descobrir a sua origem? De um modo geral, uma lei da natureza é um resumo da experiência sobre o comportamento das entidades. São sofisticações do folclore, como «o que sobe tem de descer» e «uma panela vigiada nunca ferve». O folclore está quase sempre errado em algum grau. O que sobe não desce se for lançado com velocidade suficiente para entrar em órbita. Panelas vigiadas acabam por ferver. As leis naturais são tipicamente melhorias do folclore, pois foram construídas a partir de observações realizadas em condições controladas, isolando o fenómeno que pretendem explicar de influências externas (a lama do carro de Aristóteles, por exemplo, o ar em torno da sua flecha).

Acredita-se que as leis naturais sejam universalmente válidas no espaço e intemporais. A ubiquidade e talvez a persistência eterna das leis significam que qualquer lei da natureza é considerada válida não apenas de um canto de um laboratório para outro, mas através dos continentes e para além deles, em todo o universo. Talvez falhem em regiões onde os próprios conceitos de espaço e tempo falham, como no interior de buracos negros, mas onde o espaço e o tempo são benignos, uma lei válida aqui e agora é uma lei válida ali e então.

As leis são estabelecidas em laboratórios que ocupam alguns metros cúbicos do universo, mas acredita-se que se apliquem ao universo inteiro. São formuladas num período comparável a uma vida humana, mas acredita-se que se apliquem a algo como a eternidade. Existem fundamentos para estas crenças, mas é necessário cautela ao abraçá-las sem reservas.

À pequena escala da experiência humana direta — uma fração ínfima do tempo durante o qual o universo existiu e uma fração evanescente do seu volume — as leis revelaram-se idênticas onde e quando quer que tenham sido testadas, pelo menos na Terra. À escala maior da experiência humana, foram testadas pela capacidade dos astrónomos de observar fenómenos a distâncias enormes da Terra, noutras galáxias, e igualmente muito atrás no tempo. A menos que as distâncias no espaço e no tempo conspirarem para nos enganar, cancelando de algum modo desvios, não foram detetados desvios relativamente às leis estabelecidas na Terra. Na curta escala temporal de alguns milhares de milhões de anos no passado, não há razão para suspeitar que, por um período

semelhante no futuro, as nossas leis atuais venham a mudar. É claro que pode acontecer que, nos próximos bilhões de anos — ou mesmo à meia-noite de amanhã — dimensões atualmente ocultas, suspeitas, mas ainda escondidas no espaço-tempo, se desenrolem, aumentando o nosso pequeno conjunto de dimensões familiares e transformando as leis bem conhecidas em algo irreconhecível. Não sabemos se isso acontecerá, mas um dia — tal é o poder de uma lei da natureza — poderemos talvez prevê-lo com base nas leis que estamos agora a estabelecer. As leis contêm em si as sementes da sua própria substituição.

Quase todas, mas não todas, as leis são aproximações, mesmo quando se referem a entidades que foram isoladas de influências externas e acidentais (a lama). Permita-me trazer aqui à cena uma figura da história e a primeira das pequenas leis simples que utilizarei para introduzir diversos pontos. (Mais adiante distinguirei entre leis grandes e leis pequenas; esta é uma lei pequena.) O muito inteligente, inventivo e laborioso Robert Hooke propôs uma lei relativa ao alongamento de molas². Como não era invulgar na época, expressou-a sob a forma de um anagrama, de modo a reivindicar prioridade e, ao mesmo tempo, ganhar tempo para explorar as suas consequências sem receio de ser ultrapassado por outros. Assim, em 1660, escreveu de forma críptica e alfabética *ceiinosssttuv*, revelando apenas mais tarde que pretendia dizer *Ut tensio, sic vis*. Na linguagem mais direta de hoje, a sua lei afirma que a força restauradora exercida por uma mola é proporcional ao quanto ela é esticada ou comprimida. Esta lei descreve muito bem o comportamento das molas, não apenas das molas propriamente ditas, mas também das ligações de tipo elástico entre átomos nas

moléculas, e tem consequências fascinantes totalmente insuspeitadas por Hooke ou mesmo pelo seu contemporâneo Newton. No entanto, é apenas uma aproximação, pois se esticar a mola em demasia, a proporcionalidade falha, mesmo que pare antes de a partir: *ceiinnnoosssttuv*. Ainda assim, a lei de Hooke é um bom guia para o comportamento das molas, desde que se tenha presente que é válida apenas para pequenas deformações.

Pode haver algumas leis que sejam exatas. Um candidato é a «lei da conservação da energia», que afirma que a energia não pode ser criada nem destruída: pode ser convertida de uma forma noutra, mas aquilo que temos hoje é, no total, o que teremos sempre e o que sempre tivemos no passado. Tão poderosa é esta lei que pode ser usada para fazer descobertas. Na década de 1920 observou-se que a energia parecia não ser conservada num certo decaimento nuclear, e um conjunto de propostas girou em torno da sugestão de que talvez, em eventos tão novos e até então não estudados, a energia não se conservasse. A visão alternativa, proposta pelo físico teórico austríaco Wolfgang Pauli em 1930, foi a de que a energia é conservada, mas que uma parte teria sido transportada por uma partícula ainda desconhecida. Assim se estimulou a procura e, mais tarde, a deteção bem-sucedida da partícula elementar hoje conhecida como neutrino. Como veremos, a lei da conservação da energia está no cerne da inteligibilidade do universo, no sentido em que está na raiz da causalidade — de que um acontecimento pode causar outro — e, portanto, no coração de toda a explicação. Terá um papel central no que se segue.

Existem outras leis que parecem ter um estatuto semelhante ao da lei de Hooke (isto é, são aproximações e agradáveis de conhecer porque nos ajudam a fazer previsões e a compreender a matéria) e outras que se assemelham à conservação da energia (não são aproximações, mas estão profundamente enraizadas na estrutura da explicação e do entendimento). Isto sugere-me que existem duas classes de leis, que designarei por *leis-raiz* e *leis-derivadas*. As leis-raiz são as leis estruturais muito profundas do universo, a legislação primária, o fundamento do entendimento, o leito rochoso da compreensão. A conservação da energia é, na minha opinião, uma lei-raiz e, embora hesite em dizê-lo, talvez a mãe de todas as leis-raiz. As leis-derivadas são as suas parentes menores, como a de Hooke e outras que em breve surgirão. São legislação secundária, pouco mais do que elaborações das leis-raiz. Não podemos prescindir delas, e em muitos casos a ciência progrediu através da sua descoberta, aplicação e interpretação. Mas são cabos num exército comandado por generais.

Existe um tipo especial de lei que preciso de reconhecer e trazer à sua atenção: uma lei que se aplica a absolutamente nada e que, ainda assim, é muito útil. Preciso de desfazer esta observação desconcertante. Como já disse, as leis-derivadas são tipicamente aproximações. Contudo, em alguns casos, a aproximação torna-se cada vez melhor à medida que o material que a lei pretende descrever se torna cada vez menos abundante. Levando esta progressão de diminuição da abundância ao extremo, a lei torna-se exata (talvez rigorosamente exata) quando a quantidade de material descrita é reduzida a zero. Eis, então, uma chamada «lei

limite», uma lei que atinge exatidão completa no limite de não ter nada para descrever.

É claro que, tal como a apresentei, parece que a lei é vazia, aplicável apenas ao nada. Mas tais leis limite são de enorme utilidade, como verá, pois, na prática raspam a lama do seu próprio funcionamento interno. Permita-me dar um exemplo para esclarecer o que tenho em mente.

O aristocrata anglo-irlandês Robert Boyle, trabalhando por volta de 1660 num barracão mesmo ao lado da High Street, em Oxford (onde hoje se encontra o University College, mas possivelmente em terreno pertencente ao meu próprio colégio, Lincoln), atuando talvez sob a sugestão do seu laborioso assistente Richard Towneley e em colaboração com o já referido e intelectualmente onnipresente Robert Hooke, investigou aquilo que então se designava como a «mola do ar», a sua resistência à compressão. Estabeleceu uma lei da natureza que parecia explicar o comportamento do gás que conhecemos como ar³. Ou seja, descobriu que, para uma dada quantidade de ar, o produto da pressão que exerce pelo volume que ocupa é uma constante. Aumente a pressão e o volume diminui, mas o produto da pressão pelo volume mantém o seu valor inicial. Aumente novamente a pressão e o volume diminui ainda mais; e o produto dos dois conserva o seu valor inicial. Assim, a lei de Boyle (que os franceses chamam lei de Mariotte) afirma que o produto da pressão pelo volume é uma constante para uma dada quantidade de gás e, acrescentaríamos hoje, a uma temperatura fixa.

A lei é, de facto, uma aproximação. Introduza mais gás e a lei é menos bem respeitada. Retire algum e melhora. Retire mais um pouco e melhora ainda mais. Retire quase tudo e torna-se praticamente perfeita. Veja para onde isto aponta: retire tudo e ela é exata. Assim, a lei de Boyle é uma lei limite, exatamente aplicável quando há tão pouco gás presente que ele pode ser considerado ausente.

Há dois pontos que preciso de salientar a este respeito. Primeiro, compreendemos hoje (Boyle não compreendia — não podia compreender, pois isso depende do conhecimento das moléculas, que estava no seu futuro) por que razão a exatidão da lei melhora à medida que a abundância do material diminui. Não entrarei em detalhes, mas essencialmente os desvios à lei resultam das interações entre as moléculas. Quando estas estão muito afastadas umas das outras (como numa amostra que consiste apenas num ténue fio de gás), essas interações são negligenciáveis e as moléculas movem-se caoticamente, de forma independente umas das outras (as palavras «caos» e «gás» têm a mesma raiz: são primas etimológicas). As interações são a lama interna que a redução da quantidade de material elimina, deixando a perfeição limpa do caos.

O segundo ponto é ainda mais importante, embora intimamente relacionado com o primeiro. Uma lei limite identifica a essência da substância, não a lama que os seus pés acumulam ao atravessar a realidade. A lei de Boyle identifica a essência da gasosidade perfeita, a eliminação das interações entre moléculas que confundem a questão nos gases reais da vida quotidiana. Uma lei

limite é o ponto de partida para compreender a natureza da própria substância, livre de qualquer acumulação de pormenores que distraiam e confundam. As leis limite identificam a perfeição do comportamento dos materiais e serão o ponto de partida para várias das nossas investigações.

Outro aspeto inicial importante das leis naturais é que algumas são intrinsecamente matemáticas e outras são adequadamente verbais. Sempre que precisar de mostrar uma equação para fundamentar um ponto, remetê-la-ei para as Notas no final do livro, de modo a disponibilizá-la aos que gostam de ver o mecanismo a funcionar por baixo das palavras. A vantagem de uma lei expressa matematicamente (a de Einstein da relatividade geral, para escolher um exemplo extremo) é que acrescenta precisão ao argumento. No seu lugar, farei todos os esforços para destilar a essência do raciocínio. Pode mesmo argumentar-se que a extração do conteúdo verbal, a interpretação física de uma equação, é uma parte essencial da compreensão do seu significado. Por outras palavras, uma possível posição é que não ver a equação constitui uma forma mais profunda de compreensão.

Nem todas as leis naturais são matemáticas, mas mesmo aquelas que não o são adquirem maior poder quando expressas matematicamente. Uma das questões mais profundas que se podem colocar acerca das leis naturais, para além da sua origem, é por que razão a matemática parece ser uma linguagem tão perfeita para a descrição da Natureza. Por que é que o mundo real dos fenómenos se mapeia tão bem neste produto extremo da mente humana? Explorei esta questão noutros contextos, mas ela é tão

central para a nossa percepção e compreensão do mundo que a retomarei mais adiante (no Capítulo 9). Suspeito que todas as questões verdadeiramente profundas sobre a natureza da realidade física (a única realidade, para além das invenções dos poetas), como a viabilidade da matemática como descrição da Natureza, provavelmente têm respostas ligadas entre si numa fonte comum e precisam de ser consideradas em conjunto.



Por fim, preciso de dizer algumas palavras sobre um conjunto de auxiliares das leis da natureza. Como já disse, uma lei da natureza é um resumo de observações sobre o comportamento das entidades. Há então dois passos na explicação de uma lei. Primeiro, pode ser proposta uma *hipótese*. Uma hipótese (do grego para «fundamento», no sentido de alicerce) é simplesmente um palpite acerca da razão subjacente ao comportamento observado. Esse palpite pode receber apoio de outras observações e, assim, amadurecer gradualmente numa teoria (do grego para «especulação aliada à contemplação»; a palavra partilha a sua origem com «teatro»). Uma teoria é uma hipótese plenamente desenvolvida, com fundamentos talvez embebidos noutros domínios do conhecimento e formulada de modo a poder ser testada por comparação com novas observações. Em muitos casos, a teoria sugere previsões, que são depois submetidas a verificação. Em muitos casos, a teoria é expressa matematicamente e as suas consequências são extraídas por dedução lógica e manipulação (e interpretação) dos símbolos. Se, em qualquer fase, uma hipótese

ou uma teoria entrar em conflito com a observação, volta-se à estaca zero, sendo necessária uma nova hipótese, que depois é desenvolvida numa nova teoria.

Embora tenha apresentado este procedimento — o ciclo de observação seguido de hipótese a amadurecer em teoria, testada contra a experiência — como uma espécie de algoritmo seguido pelos cientistas, a prática é bastante diferente. O método científico é uma pólis liberal e o instinto desempenha um papel importante nas fases iniciais da compreensão. Os cientistas têm palpites, dão saltos intelectuais, cometem certamente erros, apropriam-se de ideias de outros, atrapalham-se, e só muito ocasionalmente veem a luz. Esse é o método científico real, apesar das idealizações dos filósofos da ciência. A idealização deles é como uma lei-limite, identificando a essência do método científico despida da sua lama humana: uma atividade humana praticada no limite da ausência de humanos e das suas fragilidades. O critério central de aceitabilidade, porém, mantém-se dentro do caldeirão de procedimentos: há quase invariavelmente a comparação entre o resultado esperado e a observação experimental. Como Max Planck disse uma vez: «o único meio de conhecimento de que dispomos é a experiência; o resto é especulação».

Em alguns casos há desvios acentuados mesmo em relação a esse percurso ideal. Uma das teorias mais poderosas da ciência é a teoria da evolução por seleção natural de Darwin. Não há matemática na sua formulação desta teoria extraordinária, mas o seu poder é ampliado pela posterior elaboração matemática que ocorreu. Também não tenho a certeza de que a teoria assente numa

lei da natureza. Está certamente baseada na observação — do registo fóssil e da diversidade das espécies —, mas as próprias observações são talvez demasiado diversas para serem resumidas numa lei formulada de modo sucinto, salvo talvez o aforismo de Herbert Spencer, «a sobrevivência do mais apto» (em *Principles of Biology*, 1864), ou, de forma menos controversa e mais incisiva, que «proliferam os organismos que alcançam sucesso reprodutivo nos nichos que lhes estão atualmente disponíveis». Enquanto lei, seria certamente candidata a ser classificada como uma poderosa lei-raiz.



Esse é o pano de fundo do que se segue. Defendi que existem leis-raiz e leis-derivadas: leis de enorme importância e leis menores, respetivamente. A minha primeira tarefa é identificar e examinar algumas de cada tipo e interrogar-me sobre a sua origem. Afirmar que procurarei essa origem na indolência de não ter acontecido grande coisa na Criação e, onde isso falhar, na anarquia que se seguiu. A seu tempo, terei de o levar num voo de especulação para descobrir a fonte da relevância da matemática para a explicação física do mundo real, mas isso está ainda bastante mais adiante no caminho.

CAPÍTULO 2

Muito Barulho por Nada — Como As Leis Podem Emergir Do Nada

O nada é extraordinariamente fecundo. Dentro do compasso infinito do nada reside potencialmente tudo, mas é um tudo que permanece inteiramente por realizar. Observações deste tipo são, naturalmente, deliberadamente enigmáticas, pois nesta fase quero captar a sua atenção e aguçar a sua curiosidade. São talvez aparentadas com a definição irrefutável, mas insatisfatória da filosofia hindu do ser como ausência de não-ser. Para evitar ficar marcado pela mesma insuficiência, preciso de desenvolver esta observação para ilustrar a potência fecunda do nada, mostrar que contemplar o nada não é um exercício de vacuidade — mais próximo da teologia do que da física —, demonstrar que é possível extrair conclusões testáveis a partir do nada no quadro da ciência, e levá-lo a acreditar que o enigma pode ser desembrulhado e o seu conteúdo tornado significativo. Quero mostrar que o nada é a ideia central, o fundamento da possibilidade de compreender as leis da natureza e, portanto, tudo o que é e tudo o que acontece. Em suma, quero mostrar que o nada é o fundamento de tudo.

Para o conduzir à compreensão do nada em absoluto e das suas consequências, e para ajudar a perceber por que razão a indolência

desempenha um papel tão importante no estabelecimento da infraestrutura mecanicista do mundo, gostaria inicialmente que pensasse no nada de uma forma muito primitivamente ligada ao senso comum. Mais tarde, essa forma terá de ceder lugar a algo mais sofisticado, e então o senso comum terá de definhar um pouco; mas, para o início desta viagem, pode pensar com toda a segurança no nada como sendo espaço vazio. Até que lhe diga o contrário, limite-se a recostar-se e a imaginar quilómetros e quilómetros de espaço vazio e uniforme, e anos e anos de tempo a estenderem-se desde um passado distante até um futuro insondável. Pense numa uniformidade intemporal de vazio, em todo o lado e em todo o tempo.

Nesta imagem, semelhante a uma pradaria, de esterilidade espacial e temporal plana, preciso de introduzir uma única figura. A extraordinária e imaginativa matemática alemã Emmy Noether nasceu em Erlangen, ensinou (suportando o peso das atitudes misóginas da época) em Göttingen e, mais tarde, escapou à perseguição nazi para o Bryn Mawr College, na Pensilvânia. Aí morreu demasiado jovem, deixando um legado riquíssimo de conceitos e teoremas matemáticos abstratos. Noether foi chamada (por Norbert Wiener, ele próprio um matemático notável, em 1935) «a maior mulher matemática que alguma vez viveu» e fascinou Albert Einstein. É absolutamente central para o meu argumento — e para a física teórica — em virtude de um teorema que desenvolveu em 1915 e publicou alguns anos depois. Não posso, evidentemente, repetir aqui o seu argumento técnico, mas a sua conclusão é muito simples. Estabeleceu que, sempre que existe uma simetria, existe uma lei de conservação correspondente.¹ Irei desembrulhar esta

afirmação e explicar o que se entende por lei de conservação, por simetria, e a relação que ela estabeleceu entre ambas.

Por «lei de conservação» entendo uma lei da natureza que afirma que, embora os acontecimentos tenham lugar, uma certa quantidade permanece inalterada («é conservada»). Já mencionei uma dessas leis, a conservação da energia, e inicialmente voltarei a concentrar-me nela aqui.



A energia é um daqueles conceitos usados ampla e frequentemente no discurso quotidiano, mas que é muito difícil de fixar e dizer exatamente o que é. Toda a gente compra grandes quantidades dela, mas dificilmente saberia dizer o que comprou. O termo entrou na física no início do século XIX e revelou-se tão valioso que se espalhou por toda a ciência, deslocando argumentos como os de Newton, centrados no conceito mais tangível de força. A sua introdução e o reconhecimento da sua potência chegaram mesmo a levar à reescrita de manuais inteiros. A força é quase literalmente tangível; a energia é abstrata. É daí que vem a sua importância, pois os conceitos abstratos são, em geral, mais amplamente aplicáveis do que os conceitos concretos. Os conceitos abstratos são o esqueleto conceptual ao qual diferentes carnes observacionais podem ser ligadas; os conceitos concretos são ilhas intelectuais.

A etimologia da palavra «energia», a partir dos termos gregos para «trabalho interior», dá uma pista sobre o seu significado. A energia é a capacidade de realizar trabalho. Essa «definição operacional» pode não lhe dar uma compreensão profunda do que a energia realmente é, mas pelo menos permite reconhecê-la: o trabalho é facilmente reconhecível porque nos reconduz ao contato com o conceito tangível de força. Trabalho é o processo de produzir movimento contra uma força oposta, como ao elevar um peso contra a atração da gravidade ou ao usar uma bateria para empurrar uma corrente elétrica através de um circuito. Quanto maior for a energia disponível, maior é a quantidade de trabalho que pode ser realizada. Uma mola enrolada possui mais energia do que quando está desenrolada: quando enrolada pode realizar trabalho; desenrolada não pode. Um tanque de água possui mais energia quando está quente do que quando está frio. Pode conceber-se algum tipo de motor para realizar trabalho usando a energia armazenada na água quente, mas não depois de a água arrefecer.

Existem vários tipos de energia. Entre eles incluem-se a energia cinética, a energia devida ao movimento, como no movimento de uma bola; a energia potencial, a energia que resulta da posição, como a energia devida à atração gravitacional da Terra sobre um peso; e a energia radiante, a energia transportada pela radiação, como o calor que sentimos proveniente do Sol, o seu papel no impulso da fotossíntese e a cascata de consequências que chamamos biosfera². Cada tipo de energia pode ser convertido em qualquer outro. Ainda assim, parece ser uma lei estrita da natureza que *a quantidade total de energia no universo é constante*. Se a

energia se perde numa forma, então tem de estar presente talvez noutra forma ou na mesma forma noutro lugar. Um exemplo familiar desta constância ocorre quando uma bola é lançada ao ar. Imediatamente após ser libertada, possui muita energia cinética. À medida que sobe contra a atração da gravidade, a sua energia potencial aumenta e a sua energia cinética diminui. No ponto mais alto da sua trajetória, está momentaneamente em repouso. A sua energia cinética é então zero e toda a sua energia inicial está agora presente como energia potencial. À medida que cai de novo em direção à Terra, acelerando, a sua energia potencial diminui e a sua energia cinética aumenta. Em cada etapa do seu voo, do início ao fim, a sua energia total — a soma das energias cinética e potencial — permanece constante. A lei da conservação da energia resume esta constância ao afirmar que *a energia não pode ser criada nem destruída*.

No Capítulo 1 referi que a lei da conservação da energia é tão poderosa que tem sido usada para prever a existência de novas partículas fundamentais, ao notar onde aparentemente falhou. O físico teórico dinamarquês, e autor de uma versão inicial da mecânica quântica, Niels Bohr, interrogara-se, ao considerar algumas observações intrigantes, se em certos processos nucleares recentemente investigados ela teria falhado. Mas, na realidade, não falhara, pois, a energia tinha sido transportada por uma partícula até então desconhecida: o neutrino, o «pequeno neutro». A existência do neutrino foi proposta por Wolfgang Pauli em 1930, depois procurada e finalmente detetada numa experiência realizada em 1956³. Este episódio sugere que a lei é semelhante à atitude de David Hume perante os milagres: é mais razoável não acreditar no

relato do que acreditar no acontecimento relatado. Assim, os cientistas tratam qualquer alegação de falha da conservação da energia com extremo ceticismo. Como neste caso, a lei foi questionada, sondada e saiu aprovada. Não é, claro, inconcebível que existam dragões escondidos em regiões inexploradas do cosmos e que, em acontecimentos ainda desconhecidos, a conservação venha a falhar.

A seu tempo — no Capítulo 8, de facto — terei de regressar a este ponto em ligação com esse grande e amplamente mal representado clarificador do pensamento humano, o princípio da incerteza de Heisenberg, que alguns consideram abrir uma brecha, permitindo que a energia flutue em escalas de tempo muito curtas. Numa escala mais ampla e de forma mais convencional, a conservação da energia está subjacente à impossibilidade do movimento perpétuo, a realização de trabalho sem consumo de combustível. Com efeito, um dos componentes da evidência a favor da lei é o fracasso em alcançar o movimento perpétuo, apesar de tentativas cada vez mais desesperadas para o conseguir. Na realidade, apesar das repetidas alegações de charlatães, o movimento perpétuo nunca foi alcançado e é hoje considerado inalcançável. Esta observação — uma certa resignação face à esperança de energia infinita sem esforço e, como consequência, a renúncia à perspectiva de riqueza infinita — é um dos fundamentos de um grande conjunto de leis naturais: as leis da termodinâmica, às quais regressarei no Capítulo 5. Os sucessores desses charlatães, estimulados pela extremidade afiada da cenoura da perspectiva de riquezas abundantes, continuaram, naturalmente, a apresentar máquinas elaboradas que pretendem produzir trabalho a partir do

nada; mas tudo o que conseguiram foi exposição, escárnio e uma maior confiança na lei. Em certa medida, devemos estar-lhes gratos (e certamente aos seus homólogos honestos e laboriosos que se deram ao trabalho de refutar as falsas alegações), pois o fracasso de ataques persistentes, agressivos e determinados à lei conduziu ao reforço da aceitação da sua validade.

Existe ainda muita outra evidência. Todos os cálculos do movimento de partículas baseados na mecânica newtoniana dependem dela e, embora existam desvios às previsões de Newton em certos casos, esses desvios devem-se a outras causas bem conhecidas. Os cálculos da mecânica quântica também dependem da sua validade e são invariavelmente bem-sucedidos no sentido de fornecerem um acordo muito preciso com a observação. Há pouca margem para dúvida de que a energia é conservada, e conservada de forma exata.

A conservação da energia, embora seja claramente de enorme importância tecnológica e económica, bem como para a resolução de problemas em manuais, é na verdade ainda mais importante do que pode parecer. É o fundamento da «causalidade», a observação aparentemente inegável de que todo o acontecimento é causado por um acontecimento precedente. Sem causalidade, os acontecimentos seriam caprichosos e o universo seria um amontoado de ocorrências desconexas. Com a causalidade, existe a possibilidade de compreensão, ao seguir a causa até ao efeito e o efeito até à causa. A causalidade traz a perspectiva de descobrir ordem e comportamento sistemático. Esse comportamento sistemático é expresso pelas leis da natureza e, por conseguinte,

permite o surgimento da forma de compreensão corporizada na ciência. A conservação da energia desempenha um papel central na causalidade ao impor uma restrição poderosa ao que pode acontecer: a energia tem de ser conservada em qualquer acontecimento. A conservação da energia é como uma força policial séria, vigilante e incorruptível, que proíbe qualquer desvio à lei que constrange a energia a um único valor fixo e, presumivelmente, cosmicamente imutável. Se a energia não fosse conservada, haveria menos restrições às ações causadas por um acontecimento anterior e bem poderíamos abdicar da causalidade. Sim, existem outras restrições, mas a energia é tão central ao comportamento e tão universalmente aplicável que a sua conservação é de importância primordial. Como observei no Capítulo 1, é um imperador entre as leis, a mãe de todas as leis-raiz.



Então, por que razão a energia é conservada? Qual é a origem desta mais imperial das leis? É aqui que entra Emmy Noether, iluminando com o seu brilhante teorema a esterilidade vazia que o convidei a contemplar. O ponto central, segundo a sua demonstração de que a conservação de algo decorre da simetria de algo relacionado, é que o caso específico *da conservação da energia* — o nosso foco atual — decorre da *uniformidade do tempo*. Essa uniformidade é o aspeto de «simetria» do teorema de Noether neste contexto.

O que significa, na prática, essa uniformidade? À superfície, a uniformidade temporal significa simplesmente que se obtém o mesmo resultado se realizar a mesma experiência à segunda-feira ou à quinta-feira (ou quando quer que seja). Isto é, a oscilação de um pêndulo ou o voo de uma bola são os mesmos, desde que todas as condições sejam idênticas, independentemente do momento em que são observados. Para exprimir essa uniformidade — essa independência relativamente ao momento de aplicação — diz-se que as leis da natureza são temporalmente invariantes. Na prática, a invariância temporal significa que, se existe uma equação que descreve um processo num determinado instante, a mesma equação se aplica em qualquer outro instante. Por outras palavras, as leis da natureza não mudam com o tempo. As *consequências* dessas leis podem mudar — um planeta pode ter derivado para uma órbita ligeiramente diferente, ou pode ter lançado a bola mais depressa do que pretendia —, mas as próprias leis permanecem invariantes.

Olhemos agora por baixo dessa superfície. Para que as leis da natureza sejam temporalmente invariantes, *o próprio tempo tem de fluir de forma uniforme*. Isto é, o tempo não pode abrandar, depois acelerar, e logo a seguir quase parar. Pense no que isso significaria para o voo de uma bola ou, numa escala maior, para a órbita de um planeta, com o tempo comprimido numa parte da trajetória e dilatado noutra: é difícil imaginar que pudesse alguma vez ter sido construída uma teoria da dinâmica do seu movimento. A bola pareceria acelerar, desacelerar, suspender-se no ar sem que forças aparentes atuassem. À segunda-feira haveria uma lei, à terça-feira outra. Mesmo que o tempo não fosse aleatoriamente

caprichoso, mas ondulasse regularmente, esticando-se e encolhendo-se de forma periódica, o voo da bola seria desconcertante e dificilmente decifrado, mesmo por um Isaac Newton, e o mundo permaneceria provavelmente um lugar dinamicamente confuso. Para que as leis da natureza sejam independentes do momento em que são aplicadas, o tempo tem de fluir de modo uniforme: tic, tic, tic... continuamente, com um ritmo firme e ininterrupto.

Posso antecipar alguns dos argumentos que poderá avançar para minar esta justificação da uniformidade do tempo. Um deles poderia ser que os nossos instrumentos de medição também se esticam e encolhem em sincronia com o tempo experimentado pela bola. Nesse caso, talvez não notássemos — ou não pudéssemos notar — a sua não-uniformidade, e, de algum modo, a física dos nossos instrumentos de medição (incluindo os nossos olhos e ouvidos) variaria sincronamente, tornando-nos cegos e surdos a essa variação. Penso que uma resposta a esta objeção é que as equações que escrevemos e resolvemos para descrever o movimento não são entidades que se estiquem e encolham (no sentido em que o «tempo» presente nelas como parâmetro não varia dessa forma) enquanto as observamos; nesse sentido, constituem uma descrição objetiva, não subjetiva, do movimento. De facto, embora o inverso do teorema de Noether (que implicaria que, se existe uma propriedade conservada, então tem de existir uma simetria relacionada) não seja tão seguro como a sua versão «direta» (que, se existe uma simetria, então existe uma propriedade conservada correspondente), um argumento adicional é que, como sabemos

que a energia é conservada, podemos, com cautela, inferir que o tempo tem de ser uniforme.

Poderá também objetar que, quando Albert Einstein subiu aos ombros de Newton, a sua visão do cosmos revelou um em que o tempo é localmente distorcido (isso é o conteúdo da relatividade geral: o reconhecimento da distorção do espaço-tempo pela presença de corpos massivos, como os planetas). Assim, o tempo não é localmente uniforme e, portanto, o teorema de Noether nada diz sobre a conservação local da energia. Trata-se de uma objeção importante, e está em boa companhia ao levantá-la. Foi, ao que parece, a sugestão do extraordinariamente perspicaz e influente matemático alemão David Hilbert — de que esta objeção precisava de ser considerada — que levou Noether a procurar a sua demonstração. Ela acabou por formular um teorema suplementar («o segundo teorema de Noether») para abordar o problema. Irei contornar a objeção apresentando duas desculpas (o que é sempre um procedimento suspeito, na ciência como na vida, pois tenho de admitir que duas desculpas não fazem uma explicação).

Em primeiro lugar, e tal como no seu teorema original, limitarei a aplicação do teorema de Noether ao global, ao universo como um todo. Embora, quando a matéria se formou e se condensou em planetas, sistemas solares e galáxias, tenha distorcido o espaço-tempo à sua volta, no conjunto, globalmente, existe uniformidade, com alongamentos aqui compensados por encurtamentos noutros locais. Visto como um todo, o espaço-tempo — e a sua componente temporal — é quase certamente plano. Em segundo lugar, qualquer

região suficientemente pequena do espaço-tempo é localmente plana, e a conservação da energia aplica-se-lhe⁴.

Espero que agora aceite, com cautela, que o tempo é uniforme à escala global (e localmente, em domínios suficientemente pequenos) e que, como consequência — de acordo com o primeiro teorema de Noether —, a energia é conservada. Como observei, se pudéssemos ouvir a passagem do tempo, então tic, tic, tic... ocorreria para sempre. Se, pelo contrário, o tempo decorresse tic tic...tic...tic tic, e assim sucessivamente, o tempo não seria uniforme e, como consequência, a energia não seria conservada, o mundo tornar-se-ia incompreensível e a ciência, inútil.



Mas por que razão é o tempo uniforme? É aqui que, pela primeira vez neste capítulo, invoco a indolência e a minha sugestão provisória de que não aconteceu grande coisa na Criação. Preciso de o levar de volta ao momento de início do universo, o momento da cosmogénese, mas antes de o fazer há vários pontos de arrumação doméstica que preciso de mencionar, simplesmente para que fiquem de lado antes que os levante.

Em primeiro lugar, é concebível que este universo seja filho de um universo precedente, que poderá, por sua vez, ser filho de um universo-avó, e assim sucessivamente, recuando para o abismo do tempo. No entanto, presumivelmente, algures no passado distante existiu um primeiro universo — chamemos-lhe o Ur-universo — que

emergiu do absolutamente nada. Este universo atual poderá ser esse Ur-universo e ter ele próprio surgido do nada (e poderá vir a ter, e talvez já tenha tido, descendência). O meu foco é o Ur-universo, quer seja este ou um antepassado deste universo. O ponto essencial é que, em algum momento, parece ter de ter ocorrido um acontecimento em que o nada se transformou em algo, mesmo que esse acontecimento tenha ocorrido há várias gerações de universos. Mesmo que esse número de gerações seja infinito, continua a ser possível que o Ur-universo tenha emergido há um tempo finito⁵. Descarto aqui essa possibilidade, não tendo evidência nem num sentido nem noutro; o meu instinto está silencioso quanto a esta questão — e, além disso, é irrelevante. Nem sequer importa realmente para a presente exposição.

Em segundo lugar, posso estar a ser simplista ao extrapolar o tempo tão para trás. Pode ser que o tempo seja um grande círculo do ser, curvando-se sobre si próprio de tal modo que, tal como a superfície da Terra, não tenha um começo. Muito no futuro poderemos dar por nós muito no passado, com o presente — talvez numa forma modificada — ainda por vir. Temos atualmente experiência de alguns milhares de milhões de anos e formámos a ideia de que o tempo migra essencialmente para a frente em linha reta. Não temos qualquer evidência a favor ou contra a possibilidade de essa linha reta ser apenas um pequeno fragmento de um vasto círculo. Está hoje na moda ridicularizar os defensores da Terra plana, cegos à curvatura da Terra. No devido tempo, nós, os tempos-planos, poderemos ser alvo de derisão semelhante. Em suma, posso estar a ser ingénuo ao supor que ocorreu um começo, em qualquer sentido. Também aqui o meu instinto irrelevante

permanece silencioso, e tudo o que posso fazer é assinalar a possibilidade e pô-la de lado. A possibilidade, contudo, existe, embora não haja evidência para ela, e poderá revelar-se um exemplo do que referi no Capítulo 1: a capacidade da ciência para progredir demonstrando que uma pergunta é destituída de sentido. Neste caso, se o tempo for circular, então não teria havido um começo identificável — pois onde começa um círculo? Nessa situação, a ciência poderá reclamar algo como uma vitória pírrica, pois poderá afirmar que eliminou o enigma do que aconteceu no início de tudo, precisamente porque não houve início. Dito de outro modo: o seu começo está no seu fim.

O tempo circular não é o único problema possível. Em tempos extremamente curtos após o início do universo — Ur ou não — e também numa análise muito próxima do presente, o próprio conceito de tempo poderá desmoronar-se. Há várias formas pelas quais isso poderia acontecer. Uma possibilidade é que, a distâncias extremamente pequenas, à chamada escala de Planck, a presumida e familiar suavidade do espaço se quebre e a distinção entre espaço e tempo se perca⁶. A física contemporânea perde-se igualmente, e ninguém tem ainda qualquer ideia de como lidar com este cenário. Nessa escala, o espaço-tempo deixa de ser um meio suave e contínuo, semelhante a um fluido, passando a assemelhar-se mais a uma caixa de areia ou a uma espuma. Este é também um problema que tenho de pôr de lado.



Na ausência de qualquer perspectiva convincente em contrário, suponhamos que houve um começo e concentremo-nos no momento em que o nada se tornou algo — esse foco de tanta especulação filosófica, mitológica e teológica (e aqui tenho de admitir, de especulação quase científica). Até agora, encorajei-o a pensar o nada como sendo apenas espaço vazio e tempo vazio, em suma, como espaço-tempo vazio. Preciso agora de o afastar dessa visão primitiva. A partir deste ponto, por nada entenderei absolutamente nada. Entenderei menos do que espaço vazio, menos do que um vácuo. Se quiser, a ausência de ser da filosofia hindu. Para sublinhar a absoluta vacuidade do nada que quero que tenha em mente, chamar-lhe-ei Nada. Este Nada não tem espaço nem tempo. Este Nada é verdadeiramente nada absoluto. Um vazio desprovido de espaço e de tempo. Vacuidade total. Vazio para além do vazio. Tudo o que tem é um nome⁷.

No momento inicial do universo (especificamente, do Ur-universo, mas deixarei de usar esse termo por simplicidade), este Nada rolou para algo e o nosso universo incipiente ficou dotado de espaço e de tempo. A consequência imediata desse rolamento é aquilo a que chamamos o Big Bang, mas quero, nesta fase, evitar a impressão de explosividade e gostaria de considerar o Big Bang como um acontecimento que ocorre depois desse rolamento. O rolamento acabou por possibilitar o Bang em certo sentido. A ciência está atualmente silenciosa quanto ao mecanismo que levou o Nada a rolar para algo, e poderá permanecer assim para sempre, embora tenham existido especulações. Aqueles de inclinação mais religiosa ou mesmo poeticamente secular poderão talvez contentar-se com a visão de um criador, ele próprio fora do Nada, a encostar o ombro

ao Nada, talvez a esbarrar nele acidentalmente, fazendo-o rolar (e talvez, se acidental, agora horrorizado com as consequências); mas esse não é o caminho da ciência.

Concentremo-nos no rolamento e deixemos de lado a questão de como ocorreu, remetendo essa importante pergunta para outro dia⁸. A minha especulação, como sabe, é que não aconteceu grande coisa quando o Nada se tornou algo. Pode provavelmente aceitar que o Nada é absolutamente uniforme: não pode ter saliências, lacunas, fossas, alongamentos ou compressões, pois de outro modo não seria Nada. Assim, quando o universo entra em existência, com não acontecendo grande coisa, parece-me plausível que essa uniformidade do Nada seja preservada. Portanto, o espaço e o tempo, quando emergem desse acontecimento, são uniformes. O tempo, em particular, é uniforme e, portanto (diz Noether), a energia é conservada. Deste modo emerge esta lei primária da natureza, seguida pela causalidade, pela ciência e pela compreensibilidade do mundo físico. Abre-se também a perspectiva da emergência, a seu tempo, da agricultura, da guerra, do seu microcosmo — o desporto — e dessas excitações e estímulos dos sentidos e do intelecto: a literatura, a música e as artes visuais. Esse “não muito” foi maravilhosamente fecundo.

A visão de que “não aconteceu grande coisa” no início do universo é, reconhecidamente, uma hipótese, uma especulação, que parece ir contra o grão da realidade. Mas as hipóteses ganham validade quando se verifica que têm consequências em concordância com a observação. A hipótese de que “não aconteceu grande coisa” conduziu, neste caso, a uma observação verificada,

pelo que talvez não esteja incorreta. Ser não incorreta, contudo, está longe de garantir que seja correta, pois alternativas podem conduzir à mesma conclusão. A ciência é severa: uma única consequência em conflito com a observação basta para condenar uma hipótese ao ferro-velho crescente das ideias descartadas da história; uma única consequência verificada é apenas um incentivo para prosseguir a investigação, não uma garantia de validade. As hipóteses amadurecem em teorias à medida que as suas implicações — e, nalguns casos, previsões — se multiplicam e se conformam à observação; mas mesmo uma teoria que tenha sobrevivido até à tranquila meia-idade pode ser condenada por uma única consequência falsa.

Todas as ideias na ciência vivem no limite. Uma consequência bem-sucedida é certamente, talvez apenas, satisfatória; mas haverá outras consequências da minha especulação que sejam apoiadas pela observação e que ajudem a mantê-la, pelo menos temporariamente, a salvo do corredor da morte da ciência?



Usei o espaço para discutir o tempo; agora tomarei o tempo para discutir o espaço. O espaço que é a nossa casa e o nosso recreio de ação surgiu no momento inicial do universo. No mesmo espírito do argumento que acabei de apresentar, à medida que o Nada rolou para se tornar algo, a uniformidade desse Nada foi herdada pelo espaço recém-nascido que emergiu. A simetria invocada na intuição

de Emmy Noether sobre a ligação entre simetria e conservação é, neste caso, a uniformidade herdada do espaço.

O que se entende por uniformidade do espaço? No mesmo espírito da interpretação da uniformidade do tempo que referi anteriormente, a interpretação da uniformidade do espaço significa que uma experiência realizada aqui terá o mesmo resultado que uma experiência realizada noutro lugar. Experiências efetuadas em diferentes laboratórios terão o mesmo resultado. As leis da natureza não dependem de onde se está. As consequências dessas leis podem ser diferentes, porque as condições podem não ser idênticas, mas as próprias leis são as mesmas. Assim, embora a lei que rege a oscilação de um pêndulo seja a mesma, o mesmo pêndulo oscila com períodos diferentes ao nível do mar e no cimo de uma montanha, onde a gravidade é mais fraca. Se se deslocar de um local para outro, não tem de alterar a equação que exprime a lei. As leis são espacialmente uniformes.

Como não existe uma distinção intrínseca entre espaço e tempo (de acordo com a relatividade, são duas faces do espaço-tempo), quaisquer observações que se apliquem ao tempo aplicam-se também ao espaço. Tal como no caso do tempo, para que uma lei não seja diferente de lugar para lugar, o próprio espaço tem de ser uniforme. Isto é, o espaço não está comprimido aqui, dilatado ali, e assim sucessivamente. Tal como aconteceria com uma distorção semelhante do tempo, é difícil imaginar que se pudesse ter construído uma teoria da dinâmica do voo de uma bola se isso não fosse assim. Aplicam-se as mesmas reservas que para o tempo, com a nossa atenção dirigida para a uniformidade global do

universo ou para algum pequeno domínio localmente plano, de modo a podermos usar o primeiro teorema de Noether (o que liga simetria e conservação) em vez de recorrer ao seu segundo teorema (o que diz respeito ao espaço-tempo distorcido). Globalmente, presumo que o espaço seja plano.

O teorema de Noether entra agora em cena no contexto da uniformidade do espaço. De acordo com o seu teorema, uma implicação da uniformidade do espaço é a conservação do momento linear. Preciso de dizer uma ou duas palavras sobre o conceito de momento linear e a sua conservação.

O momento linear é o produto da massa de um corpo pela sua velocidade⁹. Assim, uma pesada bala de canhão a mover-se rapidamente tem um elevado momento linear. Uma bola de ténis leve a mover-se lentamente tem um baixo momento linear. Apesar de serem sinónimos na linguagem corrente, em ciência a velocidade não é o mesmo que a rapidez, pois, além de depender da taxa a que a posição muda (rapidez), inclui a noção de direção. Assim, um corpo que se desloca com rapidez constante, mas muda de direção (por exemplo, um planeta em órbita em torno do Sol) está continuamente a mudar a sua velocidade. Uma bola atingida por um taco pode afastar-se com a mesma rapidez com que chegou, mas a sua velocidade — e, portanto, o seu momento linear — foi invertida. Ao pensar em momento linear, pense sempre em direção, bem como em taxa. Isso complica ligeiramente a noção de conservação do momento linear — a lei segundo a qual o momento linear total permanece inalterado em qualquer acontecimento — porque é necessário ter em conta várias direções de movimento

(questão que não se colocou ao considerar a conservação da energia, onde a direção não desempenha qualquer papel), mas continua a ser bastante fácil de visualizar.

O momento linear é conservado nas colisões de partículas. Um exemplo simples é o de duas bolas de bilhar idênticas a rolar uma na direção da outra com a mesma rapidez. O momento total é zero (porque as velocidades das duas bolas são iguais, mas opostas, somando zero quando adicionadas), e quando se encontram ambas ficam imobilizadas, mantendo-se o seu momento linear total igual a zero. Se se moverem uma na direção da outra formando um ângulo, o seu momento linear total deixa de ser zero. Quando colidem, afastam-se uma da outra. Acontece que as suas novas trajetórias são tais que o seu momento linear total permanece inalterado. Seja qual for o ângulo de colisão, sejam quais forem as suas massas possivelmente diferentes, sejam quais forem as suas velocidades iniciais possivelmente diferentes, e por mais bolas que estejam envolvidas numa colisão, o momento linear total após a colisão é o mesmo que era antes. O momento linear é, de facto, conservado, e é conservado por causa da uniformidade do espaço.

A conservação do momento linear, prevista pela hipótese de que “não aconteceu grande coisa” em combinação com o teorema de Noether, é uma observação verificada e de grande alcance. Ela sustenta todo o sistema de mecânica de Isaac Newton, hoje designado por mecânica clássica. Assim, sustenta as trajetórias dos corpos e as mudanças que ocorrem quando colidem ou se influenciam mutuamente através das forças que exercem; sustenta a pressão exercida por um gás quando as suas moléculas colidem

com as paredes do recipiente.; sustenta a descrição clássica dos planetas, das estrelas e das galáxias. Sustenta também o funcionamento dos motores a jato e dos foguetões.

Uma aliança entre o Nada e a indolência levou-nos muito longe. Forneceu a infraestrutura da causalidade, ao explicar a conservação da energia, e começou a acomodar a emergência da atividade na arena do espaço-tempo ao fornecer uma explicação para os pressupostos subjacentes da física newtoniana. Poderá haver mais? Existirão outras consequências da indolência primordial?



Tenho falado do momento linear, o momento associado a corpos que se deslocam pelo espaço com as suas diversas massas e velocidades. Preciso agora de salientar que existe outro tipo de momento, a saber, o momento angular, o momento associado ao movimento de rotação. A Terra em rotação possui momento angular em virtude do seu movimento diário em torno dos polos, a sua rotação em torno do próprio eixo. Possui também momento angular devido ao seu movimento anual em órbita em torno do Sol. Todos os corpos em rotação possuem momento angular, mesmo que não se desloquem pelo espaço. A Lua possui o seu momento angular mensal ao orbitar a Terra; possui igualmente o seu próprio momento angular de rotação ao girar sobre o seu eixo a uma taxa coincidente (e, assim, apresentar sempre a mesma face à Terra ao longo do mês lunar). Ora, outra lei da natureza bem estabelecida é que, tal como o momento linear, o momento angular é conservado.

Embora possa fazer um corpo rodar — como ao fazer girar uma roda ou ao conseguir pôr uma bola a rodar — e assim dotá-lo de momento angular, algures terá sido produzido um momento angular compensatório. Sempre que pedala para Leste, diminui o momento angular da Terra em torno do seu eixo numa quantidade correspondente, mas totalmente desprezável. Apesar de desprezável, a polícia do momento angular dá por isso. Sempre que pedala para Oeste, aumenta a rotação da Terra e o dia encurta — de forma totalmente impercetível, exceto para a polícia do momento angular, que não tolera transgressões.

Tal como no caso do momento linear, ao avaliar e compreender o significado da conservação do momento angular, a rapidez deve ser considerada em conjunto com a direção. Mas o que se entende por direção do movimento rotacional, se a direção do deslocamento de um planeta, ou de qualquer objeto em rotação, está continuamente a mudar à medida que percorre a sua trajetória? Para atribuir uma direção à rotação, pense no movimento como ocorrendo num plano e, em seguida, associe uma seta no centro do movimento circular, perpendicular ao plano. Se o movimento for no sentido dos ponteiros do relógio quando visto de baixo, a seta aponta para cima, afastando-se do plano. Se o movimento for no sentido dos ponteiros do relógio quando visto de cima, a seta aponta para baixo, entrando no plano. Encare esta convenção como um saca-rolhas comum: à medida que o saca-rolhas roda, avança na direção da seta na nossa convenção — o saca-rolhas enterra-se na rolha quando o roda no sentido dos ponteiros do relógio. Quando se desloca para a frente num automóvel, todas as rodas rodam no sentido dos ponteiros do relógio quando vistas do lado direito do

carro, pelo que cada roda pode ser imaginada como tendo uma seta presa ao cubo e apontando para a esquerda. À medida que o carro acelera, essas setas imaginárias alongam-se proporcionalmente ao aumento da velocidade. Ao travar, encolhem. Se parar e inverter a marcha, emergem e crescem para a direita. Vêm-me à mente imagens de Boadiceia e do seu carro de guerra: ela ceifaria os seus inimigos à esquerda quando avançava e ceifá-los-ia à direita quando recuava.

Há ainda outro ponto que precisa de ter em mente. O momento linear está relacionado com a velocidade através da multiplicação pela massa, uma medida da resistência a alterações do movimento linear, em linha reta. Quanto maior a massa, maior é a inércia à mudança imposta. A resistência a alterações do momento angular está de modo semelhante relacionada com uma grandeza conhecida como momento de inércia, um nome destinado a transmitir não a ideia de uma breve sesta, mas o sentido de resistência ao movimento rotacional, em vez de linear. (O termo “momento” em física é usado para denotar um tipo de influência de alavanca, em vez de uma influência ao longo de uma linha de deslocamento; como ao usar uma chave para apertar uma porca, exercendo um momento de força, uma torção.) Dois corpos podem ter a mesma massa, mas momentos de inércia diferentes. Por exemplo, pense em duas rodas com a mesma massa, mas numa delas o peso está concentrado perto do eixo e na outra está concentrado perto do aro. É mais fácil acelerar a primeira até à rotação do que a segunda, pelo que a primeira tem um momento de inércia menor. Um corpo a rodar rapidamente com um grande momento de inércia (como um volante de inércia) tem um momento angular maior do que um com

um pequeno momento de inércia a rodar à mesma taxa. Os volantes de inércia desempenham a sua função de manter um movimento rotacional suave simplesmente porque, tendo momentos de inércia elevados, são difíceis de parar. A magnitude do momento linear de um corpo em movimento é o produto da sua massa pela taxa a que a sua posição muda ao longo de uma linha reta. Por analogia, a magnitude do momento angular de um objeto em rotação é o produto do seu momento de inércia pela taxa a que está a rodar¹⁰.

A lei da conservação do momento angular exprime a observação experimental de que *o momento angular não pode ser criado nem destruído*. Pode ser transferido de um objeto para outro, como na colisão de duas bolas em rotação ou quando acelera numa bicicleta. Mas o momento angular total de todos os corpos do universo é uma constante (e provavelmente zero). O momento angular é conservado. Se um objeto adquire momento angular através da aplicação de alguma torção aceleradora, então outro objeto ligado perde-o. Se, numa colisão, uma bola sai a rodar, algures existe uma alteração correspondente e compensatória no momento angular de outro objeto (por exemplo, da Terra). Talvez a imagem mais visual seja a de um patinador no gelo em rotação, que gira mais rapidamente à medida que reduz o seu momento de inércia ao recolher os braços, mantendo, contudo, o seu momento angular constante.

E por que razão o momento angular é conservado? Sabe agora que, para compreender a origem de qualquer lei de conservação, é necessário recorrer ao teorema de Noether e procurar a simetria subjacente associada à grandeza conservada. Neste caso, o seu

teorema identifica a simetria relevante como a isotropia do espaço como fonte da lei. Por isotropia (do grego para “mesma rotação”) entende-se a uniformidade do espaço ao percorrer uma volta em torno de um ponto. Pense num ponto, afaste-se um pouco e depois percorra um círculo em torno do ponto original. Se verificar que não há variação no espaço (seja lá o que isso signifique) ao longo do percurso, então ele é isotrópico na vizinhança do ponto escolhido. Uma bola de ténis de mesa é isotrópica; uma bola de golfe também o é, se ignorar as covinhas. Assim, para explicar a conservação do momento angular, é preciso detetar a razão pela qual o espaço possui uniformidade rotacional.

Nesta altura, provavelmente já consegue ver como decorre o argumento. É preciso voltar a pensar no Nada. O Nada é isotrópico. Isto é, o nada absoluto que precedeu o nosso universo (ou o Ur-universo) tinha de ser isotrópico. Se não o fosse, se tivesse saliências e depressões, não seria Nada. Quando o Nada rolou para se tornar algo, não aconteceu grande coisa (segundo a minha hipótese). Assim, a isotropia do Nada foi preservada à medida que o espaço e o tempo emergiram e, conseqüentemente, o espaço que hoje possuímos é isotrópico. A implicação dessa isotropia é que o momento angular é conservado. Eis mais uma lei básica da natureza que emergiu, sem necessidade de ser imposta.

Incidentalmente, há ainda outro aspeto do Nada que vale a pena mencionar. Quando nós (e aqui o “nós” refere-se coletivamente a astrónomos e cosmólogos, os nossos observadores) observamos as galáxias que povoam o universo visível — não apenas as estrelas individuais — verificamos que elas rodam e, portanto, possuem

momento angular. Contudo, as taxas de rotação variam e as orientações das setas que representam as direções de rotação são aparentemente aleatórias. Quando se avalia o momento angular total do universo visível (permitindo que momentos angulares opostos se anulem), verifica-se que o resultado é zero. Globalmente, o universo visível tem momento angular nulo, apesar de os seus componentes individuais rodarem. Isso é exatamente o que se deveria esperar quando o Nada rolou para se tornar aparentemente algo. O Nada (isto é, o Nada; não apenas um nada que se faz passar por Nada no início de uma frase) não possui momento angular, pelo que não é surpreendente que o algo emergente herde o mesmo momento angular. Nenhum momento angular foi criado na Criação; não existe (no total) hoje. O nosso algo atual limitou-se a herdar as propriedades do seu progenitor Nada.



Onde estamos? Espero que esteja claro que a suposição de que não aconteceu grande coisa quando o universo surgiu conduziu a três grandes leis da natureza: a conservação da energia e a conservação de dois tipos de momento — o momento linear e o momento angular. Em cada caso, o nosso universo parece simplesmente ter herdado a uniformidade do seu vazio precursor e manifestar essa herança sob a forma de três leis fundamentais da natureza. As leis não precisaram de ser impostas: são apenas as consequências da herança proveniente de um estado primordial precursor, o estado do Nada absoluto. Existem outras leis de

conservação (a da carga elétrica, por exemplo), e cada uma delas tem uma simetria associada. Para as desenterrar, precisamos de olhar mais profundamente para a própria natureza da Natureza, e regressarei a este ponto num capítulo posterior. Existem também muitas leis pequenas, aquilo a que chamei leis-derivadas, e ainda não explorei a sua origem. O seu tempo há-de chegar. Por agora, é tempo de encontrar espaço para que a indolência descanse e para que a anarquia reine.

CAPÍTULO 3

A Anarquia Reina

— Como Algumas Leis Emergem da Ausência de Leis

Tal como na política, assim na ciência: quando a indolência fica exausta do esforço, a anarquia entra em cena e assume o controle. Uma vez que um universo tenha rolado para a existência, preservando as várias uniformidades características do Nada e dando assim origem às grandes leis de conservação da natureza, algumas das leis que então o povoam e governam são consequências da anarquia. Neste capítulo, defendo que a mecânica clássica desenvolvida no final do século XVII por Isaac Newton (1642–1727) e elaborada com tão grande eficácia nos dois séculos seguintes, bem como a mecânica quântica de Erwin Schrödinger (1887–1961) e de Heisenberg, são manifestações da anarquia. O sistema de leis desenvolvido por Schrödinger e Heisenberg no início do século XX, e posteriormente elaborado por outros, está subjacente à extraordinária formulação anterior de Newton. Essas leis permanecem profundamente desconcertantes, mas são o resultado da Natureza solta, deixada a correr livremente, exceto pela restrição imposta pelas grandes leis de conservação que emergiram da *indolência estrutural primordial (ausência de interferência estrutural)*. Mostrei que a *Indolência Estrutural Primordial* policia a arena; aqui pretendo mostrar que é da *Anarquia*

Operativa Primordial que brotam as leis do comportamento dentro dessa arena.

Pretendo falar tanto de mecânica clássica como de mecânica quântica. Fá-lo-ei descrevendo o seu conteúdo de forma muito geral, apresentando o pequeno conjunto de conceitos centrais de ambas, sem — no caso da mecânica quântica — me deixar enredar na fascinante, mas espinhosa questão da interpretação. De facto, a interpretação é secundária em relação às leis que governam o comportamento das entidades, pois a interpretação não é mais do que uma tentativa de estabelecer uma forma de pensar as consequências das leis em termos de conceitos relacionados com a experiência quotidiana. As leis da natureza, ao contrário das leis que regem as sociedades, podem existir sem interpretação; mas, como já referi, é inerente à ciência procurar interpretar aquilo que as leis implicam e, assim, iluminar-nos acerca da realidade física.

A interpretação da mecânica clássica é fácil porque essa descrição do comportamento está em contato direto com a experiência quotidiana através da nossa percepção e compreensão da posição e da velocidade. A revolução do pensamento associada à mecânica quântica gerou uma perplexidade que permanece por resolver, apesar de as previsões numéricas e observacionais — resultantes da aplicação das suas leis (isto é, simplesmente fazendo os cálculos que elas exigem), algo distinto da sua interpretação — estarem invariavelmente (até agora, pelo menos) de acordo com a observação, e de forma notavelmente precisa. Pode ser que os cérebros humanos estejam mal equipados para se libertarem da herança de observar o movimento de objetos familiares e

simplesmente não consigam lidar com as contraintuitividades da mecânica quântica. Temos cérebros construídos — evoluídos por razões de sobrevivência — para lidar com a savana e a selva, e pode haver razões estruturais neurológicas fundamentais pelas quais não conseguimos compreender a mecânica quântica. Um dia, no futuro, suponho que um computador quântico possa afirmar compreender os seus próprios princípios, deixando-nos ainda perplexos, apesar de termos sido nós a construí-lo. A dificuldade, talvez impossibilidade intrínseca, de compreender a mecânica quântica é, para o meu propósito, irrelevante. O que me interessa são as próprias leis, não a dificuldade de um ser humano em lidar com o que elas significam.

O meu objetivo, então, é simplesmente mostrar que um pilar central da mecânica quântica emerge de forma muito natural da anarquia e que esse pilar é um trampolim para a emergência da mecânica clássica de Newton de um modo muito direto. Para o fazer, preciso primeiro de o levar por aquilo que poderá parecer um desvio rural, mas que acabará por se revelar uma autoestrada para a compreensão. Esse desvio irá apresentar-lhe uma lei-derivada que conduzirá à identificação de uma lei-raiz.



Uma pequena lei da natureza, uma lei-derivada na minha classificação, é a de que *a luz se propaga em linhas retas*. Associada a esse comportamento está a lei-derivada consequente que rege a reflexão em espelhos, segundo a qual o ângulo de reflexão é igual

ao ângulo de incidência. Isto é, um espelho plano reflete o mundo com fidelidade: não distorce a imagem nem a torna ininteligível ou, pelo menos, enganadora. Poderá também recordar a lei de Snell, descoberta e redescoberta por várias pessoas, incluindo, em 1621, o astrónomo neerlandês Willebrord Snellius (1580–1626), relativa ao ângulo de refração quando um raio de luz se desvia na interface entre o ar e a água (ou qualquer outro meio transparente), dependendo desse ângulo dos valores relativos dos índices de refração do ar e do meio (voltarei a isto mais adiante).¹ Qual é a origem deste comportamento?

Em cada caso — e a um nível de integração e de compreensão ligeiramente mais profundo, que combina tanto as leis da reflexão como as da refração — a lei geral é que *a luz adota o percurso entre a sua fonte e o seu destino que demora menos tempo*. (Não pergunte ainda porquê: isso virá mais tarde.) Este é o princípio do menor tempo, proposto e explorado pelo matemático francês Pierre de Fermat (1601?–1665), embora possa ser recuado até ao matemático grego, inventor e praticante precoce do método científico — entendido como a aliança entre experiência e teoria —, Herão de Alexandria (c. 10–c. 70 d.C.), por volta de 60 d.C. Pode também ser atribuído ao matemático e astrónomo árabe, igualmente pioneiro do método científico, Ali-Hassan ibn al-Hasan ibn al-Haytham (mais comodamente, evitando a enumeração da ascendência e na forma latinizada, Alhazen; c. 965–c. 1040), em 1021, no seu tratado de óptica.

De passagem, gostaria que tivesse presente — pois isso tornar-se-á relevante em breve — uma proposta semelhante, o princípio

da menor ação, onde “ação” tem um significado técnico que esclarecerei mais adiante, mas que por agora pode tomar ao pé da letra como “esforço”. Este princípio foi proposto pelo filósofo francês e pensador polivalente Pierre Maupertuis (1698–1759), no início da década de 1740, e é relevante para a propagação de partículas (o que inclui tudo, desde uma ervilha até um planeta). Segundo ele, *as partículas adotam trajetórias que minimizam a ação associada ao percurso entre dois pontos fixos*. Uma bola em voo descreve, depois de abandonar o taco, uma trajetória que minimiza a ação. A Terra encontrou uma órbita que, dada a sua velocidade orbital e a sua distância ao Sol, envolve a menor ação. Os outros planetas e os fragmentos de detritos cósmicos também o fizeram, cada um à sua maneira característica. Um projétil em voo percorre uma trajetória que envolve a menor ação entre a origem e o alvo. Qualquer outro percurso corresponderia a uma ação maior, a mais esforço. É como se o princípio da indolência estivesse novamente em funcionamento, com a menor ação a corresponder à maior preguiça; mas mostrarei que aqui opera um princípio diferente daquele que esta alusão antropomórfica poderia sugerir.

Há ainda uma sugestão profunda latente nestes dois princípios — um aplicável à luz e o outro às partículas: talvez, por terem uma forma semelhante, tenham uma origem comum e possam ser unificados; voltarei a desenvolver esta ideia mais adiante. Um meta-princípio da ciência é, talvez, o de que as coincidências na Natureza são sempre dignas de exploração, pois podem brotar de analogias estruturais e apontar para relações profundas. As coincidências, sempre suspeitas, quando perseguidas podem fornecer compreensão. Os primeiros proponentes destes princípios

não estavam totalmente certos se o menor tempo e a menor ação eram obrigações morais impostas por uma Natureza austeramente mandatária. São, naturalmente, princípios físicos e ambos emergem — como defenderei — da Anarquia Operativa Primordial, essa que é a menos moralmente constrangida de todas as ordens políticas.



Voltemos a concentrar-nos no espelho. É um exercício relativamente simples de geometria mostrar que a menor distância entre a fonte e o olho, passando por um espelho plano — assumindo que ambos os troços do percurso são linhas retas — corresponde à situação em que o ângulo de incidência é igual ao ângulo de reflexão. Esse percurso é também o que demora menos tempo a ser percorrido pela luz, que se propaga, como sabemos, a velocidade constante. Assim, de acordo com o princípio do menor tempo, o trajeto seguido pela luz quando se reflete num espelho é tal que o ângulo de reflexão é igual ao ângulo de incidência.

Passemos agora a Snell e à sua lei da refração. A luz desloca-se mais lentamente no vidro ou na água (ou em qualquer outro meio transparente denso) do que no vácuo ou no ar. A razão entre a sua velocidade no vácuo e a sua velocidade no meio chama-se índice de refração do meio. No caso da água, o índice de refração é cerca de 1,3, o que significa que a luz viaja aproximadamente 30% mais devagar na água do que no vácuo (ou no ar, que para estes efeitos é praticamente equivalente). Provavelmente consegue atravessar

água relativamente profunda a uma velocidade cerca de dez vezes inferior àquela a que caminha no ar; se atribuíssemos à água um índice de refração para o seu próprio movimento, esse valor seria da ordem de 10.

Considere agora qual será o trajeto correspondente ao menor tempo de percurso quando a luz parte de uma fonte no ar e tem de atingir um destino debaixo de água. Um problema análogo é o de chegar o mais rapidamente possível a alguém que se está a afogar num lago. Para manter a imagem simples, fiquemo-nos pelo caminhar e pelo vadear. Poderia seguir uma linha reta desde a espreguiçadeira até à vítima, o que implicaria algum percurso a caminhar e depois a vadear. Em alternativa, poderia reduzir o tempo passado a vadear ao mínimo, caminhando até um ponto diretamente em frente da vítima e entrando na água apenas no último instante. O problema é que, nesse caso, o tempo total a caminhar aumentou, porque o percurso em terra é mais longo, e a breve travessia a vau não compensa totalmente esse tempo adicional. Como provavelmente suspeita, existe um trajeto intermédio que minimiza o tempo total: caminhar obliquamente pela terra e depois vadear também obliquamente pela água até à vítima. O ângulo exato com que terá de mudar de direção ao entrar na água depende das suas velocidades relativas nos dois meios e, portanto, dos respetivos índices de refração. De facto, como um pouco de raciocínio geométrico permite mostrar, os ângulos que tem de adotar ao correr pela terra e depois ao vadear pela água — e que a luz, por analogia, adota ao passar de um meio para outro — são dados pela lei da refração de Snell.² Da lei de Snell decorrem então todas as propriedades dos prismas e das lentes, que

dependem da refração para o seu funcionamento, e daí toda a área conhecida como óptica geométrica. Esta designação deve-se ao facto de tratar os percursos da luz como uma sucessão de linhas retas que fazem entre si os ângulos apropriados: o trajeto global é construído geometricamente.

Neste ponto do argumento, apresentei-lhe duas leis-derivadas, a da reflexão e a da refração, e mostrei que ambas são manifestações de uma lei mais profunda: a de que a luz adota o percurso de menor tempo entre a fonte e o destino. Mas por que razão a luz se comporta assim? Qual é a origem desta lei fundamental da óptica geométrica? E como “sabe” a luz, aparentemente de antemão, qual é o percurso correspondente ao menor tempo de propagação? Se começasse por seguir uma direcção que afinal se revelasse errada, deveria voltar atrás e recomeçar (o que acrescentaria ainda mais tempo) ou simplesmente avançar e esperar pelo melhor? Qual é a origem desta sua aparente presciência?



A origem da lei do menor tempo é a anarquia. Para perceber como isso funciona, precisamos de partir do facto de que a luz é radiação eletromagnética: uma onda de campos elétricos e magnéticos oscilantes que se propaga àquilo a que, não surpreendentemente, se chama a “velocidade da luz”. Para evitar esta aparente tautologia, penso que essa velocidade deveria chamar-se algo como “velocidade de Maxwell”, em homenagem a

James Clerk Maxwell (1831–1879), cuja vida foi curta, mas que é um dos imortais da ciência — o físico escocês do século XIX que identificou pela primeira vez a natureza eletromagnética da luz.

Para visualizar a luz, pense numa onda como um comboio de cristas e vales a avançar pelo espaço, de tal modo que cada crista se desloca à velocidade da luz. A distância entre cristas vizinhas é o comprimento de onda da luz. As alturas das cristas e as profundidades dos vales indicam a intensidade do campo elétrico em cada ponto e, portanto, o brilho da luz: uma crista indica que o campo está orientado para cima relativamente à direção de propagação; um vale indica que o campo mudou de direção e aponta para baixo. Se pudesse estender um dedo e sentir a passagem do raio, sentiria uma sequência de impulsos de campo elétrico a mudar rapidamente de direção à medida que a luz passasse. A frequência com que esses impulsos mudam de direção corresponde à cor da luz: se mudarem relativamente devagar (ainda assim muito rapidamente), a luz seria vermelha; se mudarem mais depressa, a luz seria amarela, azul ou mesmo violeta. A luz branca é uma mistura de raios de todas as cores, do vermelho ao violeta. O infravermelho (abaixo do vermelho) e o ultravioleta (para além do violeta) dispensam explicações. Se abrandar a alternância dos impulsos, obtém ondas de rádio; se acelerar, obtém raios X. Um ponto final importante: como todas as cores da luz viajam exatamente à mesma velocidade, quanto mais rapidamente os impulsos alternam (maior frequência), menor é a separação entre cristas (menor comprimento de onda). Assim, a luz azul tem um comprimento de onda menor do que a vermelha, e os raios X ainda menor. O comprimento de onda da luz visível é cerca

de meio milésimo de milímetro, o que — apesar das queixas dos pessimistas que o consideram “inimaginavelmente pequeno” — talvez esteja quase ao alcance da imaginação.

Com esta descrição e a Anarquia Operativa Primordial em mente, considere uma onda de uma única cor que parte de uma fonte e segue um percurso inteiramente arbitrário até um destino (o seu olho, por exemplo): talvez se enrole, recue sobre si própria e depois entre diretamente no olho. A onda é uma sequência de cristas e vales ao longo desse trajeto caprichoso; suponha que, ao entrar no olho, coincide com uma crista. Considere agora um percurso que se assemelha muito ao primeiro, mas não é exatamente o mesmo — uma curvatura ligeiramente diferente, uma aproximação um pouco mais tardia ao olho. A onda é a mesma sequência de cristas e vales, mas ao chegar ao olho pode coincidir não com uma crista, mas com um vale, ou algo próximo de um vale. Esse vale anula a crista do primeiro percurso (os campos elétricos correspondentes apontam em direções opostas e cancelam-se). O cancelamento pode não ser perfeito, mas imagine um grande número de percursos vizinhos do primeiro, cada um com um comprimento ligeiramente diferente. No conjunto, as cristas e vales dessa multidão de vizinhos anulam-se no olho. Por outras palavras: embora esses percursos sejam permitidos, não verá luz que tenha viajado por eles. A anarquia anulou-se a si própria.

Considere agora um percurso retilíneo entre a fonte e o olho. Mais uma vez, a onda que o percorre pode chegar ao olho como uma crista. Considere depois um percurso vizinho, quase, mas não exatamente igual ao retilíneo. Ele chegará como quase uma crista,

porque a distância percorrida é quase a mesma. Em particular, por os percursos serem tão semelhantes, não chega como um vale que cancele a crista do primeiro. Há muitos outros percursos muito próximos do retilíneo, e todos chegam como quase cristas. Coexistem: não se anulam mutuamente. Em suma, verá luz a viajar por esses percursos quase retos. A anarquia deixou uma brecha.

Pode parecer que o argumento foi preparado de forma tendenciosa, pedindo-lhe que aceite que percursos sinuosos têm vizinhos destrutivos e os retos não. É verdade; mas esta distinção assenta em resultados matemáticos rigorosos. Sei que isto soa a “confie em mim”, mas trata-se de um resultado padrão da óptica física (ou óptica ondulatória), a versão da óptica que reconhece a natureza ondulatória da luz em vez de tratar de imediato os seus percursos como geometricamente retos. As Notas esboçam o essencial do argumento relevante.³

Deverá agora ser claro onde a discussão nos trouxe. Nenhuma lei foi imposta a reger a propagação da luz (a luz pode seguir qualquer percurso entre a fonte e o destino), mas o resultado desse comportamento *laissez-faire* (não-intervenção) é uma lei: a luz propaga-se por percursos que minimizam o tempo de voo. A luz não sabe antecipadamente qual será o percurso de menor tempo; não tem presciência. Experimenta todos os percursos simultaneamente, e verifica-se que apenas o percurso de menor tempo sobrevive à destruição pelos seus vizinhos. Da anarquia emerge a lei.

Até aqui falei apenas da luz a propagar-se num meio uniforme, como o vácuo ou (com boa aproximação) o ar. Como emergem a

lei de Snell e, na verdade, toda a óptica das lentes em câmaras e microscópios, onde a luz é desviada de várias formas ao atravessar sucessões de meios transparentes com diferentes formas? Nestes casos é necessário considerar o índice de refração do meio e o seu efeito no comboio de cristas e vales. A luz desloca-se mais lentamente num meio denso; assim, embora a direção do campo elétrico alterne à mesma taxa (a luz não muda de cor ao entrar na água ou no vidro), as cristas e os vales ficam mais próximos: o comprimento de onda é menor. Essa modificação do espaçamento acaba por alterar qual o percurso que tem vizinhos não destrutivos. O percurso sobrevivente já não é a linha reta entre fonte e destino, mas um trajeto que é retilíneo num meio e depois se encurva para outra reta no segundo meio. Seguindo a análise até ao fim, verifica-se que o percurso sobrevivente — aquele cujos vizinhos não o anulam — é exatamente o que corresponde ao menor tempo de passagem e é dado pela lei de Snell. A anarquia é o fundamento de toda a óptica.

Preciso de intercalar uma observação importante. A anulação dos percursos vizinhos é tanto mais eficaz quanto menor for o comprimento de onda (quanto maior a frequência) da radiação. Quando o comprimento de onda é muito curto, até desvios minúsculos entre percursos produzem diferenças grandes nas posições relativas de cristas e vales, e o cancelamento é completo. O argumento apresentado é menos rigoroso quando o comprimento de onda é grande; então, desvios consideráveis entre percursos podem não resultar em anulação. A óptica geométrica trata a luz como tendo um comprimento de onda tão curto que mesmo diferenças infinitesimais entre percursos são eficazes, com o

resultado de a luz poder ser tratada como viajando numa sequência de linhas geometricamente retas. Na prática, a luz tem um comprimento de onda mensurável e a óptica geométrica é uma aproximação imperfeita. Os pequenos desvios permitidos às linhas retas manifestam-se como aberrações das lentes e outros fenómenos ópticos. As ondas de rádio têm comprimentos de onda da ordem de metros ou mais; para elas, a óptica geométrica é uma aproximação pobre e mesmo objetos grandes não bloqueiam o seu caminho — os rádios funcionam à volta de esquinas.

Embora o som não seja eletromagnético (é uma propagação ondulatória de diferenças de pressão), segue as mesmas regras de propagação que a luz. No entanto, os comprimentos de onda típicos são da ordem de metros (o comprimento de onda do dó central é cerca de 1,3 m), pelo que uma “acústica geométrica” seria uma aproximação muito pobre num mundo cheio de objetos do tamanho humano. É por isso que conseguimos ouvir à volta das esquinas.



Poderá, com toda a legitimidade, objetar que a “lei” que escolhi para demonstrar o papel autorrestritivo da anarquia é bastante trivial, uma lei-derivada juvenil. À medida que este capítulo se desenvolve, porém, acabará por ver que, pelo contrário, ela é extraordinariamente rica e que há muito mais nesta discussão sobre a propagação da luz do que aquilo que salta imediatamente à vista. Exatamente a mesma anarquia operativa primordial que aprisiona a luz pela sua ausência de constrangimentos aprisiona igualmente

a matéria e, desse modo, dá conta da mecânica quântica e, por extensão, até da própria mecânica clássica.

O conceito crucial de que necessito para estabelecer esta ligação é o de que as partículas são ondulatórias. Aqui tenho de o conduzir ao centro de mais uma dessas grandes revoluções da ciência que substituíram o senso comum por algo que, em vários aspetos, se revela mais simples, embora à primeira vista possa parecer chocante e improvável. Como referi no Capítulo 1, a física clássica mantinha as partículas arrumadas num compartimento conceptual e as ondas noutro. Os seus caracteres eram inteiramente distintos. As partículas eram entidades pontuais, localizadas no espaço; as ondas eram entidades ondulantes que se espalhavam, talvez indefinidamente, pelo espaço. O que poderia ser mais diferente? Quem poderia confundir uma coisa com a outra?

A própria Natureza mostrou que não era assim. Primeiro, tornou-se claro para os cientistas que certas experiências realizadas com a partícula mais leve então conhecida, o eletrão, mostravam que ela se comportava como se fosse uma onda. Já mencionei a minha imagem de J. J. Thomson e do seu filho George ao pequeno-almoço, com uma rigidez gelada entre ambos. J. J. tinha demonstrado de forma incontestável que o eletrão, que descobrira em 1897, é uma partícula com massa e carga bem definidas. O seu filho, G. P., surgiu depois e, em 1927, demonstrou com igual convicção que o eletrão é uma onda. Sensivelmente na mesma época, Clinton Davisson (1881–1958) e Lester Germer (1896–1951) chegaram à mesma conclusão ao mostrarem que os eletrões sofrem

exatamente o mesmo tipo de difração característico da luz ondulatória.⁴

Havia ainda outros enigmas perturbadores. Todos sabiam que a luz é uma onda (usei esse conceito ao longo deste capítulo, naturalmente), e miríades de experiências realizadas ao longo de décadas tinham-no confirmado. Depois surgiram dois incómodos. O primeiro foi o funcionário de patentes germano-suíço-austríaco-americano Albert Einstein, que em 1905 mostrou que um certo efeito envolvendo a ejeção de eletrões de metais quando iluminados por radiação ultravioleta (o chamado efeito fotoelétrico) podia ser explicado de forma simples se os raios de luz fossem, afinal, correntes de partículas. (Essa explicação valeu-lhe o Prémio Nobel em 1921; a relatividade, embora intelectualmente mais rica e de maior alcance, era então vista com maior desconfiança.) Essas partículas vieram a ser conhecidas como fotões. O outro incómodo foi o físico americano Arthur Compton (1892–1962), que em 1922 mostrou que, para explicar a forma como a luz ricocheteia nos eletrões, a luz tinha de ser tratada como um fluxo de fotões a atuar como pequenos projéteis. (Ele também recebeu um Nobel, em 1927.) A física encontrava-se num impasse: aquilo que se sabia serem partículas comportava-se como ondas, e aquilo que se sabia serem ondas comportava-se como partículas. O que se estava afinal a passar?

Do reconhecimento deste carácter duplo do eletrão e do fotão — e mais tarde do carácter duplo de toda a matéria e de toda a radiação — emergiu o conceito de dualidade, segundo o qual as entidades exibem características tanto de partícula como de onda,

dependendo do tipo de observação realizada. O reconhecimento da dualidade é um dos pilares da mecânica quântica, que emergiu pelas mãos (mais precisamente, pelos cérebros) sobretudo de Werner Heisenberg, isolado numa ilha; de Erwin Schrödinger, numa montanha com uma amante; e de Paul Dirac (1902–84), num mundo só seu, entre 1925 e 1927. A mecânica quântica e os seus desenvolvimentos sofisticados, como a eletrodinâmica quântica, produziram previsões numéricas de precisão extraordinária e foram testadas de forma aparentemente exaustiva até inúmeras casas decimais, sem que jamais se tenha detetado qualquer desvio ou discrepância. Por outras palavras, a mecânica quântica, baseada no reconhecimento da dualidade, pode muito bem-estar correta.

Com estes preliminares estabelecidos, e reconhecida a dualidade da matéria e da radiação, é agora tempo de ver como este conceito abre a porta à compreensão da origem das leis que governam o movimento das partículas. Desde que as partículas se comportem como ondas, já conhece a resposta quanto à existência de regras: não há regras. As ondas propagam-se entre a fonte e o destino sem constrangimento, seguindo todos os percursos possíveis. No entanto, todos os percursos exceto um — o percurso que envolve o menor tempo (precisarei isto em breve) — têm vizinhos destrutivos e são anulados. Apenas esse percurso especial tem vizinhos que não o destroem, e é esse o percurso que concluímos que a partícula segue.

Assim, consideremos como um carácter intrínseco de uma partícula (não apenas do eletrão) o facto de ser ondulatória. Da anarquia operativa primordial decorre que uma partícula viajará em

linha reta num meio uniforme, tal como a luz, à luz dos argumentos que já descrevi. Isto é: permita-se a uma partícula que percorra qualquer trajeto que seja através do espaço vazio e, em virtude do seu carácter ondulatório, o único percurso que terá vizinhos não destrutivos será o percurso retilíneo entre a fonte e o destino.



Movimento em linhas retas não é a única característica das partículas, e o mundo seria um lugar entediante e excessivamente previsível se fosse apenas isso. E quanto a comportamentos mais complexos, como órbitas e trajetórias curvas: será que a anarquia operativa primordial também os governa? Sim — tendo em conta a noção de ação.

Aqui encontro alguma dificuldade, porque “ação” é um termo técnico que apenas reflete de forma ténue a noção quotidiana associada à palavra. Já lhe pedi que pensasse em ação como esforço, dispêndio, fôlego físico. Em física existe, naturalmente, uma definição formal, enterrada nas equações da mecânica clássica, mas introduzi-la aqui em todo o seu rigor seria uma maratona técnica excessiva⁵. Assim, manter-me-ei na ideia de ação como dispêndio.

Quando uma partícula clássica percorre a sua trajetória, isso envolve uma certa quantidade de dispêndio. De acordo com o princípio da ação mínima de Maupertuis, o percurso que uma partícula efetivamente adota é aquele que corresponde ao menor

dispêndio, tal como nós próprios tenderíamos a escolher. Segue-se que as trajetórias das partículas, que podem ser calculadas a partir da mecânica clássica de Newton na sua formulação original baseada nas forças que atuam em cada ponto, podem igualmente ser obtidas determinando qual a trajetória que envolve a menor ação.

Surge então de imediato a questão: como sabe uma partícula, de antemão, qual é o percurso de menor ação? Poderia iniciar-se por um trajeto aparentemente sedutor e depois deparar-se com uma colina de força contrária, exigindo um esforço considerável para a ultrapassar. Nessa altura já seria tarde demais para voltar atrás e recomeçar por um caminho mais fácil, menos exigente em dispêndio.

Exatamente o mesmo problema surgiu na discussão da propagação da luz, e exatamente a mesma solução se apresenta aqui. Nesse caso, a resolução residiu na natureza ondulatória da luz em colaboração com a anarquia operativa primordial. Neste caso, a resolução reside na natureza ondulatória da matéria, igualmente em colaboração com a anarquia. Em vez do índice de refração do meio através do qual a luz se propaga — que abrandava o seu movimento e altera a sua fase (se termina num pico, num vale ou algures entre ambos) no destino — temos agora a ação ao longo da trajetória (o dispêndio associado ao percurso), que afeta a passagem da onda através de regiões com diferentes energias potenciais e, desse modo, controla a fase da onda de matéria no destino. Estas ideias foram exploradas por esse grande explorador e expositor científico, Richard Feynman, que descobriu que podia utilizá-las para chegar à formulação da mecânica quântica⁶.

Eis o argumento. Tratamos a partícula em questão como uma onda, e deixamos a anarquia governar, abandonando qualquer controlo. As ondas propagam-se por todos os percursos possíveis entre a origem e o destino, cada um terminando com uma certa fase que depende da ação envolvida nesse percurso⁷. Em geral, todos esses caminhos têm vizinhos que terminam com fases significativamente diferentes — alguns em pico, outros em vale — e anulam-se mutuamente. Apenas o percurso correspondente à ação mínima possui vizinhos benignos que não o anulam. Nós, observadores, não vemos este submundo de conflitos: apenas constatamos que o único percurso que sobrevive é o percurso de menor ação.

Quando a ação é muito grande — como acontece com partículas pesadas, como bolas ou planetas — a anulação provocada pelos vizinhos é extremamente eficaz e, tal como na óptica geométrica, as partículas podem ser tratadas como se viajassem ao longo de uma trajetória única, bem definida. A mecânica clássica emerge assim da natureza ondulatória das partículas, tal como a óptica geométrica emerge da óptica física e da natureza ondulatória da luz. Quando as partículas são pequenas, como os eletrões, a ação é diminuta e a anulação pelos vizinhos é largamente ineficaz — exatamente como acontece com o som — e a mecânica clássica falha, tal como a acústica geométrica é inaplicável ao som. A mecânica quântica, sempre presente, mas ocultada quando a ação é grande, passa agora a ser indispensável para avaliar o movimento das partículas.



Finalmente, para encerrar esta discussão, preciso de introduzir o conceito de equação diferencial, pois muitas leis da Natureza — e em especial as da mecânica clássica e da mecânica quântica — são expressas sob essa forma.⁸ Este tema é importante porque se afirma frequentemente que a característica matemática central da física é precisamente uma equação diferencial. Uma equação comum, familiar, diz-lhe como uma propriedade depende de outra, como em

$$E = mc^2$$

, que indica como a energia E depende da massa m . Uma equação diferencial, pelo contrário, diz-lhe como a variação de uma propriedade (daí a alusão à “diferença” em diferencial) depende de várias propriedades, incluindo da própria propriedade. A segunda lei de Newton — segundo a qual *a taxa de variação do momento linear é proporcional à força aplicada* — é um exemplo de uma equação diferencial expressa verbalmente.

Há aqui, contudo, um ponto adicional importante: as equações diferenciais exprimem variações infinitesimais de uma propriedade. A razão (e a vantagem) dessa restrição é que as condições podem variar de ponto para ponto. Por exemplo, a força na segunda lei de Newton pode mudar de lugar para lugar e de instante para instante, e para determinar o seu efeito global na trajetória de uma partícula é necessário considerar o efeito cumulativo de muitos pequenos

passos — na verdade, de um número infinito de passos infinitesimais. Dizemos então que o efeito global de uma força deve ser encontrado por integração (ou seja, combinando todos esses pequenos passos), ou, de forma equivalente, que a equação diferencial deve ser resolvida por integração. Assim, se num ponto do espaço uma força exerce uma determinada influência e num ponto vizinho exerce uma influência diferente, há um pequeno empurrão no primeiro ponto e outro no segundo; no conjunto, o resultado das forças é a soma desses empurrões. Como espero que agora se torne claro, uma equação diferencial é usada para avaliar como uma partícula, por exemplo, avança tateando o caminho até um destino e, desse modo, traça uma trajetória.

As equações diferenciais — tão centrais, fundamentais e onnipresentes na física — são filhas da anarquia operativa primordial. Como vimos, a anarquia conduz às leis do menor tempo na óptica e da menor ação na mecânica, mas estes mínimos referem-se ao percurso completo, não a um fragmento infinitesimal desse percurso. Ainda assim, um resultado matemático notável é o *facto de que avançar passo a passo de acordo com a equação diferencial apropriada garante que acabará por se encontrar a seguir o percurso globalmente mínimo*. Tudo o que uma equação diferencial faz é fornecer instruções locais sobre o que deve ser feito em cada ponto e em cada instante — se deve desviar-se para a esquerda ou para a direita, se deve acelerar ou não, e assim por diante — de modo a assegurar que se chega ao destino tendo percorrido o caminho de menor tempo ou de menor ação.⁹ Um critério global foi decomposto numa série de instruções locais. Assim, embora as equações diferenciais sejam frequentemente

tomadas como o núcleo da física, é defensável — e gosto de pensar que é verdadeiro — que, pelo menos na mecânica clássica e na mecânica quântica, elas são construções secundárias. As características centrais são, na realidade, os comportamentos globais baseados na anarquia, sendo as equações diferenciais apenas a forma prática de indicar como agir localmente — uma espécie de guia de bolso para a trajetória.



Onde nos trouxe a anarquia? Deixámos a luz livre para encontrar o seu próprio destino, sem impor uma regra, e descobrimos que ela acaba por obedecer a uma regra: percorre o caminho de menor tempo. Aceitámos a evidência experimental da dualidade das partículas, em particular do seu carácter ondulatório, e, deixando-as livres sem impor uma regra, descobrimos que também elas acabam por obedecer a uma regra: percorrem o caminho de menor ação. Tal como a óptica geométrica emerge da óptica física à medida que o comprimento de onda da luz diminui, também a mecânica clássica emerge da mecânica ondulatória (o nome mais antigo da mecânica quântica) quando o corpo em movimento deixa de ser diminuto e passa a ter uma dimensão familiar do quotidiano. Vimos ainda que as equações diferenciais, fundamentais e onnipresentes na mecânica clássica e quântica, são instruções locais para encontrar e percorrer o caminho que satisfaz critérios globais: o menor tempo para a luz e a menor ação para as partículas. A anarquia conduziu-nos à física.

CAPÍTULO 4

O Calor Do Momento

— Leis Relacionadas com a Temperatura

Se eu estivesse abandonado numa ilha deserta, com apenas uma palmeira como companhia e água, água por todo o lado, há um conceito que gostaria que me acompanhasse. É extraordinariamente fértil em consequências, oferece uma visão profunda sobre a natureza da matéria e sobre as transformações que esta sofre, e esclarece um dos conceitos mais esquivos da ciência, embora ligado a aspetos familiares da vida quotidiana: a temperatura. Dentro de momentos, apresentar-lhe-ei esse conceito tão companheiro.

O facto de a temperatura desempenhar um papel nas propriedades da matéria e nas leis que as governam garante-lhe um lugar neste relato. Este capítulo serve também como uma introdução inicial à termodinâmica, o corpo de leis de enorme importância que trata das transformações de energia, como a relação entre calor e trabalho e a razão pela qual algo acontece, pura e simplesmente.

A temperatura entra na descrição das propriedades da matéria de duas formas: uma proveniente do mundo dos fenómenos observáveis, a que chamamos o aspeto fenomenológico, e outra

oriunda do submundo dos átomos e das moléculas, a que chamamos o aspecto microscópico ou, porque este mundo está para além do alcance dos microscópios convencionais, o aspecto molecular. Todos estamos familiarizados, de um modo geral, com a temperatura e com as várias escalas em que é medida, e conhecemos a sua importância para o nosso bem-estar fisiológico. Sabemos que há objetos quentes e objetos frios, e sabemos que elevar a temperatura é essencial para provocar as transformações de que a indústria e os cozinheiros necessitam. Mas o que é, afinal? E poderá a temperatura ser um aspecto da anarquia, em aliança com a indolência estrutural primordial?

Para tratar de todas estas questões, gostaria que conhecesse um dos meus heróis: um cientista bem conhecido entre cientistas, mas cujo nome não surge com facilidade ou frequência nos lábios do público em geral. Trata-se do físico teórico vienense Ludwig Boltzmann (1844–1906), que, apesar de míope, viu mais longe na estrutura da matéria do que a maioria dos seus contemporâneos e que, angustiado pela receção desfavorável das suas ideias, acabou por se suicidar. As ideias de Boltzmann forjaram a ligação entre o microscópico e o fenomenológico, clarificaram o conceito de temperatura e estabeleceram uma forma de pensar as propriedades da matéria macroscópica em termos do comportamento dos átomos que a compõem. As suas ideias explicam por que razão a matéria persiste no mundo quotidiano e por que motivo, quando aquecida, a matéria se liberta para o mundo da transformação química. Um excelente companheiro *conceptual*, portanto, para quem estivesse isolado numa ilha deserta do ponto de vista intelectual — embora talvez não um companheiro alegre para uma ilha deserta real.



Não faço ideia se Boltzmann pensava da forma que estou prestes a descrever e, na verdade, estou confiante de que é improvável que o tenha feito; ainda assim, eis uma imagem que capta a essência da sua abordagem.

Imagine que está deitado em frente a uma estante com muitas prateleiras, rodeado por pilhas dos seus livros (para esta discussão, são todos idênticos). Ali deitado, talvez com os olhos vendados, atira os livros ao acaso para as prateleiras. Retira a venda e observa a distribuição dos livros pelas prateleiras. Haverá alguns nas prateleiras superiores, outros nas intermédias e outros nas inferiores. Não há qualquer padrão particular. Esvazia as prateleiras, retoma a posição inicial e volta a atirar os livros. Abre os olhos e vê outro padrão aparentemente aleatório. É muito improvável que todos os livros fiquem na prateleira mais alta, ou todos numa prateleira específica qualquer.

Você — e esta visão é obviamente fantasiosa — repete o procedimento milhões de vezes, registando a distribuição após cada episódio. Algumas distribuições de livros (por exemplo, todos numa única prateleira) quase nunca ocorrem; outras surgem com bastante frequência. Contudo — e isto chama-lhe a atenção — existe uma distribuição esmagadoramente mais provável, uma que aparece repetidamente. Nessa distribuição, a maioria dos livros surge na prateleira mais baixa, menos na seguinte, ainda menos na seguinte, e assim sucessivamente, até à prateleira mais alta, que

pode até não ter livro algum. Esta distribuição mais provável das populações das prateleiras é a distribuição de Boltzmann, o conceito central deste capítulo e o meu companheiro conceptual preferido na ilha deserta intelectual.

A distribuição real de Boltzmann não se aplica a livros em prateleiras, mas a moléculas e átomos. Como hoje é bem sabido, uma consequência da mecânica quântica é que a energia que qualquer objeto pode possuir está limitada a valores discretos. Uma molécula não pode vibrar ou rodar com energia arbitrária: só pode aceitar energia em passos ("quanta"). Até você, na sua bicicleta, acelera aos solavancos, mas esses solavancos são tão pequenos que, para todos os efeitos práticos, acelera de forma suave. Para átomos e moléculas, porém, esses solavancos estão longe de ser desprezáveis. Estes níveis de energia, as energias permitidas, são as prateleiras da analogia. Os livros da analogia são os átomos e as moléculas. O seu lançar displicente corresponde ao empurrão aleatório que faz com que átomos e moléculas passem de um nível de energia para outro. Os resultados desses lançamentos aleatórios são as populações das moléculas pelos níveis de energia disponíveis. Quase nunca encontrará todas as moléculas no mesmo nível de energia. A distribuição mais provável das moléculas, resultante desta dispersão aleatória pelos níveis de energia disponíveis, é a distribuição de Boltzmann, com a maioria das moléculas no estado de energia mais baixo, menos no nível seguinte, ainda menos no seguinte, e muito poucas — talvez nenhuma — em níveis de energia muito elevados.

Agora, tenho de admitir que a distribuição de Boltzmann não é apenas o resultado de um comportamento aleatório e anárquico. A indolência estrutural primordial também intervém. A energia total das moléculas é fixa (isto é uma consequência da indolência e da sua implicação, a conservação da energia, como argumentei no Capítulo 2). Assim, nem todas as moléculas podem acabar num único nível de energia elevado, pois a energia total excederia a disponível. Nem, em geral, podem todas acabar no nível de energia mais baixo, porque então a energia total não corresponderia à energia disponível. (“Em geral” é sempre uma expressão escorregadia, mas destina-se a indicar que podem existir casos especiais que permitem exceções à regra geral; voltarei a este ponto dentro de alguns parágrafos — aceite-o por agora.) A derivação de Boltzmann teve esta restrição em conta, e a distribuição que descrevi, com sucessivamente menos moléculas em níveis de energia crescente, é o resultado efetivo. Em suma, a distribuição de Boltzmann é o resultado da anarquia operativa primordial em aliança com a indolência estrutural primordial: o quase aleatório povoamento dos níveis de energia, comportamento anárquico, sujeito à conservação da energia, essa consequência da indolência.

É aqui que preciso de introduzir outra imagem verbal, desta vez sob a forma da expressão matemática real da distribuição de Boltzmann. Acontece que a diminuição gradual das populações com o aumento da energia dos níveis é descrita por uma expressão matemática muito simples.¹ Além disso, essa expressão depende do valor de um único parâmetro. Quando esse parâmetro tem um valor pequeno, as populações diminuem muito rapidamente com o

aumento da energia e apenas os níveis de energia mais baixos são ocupados (embora ainda com a diminuição típica à medida que se consideram níveis de energia mais elevados). Quando o parâmetro tem um valor elevado, as populações estendem-se até níveis de energia altos: embora continue a haver mais moléculas no nível de energia mais baixo, menos no seguinte, e assim por diante, passam agora a existir moléculas com energias muito elevadas. A expressão e o parâmetro em questão são universais, no sentido de que se aplicam a qualquer tipo de substância e a qualquer tipo de movimento. Isto é, para um dado valor do parâmetro, a população relativa de um nível de determinada energia é a mesma, quer se trate de vibrações ou rotações de moléculas, quer se trate das vibrações de átomos em sólidos, quer a substância envolvida seja chumbo ou lítio, giz ou queijo, ou qualquer outra.

O nome dado a este parâmetro universal que controla as populações é temperatura. Espero que agora tenha alguma intuição sobre a sua natureza. Uma temperatura baixa descreve uma distribuição de Boltzmann em que apenas os níveis de energia baixos estão ocupados, com populações decrescentes à medida que se sobe para níveis de energia mais elevados. Uma temperatura elevada descreve uma distribuição de Boltzmann em que as populações se estendem para níveis de energia altos; quanto mais elevada a temperatura, maior esse alcance.

Antes de abandonar este ponto, preciso de desmontar aquele “em geral” de há alguns parágrafos. Suponha que o valor do parâmetro, a temperatura, é fixado em zero. Neste caso, de acordo com a forma da distribuição de Boltzmann para este valor da

temperatura, todas as moléculas estarão no nível de energia mais baixo; não haverá qualquer molécula em qualquer outro nível de energia. Todos os livros estão na prateleira inferior. Este é o zero absoluto de temperatura, não sendo fisicamente significativo qualquer valor inferior, pois como poderiam as moléculas ocupar um nível de energia ainda mais baixo do que o mais baixo? Naturalmente, esta distribuição particular ainda tem de respeitar a conservação da energia, pelo que só é atingível quando toda a energia foi retirada da amostra e a energia total é, na prática, zero. (“Na prática” é outra expressão extremamente útil e escorregadia; deixá-la-ei passar. Coloquei-a ali, como um pedante não reformado, para dar a entender a outros pedantes que sei o que eles estão — ou deveriam estar — a pensar.²⁾)



Tanto quanto basta, por agora, para a interpretação molecular da temperatura e para a perspetiva que a distribuição de Boltzmann traz ao seu significado. A medição da temperatura estava bem estabelecida — mas o conceito permanecia obscuro — muito antes de Boltzmann se ter suicidado, e as escalas quotidianas familiares (em particular a de Fahrenheit e a de Celsius) já estavam há muito consolidadas de forma pragmática, com cada inventor a escolher “pontos fixos” facilmente reproduzíveis e transportáveis para definir a escala. Assim, Daniel Fahrenheit (1686–1736) fixou o zero da sua escala na temperatura mais baixa então facilmente atingível (e bastante acima do zero absoluto de que tenho vindo a falar), concretamente o ponto de congelação de uma mistura de sal

comum e água, e tomou como 96 (de modo algo enigmático, não 100) a temperatura da sua axila facilmente transportável — ou, pelo menos, da axila média e ubíqua. As 96 graduações entre estes dois pontos vagamente fixos fazem com que o ponto de congelação da água pura surja como 32 na sua escala e o ponto de ebulição da água como 212, bem acima da temperatura da sua axila. Anders Celsius (1701–1744), de forma bastante mais sensata, centrou-se nas propriedades da própria água para especificar os pontos fixos, colocando 100 no seu ponto de congelação e 0 no seu ponto de ebulição. A sua escala foi, entretanto, invertida (exploro a sabedoria dessa inversão no Capítulo 9), de modo que os objetos mais quentes tenham temperaturas mais elevadas do que os mais frios. É um pormenor curioso que, pelas suas definições, ambas sejam escalas “centígradas”, no sentido de terem cerca de 100 graduações, ou graus, entre os seus pontos fixos; mas a sociedade moderna, vendo 32 e 212 em vez do 0 original de Fahrenheit (a sua mistura de sal) e 96 (a sua axila), tende a considerar apenas a escala Celsius como sinónima de centígrada.

Para completar esta referência às escalas de temperatura, a escala que sensatamente fixa o zero absoluto em 0 chama-se “escala de temperatura termodinâmica” ou, de forma mais coloquial, “temperatura absoluta”. Se as graduações desta escala tiverem o mesmo tamanho que as da escala Celsius, então é conhecida como “escala Kelvin”, em homenagem a William Thomson, Barão Kelvin de Largs (1824–1907), um pioneiro da termodinâmica.³ Se as graduações tiverem o mesmo tamanho que as da escala Fahrenheit, então a escala termodinâmica de temperatura é conhecida como “escala Rankine”, em homenagem

ao engenheiro escocês John Rankine (1820–1872), hoje — mas não então — um teórico menos conhecido das máquinas a vapor e compositor de canções humorísticas. Quase ninguém, tanto quanto sei, usa hoje a escala Rankine: talvez alguns engenheiros ainda o façam na América, onde no quotidiano Fahrenheit se recusa teimosamente a ceder terreno a Celsius. Para ser completo, o zero absoluto situa-se em $-273,15\text{ }^{\circ}\text{C}$ ou $-459,67\text{ }^{\circ}\text{F}$.

Depois desta incursão no pragmático, a questão que agora preciso de explorar é como o conceito e a importância da temperatura entraram na ciência — e especificamente na termodinâmica — como uma propriedade observável antes de se aceitar a realidade das moléculas e sem qualquer antecipação de que os seus níveis de energia fossem discretos. Isto é: o que era a temperatura antes de Boltzmann?

A temperatura entrou formalmente na termodinâmica como um acrescento tardio. É preciso saber que uma característica da termodinâmica é que cada uma das suas leis tipicamente (lá está outra vez a palavra escorregadia) introduz uma nova propriedade relacionada com a energia. Assim, a primeira lei da termodinâmica introduz a propriedade que conhecemos como energia; a segunda lei da termodinâmica (que fará a sua entrada no Capítulo 5) introduz a propriedade que conhecemos como “entropia”. Ambas fazem uso do conceito de temperatura de várias formas, e foi-se tornando claro para os fundadores da termodinâmica que, embora tivessem formulações rigorosas da primeira e da segunda leis — e, portanto, definições de energia e de entropia —, a própria temperatura não tinha sido definida pela enunciação de uma lei.

Tinha de ser formulada uma nova lei, mais fundamental do que a primeira ou a segunda, uma lei que formalizasse a definição de temperatura. Nessa altura, com a 1 e a 2 já ocupadas, os fundadores da termodinâmica tiveram simplesmente de cerrar os dentes e recorrer a chamar a essa nova lei — que logicamente precedia a primeira e a segunda — a “lei zero da termodinâmica”. (Não tenho conhecimento de qualquer outro ramo da ciência que tenha precisado de introduzir uma lei zero como acrescento tardio; talvez exista alguma à espreita, não enunciada, na mecânica clássica newtoniana.) Em suma, a lei zero introduz a temperatura de forma formal, e preciso agora de explicar o seu conteúdo aparentemente banal e como ela cumpre a sua função.

Suponha que tem três objetos, a que chamarei A (por exemplo, um bloco de ferro), B (um balde de água) e T (se estava à espera de C, aguarde). Uma característica que agora se tornará clara é o carácter peculiar dos termodinamicistas, os praticantes da termodinâmica: ficam genuinamente entusiasmados quando reparam que nada acontece. Talvez tenha notado isso na discussão da conservação da energia no Capítulo 2: ficaram verdadeiramente extasiados (à sua maneira abstrata) ao notar que a energia total do universo não se alterava. Esse entusiasmo resultou na primeira lei da termodinâmica, que é uma espécie de elaboração da lei da conservação da energia. Eis outro cenário que, para eles, é excitante. Suponha que coloca A e T em contato e observa que nada acontece. Agora suponha que, separadamente, coloca B e T em contato e nada acontece. A lei zero afirma que, *se agora colocar A e B em contato um com o outro (colocar o bloco de ferro no balde de água), então nada acontecerá*. Esta é uma observação universal:

seja qual for a natureza de A e B, se nada acontece quando cada um, por sua vez, é colocado em contato com T, então nada acontecerá quando A é colocado em contato com B. Essa observação é, para um termodinamicista, quase esmagadoramente orgástica e deixa-o inundado de alegria.

Espero que agora seja claro que o objeto T está a desempenhar o papel de um termómetro — daí o T — e, de certo modo, está a medir a temperatura. Assim, quando A é colocado em contato com T e nada acontece (como, por exemplo, um fio de mercúrio permanecer com o mesmo comprimento no interior do tubo de vidro de T), isso significa que A tem a temperatura representada pelo comprimento do fio de mercúrio. Quando B é colocado em contato com T e nada acontece, isso significa que B tem a temperatura registada por T, a mesma que a de A. Assim, A e B “têm a mesma temperatura” e podemos estar confiantes de que, quando colocados em contato entre si, nada acontecerá. Este ciclo de “nadas a acontecer” é a forma como a lei zero introduz o conceito de temperatura.

Preciso agora de ligar o conceito de temperatura, tal como é introduzido pela lei zero, à sua interpretação molecular em termos da distribuição de Boltzmann. O ponto-chave, como sublinhei anteriormente, é que a temperatura é o parâmetro que caracteriza a distribuição das moléculas pelos níveis de energia disponíveis e é universal (especificamente, independente da substância em causa). O objeto A (o ferro) tem um conjunto de níveis de energia e os seus átomos distribuem-se por eles de acordo com a distribuição de Boltzmann para a temperatura vigente. O objeto B (a água) tem

um conjunto de níveis de energia e as suas moléculas distribuem-se por eles de acordo com a mesma distribuição de Boltzmann (o parâmetro, a temperatura, é o mesmo que para A). Quando A e B são colocados em contato (o ferro é imerso na água), os seus níveis de energia entrelaçam-se como os dedos das suas duas mãos quando estas são aproximadas. A distribuição das moléculas mantém-se a mesma, a temperatura mantém-se a mesma e, em suma, nada acontece.

A distribuição de Boltzmann capta — e, como todas as interpretações moleculares dos fenómenos, enriquece — o conceito de temperatura. Pode agora começar a perceber por que razão ela também capta a estabilidade da matéria no mundo quotidiano e a capacidade da matéria para sofrer transformações quando aquecida. A temperaturas normais, a distribuição das populações não se estende muito para níveis elevados de energia e a maioria das moléculas encontra-se em níveis de energia baixos, capazes de pouco mais do que, por exemplo, vibrar languidamente. Assim, a matéria é duradoura. Quando a temperatura aumenta, cada vez mais moléculas ocupam níveis de energia elevados e, tal como na vida, para as moléculas, energia elevada significa que as coisas podem acontecer. Em particular, átomos podem ser libertados e novas ligações formadas: pode ocorrer reação química. Na cozinha, cozinhar é o processo de usar um forno ou uma placa para empurrar moléculas para níveis de energia mais elevados até que um número suficientemente grande delas tenha energia bastante para reagir. A refrigeração faz descer as moléculas para os seus níveis de energia mais baixos e acalma-as. Diríamos que as preserva.

Existe uma lei em química sobre as velocidades a que as reações ocorrem e que nasce da distribuição de Boltzmann. O químico sueco Svante Arrhenius (1859–1927), que recebeu um dos primeiros Prêmios Nobel (em 1903), prêmio esse que ele próprio ajudou a instituir, propôs aquilo que hoje é conhecido como a “lei das velocidades de Arrhenius”, segundo a qual a velocidade de uma reação química aumenta com a temperatura de uma forma específica que depende de um único parâmetro conhecido como “energia de ativação”, o qual varia de reação para reação.⁴ Em termos gerais (o que significa que há muitas exceções), a velocidade de uma reação química duplica tipicamente por cada aumento de 10 graus na temperatura. A explicação reside na distribuição de Boltzmann, pois a energia de ativação é simplesmente a energia mínima que as moléculas devem possuir para reagir, e o número de moléculas que a possuem aumenta com a temperatura à medida que a distribuição de Boltzmann se estende para níveis mais elevados. O arrefecimento (refrigeração) tem o efeito oposto: à medida que a distribuição de Boltzmann desce para níveis mais baixos, menos moléculas têm energia para reagir e a reação abrandam.

A lei de Arrhenius tem inúmeras consequências no quotidiano. Cozinhamos alimentos elevando as moléculas para níveis acima da sua energia de ativação ao aumentar a temperatura em muitas dezenas de graus e, desse modo, acelerando as reações que decompõem os alimentos. Preservamos alimentos encolhendo a distribuição de Boltzmann, de modo que poucas moléculas tenham energia para reagir. O corpo combate doenças com febre, elevando a sua temperatura para perturbar o delicado equilíbrio das

velocidades das reações químicas que mantêm vivos tanto nós como as bactérias invasoras (há claramente aqui a necessidade de um equilíbrio delicado!). Os pirilampos piscam mais depressa em noites quentes do que em noites frias. A indústria recorre à lei das velocidades para provocar as reações de que necessita para transformar matérias-primas em materiais. À nossa volta há um coro de reações químicas a zumbir de acordo com a lei de Arrhenius, exibindo, com as suas velocidades variáveis, a modificação da distribuição de Boltzmann — com o seu fundamento na anarquia — que as mudanças de temperatura produzem.



A questão que talvez se esteja a formar na sua mente é o que acontece quando dois objetos que não têm a mesma temperatura são aproximados e colocados em contato. Aqui abandonamos o domínio da lei zero e o entusiasmo do “nada acontecer”, e entramos no território das leis do “nada não acontecer”. Mas, antes de lá chegarmos de forma formal, já sabe, pela experiência quotidiana, o que ocorre: a energia flui do corpo mais quente para o corpo mais frio (o ferro quente arrefece e a água fria aquece) e, com o tempo, os dois corpos atingem uma temperatura intermédia e retomam a sua vida de equilíbrio térmico, com nada aparentemente a acontecer. Desenvolverei esta característica familiar da Natureza e usá-la-ei para introduzir outra lei de grande alcance que, desnecessário será dizer, também brota da anarquia.

Preciso de o introduzir à noção de “calor”. Num certo sentido, isso é fácil, porque não existe tal coisa. Apesar do uso corrente, um corpo quente não contém calor; não perde calor ao arrefecer, pois nunca o teve. Ao contrário do que muitos dizem — e talvez do que já pensa — o calor não é uma forma de energia. Em ciência, o calor não é uma coisa; é um processo. *Calor é energia em trânsito em resultado de uma diferença de temperatura*. Aquecer não é aumentar o calor de um objeto; é um processo que pode aumentar a sua temperatura (fazer trabalho — por exemplo, agitar vigorosamente um líquido — também pode elevar a temperatura). Para o cientista, aquecer é transferir energia para um objeto explorando uma diferença de temperatura; não é transferir algo chamado “calor”. O calor foi outrora pensado como um fluido, o calórico (do latim *calor*, significando calor, com etimologia que recua ao sânscrito *carad*, colheita, o “tempo quente”), e o seu escoamento tem muitos atributos de um fluido; mas isso pertence ao início do século XIX e essa interpretação foi ultrapassada. Estas observações minuciosas corroem o significado quotidiano de “calor”, mas a ciência costuma avançar apropriando-se de um termo comum e refinando o seu significado. Aqui, o termo quotidiano “calor”, que como substantivo sugere posse como propriedade (“este forno liberta muito calor”), é destilado no espírito fino de processo: um processo em que a energia é trocada devido a uma diferença de temperatura. A pedanteria pode limpar o cérebro ou entupi-lo; espero o primeiro caso, pois a clareza vocabular é crucial — em lado nenhum mais do que na ciência, onde a verdade depende da precisão. O Capítulo 9, sobre o papel da matemática na expressão das leis da Natureza, leva esta atitude ao extremo e mostra como ela sustenta o extraordinário poder da

matemática na ciência. Sigamos, então, com o significado de calor bem presente.

Primeiro, consideremos a interpretação da transferência de energia como calor à luz da distribuição de Boltzmann. Seja A (o ferro) e B (a água) com temperaturas diferentes, sendo A mais quente (isto é, com temperatura mais elevada) do que B. Observacionalmente, isso significa que, quando cada um é colocado em contato com corpos T distintos (termômetros), nada acontece desde que o comprimento do fio de mercúrio em cada T seja diferente, correspondendo às temperaturas de A e de B. Agora olhe por baixo da superfície, para o submundo dos átomos que compõem A e B. Como o parâmetro a que chamamos “temperatura” é diferente em cada caso, os átomos de A exibem uma distribuição de Boltzmann diferente sobre os seus níveis de energia da que as moléculas de B exibem: em A, mais quente, os átomos povoam níveis de energia mais elevados do que em B, mais frio.

Quando os dois corpos são colocados em contato, todos os níveis de energia de A e de B ficam disponíveis para todos os átomos (pense, como antes, nos dedos das mãos entrelaçados, representando dois conjuntos de níveis que se fundem num único; ou numa única estante formada pela junção de duas). Uma vez restabelecido o equilíbrio, há uma única distribuição de Boltzmann dos átomos por este conjunto único de níveis. Para a atingir, alguns átomos têm de “cair” dos níveis altos providos por A para níveis mais baixos, sejam de A ou de B, até que as populações coincidam com a distribuição mais provável de Boltzmann. Como resultado dessa queda, os níveis do outrora frio B tornam-se mais povoados

à custa das populações do outrora quente A, incluindo alguns níveis mais altos de B que antes estavam muito pouco povoados. O sistema conjunto passa, assim, a ser caracterizado por uma única temperatura, intermédia entre as duas iniciais. O ferro arrefeceu e a água aqueceu.

Eis um ponto ligeiramente pedante (não resisto). Quando deixa uma chávena de café quente em repouso, ela arrefece até à temperatura do ambiente, não até um valor intermédio. O que se passa, aos olhos de Boltzmann? Mesmo o bloco quente de ferro considerado até aqui arrefeceria sem efeito discernível no meio envolvente. A palavra-chave é discernível. O ambiente (a mesa, a sala, a Terra, o universo...) é tão vasto que, para acomodar a energia recebida, os seus miríades de níveis de energia sofrem uma redistribuição quase totalmente negligenciável. Ou seja, embora a energia seja acomodada e ocorra uma redistribuição diminuta, ela é impercetível e, portanto, a temperatura do ambiente permanece, para todos os efeitos práticos, inalterada. É como notar que uma enorme folha de mata-borrão permanece branca apesar de ter absorvido uma gota de tinta.

Ainda não falei do tempo, em particular do tempo que um objeto leva a arrefecer até à temperatura do ambiente. Esta mudança de foco — de temperatura para tempo — introduz a lei a que quero chegar e as suas importantes ramificações. A anarquia continuará a ser a origem. Para mostrar o seu papel, delineio primeiro a regularidade do fenómeno, a lei do arrefecimento, e depois descrevo a anarquia subjacente enquanto os átomos seguem a sua vida, alheios a regras.

A taxa a que um corpo arrefece até à temperatura do ambiente é resumida pela *lei do arrefecimento de Newton* (o Newton de sempre), publicada, ao que parece anonimamente, em 1701. A lei sintetiza observações ao afirmar que *a taxa de variação da temperatura de um objeto quente é proporcional à diferença de temperatura entre ele e o ambiente*.⁵ Um objeto muito quente (comparado com o ambiente) arrefece rapidamente no início; à medida que a sua temperatura desce, a taxa de arrefecimento diminui, anulando-se quando atinge a mesma temperatura do ambiente. Este comportamento — em que o valor de uma propriedade decai a uma taxa proporcional ao seu valor corrente (aqui, a “propriedade” é a diferença de temperatura) — chama-se decaimento exponencial. O termo “exponencial” é muitas vezes mal-usado no discurso comum; aqui uso-o no seu sentido preciso. Um decaimento exponencial pode ser extremamente lento, como quando a temperatura de um objeto é quase igual à do ambiente. E não exagero: o decaimento exponencial tem análogos por toda a ciência; apoiar-me-ei no arrefecimento de Newton para introduzir outro deles.

A primeira característica importante que tenho de introduzir, e que ignorei até agora, é que as moléculas numa distribuição de Boltzmann não estão paradas nos seus níveis-prateleira: saltam incessantemente entre níveis. Tal como um Dickens pode cair subitamente para uma prateleira mais baixa e um Trollope ocupar o lugar ao ser promovido de baixo, mantendo-se globalmente a distribuição de Boltzmann, assim as moléculas migram continuamente. No submundo atómico, tudo é movimento, migração e reacomodação. Isto é crucial: a distribuição de

Boltzmann é dinâmica, um ser vivo pulsante de mudança interna. A calma observada externamente esconde a tempestade interior.

A segunda característica é a taxa a que uma molécula individual salta entre níveis devido ao empurrão constante das colisões. Essa taxa varia amplamente: algumas moléculas permanecem muito tempo num nível e depois saltam rapidamente por vários. Pense em cada molécula ocupando um nível por durações variáveis, com um tempo médio de vida de uma fração ínfima de segundo, antes de transitar. O ponto essencial é que o comportamento de uma molécula individual (o seu tempo de permanência num estado) é independente do que as outras fazem: cada molécula é uma ilha.

Agora imagine juntar dois objetos (o ferro A e a água B). A redistribuição ocorre como descrito, mas precisamos de integrar o facto de as moléculas migrarem a uma taxa média constante. O número médio que salta num dado intervalo depende do tempo médio de vida (quanto menor, mais saltos) e do número de moléculas prontas a saltar (quanto maior a população, mais saltos). Assim, a taxa a que a população de um nível muda depende do tempo médio de vida e da população. Eis o ponto crucial: quando A é muito mais quente do que B, há muitas moléculas em níveis altos prontas a redistribuir-se, logo a redistribuição é rápida. Quando as temperaturas são quase iguais, poucas moléculas precisam de redistribuir-se, e o processo é lento. Em suma, a taxa de redistribuição é proporcional à diferença entre as distribuições; como estas dependem das temperaturas, a taxa de variação da temperatura é proporcional à diferença de temperatura. Essa

proporcionalidade implica arrefecimento exponencial — o conteúdo da lei de Newton.

O ponto crucial é que, se as moléculas podem saltar entre níveis sem constrangimento, o resultado é uma lei: a do decaimento exponencial. Mais uma vez, da anarquia emerge a lei. Decaimentos exponenciais (e por vezes crescimentos exponenciais) são comuns na física e na química, e todos derivam de comportamentos subjacentes anárquicos em que indivíduos mudam aleatória e independentemente.

Um exemplo importante é a *lei do decaimento radioativo*, em que a atividade de um isótopo decai exponencialmente no tempo.⁶ A radioatividade resulta da fragmentação de um núcleo (por emissão de partículas alfa ou beta) ou de colapsos internos que geram fótons gama, com cada núcleo tendo uma probabilidade constante de fragmentação por unidade de tempo. Estes processos são independentes do que ocorre no núcleo vizinho e produzem decaimento exponencial.

Por exemplo, um núcleo de carbono-14 (seis prótons e oito neutrões) tem uma certa probabilidade de emitir uma partícula beta — um eletrão rápido — por segundo (cerca de uma em 250 mil milhões). Essa probabilidade individual é a mesma para todos os núcleos de carbono-14 numa amostra e independe das condições externas e do que acontece ao lado. Após a emissão, o núcleo torna-se azoto-14 (sete prótons e sete neutrões) e fica inativo. A taxa global de emissões na amostra decresce com o tempo, proporcional ao número de núcleos ainda disponíveis para decair:

alta no início, decrescendo depois, exatamente como a diferença de temperatura no caso de Newton — um decaimento exponencial.

Há consequências importantes. Uma positiva é a datação por carbono-14, que permite estimar a idade de artefatos orgânicos a partir da razão entre carbono-14 e carbono-12. Menos benigno é o decaimento lento de muitos isótopos radioativos, especialmente os remanescentes da fissão nuclear em centrais e explosões. Uma consequência matemática do decaimento exponencial é que o tempo para reduzir a metade a abundância inicial é sempre o mesmo, repetidamente: a meia-vida (para o carbono-14, 5730 anos). Alguns isótopos têm meias-vidas de frações de segundo; outros, de anos ou milhares de anos. Nada podemos fazer para alterar isso, exceto mudar a identidade do isótopo por outro processo nuclear e transformá-lo num de vida curta.



Este capítulo cobriu muito terreno e, mais uma vez, um resumo pode ser útil. Argumentei que o resultado esmagadoramente mais provável da distribuição aleatória das moléculas pelos níveis de energia disponíveis (sujeita à restrição da conservação da energia) é a distribuição de Boltzmann: uma distribuição dinâmica de populações que depende de um único parâmetro universal, a temperatura. Esta distribuição comporta-se como seria de esperar a partir da noção comum de temperatura e ajuda a explicar por que razão a matéria é estável em condições normais, mas, à medida que a temperatura aumenta, se torna progressivamente capaz de

se transformar em substâncias diferentes. Salientei também que, se não se impuser qualquer restrição ao comportamento de moléculas individuais e independentes, permitindo-lhes agir de forma aleatória, o resultado é um tipo de comportamento que ocorre amplamente na Natureza: o decaimento exponencial. Essa discussão conduziu à explicação de duas leis da natureza: a *lei do arrefecimento de Newton* e a *lei do decaimento radioativo*.

A indolência e a anarquia ergueram-se por detrás de vários parapeitos ao longo desta discussão. Elas (através da mecânica quântica) explicam a existência de níveis de energia. Os pormenores da derivação da distribuição de Boltzmann dependeram da indolência estrutural primordial, expressa na conservação da energia, e da anarquia operativa primordial, manifestada na distribuição aleatória das moléculas pelos seus níveis de energia. As velocidades a que as temperaturas se igualam — representativas de diversos tipos de mudança — assentam igualmente na anarquia individual do comportamento dos indivíduos, que, inconscientemente, conspiram para gerar, ou talvez apenas tropeçam, em uma lei.

CAPÍTULO 5

Para Além da Anarquia — Por Que é Que Algo Acontece

Mencionei, no Capítulo 4, que os termodinamicistas — as pessoas que estudam e aplicam a termodinâmica — ficam extremamente entusiasmados quando nada acontece. Se, para sua decepção, algo acaba por acontecer, então retiram prazer do facto de observar que as coisas invariavelmente pioram. Essa observação, a de que as coisas pioram, é a segunda lei da termodinâmica, uma das minhas leis da natureza preferidas. Claro que, em ciência, esta formulação popular da lei é vestida com trajes formais e ganha poder ao ser expressa de modo mais preciso e matemático, mas “as coisas pioram” é a sua essência. Como observação preliminar adicional, referi também que cada lei da termodinâmica introduz uma nova propriedade relacionada com a energia e com os aspetos da sua transformação: essa reflexão tardia, a lei zero, introduziu a temperatura, e a primeira lei introduziu a energia. A segunda lei introduz uma terceira grande propriedade, a entropia. O meu objetivo aqui é mostrar que a segunda lei é mais uma manifestação da indolência e da anarquia, mas que as propriedades da entropia explicam a emergência de estruturas, acontecimentos e opiniões que podem ser requintadas.

As coisas pioram. Preciso de desenvolver esta afirmação e levá-la ao ponto em que se torne claro como ela permite — e até provoca — a emergência do requintado. A observação de que as coisas pioram é, claro, uma interpretação ligeiramente jocosa da formulação formal da segunda lei, que afirma que, *num processo espontâneo, a entropia de um sistema isolado tende a aumentar*. Existem vários termos nesta formulação mais austera que preciso de explicar, mas não permita que eles obscureçam a impressão global do conteúdo da lei: o universo deriva inexoravelmente para um estado de maior deterioração.

Preciso de explicar, espero que sem excessiva turgidez, o que se entende por processo espontâneo, por entropia, naturalmente, e por sistema isolado. Um processo espontâneo é um acontecimento que pode ocorrer sem intervenção externa, sem ser forçado; é uma mudança natural, como a água a correr encosta abaixo ou um gás a expandir-se para o vazio. Espontâneo não significa rápido: alguns processos podem ser espontâneos, mas demorar imenso tempo, até Eras, a desenrolar-se, como o melaço a escorrer encosta abaixo ou o movimento dos glaciares. Outros processos espontâneos podem ocorrer num piscar de olhos, como a expansão de um gás para o vazio. A espontaneidade, neste contexto, diz respeito à tendência, não à velocidade de realização dessa tendência.

A palavra entropia vem do grego, significando “virar-se para”, e foi cunhada em 1856 pelo físico alemão Rudolf Clausius (1822–88), que voltará a surgir mais adiante neste relato. A entropia é uma medida — que pode ser definida com precisão — da desordem: dito de forma simples, quanto maior a desordem, maior é a entropia.

Várias pessoas chegaram à definição quantitativa de entropia; entre elas estava Boltzmann, o herói do Capítulo 4. A sua fórmula para expressar a entropia como uma medida de desordem — não confundir com a fórmula da sua distribuição — está gravada na sua lápide em Viena.¹ Não precisamos de a conhecer: trato aqui de interpretações, não de equações. O “aumento da desordem” é muitas vezes fácil de identificar, mas por vezes apresenta-se disfarçado sob subtis enganos. Darei exemplos mais adiante.

Por fim, um sistema isolado significa a parte do mundo que nos possa interessar (o “sistema”), mas separada de qualquer interação com o seu exterior. Nenhuma energia pode sair ou entrar num sistema isolado, nem matéria. Imagine algo dentro de um recipiente opaco (para impedir a entrada ou saída de radiação), rígido (para impedir que a energia seja usada para realizar trabalho de expansão), selado (para impedir a entrada ou saída de matéria) e em vácuo (para impedir a entrada ou saída de energia sob a forma de calor). Ao aplicar a segunda lei, o conceito de sistema isolado desempenha um papel crucial. Se quiser pensar em grande, então o universo inteiro é um sistema isolado (ou pelo menos assim o tomamos). Deve ter presente que um termodinamicista pode, com alguma modéstia, pensar em pequeno e considerar um frasco tapado, imerso num banho de água, como sendo o seu universo inteiro.

À medida que desenvolvo esta exposição da segunda lei e mostro como a perda de forma pode gerar forma, como ela sustenta a marcha majestosa da evolução e como explica tanto a emergência do horrível como do requintado, preciso que aceite que existe uma

tendência natural para a matéria e a energia se dispersarem de forma desordenada. Há aqui questões profundas às quais terei de regressar, mas, por agora, espero que possa aceitar como “evidente” que, se átomos e moléculas puderem vaguear livremente — sendo o “livremente” entendido como não sendo dirigidos a deslocar-se numa direção particular ou a organizarem-se de um modo específico —, então é muito mais provável que uma estrutura decaia em desordem do que a desordem se organize espontaneamente numa estrutura. Assim, as moléculas de um gás injetadas num canto de um recipiente têm muito mais probabilidade de se espalharem até preencherem todo o recipiente do que as moléculas de um gás que já o preenche uniformemente se agruparem, sem intervenção externa, num único canto. Claro que poderia comprimi-las num canto usando algum tipo de pistão, mas isso constituiria uma intervenção externa, o que é proibido num sistema isolado. De modo semelhante, a energia dos átomos que vibram vigorosamente num bloco de ferro quente tem muito mais probabilidade de agitar as moléculas vizinhas do meio envolvente e, assim, dispersar a sua energia por essas moléculas do que de a energia se acumular no bloco por choques aleatórios das moléculas externas, fazendo com que o bloco aqueça à custa de um meio mais frio. Mais uma vez, poderia conceber um modo de aquecer o bloco usando energia do exterior, mas tais artifícios são intervenções externas e, por isso, não são permitidos num sistema isolado.



A direção da mudança natural é a de a matéria e a energia se dispersarem em desordem, sem serem travadas por regras, exceto pela regra suprema, originada na indolência estrutural primordial, da conservação da energia. Esta afirmação pode ser expressa de outra forma: embora a quantidade de energia no universo permaneça constante, a sua qualidade tende a degradar-se. A energia que está concentrada num único local é de alta qualidade no sentido em que pode ser usada para realizar todo o tipo de ações (pense num litro de combustível); uma vez libertada e dispersa (por exemplo, pela combustão), essa energia continua algures presente, mas torna-se muito menos útil. Um gás a alta pressão num cilindro representa energia localizada e de alta qualidade, pois as suas moléculas se movem rapidamente num espaço confinado. Essa qualidade degrada-se se o gás for deixado escapar e a energia das suas moléculas se dispersar. Aqui está a termodinâmica numa casca de noz: a quantidade de energia conserva-se; a sua qualidade decai.

A entropia é simplesmente uma medida dessa qualidade, sendo que entropia elevada significa baixa qualidade. A entropia de um combustível é baixa; a dos seus produtos de combustão é elevada. A entropia de um gás comprimido é baixa; depois de se expandir, a sua entropia é elevada. Assim, “a quantidade conserva-se, a qualidade decai” transforma-se em “a energia conserva-se, a entropia aumenta”. Do mesmo modo, o jocoso “as coisas pioram” torna-se, numa formulação mais formal, “a entropia tende a aumentar”.

Quando a entropia foi introduzida no mundo da ciência, na década de 1850, houve grande perplexidade quanto à sua origem. Os vitorianos sentiam-se confortáveis com a constância da energia, pois (na sua perspectiva) ela não precisava de vir de lado nenhum depois de o Criador ter dotado suficientemente o universo com aquilo que, na Sua infinita sabedoria, considerara ser exatamente o bastante para servir as nossas necessidades para sempre. Mas a entropia parecia estar a surgir do nada. Estaria a Criação ainda em curso? Existiria um poço profundo e inesgotável de entropia, ainda por descobrir, a ser lentamente bombeado para a nossa percepção a um ritmo considerado adequado por esse mesmo Criador infinitamente sábio? A ciência veio em socorro desta visão culturalmente apropriada, mas simplista, da natureza da realidade, como em tantas outras ocasiões, sob a forma de uma compreensão molecular da entropia.



Sempre que ocorre uma mudança, a desordem do universo aumenta, a qualidade da sua energia degrada-se, a sua entropia aumenta. O curioso é que, tal é a interligação dos acontecimentos no mundo, que esta degradação não corresponde a um deslizamento cósmico uniforme para a desordem, a uma eliminação geral da estrutura, a uma dispersão global da energia, a um colapso da matéria numa espécie de lodo indistinto. Podem emergir abatimentos locais do caos — nós incluídos. A única exigência da segunda lei é que a entropia total de um sistema isolado (o universo, ou uma parte isolada dele, como aquele pequeno banho

de água com o seu frasco) aumente numa mudança espontânea: em bolsões localizados a entropia pode diminuir e uma estrutura pode emergir, desde que, no conjunto, haja um aumento da desordem.

Vejamos com mais detalhe o que isto significa. Considere um motor de combustão interna. O combustível é uma concentração compacta de energia armazenada. Quando arde, as suas moléculas fragmentam-se (se for um hidrocarboneto, transformam-se em muitas pequenas moléculas de dióxido de carbono e água) e dispersam-se. A energia libertada na combustão espalha-se pelo meio envolvente. A disposição de pistões e engrenagens do motor foi concebida para responder a essa dispersão e propagação e, na prática, para a capturar. O motor pode fazer parte de uma grua usada para construir uma catedral, erguendo blocos para a sua posição. Assim, enquanto ocorre dispersão no próprio motor, uma estrutura emerge noutro local. No conjunto, o universo tornou-se mais desordenado; localmente, na catedral, emergiu uma estrutura. Globalmente há um aumento da desordem; localmente houve um abatimento.

Esta produção de ordem impulsionada por um colapso para uma desordem maior pode ser encontrada onde quer que se olhe. Há frequentemente cadeias ligadas em que um aumento de desordem gera ordem noutro ponto, que depois se degrada e, ao degradar-se, gera ordem noutro lugar. O que importa em toda esta concatenação de acontecimentos é que surja mais desordem algures do que a desordem que é destruída noutro sítio. Essa destruição de desordem é a geração de ordem.

O Sol é o grande dispersor no céu e, através desta concatenação de dispersões, impulsiona acontecimentos — incluindo a evolução — na Terra. As reações de fusão nuclear que ocorrem no interior do Sol libertam energia que se dispersa pelo espaço. Uma fração minúscula dessa energia é capturada pela vegetação verde na Terra e usada para construir estrutura orgânica. Neste caso, os materiais iniciais desordenados são o dióxido de carbono e a água, e as estruturas altamente organizadas que passam a ser geradas são os hidratos de carbono que cobrem a litosfera áspera com uma biosfera benévola. Estruturas orgânicas — plantas, árvores, o que quiser — foram conduzidas à existência; mas o Sol morreu um pouco e o sistema solar tornou-se mais desordenado, apesar do crescimento da vegetação na Terra.

Essa vegetação é alimento para os animais. O alimento é combustível para a combustão interna que nos alimenta a nós e a eles. A combustão do alimento é muito mais subtil do que a combustão de um combustível num motor — não há chamas dentro de nós — mas é análoga na medida em que a digestão degrada moléculas complexas em moléculas pequenas, incluindo dióxido de carbono e água, libertando energia. Os organismos não são arranjos de engrenagens e rodas dentadas, mas os processos metabólicos no seu interior (e no nosso) são os análogos orgânicos dessas engrenagens e transmitem o poder organizador da digestão inicial para os locais análogos às gruas. Aí, aminoácidos — alguns dos pequenos blocos moleculares de construção da Natureza — são içados para as estruturas que conhecemos como proteínas, pequenas catedrais moleculares intrincadas, e o organismo cresce. No conjunto, considerando a digestão da refeição, houve um

aumento da desordem; mas a bioquímica do organismo explorou, à maneira de um motor, essa dispersão e uma estrutura — talvez você — emergiu. Como indiquei, somos abatimentos locais da desordem; somos filhos do caos.

Não somos apenas você e eu que somos conduzidos à existência pela geração de maior desordem. Todo o ecossistema é um descendente da desordem e uma consequência do caos. Nenhum organismo é uma ilha. A seleção natural é a forma extraordinariamente complexa e maravilhosa que a Natureza encontrou para se acomodar à segunda lei. A biosfera é uma reticulação extraordinária de entidades interdependentes, umas alimentando-se das outras — literalmente vivendo da dispersão que a devoração desenvolve. O combustível, o que é comido, é escasso e essencial à sobrevivência e, mais tarde, à propagação, pois a vida é estrutura e tem de ser sustentada aumentando a desordem do universo. Os seres vivos simplesmente não poderiam ter evitado viver uns dos outros, e a seleção natural, resultando na evolução, é uma consequência disso.

Ainda há quem não consiga reconciliar o colapso para a desordem, que a segunda lei identifica como a mola motriz da mudança, com a emergência das estruturas elaboradamente organizadas conhecidas como organismos. Não conseguem ver que a dispersão engendra estrutura. A resolução desta dificuldade é o ponto que tenho sublinhado repetidamente. A única necessidade é que a desordem global esteja a aumentar. Ligado um acontecimento a outro, um foco local de aumento de desordem (o consumo de combustível, a ingestão de um antílope, e miríades de

outras possibilidades — incluindo essa elaboração civilizada de geração de desordem a que chamamos um jantar) pode conduzir outra região do universo da desordem para a ordem. Tudo o que é necessário, para além do mecanismo que liga essas regiões, é que o aumento de desordem supere a diminuição de desordem, de modo que, no conjunto, haja um acréscimo de desordem. Incontáveis processos exemplificam esta interação inscrita na segunda lei; a evolução por seleção natural é apenas a mais entusiasmante.



Tenho-me concentrado sobretudo nos organismos, pois é aí que a segunda lei ilumina de forma brilhante e talvez surpreendente. Existem, contudo, muitas outras manifestações puramente inorgânicas e tecnológicas desta lei. A maioria das aplicações da segunda lei na tecnologia não decorre da fórmula da entropia inscrita na lápide de Boltzmann, mas de uma expressão alternativa proposta por Clausius em 1850. Ignorando a interpretação molecular da entropia, Clausius propôs aquilo que, à primeira vista, parece ser uma expressão totalmente independente para a variação da entropia associada a um processo, em termos de propriedades observáveis (em contraste com a dispersão desordenada de energia e moléculas). Propôs que a variação da entropia fosse calculada acompanhando a quantidade de energia transferida sob a forma de calor para dentro ou para fora do sistema e dividindo esse valor pela temperatura a que a transferência ocorreu.²

Clausius não associou o resultado do seu cálculo à desordem, mas nós podemos fazê-lo. A transferência de energia como calor recorre ao choque aleatório de moléculas vizinhas, como as de uma chama quente ou os átomos a vibrar vigorosamente num aquecedor elétrico. Esse choque agita as moléculas do sistema de interesse num movimento desordenado e, assim, aumenta a entropia. Até aqui, tudo bem: transferir energia como calor aumenta a entropia. Mas por que é que a temperatura importa? A analogia de que gosto é a de espirrar numa rua movimentada ou numa biblioteca silenciosa. Uma rua movimentada é o análogo de um objeto quente, com grande agitação térmica. Uma biblioteca silenciosa é o análogo de um objeto frio, com átomos pouco agitados. O espirro é o análogo de uma injeção de energia sob a forma de calor. Quando se espirra numa rua movimentada, o aumento de desordem é relativamente pequeno. Quando se espirra numa biblioteca silenciosa, o aumento de desordem é substancial. Assim sucede na definição de Clausius: a energia transferida como calor para um objeto quente não aumenta muito a desordem, pelo que a variação de entropia é pequena. Quando a mesma quantidade de energia é adicionada como calor a um objeto frio, a variação de entropia é grande. A temperatura na fórmula de Clausius capta a diferença entre a rua e a biblioteca.

A abordagem de Clausius capta também um resultado muito importante estabelecido nos primórdios da termodinâmica pelo engenheiro francês Sadi Carnot (1796–1832), num trabalho que permaneceu largamente ignorado durante décadas, por as suas conclusões serem tão inesperadas e incompatíveis com o senso comum dos engenheiros da época. Carnot argumentou (usando

conceitos que hoje considerariamos incorretos, como tratar o calor como “calórico”, um fluido imponderável, sem peso, que geraria trabalho à semelhança de uma roda de água ao atravessar um motor) que a eficiência de uma máquina a vapor ideal depende apenas das temperaturas da fonte quente, de onde a energia é retirada sob a forma de calor, e do reservatório frio, para onde a energia é descarregada.³ Mostrou, de forma talvez ainda mais notável, que a eficiência é independente da identidade e da pressão da substância de trabalho (tipicamente vapor).

O resultado de Carnot não é uma nova lei da Natureza, mas ilustra como uma lei — neste caso, a segunda lei da termodinâmica — pode lançar uma rede ampla e capturar comportamentos muito diversos. Eis o argumento. Imagine um motor constituído por uma fonte quente de energia, um reservatório frio para onde a energia pode ser descartada e, entre ambos, um dispositivo para usar energia para realizar trabalho (pense-se numa espécie de turbina). Imagine agora retirar alguma energia, na forma de calor, da fonte quente. A entropia da fonte diminui, mas, como a temperatura é elevada, a fórmula de Clausius implica que essa diminuição não é grande (a fonte quente é como uma rua movimentada). A energia extraída é convertida em trabalho por meio de um dispositivo mecânico. Neste ponto, deve ser claro que nem toda essa energia pode ser convertida em trabalho. Se o fosse, não haveria qualquer outra variação de entropia e, globalmente, a entropia diminuiria. Isso significaria que o motor não funcionaria, pois, para que ocorra uma mudança natural, a entropia tem de aumentar.

Para que o motor funcione, parte da energia retirada da fonte quente tem necessariamente de ser transferida para um reservatório frio (que pode ser a atmosfera ou um rio), na forma de calor. Essa transferência de energia para o reservatório frio provoca um aumento da sua entropia. Mesmo uma quantidade pequena de energia transferida tem um efeito significativo na entropia do reservatório, porque a sua temperatura é baixa — tal como um espirro causa muito mais perturbação numa biblioteca silenciosa do que numa rua movimentada. A energia retirada da fonte quente pode ser parcialmente convertida em trabalho, mas não pode sê-lo na totalidade. Se toda a energia fosse convertida em trabalho, a entropia global diminuiria, o que violaria a segunda lei da termodinâmica. Assim, uma fração da energia tem de ser inevitavelmente descartada no reservatório frio, tornando-se indisponível para realizar trabalho. A quantidade mínima de energia que tem de ser descartada é exatamente a necessária para garantir que o aumento de entropia do reservatório frio compensa a diminuição de entropia associada à retirada de energia da fonte quente. Como esse balanço depende apenas das temperaturas da fonte quente e do reservatório frio, a eficiência máxima possível de um motor depende exclusivamente dessas duas temperaturas, e de mais nenhum detalhe da sua construção.

O motor realiza mais trabalho quanto menor for a fração da energia retirada da fonte quente que tem de ser transferida para o reservatório frio na forma de calor. No entanto, esse valor mínimo não é arbitrário: tem de ser suficiente para assegurar que o aumento de entropia do reservatório frio compensa exatamente a diminuição de entropia associada à retirada de energia da fonte

quente. Como o reservatório frio se encontra a baixa temperatura, uma quantidade relativamente pequena de energia transferida na forma de calor é suficiente para produzir um grande aumento de entropia. A quantidade mínima necessária é determinada exclusivamente pelas temperaturas da fonte quente e do reservatório frio, e não depende de qualquer outro aspecto do motor. A consequência é que a eficiência do motor — definida pela fração da energia retirada da fonte quente que pode ser convertida em trabalho — depende apenas dessas duas temperaturas e é independente da substância de trabalho, da pressão, do desenho mecânico ou de quaisquer outros detalhes da construção e da operação do motor. Para atingir a eficiência máxima, é, portanto, necessário dispor de uma fonte tão quente quanto possível, de modo que a redução de entropia associada à extração de energia seja mínima, e de um reservatório tão frio quanto possível, para que até uma pequena quantidade de energia descartada gere um aumento suficiente de entropia. Foi precisamente esta a conclusão a que Carnot chegou no início do século XIX, perante o ceticismo dos seus contemporâneos — e estava certo.

Existem várias formulações alternativas e equivalentes da segunda lei da termodinâmica que não mencionam a entropia, mas que agora podem ser compreendidas à luz que a entropia lhes lança, bem como desta breve discussão do trabalho de Carnot. Ambas as formulações que apresentarei ilustram uma das minhas citações favoritas, originalmente feita pelo bioquímico húngaro Albert Szent-Györgyi (1893–1966), segundo a qual ser cientista envolve ver o que todos os outros viram, mas pensar o que mais ninguém pensou. William Thomson (Lord Kelvin de Largs, 1824–

1907; tomou o seu título, em 1892, do rio Kelvin, que passa perto do seu laboratório em Glasgow) viu, como muitos antes dele, que *uma máquina a vapor não funcionaria sem um reservatório frio*. Pensou nisso e usou-o como base da chamada “formulação de Kelvin” da segunda lei da termodinâmica, a partir da qual teceu toda uma rede de termodinâmica.⁴ Hoje sabemos porquê: se não houver um reservatório frio, não pode haver aumento de entropia e, portanto, tal motor seria impotente. Rudolf Clausius foi também um cientista que viu e depois pensou. Reparou (fantasio aqui anacronicamente) que, para funcionar, um frigorífico tinha de estar ligado à corrente. Mais precisamente, notou — como todos sabiam, mas ninguém tinha pensado a fundo — que o calor não flui de um corpo frio para um corpo quente sem que se realize trabalho para o provocar. Desenvolveu esta observação no que hoje chamamos a “formulação de Clausius” da segunda lei e construiu a sua versão da termodinâmica a partir daí.⁵ Hoje sabemos porquê: se a energia é transferida de um objeto frio para fora na forma de calor, ocorre uma grande diminuição da sua entropia; quando essa mesma energia entra num objeto quente, o aumento de entropia aí produzido é apenas pequeno. No balanço global, a entropia diminui, e por isso o processo não é espontâneo. Para que tal transferência ocorra, é necessário realizar trabalho externo: o frigorífico tem de estar ligado. O fluxo de energia retirado ao objeto frio tem de ser complementado por trabalho, de modo que, ao entrar no objeto quente, a energia transferida seja suficiente para gerar um aumento de entropia que compense — e supere — a diminuição de entropia ocorrida no objeto frio. Um aspeto notável é a forma como duas formulações aparentemente distintas de uma mesma lei da Natureza — a de Kelvin e a de Clausius — se fundem numa única

formulação coerente quando se introduz aquilo que inicialmente parecia ser uma noção abstrata: a entropia. A abstração revela-se aqui uma ferramenta extraordinariamente poderosa para unificar o aparentemente disperso, promover o progresso conceptual e aprofundar a compreensão.

Eis outro resultado talvez surpreendente desta análise, mais uma inversão do senso comum. Pode defender-se que o componente mais determinante de um motor não são os seus elementos fortemente engenheirados, mas sim o seu meio envolvente natural — a atmosfera, um rio ou qualquer outro reservatório frio disponível. Como vimos, o que efetivamente impulsiona a mudança é um aumento de entropia, e num motor esse aumento é alcançado através da transferência de energia na forma de calor para um reservatório frio. Se tal aumento não ocorrer, o motor é incapaz de funcionar; assim, o elemento essencial é precisamente o local onde a entropia cresce — o meio envolvente. É claro que, para que esse aumento de entropia aconteça, tem de existir um fornecimento de energia sob a forma de calor, proveniente da fonte quente, que após acionar a turbina ou o pistão é parcialmente descarregado no reservatório frio. No entanto, esse fornecimento é, num certo sentido, secundário e até atua contra o objetivo principal, pois a retirada de energia da fonte quente provoca uma pequena diminuição de entropia, o que é exatamente o oposto do que impulsiona o funcionamento do motor. A potência de um motor é, portanto, fundamentalmente um atributo do seu meio envolvente, sendo a fonte quente um mal necessário — indispensável, mas secundário.

Os engenheiros utilizam a conclusão de Carnot na sua busca por melhorar a eficiência dos motores e de uma gama de máquinas relacionadas, como frigoríficos e bombas de calor. Estas aplicações tecnológicas dependem, em última análise, da visão da Natureza que apresentei, segundo a qual o universo afunda gradualmente na desordem, sem princípio orientador algum, exceto leis profundas baseadas na indolência, como a conservação da energia.



Há várias pontas soltas a esvoaçar neste capítulo e preciso de tratar daquelas que mais me chamaram a atenção. Uma dessas pontas soltas é saber se haverá um fim para os acontecimentos. Quando a compreensão termodinâmica da mudança natural emergiu na consciência dos seus criadores, estes depararam-se com a perspectiva do fim do mundo e com a visão da sua chamada «morte térmica». Não se trata apenas de alterações climáticas. A morte térmica, do universo inteiro, ocorre quando a desordem cresce até se tornar completa. Nos seus últimos estertores térmicos, toda a energia do universo terá, segundo esta visão, degradado em movimento térmico caótico (coloquialmente, «calor»), e a possibilidade de um aumento adicional da desordem — e, portanto, a possibilidade de mudança natural — terá desaparecido. Todas as nossas estruturas, processos, realizações e aspirações serão como se nunca tivessem existido, e não haverá qualquer perspectiva de a segunda lei ter uma segunda oportunidade.

Tal cenário pode, de facto, vir a revelar-se o nosso futuro distante e sem características distintivas. Provavelmente está mais longe do que os seus primeiros proponentes temiam, pois hoje sabemos que o universo não é apenas uma esfera finita, mas um balão em expansão — e, ao que tudo indica, em expansão acelerada. A cada dia que passa, à medida que o universo se expande, há mais espaço para o caos. Houve um tempo em que se pensou que o universo poderia um dia colapsar novamente sobre o seu ponto inicial, semelhante a um ovo primordial carregado de potencial, o que levantava preocupações quanto à trajetória futura da entropia. Contudo, essa possibilidade já não é atualmente considerada plausível (embora possamos estar enganados quando pensamos em escalas de tempo de bilhões de bilhões de anos). Já referi que a nossa investigação do universo cobre apenas alguns milhares de milhões de anos, e que uma visão profundamente diferente poderá ser necessária em escalas temporais muito superiores a essa (mencionei também a possível circularidade do tempo). Por agora, ninguém tem ainda qualquer pista segura sobre estas questões.

Existe também a outra extremidade das coisas: o seu início. Pode dizer-se algo acerca da entropia — a medida da desordem — no começo de tudo? Se aceitar a minha premissa inicial de que pouco aconteceu no momento da Criação, então há uma resposta.

Antes do início (deste universo, ou talvez de algum universo precursor original, um Ur-universo), não havia absolutamente nada. Esse Nada tinha necessariamente uma uniformidade perfeita, pois, se não a tivesse, não seria Nada. Se pouco aconteceu quando

o Nada se transformou em algo, então essa uniformidade perfeita ter-se-á mantido (é isso que tenho vindo a supor), e a minha tese desde o início é que o universo recém-nascido teria herdado essa uniformidade do Nada. Na ausência de desordem caótica, a entropia inicial teria sido zero.

O resto é, literalmente, história. À medida que o tempo avançou, ocorreram eventos naturais que reconheceríamos como uma degradação em direção à desordem, globalmente, embora não necessariamente a nível local. Formaram-se estrelas, e também galáxias. Planetas surgiram e desapareceram, biosferas também, tal como batalhas. O pensamento, a arte e a compreensão emergiram — certamente aqui, e esperemos que também noutros lugares, pois são demasiado preciosos para ficarem confinados apenas a nós. Ainda estamos no meio deste desenrolar da Criação, no meio do aumento da desordem, com abatimentos locais que chamamos atributos e artefatos da civilização.

A seta do tempo entra inevitavelmente nesta discussão. A maré imparável do aumento da entropia anda de mãos dadas com a aparente irreversibilidade do tempo, oferecendo-nos um futuro e impedindo-nos de visitar e manipular o passado. Todos os nossos ontens correspondem a épocas de menor entropia global e são irrecuperáveis (felizmente, em muitos casos). Dado que existe tempo, e dado que os acontecimentos ocorrem de forma inexoravelmente acompanhada por um aumento de entropia, apenas o futuro nos está reservado; o passado é história intocável, congelada para sempre na sua imutabilidade. Sim, o aumento da entropia e, em particular, a acumulação local de acontecimentos

vivididos, contribuem para as nossas memórias e para a experiência da passagem do tempo, mas nisso não há nada de particularmente misterioso, à luz dos argumentos que apresentei sobre o processo contínuo e imparável de afundamento no caos.



Ou talvez não. No pano de fundo de tudo o que tenho dito esconde-se um mistério relacionado com as leis da Natureza que seria impróprio ocultar. Todas as leis fundamentais da Natureza parecem ser simétricas relativamente à direção do tempo. Isto é, em profundidade, a Natureza parece ignorar o sentido do tempo; no entanto, à superfície — na nossa experiência — ela parece plenamente consciente dessa direção. Com efeito, as leis da Natureza ou não mencionam o tempo, ou não incorporam qualquer direção temporal privilegiada. Ou os resultados são intemporais (como a lei da conservação da energia) ou, quando as equações são resolvidas, funcionam igualmente bem em ambas as direções do tempo. Um exemplo é a solução das equações de Newton para o movimento de um planeta: é possível seguir a sua órbita para o futuro ou, invertendo o sinal do tempo, para o passado. Não há nada nas equações que imponha que se viaje apenas para a frente no tempo. Como é que, então, o nosso compromisso com o futuro emerge desta aparente indiferença da Natureza relativamente à direção do tempo?

Boltzmann, o herói trágico do Capítulo 4, é aqui relevante, mas num contexto bastante diferente — e a sua contribuição está longe

de ser tão inequívoca como ele próprio supôs. Boltzmann acreditava ter demonstrado que, se se começar com um conjunto arbitrário de moléculas em movimento segundo leis simétricas no tempo, então, independentemente das posições e velocidades iniciais, elas acabarão por se acomodar na distribuição mais aleatória possível. Ou seja, pensava ter mostrado que a assimetria do tempo emergia da simetria temporal simplesmente ao considerar o resultado estatístico coletivo das leis, em vez de se concentrar na trajetória de uma única molécula. Atribuiu a seta do tempo ao comportamento do conjunto, não ao do indivíduo.

Para captar a essência do seu argumento, imagine uma bola numa metade de uma caixa. Está em movimento, a colidir com as paredes. Há uma probabilidade razoável de o seu movimento a trazer de volta, ainda que momentaneamente, ao ponto de partida. Agora considere duas bolas. A situação torna-se mais complexa, pois podem colidir entre si e afastar-se em direções diferentes. Ainda assim, é plausível que, após algum tempo, voltem fugazmente à configuração inicial, embora tal dependa da precisão com que se define o que significa «a mesma configuração». Também é plausível que três bolas regressem à configuração inicial, ou mesmo quatro — mas seria necessário esperar muito mais tempo. Se, porém, considerar cem bolas, mil bolas, ou biliões, então, embora em princípio tal regresso não seja impossível, na prática poderia exigir um tempo comparável à idade do universo. Surge assim uma irreversibilidade prática, apesar de a irreversibilidade não estar presente nas leis que governam a trajetória de cada partícula individual.

Uma explicação mais profunda poderá residir ainda mais fundo na própria trama da realidade. Aquilo que tomamos como uma única seta do tempo poderá, na verdade, resultar do acoplamento de duas setas temporais: uma estatística (a componente boltzmanniana) e outra cosmológica. A seta cosmológica do tempo poderá eliminar até a reversibilidade «em princípio». Enquanto espera, durante quase uma eternidade, que as cem bolas regressem à sua configuração inicial, o universo muda: expande-se. Já não existe a possibilidade de revisitar o estado inicial, pois o espaço-tempo em que esse estado foi definido tornou-se passado. Mesmo em princípio, não se pode esperar reencontrar as bolas na sua configuração original — e quanto mais tempo passa, menos provável isso se torna.

Assim, embora as leis da Natureza possam ser reversíveis no tempo, a sua manifestação no mundo real, marcado por interações complexas e encenada num palco cósmico em transformação, tornou-as irreversíveis na prática. Não há retorno possível.



Tratei de três leis da termodinâmica: a lei zero (sobre a temperatura), a primeira lei (sobre a energia) e a segunda lei (sobre a entropia). Existe ainda uma quarta lei da termodinâmica que, inevitavelmente por causa desse zero tardio, é conhecida como a terceira lei da termodinâmica. Algumas pessoas perguntam-se se ela é realmente uma lei, porque, ao contrário das outras três, não introduz uma nova propriedade física. Talvez isso signifique que se

limita a completar as outras três, encerrando de vez a termodinâmica.

A terceira lei, na forma originalmente proposta pelo químico alemão Walther Nernst (1864–1941) em 1905 — no meio de uma ligeira controvérsia sobre prioridade — afirma, em essência, que *o zero absoluto não pode ser alcançado por um número finito de etapas*. Se alguém estivesse num humor sardonicamente sombrio, poderia interpretar a primeira lei como dizendo que nada acontece, a segunda lei como dizendo que, se algo acontecer, então as coisas pioram, e esta terceira lei como implicando que, de qualquer modo, o fracasso é inevitável. A formulação de Nernst assemelha-se às da segunda lei propostas por Kelvin e Clausius, no sentido em que se refere a observações e não a uma explicação molecular subjacente. Um entendimento mais profundo foi alcançado em 1923, quando os dois químicos norte-americanos Gilbert Lewis (1875–1946) e Merle Randall (1888–1950) encontraram uma forma de expressar a lei em termos moleculares e afirmaram, em essência, que *todas as substâncias perfeitamente cristalinas têm a mesma entropia a zero absoluto*. Não posso, nestas páginas, demonstrar por que razão estas duas formulações — tão diferentes no formalismo — são equivalentes na prática, mas, em termos gerais, isso resulta do facto de que, como todas as entropias tendem para o mesmo valor, torna-se necessário realizar cada vez mais trabalho para extrair energia sob a forma de calor à medida que a temperatura se aproxima de zero, sendo finalmente necessário um trabalho infinito para lá chegar.⁶

Tudo o que a terceira lei afirma é que todas as substâncias têm a mesma entropia a zero absoluto. Não revela qual é esse valor. Contudo, a interpretação de Boltzmann da entropia como uma medida da desordem sugere um valor: zero. Como a substância é um cristal perfeito, todas as suas moléculas ou iões encontram-se dispostos em arranjos perfeitamente ordenados, pelo que não existe desordem devida a imperfeições ou a uma molécula fora do lugar. Como a temperatura é zero, todas as moléculas estão no nível de energia mais baixo possível, pelo que não há desordem resultante de uma molécula vibrar mais do que outra. Estamos perante a perfeição — e a perfeição implica que a entropia seja zero, independentemente da identidade do material. Não admira, portanto, que esse estado seja inacessível.

A terceira lei tem, obviamente, implicações para quem procura atingir temperaturas muito baixas e para a física fascinante que se espera encontrar nesse domínio. Mesmo para os mortais comuns que habitam laboratórios quentes, a lei é essencial, pois o facto de a entropia ser zero sob certas condições fornece um ponto de partida para uma vasta gama de cálculos em termodinâmica, incluindo a previsão numérica de se uma reação química “vai ocorrer” ou não. Esses cálculos dificilmente serão de grande interesse no contexto deste livro, mas convém saber que a terceira lei completa as outras três e torna-as quantitativamente mais úteis do que seriam isoladamente.

Digo que completa as outras; mas poderá haver uma quinta, sexta... lei da termodinâmica? Ninguém sabe, embora alguns afirmem que poderão existir mais. A termodinâmica convencional,

em particular a segunda lei, trata de tendências para a mudança e de sistemas em equilíbrio, sem tendência para sofrer novas alterações. Existe algum interesse em formular versões da termodinâmica que lidem com a taxa a que essa tendência inicial se concretiza, como a taxa de produção de entropia num processo longe do equilíbrio e afastado dele, como um corpo humano vivo, em que o equilíbrio é a morte. Essas estruturas dinâmicas foram estudadas pelo químico belga de origem russa Ilya Prigogine (1917–2003), trabalho que lhe valeu o Prémio Nobel em 1977; contudo, certos aspetos da sua obra — e a sua visão de que o determinismo na Natureza estaria morto — permanecem controversos e, para alguns, anátemas, verdadeiramente para além da anarquia.⁷



Procurei mostrar que a Natureza, deixada a si própria, se degrada gradualmente, mas, nesse processo, faz emergir abatimentos locais do caos que podem ser requintados. A segunda lei da termodinâmica encapsula esta tendência da matéria e da energia para se dispersarem e fornece um entendimento profundo da força motriz não dirigida que está por detrás de todos os fenómenos naturais. Considero extraordinário que um princípio quotidiano tão simples possa dar conta de toda a mudança. Mostrei que as suas implicações incluem a eficiência dos motores e, por essa via, a das economias; e, no entanto, escondido por baixo dele, permanece o problema de saber por que razão leis da Natureza simétricas no tempo dão origem ao voo singular da seta do tempo.

A segunda lei é filha da anarquia, e, ainda assim, gera o humilde e o extraordinário.

CAPÍTULO 6

O Poder Criativo da Ignorância — Como a Matéria Responde à Mudança

A ignorância é uma aliada eficaz da indolência e da anarquia. Neste capítulo quero mostrar como o não saber pode ser usado de forma construtiva para chegar ao saber. A lei da Natureza que pretendo inicialmente esclarecer foi uma das primeiras a ser expressa quantitativamente, quando os cientistas começaram a perceber a importância de atribuir números à Natureza, mas só veio a ser verdadeiramente compreendida no final do século XIX. A sua elucidação nasce, precisamente, da ignorância.

A lei em causa diz respeito à forma estruturalmente mais simples da matéria, um gás, e foi estabelecida por Robert Boyle (1627–1691), a trabalhar em Oxford no início da década de 1660, ou, como os franceses reivindicam, por Edme Mariotte (1620–1684), a trabalhar em Paris em 1679. Foi posteriormente desenvolvida por Jacques Charles (1746–1823) quando, como tantas vezes acontece, o estudo da Natureza foi impulsionado pelas exigências e oportunidades do avanço tecnológico — no seu caso, pelo crescente interesse em voar em balões. Estas leis são de interesse histórico porque estão entre os primeiros resumos quantitativos das propriedades da matéria, isto é, expressos de forma aberta ao

cálculo numérico e à previsão. São também a base do desenvolvimento da termodinâmica e da sua aplicação a fenômenos e processos químicos e de engenharia, pelo que têm uma importância fundamental e prática enorme.

Já esbocei a formulação da lei de Boyle no Capítulo 1, pelo que aqui basta recordar brevemente o seu conteúdo. Boyle — e, de forma independente, Mariotte, pois as notícias viajavam lentamente na época — estabeleceram que *a pressão exercida por um gás é inversamente proporcional ao volume que ocupa*. Assim, ao diminuir o volume ocupado por um gás, a sua pressão aumenta. Em termos concretos, se um gás for confinado a metade do seu volume inicial empurrando um êmbolo, a sua pressão duplica. Hoje em dia, não temos grande dificuldade em explicar qualitativamente esta lei. A imagem moderna de um gás é a de um enxame de moléculas que se deslocam incessantemente e de forma caótica através de um espaço praticamente vazio. Quando o gás é comprimido, o mesmo número de moléculas passa a ocupar um volume menor, mas continua a mover-se com a mesma velocidade média (porque a temperatura se mantém constante, e é a temperatura que determina a velocidade). Como consequência da maior densidade molecular, num dado intervalo de tempo mais moléculas colidem com as paredes do recipiente. A força que exercem sobre essas paredes é sentida como pressão e, como a força total é agora maior, também a pressão o é. O verdadeiro desafio, porém, é explicar o aspeto quantitativo da lei: a relação numericamente precisa entre pressão e volume quando a temperatura é mantida constante.

Charles acrescentou a esta observação o estudo do que acontece quando a temperatura é autorizada a variar. Voar, no final do século XVIII, foi inicialmente conseguido com balões de ar quente. Uma ovelha, um pato e um galo foram a carga — sem dúvida aterrorizada — do primeiro voo, em 19 de setembro de 1783, num balão de ar quente construído pelos irmãos Montgolfier, Joseph e Étienne. Algumas semanas depois, o ser humano deu o seu primeiro passo, pequeno, arriscado, mas prenhe de significado, rumo aos céus. Balões de hidrogénio (e pouco depois balões de gás de cidade, mais facilmente disponível, mas menos eficaz para a sustentação e, além disso, venenoso e inflamável) começaram rapidamente a subir, sendo então mais práticos do que os balões de ar quente, até que, na década de 1950, os balões modernos alimentados por gás engarrafado reconquistaram os céus. Os balões cheios de gás evitavam a necessidade de transportar um braseiro em chamas no cesto, como acontecia nos balões de ar quente. Além disso, estes últimos só conseguiam manter-se no ar enquanto durasse o pesado combustível. Em ambos os tipos de balão havia um grande interesse em compreender como a temperatura afetava o ar e, consequentemente, a capacidade de sustentação. Nos balões de ar quente, a elevação resultava da menor densidade do ar aquecido; nos balões de gás, a flutuabilidade dependia da temperatura do ar circundante, que varia com a altitude. Outro investigador precoce das propriedades dos gases, Joseph Gay-Lussac (1778–1850), utilizou um balão em 1804, juntamente com um colega, para estabelecer o que então foi considerado um recorde mundial de altitude — que afirmaram, com precisão suspeita, ser de 7016 metros (20.018 pés) — numa tentativa audaz de analisar a variação da composição e das propriedades da atmosfera com a altitude.

Numa série de experiências, Charles, ele próprio um pioneiro do balonismo, estabeleceu o que hoje chamamos a lei de Charles: mantendo-se constante o volume, a pressão exercida por uma quantidade fixa de gás aumenta proporcionalmente à temperatura. Isto é, se a temperatura duplicar, a pressão duplica. É necessário ter cuidado aqui, porque a “temperatura” desta lei é a temperatura absoluta, normalmente expressa na escala de Kelvin que introduzi no Capítulo 4. A lei não funciona para as escalas mais artificiais de Celsius e Fahrenheit. Assim, se o gás estiver inicialmente a 20 °C, isto é, 293 K, será necessário duplicar para 586 K (correspondentes a 313 °C) para duplicar a pressão; não basta duplicar 20 e esperar que a pressão duplique a uns amenos 40 °C.

As leis de Boyle e de Charles podem ser combinadas numa única lei, a lei do gás perfeito, que em palavras afirma que *a pressão de um gás é inversamente proporcional ao volume e diretamente proporcional à temperatura absoluta*.¹ Encontrará também esta expressão designada como lei do gás ideal; pode tratá-las como sinónimas.² A lei é “universal” no sentido em que se aplica a todos os gases, independentemente da sua identidade química, e também a misturas, como o ar. Além disso, a forma matemática da lei envolve apenas uma constante fundamental, conhecida de forma pouco imaginativa como a “constante dos gases”, e essa constante é a mesma para todos os gases. Na verdade, a constante dos gases já estava presente, embora disfarçada, no Capítulo 4, pois não passa da constante de Boltzmann sob outra forma. Esse disfarce permite que a constante dos gases apareça em muitas expressões que nada têm a ver com gases, como, por exemplo, na expressão que permite calcular a voltagem de pilhas elétricas.

No Capítulo 1 introduzi o conceito de lei limite, uma lei que se torna progressivamente mais fiável à medida que a substância descrita é removida e que se aplica exatamente quando toda a substância desapareceu. A lei do gás perfeito é uma dessas leis limite, no sentido em que se torna mais rigorosa à medida que a pressão é reduzida a zero ou, de forma equivalente, à medida que o volume ocupado pelo gás tende para o infinito. A lei do gás perfeito descreve a “gaseidade perfeita”, o comportamento que seria observado se não existissem fatores complicativos, como as moléculas se atraírem momentaneamente em vez de se moverem completamente livres, ou o facto de o espaço disponível para uma molécula durante o seu movimento caótico já estar parcialmente ocupado por outra. Quando o volume ocupado pelo gás é muito grande, as moléculas encontram-se tão raramente (no limite de um espaço infinitamente vasto entre elas, nunca) que são praticamente indiferentes à presença umas das outras. Na prática, isso significa que a lei do gás perfeito é obedecida de forma perfeita no limite de volume infinito ou de pressão zero. Em suma, a lei do gás perfeito é obedecida exatamente por todos os gases apenas quando eles... não estão lá.

Apesar desta última observação, as leis limite estão longe de ser inúteis. Como muitas leis deste tipo, a lei do gás perfeito revela-se um ponto de partida sensato para a discussão de sistemas reais — neste caso, de todos os gases em condições correntes —, sendo necessário acomodar idiossincrasias individuais apenas quando comportamentos mais complexos se tornam relevantes. É um pouco como considerar que o percurso “perfeito” entre dois pontos é uma linha reta, como se diz que os corvos voam: pode ser uma

forma sensata de começar a pensar numa viagem, embora na prática seja necessário seguir as estradas reais. A linha reta é o percurso “limite” na ausência das várias irregularidades da paisagem. Na prática, verifica-se que a lei do gás perfeito é fiável às pressões normais encontradas nas aplicações do quotidiano, mas são necessárias correções para trabalhos muito precisos, quando as pressões são anormalmente elevadas ou quando as temperaturas são tão baixas que o gás está perto de se condensar num líquido. De facto, a lei do gás perfeito é de enorme importância porque está na base do formalismo — no sentido de constituir o ponto de partida para a formulação de expressões — de grande parte da termodinâmica e das suas aplicações.

Existem várias leis limite na ciência, todas elas exprimindo a “essência perfeita” de uma determinada propriedade, e todas servindo como pontos de partida úteis para descrever comportamentos mais elaborados. As que tenho em mente, mas que aqui apenas mencionarei, dizem respeito às propriedades de misturas líquidas e ao efeito de substâncias dissolvidas nas propriedades dos solventes. A maioria foi identificada nos primórdios da química moderna e ostenta o nome dos seus descobridores, entre os quais se incluem o químico inglês William Henry (1774–1836), no que respeita à solubilidade de gases em líquidos — relevante para o fabrico de água gaseificada e champanhe e para a ocorrência das “bends” (*doença da descompressão, causada pela libertação de bolhas de gás no sangue quando a pressão diminui rapidamente*) em mergulhadores de águas profundas —, o químico francês François-Marie Raoult (1830–1901), sobre a forma como uma substância dissolvida afeta

as propriedades de uma solução, como o efeito do sal no ponto de congelação da água, e o químico neerlandês Jacobus van 't Hoff (1852–1911), no que respeita à propriedade vital da osmose (do grego “empurrar”), a aparente capacidade de um solvente atravessar uma membrana, mantendo, entre outras coisas, as células biológicas túrgidas e saudáveis, as plantas viçosas e as árvores alimentadas com nutrientes.³ Para além de ilustrarem como um académico suficientemente madrugador pode conquistar o “verme da imortalidade” ao identificar um comportamento simples e sistemático, e de evidenciarem o carácter agradavelmente internacional da empresa científica, nenhuma destas leis é de grande relevância direta para a discussão presente, apesar de serem de considerável importância para químicos físicos como eu. As três podem ser reconduzidas ao colapso anárquico para o caos capturado e expresso pela segunda lei da termodinâmica. Em poucas palavras, as plantas murchariam, as colheitas falhariam e nós morreríamos na ausência destas simples consequências da segunda lei. Seriam, de facto, leis verdadeiramente... limitantes.



Postas de lado essas apologias, é tempo de ver como a ignorância se alia à indolência e à anarquia para iluminar a lei do gás perfeito, o alvo inicial e principal deste capítulo. Já sabe que um gás é constituído por moléculas num estado de movimento caótico incessante, zigzagueando de um lado para o outro, chocando entre si e partindo depois numa nova direcção e com uma velocidade diferente. Você e eu estamos imersos numa tempestade de

moléculas mesmo nos dias mais calmos, com a nossa superfície a ser constantemente fustigada por esse turbilhão contínuo, que, em grande medida, nos mantém com forma. Para imaginar a escala da agitação num gás como o ar, pense numa molécula como tendo o tamanho de uma bola de ténis a atravessar cerca do comprimento de um campo de ténis antes de colidir com outra molécula. Toda esta atividade caótica decorre sob o regime da mecânica clássica, esse afloramento da indolência e da anarquia.

A ignorância de que o conhecimento irá brotar é a nossa total falta de informação sobre o que está a acontecer a uma molécula individual neste turbilhão de atividade incessante entre miríades de outras. Embora cada colisão individual contribua para a pressão, não há qualquer necessidade de seguir a trajetória de cada molécula em particular. Tal como em sociologia é possível prever com razoável confiança o comportamento de uma multidão, mas não o de um indivíduo dessa multidão, também aqui, porque estão presentes miríades de moléculas no conjunto a que chamamos gás, podemos confiar em ignorar a contribuição de moléculas individuais precisamente porque nada sabemos sobre elas. À semelhança dos sociólogos quantitativos, precisamos de recuar e focar-nos não no indivíduo, mas na multidão.

Quando isso é feito, e a matemática do enxame é desenvolvida usando as várias leis de Newton, emerge a lei de Boyle.⁴ Pode parecer que isto é tão óbvio que não exigiria grande aparato matemático para lá chegar. Em certo sentido, isso é verdade; mas, na realidade, o resultado matemático é mais rico do que a simples lei de Boyle, pois mostra como a pressão e o volume dependem

conjuntamente de várias características do gás, como a massa de uma das suas moléculas e a velocidade média das moléculas — uma dependência que, creio, não poderia ser inferida apenas a partir de uma interpretação pictórica da lei. Uma vez mais, a aplicação da matemática para exprimir quantitativamente uma imagem qualitativa enriquece o entendimento e aprofunda a compreensão.



E quanto à lei de Charles, segundo a qual a pressão é proporcional à temperatura absoluta? Mais uma vez, a ignorância em aliança com a matemática (discretamente usada) será útil e conduzirá a uma explicação dessa lei.

A chave está na ligação entre a velocidade molecular e a temperatura. A velocidade desempenha um papel duplo. Se as moléculas se movem rapidamente, mais delas conseguem embater na parede num dado intervalo de tempo e, quando o fazem, o impacto que exercem é maior. Assim, ao elevar a temperatura, a pressão exercida pelo gás deve aumentar por ambas as razões: tanto pela maior frequência de impactos como pela maior intensidade dos choques. Sabe-se (voltarei a este ponto dentro de momentos) que a velocidade média das moléculas num gás é proporcional à raiz quadrada da temperatura (a temperatura absoluta); como o produto de duas raízes quadradas da temperatura dá a própria temperatura, chegamos assim a uma explicação da lei de Charles: a pressão é proporcional à temperatura.⁵

A questão decisiva é perceber por que razão a velocidade média das moléculas num gás é proporcional à raiz quadrada da temperatura. Vejamos primeiro o que isso significa na prática para o ar. A velocidade média depende da massa das moléculas: moléculas leves, como o azoto, deslocam-se a cerca de 500 metros por segundo (1120 milhas por hora) a 20 °C, enquanto moléculas mais pesadas, como o dióxido de carbono, avançam mais lentamente, a cerca de 380 metros por segundo (850 milhas por hora). Estes valores oferecem, aliás, uma pista para outro fenómeno físico: a propagação do som. A velocidade do som no ar é de cerca de 340 metros por segundo (760 milhas por hora) ao nível do mar, valor próximo destas velocidades moleculares. O som é, de facto, uma onda de pressão que depende do reajustamento coletivo das posições das moléculas de ar, produzindo uma ondulação da pressão; a rapidez com que isso ocorre depende da velocidade com que as moléculas se podem mover. Não é surpreendente, portanto, que a velocidade do som seja comparável à velocidade média das moléculas. O nosso foco aqui, porém, é a dependência da velocidade molecular em relação à temperatura e, como essa velocidade varia com a raiz quadrada da temperatura, é fácil calcular que a velocidade média diminui cerca de 4% ao passar de um dia ameno a 20 °C (298 K) para um dia frio a 0 °C (273 K).

É necessário explicar por que motivo a velocidade média das moléculas é proporcional à raiz quadrada da temperatura, pois assim teremos explicado plenamente a lei de Charles. As moléculas de um gás possuem apenas energia cinética — a energia do movimento — porque, no limite em que se encontram afastadas a maior parte do tempo, praticamente não interagem entre si e,

portanto, não têm energia potencial associada às suas posições relativas. A velocidade está ligada à energia cinética, pelo que uma via indireta para calcular a velocidade média consiste em determinar a energia cinética média das moléculas e depois interpretar esse valor médio em termos de velocidade média. No Capítulo 4 viu como Boltzmann mostrou ser possível calcular propriedades médias, imaginando o lançamento de livros (moléculas) sobre prateleiras (níveis de energia) e identificando o resultado mais provável sem impor qualquer orientação (exceto a exigência de que a energia total tenha um valor fixo). Quando esse procedimento é aplicado a um gás, emerge do cálculo uma expressão para a energia cinética média das moléculas e, a partir dela, para a respetiva velocidade média. E, de facto, essa velocidade média é proporcional à raiz quadrada da temperatura, exatamente como é necessário para explicar a lei de Charles.

E não fica por aqui. Já viu que a combinação das leis de Boyle e de Charles conduz à lei do gás perfeito, a lei-limite que serve de ponto de partida para tantas aplicações da termodinâmica. Conhece agora a sua origem: o contributo da lei de Boyle resulta da consideração do número de impactos das moléculas do gás, e o contributo da lei de Charles resulta do papel da velocidade molecular e da sua dependência da temperatura.

Espero que consiga apreciar o carácter vertiginoso desta conclusão. Da ignorância — neste caso, de não conhecermos quaisquer detalhes do comportamento individual — destilou-se uma lei da Natureza, a lei do gás perfeito. Pelo caminho, emergiu uma compreensão da velocidade média das moléculas num gás e da sua

variação com a temperatura e (embora não o tenha mencionado aqui, exceto nas Notas) com a massa das moléculas que constituem o gás. A ignorância, quando devidamente organizada e aplicada, pode ser uma poderosa fonte de compreensão.



A lei dos gases perfeitos é apenas uma das pequenas leis da Natureza a que me referi no Capítulo 1. São leis dependentes, estas leis marginais, pendendo como frutos das árvores das grandes leis, as leis-mãe. Talvez existam outras leis que, uma vez aceites as leis-mãe como consequências da indolência e da anarquia, emergem de uma aplicação adicional da ignorância.

Uma dessas leis marginais é a lei de Hooke, que também mencionei no Capítulo 1. Robert Hooke (1635–1703) foi um dos pensadores verdadeiramente imaginativos do século XVII, numa época em que a confusão e a obscuridade iam sendo afastadas pela maré crescente do pensamento racional representado pelo Iluminismo. Influenciou profundamente o pensamento de Newton e produziu ideias notáveis por mérito próprio. Como vimos no Capítulo 1, a lei de Hooke afirma que, ao esticar uma mola, a força restauradora é proporcional ao deslocamento. Assim, puxe uma mola um centímetro a partir da posição de repouso e sentirá uma resistência; puxe-a duas vezes mais e sentirá uma resistência duas vezes maior.⁶ Uma consequência desta lei — consequência que decorre diretamente da mecânica newtoniana, fundada na anarquia — é que as molas oscilam de forma regular, tal como os pêndulos,

e que, por esse motivo, o mundo mecânico consegue medir o tempo.

A lei de Hooke é outro exemplo de uma lei limite, sendo estritamente válida apenas quando não há deslocamento em relação ao equilíbrio. É exatamente verdadeira para molas esticadas, desde que não estejam esticadas, e exatamente verdadeira para pêndulos oscilantes, desde que não estejam a oscilar. Todas as molas e pêndulos exibem desvios da lei para extensões e oscilações mensuráveis, mas conformam-se cada vez mais com ela à medida que essas extensões e oscilações se aproximam de zero. Na maioria dos casos, o desvio é negligenciável e a lei pode ser usada para fazer previsões fiáveis — e os relógios mantêm um tempo excelente. Mas estique demasiado e, como um elástico excessivamente tensionado, a lei quebra.

Onde entra a ignorância na explicação da lei de Hooke? Se se for ignorante quanto a fatores externos, chega-se à conclusão de Hooke sempre que exista uma força que se oponha a um deslocamento. O argumento é o seguinte. Considere qualquer propriedade e pense em como ela muda quando um parâmetro é alterado. Por parâmetro entendo qualquer coisa que possa ser ajustada: a extensão de uma mola, o ângulo de um pêndulo, o comprimento de uma ligação numa molécula, a pressão aplicada a um sólido, e assim por diante. Em cada caso, pode existir uma propriedade do objeto ajustado que depende da magnitude do ajuste e que atinge — ou atravessa — um mínimo quando o ajuste é zero. Por exemplo, pode tratar-se da energia de uma mola que está a ser estendida ou comprimida. Se imaginar um gráfico da

energia da mola em função do ajuste, ele terá a forma de uma curva que sobe em ambos os lados do ajuste nulo. Assim, a energia de uma mola aumenta quer quando é estendida quer quando é comprimida, sendo mínima quando está em repouso. A menos que intervenha um comportamento muito estranho, todas essas curvas — qualquer que seja a propriedade e qualquer que seja a natureza do ajuste — começam da mesma forma a partir do valor correspondente ao ajuste zero. Essa forma inicial da dependência da energia em relação ao deslocamento é parabólica e implica (de acordo com a mecânica de Newton) que a força restauradora é proporcional ao deslocamento, exatamente como afirma a lei de Hooke.⁷ Assim, não saber como a força se comporta conduz-nos a saber como ela é mais provável comportar-se.



Gostaria de regressar ao ponto que referi há pouco acerca de como a lei de Hooke sustenta a marcação do tempo no mundo — ou, pelo menos, como o fez quando os relógios eram governados por pêndulos oscilantes e os relógios de pulso por rodas de balanço oscilantes. Será possível identificar uma ligação entre uma força restauradora proporcional ao deslocamento e a batida regular de oscilação necessária para medir o tempo, sem resolver as equações da mecânica clássica? Existirá uma explicação profunda, talvez insuspeita, subjacente à regularidade da oscilação?

A regularidade, quer no espaço quer no tempo, sugere uma simetria subjacente. Precisamos de a identificar. Neste caso, devido

à dependência da energia cinética da velocidade (e, portanto, do momento linear) do peso oscilante do pêndulo, e da energia potencial à medida que esse peso sobe e desce sob a influência da gravidade, existe um fluxo regular de uma forma de energia para a outra e de volta novamente. Uma, a energia cinética, depende do quadrado do momento linear; a outra, a energia potencial, depende do quadrado do deslocamento.⁸ Cada forma de energia alimenta a outra, e o resultado é uma troca regular. Ao observarmos o movimento de um pêndulo, vemo-lo parar momentaneamente nos pontos de inversão, nos limites da oscilação, onde a energia cinética é nula, mas a energia potencial é máxima. Essa energia potencial reconverte-se em energia cinética à medida que o pêndulo acelera. Quando passa brevemente pela vertical, o pêndulo move-se à maior velocidade: perdeu toda a energia potencial e possui agora energia cinética máxima. Em seguida, essa energia cinética reconverte-se em energia potencial à medida que o pêndulo abrande e sobe do outro lado da oscilação. A simetria desse fluxo persiste independentemente da amplitude da oscilação, e o pêndulo bate como o coração de um relógio.

Eis o ponto que pretendo salientar para aprofundar ainda mais esta simetria, mas que requer alguma preparação. Existem duas formas de descrever o mundo. Pode ser descrito em termos das posições das coisas, ou em termos dos momentos das coisas (introduzi o momento linear no Capítulo 2 como massa vezes velocidade). A minha descrição do Nada até agora foi feita em termos de posição: afirmei que o Nada é obviamente uniforme, pois a existência de saliências ou depressões localizadas no absolutamente nada é inconcebível, ou pelo menos um oxímoro.

Mas suponha que coloca uns óculos diferentes e observa o Nada em termos de momentos. Este procedimento não é tão extravagante quanto possa parecer, pois existem técnicas de investigação da matéria que fazem precisamente isso. Por exemplo, muitos dos extraordinários avanços na biologia (como a estrutura do ADN) foram obtidos desse modo. Refiro-me à técnica da difração de raios X, na qual um feixe de raios X atravessa um cristal e emerge espalhado (tecnicamente, difratado), produzindo um padrão de pontos cuja análise se traduz em informação sobre a disposição espacial. Esses pontos são, essencialmente, a estrutura da molécula observada através de "óculos de momento".⁹

Provavelmente aceitará que, ao olhar para o Nada com estes novos óculos, continua a ver nada. Esta nova visão do Nada é tão uniforme quanto a visão baseada na posição. Sendo assim, e aceitando que pouco aconteceu na Criação, imediatamente após o início do universo ele permanece uniforme também em termos de momentos lineares. Uma consequência dessa uniformidade é que as leis da Natureza são independentes das velocidades atribuídas às coisas. Desde que o observador e o observado se desloquem à mesma velocidade (de modo a não termos de considerar a relatividade), as leis são as mesmas. Por exemplo, a oscilação de um pêndulo obedece à mesma lei quer o relógio que ele aciona se desloque a 100 metros por segundo, quer esteja estacionário. (Relataria desvios se estivesse parado e o pêndulo em movimento relativamente a si, mas isso já é relatividade e um assunto completamente diferente: ver Capítulo 9.)

O universo possui uma simetria profunda quando descrito em termos de posições e de momentos lineares. Um pêndulo exibe essa simetria. A energia de um pêndulo oscilante (aliás, de qualquer oscilador, incluindo uma massa presa à extremidade de uma mola) tem contribuições equivalentes e simétricas no espaço e no momento. À medida que oscila, a energia flui do momento linear para o deslocamento e depois regressa. Se retirasse os óculos do cotidiano e colocasse os óculos do momento, não veria qualquer diferença. Onde antes via deslocamento, veria agora momento, e vice-versa. É esta simetria que está na base da oscilação rítmica e incessante do pêndulo, da massa na extremidade de uma mola ou da oscilação contínua da roda de balanço.



A ignorância levou-nos longe neste capítulo e emergiu, quando devidamente canalizada, como base do conhecimento. Um dos tipos de ignorância em causa é a ausência de conhecimento sobre o comportamento de entidades individuais, o que nos obriga, como sociólogos, a avaliar o comportamento do conjunto — no nosso caso, um gás de moléculas — e a descobrir que esse conjunto obedece a certas leis que são depois utilizadas com grande eficácia na termodinâmica. Houve ainda um segundo tipo de ignorância, relativo ao comportamento de entidades — refiro-me a pêndulos e molas, mas também a muitos outros sistemas — quando são ligeiramente afastadas do repouso e que, salvo casos especiais, tendem a comportar-se de uma forma característica comum, manifestando-se como leis que explicam esse comportamento. Esse

comportamento, como bônus, revela também outra simetria profunda do universo, que remonta à sua origem a partir do Nada uniforme.

CAPÍTULO 7

A Carga da Brigada Ligeira *— As Leis da Eletricidade e do Magnetismo*

As leis da eletricidade e do magnetismo desempenham um papel especial na Natureza, não apenas pelas formas como sustentam a nossa existência e estrutura, mas também porque constituem os alicerces da maioria das atividades, prazeres e passatempos da vida. A energia radiante do Sol chega até nós sob a forma de radiação eletromagnética e, onde incide, impulsiona a fotossíntese, construindo a biosfera, revestindo a superfície inorgânica da Terra com florestas, campos e pradarias, e esverdeando os mares. Através dessa revitalização, povoa a terra, o mar e o ar com organismos móveis e, em última instância, fornece os nossos pequenos-almoços e, através deles, a nossa criatividade e fruição. De forma ainda mais fundamental, as forças eletromagnéticas mantêm unidos os átomos e as moléculas e são responsáveis pela existência das múltiplas formas de matéria tangível. Essas forças são também responsáveis pelo transporte e pela comunicação, que vão do essencial ao trivial, mas que se tornaram uma parte inescapável do prazer humano, do massacre, do comércio, da persistência e, numa palavra, da existência.

Introduzi sub-repticiamente outro aspeto da importância destas leis ao transitar da eletricidade e do magnetismo para o

«eletromagnetismo», a sua unificação. A unificação de todas as forças — mostrar que são todas manifestações de uma única força-mãe — é um dos santos graais da física. A sua concretização, atualmente longe de estar completa, mas sedutoramente próxima, mostraria que uma coleção aparentemente heterogênea de forças possui uma unidade subjacente e que o mundo é mais simples do que as aparências, tomadas literalmente, sugerem. A correspondente coleção heterogênea de leis que descrevem a atuação dessas forças fundir-se-ia então numa única lei, tornando mais provável a descoberta da sua origem.

A simetria desempenha um papel central na procura desse graal. Já toquei no seu papel noutros contextos em capítulos anteriores: a identificação, por Emmy Noether, da ligação profunda entre simetria e leis de conservação no Capítulo 2; a minha visão de que a uniformidade do Nada sobrevive ao nascimento do universo; e, mais recentemente, a minha interpretação dos relógios como a face pública de uma simetria oculta que permeia o submundo das propriedades.

Se desejar uma analogia concreta do papel da simetria na unificação da eletricidade e do magnetismo (e, por extensão, das restantes forças), a seguinte poderá ajudar. Imagine um quadrado a representar a eletricidade e um hexágono regular a representar o magnetismo. Estas duas formas são bastante distintas e não há maneira de torcer ou rodar uma para a transformar na outra. A eletricidade quadrada e o magnetismo hexagonal parecem entidades separadas. Agora imagine um cubo. Quando o observa de frente, vê um quadrado; quando o observa ao longo de uma

diagonal do corpo (que atravessa vértices opostos), vê um hexágono. Assim, se em vez de pensarmos nas duas forças como entidades separadas ampliarmos a nossa visão, subirmos uma dimensão e as pensarmos combinadas num cubo, então a eletricidade quadrada e o magnetismo hexagonal são claramente manifestações de uma única entidade, relacionadas por rotações num espaço abstrato. O cubo é o eletromagnetismo.

Terei mais a dizer nesta linha e, à medida que este capítulo se desenvolve, poderá ser útil pensar numa extensão da analogia do cubo. Como já referi, um dos grandes impulsos da física teórica moderna é unificar todas as forças do mundo e, em particular, demonstrar que a força eletromagnética unificada é apenas uma face de uma força única. Esta unificação adicional já foi alcançada com a «força fraca», a força que atua no interior dos núcleos atômicos, separando os seus constituintes e dando origem à radioatividade — a radiação emitida por um núcleo sob a forma de raios gama (fotões de radiação eletromagnética de comprimento de onda muito curto e energia muito elevada) ou de partículas carregadas (raios alfa e beta). A atenção centra-se atualmente em certos aspetos da «força forte», a força de alcance muito curto que mantém unidos os constituintes de um núcleo, apesar da força eletromagnética tentar separar as partículas eletricamente carregadas, densamente compactadas. Devemos agradecer que a força forte não tenha um alcance tão longo como o do eletromagnetismo, pois, se o tivesse, nós e tudo o que existe seríamos sugados para um único átomo colossal. Talvez a gravitação venha um dia a integrar-se nessa unificação, embora pareça ter características enigmáticas relacionadas com a própria

estrutura do espaço-tempo. A unificação de todas as forças será alcançada quando a elaboração do cubo — e as suas rotações e outras manipulações num espaço abstrato de muitas dimensões — tiver sido identificada. A unificação é a procura de um graal; esse graal não é um cálice: é um cubo atualmente inimaginavelmente elaborado, num espaço abstrato de múltiplas dimensões.

Concentrar-me-ei neste capítulo nas leis do eletromagnetismo, sobretudo porque são as mais familiares — ou, pelo menos, as menos estranhas. (De qualquer modo, revirei alguns dos seus aspetos para destilar a sua essência.) Observações semelhantes aplicar-se-ão às forças fraca (certamente) e forte (penso eu), pelo que o meu objetivo é identificar a origem destas leis familiares e deixar à sua imaginação aceitar que observações análogas se aplicam às outras forças. Reconheço que isto é um salto de fé, pois os físicos ainda lutam para as incorporar num esquema unificado.

Para que não perca de vista para onde me dirijo, recordo que quero defender que as leis do eletromagnetismo são mais uma consequência da indolência e da anarquia, emergindo do facto de praticamente nada ter acontecido no início do universo, quando algo surgiu, essencial e presumivelmente de forma espontânea, a partir do Nada. Há vários aspetos que precisamos agora de considerar.



Uma das primeiras leis do eletromagnetismo a ser formulada foi a lei de Coulomb. O físico francês Charles-Augustin de Coulomb

(1736–1806) formulou, em 1784, a lei que hoje leva o seu nome e propôs que a força entre duas cargas elétricas enfraquece com o quadrado da distância que as separa. Trata-se da chamada lei do inverso do quadrado. A mesma conclusão tinha sido alcançada por outros (incluindo Joseph Priestley e Henry Cavendish, ambos em Inglaterra), mas Coulomb é geralmente creditado por uma investigação sistemática e pela formulação clara da lei. Já existiam indícios anteriores de que a dependência inversa do quadrado da separação seria plausível, pois a gravitação obedece à mesma dependência.¹

Suspeito que, se fosse Deus e, apesar das suas decisões aparentemente caprichosas noutros contextos, quisesse oferecer à humanidade a lei de interação elétrica mais bela que pudesse conceber — e que ela fosse capaz de apreciar —, nos daria a lei de Coulomb. É uma lei de beleza singular, mas essa beleza é mais profunda do que aparenta e passa despercebida ao olhar do observador casual.

Primeiro, e de forma algo trivial, a força que descreve é esfericamente simétrica, como uma bola — esse corpo de simetria mais perfeita em três dimensões. Explicarei este ponto de seguida. Pode quantificar-se o «mais» nessa avaliação? Sim: uma esfera possui um número infinito de eixos de simetria (qualquer diâmetro) e um número infinito de ângulos de rotação em torno de qualquer desses eixos que a deixam aparentemente inalterada. Pense agora num espelho a cortar a esfera pelo centro e a refletir um hemisfério no outro: esse espelho pode assumir um número infinito de orientações. Em três dimensões, não existe objeto com maior

simetria: a esfera é o objeto tridimensional mais simétrico, com a sua plenitude de infinitos. Se se inclinar a identificar simetria com beleza, então a esfera é a personificação da beleza — ou, pelo menos, de uma certa percepção de beleza primordial.

A lei de Coulomb é esférica no seguinte sentido: a direção de uma carga relativamente à outra é irrelevante. A intensidade da interação decresce com a distância da mesma forma e na mesma medida em qualquer direção. Isto pode não parecer particularmente empolgante, mas tem consequências profundas para as estruturas dos átomos e, por consequência, para as propriedades da matéria. Além disso, aqui surge mais uma pista sobre a importância do Nada. Uma visão simplista (mas que pode, ainda assim, ser correta) é que a simetria esférica da força de Coulomb emerge da uniformidade — e, especificamente, da simetria esférica — do absolutamente nada. Quando a força emergiu (tema que discutirei de forma mais sofisticada mais adiante neste capítulo), o Nada evoluiu para um meio de propagação da força e, ao emergir do nada absoluto, não lhe foi imposto qualquer constrangimento adicional. A indolência reside no coração de Coulomb.

Segundo, apesar de tudo o que disse sobre a esfera possuir simetria infinita, a lei de Coulomb é mais do que esfericamente simétrica. Possui uma simetria interna que não é discernível apenas ao notar a distância na sua expressão e a ausência de referência à direção. Se permitirmos que o olhar do observador casual se torne mais sofisticado e, em vez de avaliarmos a interação nas nossas habituais três dimensões, dermos um passo para uma quarta, essa

simetria esférica mantém-se e torna-se aquilo a que chamarei hiperesférica.²

Reconheço que provavelmente estou a pedir à sua imaginação visual que ultrapasse os seus limites (tal como à minha), mas tal é o poder da matemática que ela consegue dar esse passo e demonstrar simbolicamente que o que digo é verdadeiro. Posso oferecer-lhe uma pista visual do que está envolvido ao passar de três para quatro dimensões, fazendo primeiro a transição de duas para três e pedindo-lhe depois que aceite que algo semelhante ocorre ao subir para quatro. Já viu que um quadrado e um hexágono podem ser relacionados quando passamos de um plano bidimensional para um cubo tridimensional; peço-lhe agora que pense de forma análoga, embora a questão aqui seja um pouco diferente.

Eis a imagem que quero que tenha em mente. Imagine uma folha quadrada de papel com um grande círculo vermelho no centro. Agora imagine a mesma folha, mas com metade pintada de vermelho e a outra metade em branco. Claramente, não existe relação entre os dois padrões. Ou existirá? Um quadrado e um hexágono tornaram-se relacionados quando o convidei a subir uma dimensão e a pensar num cubo; poderá o mesmo acontecer se fizermos algo semelhante com o círculo e os retângulos?

Convido-o a passar de duas para três dimensões imaginando uma esfera pousada no centro da folha branca. Suponha que o hemisfério sul da esfera está pintado de vermelho e o hemisfério norte permanece branco. Imagine traçar uma linha a partir do Polo

Norte, atravessando a esfera, até ao papel. Se a linha atravessar a região vermelha, o papel por baixo fica vermelho. Como provavelmente consegue visualizar, este processo resulta num círculo vermelho centrado no ponto onde o Polo Sul toca o papel. Agora rode a esfera 90° em torno de um eixo que passa pelo seu equador, de modo que o hemisfério ocidental seja agora a metade vermelha. Repita o exercício de projeção a partir do novo Polo Norte, no topo da esfera. Como também deverá conseguir visualizar, metade do papel fica agora vermelha e a outra metade permanece branca. Vê-se assim que, embora em duas dimensões os dois padrões não estejam relacionados, eles tornam-se relacionados quando subimos para três: existe uma única esfera, e a nossa percepção bidimensional falhou em identificar a simetria subjacente. Assim acontece com a lei de Coulomb: só ao entrarmos em quatro dimensões podemos apreciar plenamente a sua simetria total, idêntica em todas as direções nesse espaço de quatro dimensões.

Não existem, tanto quanto sei, consequências *quotidianas imediatas* e particularmente interessantes desta simetria oculta de dimensão superior. Conheço uma consequência, que é um aspeto esotérico da estrutura dos átomos de hidrogénio, que hesito em abordar, apesar de esses átomos constituírem (à exceção da matéria escura) a matéria mais abundante do universo.³ Poderia forçar um pouco o argumento e dizer que a eliminação dessa simetria hiperesférica na interação de um eletrão com um núcleo, quando mais do que um eletrão está presente num átomo (como acontece com todos os elementos exceto o hidrogénio), é responsável pela estrutura da Tabela Periódica dos elementos e por

todo o mundo da química (que inclui a biologia e, por extrapolação, a sociologia). Mas isso seria pouco mais do que exibicionismo; ainda assim, vale a pena tê-lo presente. Existe, de facto, aqui uma oportunidade para introduzir outra lei da Natureza algo vaga: a *lei periódica*, que reconhece que as propriedades dos elementos se repetem em certa medida, após intervalos variados, quando organizados segundo o seu número atómico (o número de prótons no núcleo). Assim, o silício (número atómico 14) assemelha-se ao carbono (número atómico 6), oito elementos antes, e o cloro (número atómico 17) assemelha-se ao flúor (número atómico 9), também oito elementos antes. A lei periódica, a Tabela Periódica e a totalidade da química, da biologia e da sociologia são consequências diretas da simetria da interação de Coulomb em aliança com as regras que governam a forma como os eletrões, de carga negativa, se organizam em torno dos núcleos atómicos, de carga positiva.



Vou agora aprofundar a interação de Coulomb e as outras interações que, em alguns casos, são responsáveis por manter a matéria coesa e, noutros, por a afastar, procurando mais profundamente a sua origem na indolência e na anarquia.

O meu ponto de partida é a equação de Schrödinger para a propagação de ondas, que introduzi no Capítulo 3. Aí referi que um aspeto central da mecânica quântica é a “dualidade” da matéria: as partículas assumem aspetos de ondas e as ondas assumem aspetos

de partículas. O ponto que aqui preciso de salientar é uma interpretação de onda devida ao físico alemão Max Born (1882–1970; as suas várias contribuições para a formulação da mecânica quântica foram reconhecidas tardiamente com um Prémio Nobel em 1954). A “interpretação de Born” de uma onda em mecânica quântica afirma que o quadrado da sua amplitude numa região indica a probabilidade de aí encontrar a partícula. Com essa interpretação em mente, imagine uma onda de amplitude uniforme (a altura dos seus picos) e comprimento de onda constante (a separação entre picos), estendendo-se daqui até ao horizonte e para além dele. Agora imagine deslocar toda a onda um pouco, de modo que todos os seus picos e vales se movam ligeiramente. Nada de observável mudou, no sentido em que, se avaliasse a probabilidade de encontrar a partícula em qualquer ponto, obteria o mesmo resultado.⁴ Dizemos que a observação é invariante (isto é, não se altera) sob uma transformação de calibre (“gauge”) global (isto é, igual em toda a parte). Esta última expressão precisa apenas de uma breve explicação. Tudo o que significa é que, se colocasse uma régua de medição (um calibre) ao longo da onda e registasse a posição do primeiro pico, então, depois de a onda ter sido ligeiramente deslocada, teria de mover também a régua um pouco para obter a mesma leitura.

Até aqui, tudo é algo trivial, talvez como todas as verdades últimas, mas estamos no limiar de entrar no mundo das “teorias de calibre” das interações entre partículas, uma das fronteiras da física moderna. Chega agora a destrivialização do que antecede, que se revelará bastante surpreendente nas suas consequências.

Expliquei, ou pelo menos mencionei, no Capítulo 3, que as equações do movimento das partículas podem ser estabelecidas começando por formular uma expressão para a “ação” associada a um percurso e, em seguida, procurando o percurso que envolve a menor ação. Esse percurso de ação mínima é o percurso que a partícula adota, sendo o que sobrevive, não sendo eliminado pelos seus vizinhos. Acrescentei que as equações “diferenciais” de Newton podem ser vistas como uma forma de instruir a partícula sobre como tatear esse percurso, passo infinitesimal após passo infinitesimal. Essa discussão foi expressa em termos de partículas “familiares”, como os elétrons, mas aplica-se igualmente às partículas menos tangíveis da radiação eletromagnética, os fótons, porque em mecânica quântica tudo é simultaneamente partícula e onda. Assim, o princípio de que uma partícula adota o percurso de menor ação pode ser aplicado ao eletromagnetismo e à sua partícula propagante, o fóton.

No caso do eletromagnetismo, formula-se uma expressão para a ação, esta é minimizada e, dessa minimização, emergem as equivalentes das equações de Newton, mas agora descrevendo o comportamento do campo eletromagnético. Estas equações são conhecidas como as “equações de Maxwell” e foram formuladas por essa estrela fugaz brevemente incandescente, James Clerk Maxwell (1831–79), em 1861. Constituem uma versão matemática das investigações experimentais pioneiras realizadas por Michael Faraday (1791–1867), na Royal Institution de Londres, sobre eletricidade e magnetismo. As equações demonstraram a inter-relação entre a eletricidade “quadrada” e o magnetismo “hexagonal” e efetuaram a sua unificação como eletromagnetismo

“cúbico”. Uma pista sobre como visualizar essa unificação consiste em reconhecer, como explicarei com mais detalhe no Capítulo 8, que, de acordo com a teoria da relatividade restrita, o efeito do movimento é rodar aquilo que se entende por espaço em tempo e vice-versa. Quanto maior a velocidade, maior a rotação, e aquilo que começou por se apresentar como a face “elétrica” quadrada do cubo passa, cada vez mais, a parecer-se com a face “magnética” hexagonal, e vice-versa.

As equações de Maxwell são, essencialmente, um resumo das leis unificadas da eletricidade e do magnetismo; assim, uma vez compreendida a origem dessas equações, compreenderemos também a origem dessas leis.



No final do século XVIII, o matemático italiano Joseph-Louis Lagrange (1736–1813 — é sempre assim designado, embora tenha nascido Giuseppe Lodovico Lagrangia e se tenha “afrancesado” devido à sua longa permanência em Paris) formulou uma versão particularmente elegante da mecânica de Newton, que continua a ser ideal para desenvolver as equações de movimento que aqui procuramos. O seu procedimento envolve uma forma de conjectura orientada por restrições. Em primeiro lugar, sujeita a diversas considerações técnicas que não é necessário detalhar, propõe-se uma função matemática conhecida, de forma apropriada, como o lagrangiano. Existem várias regras para escrever um lagrangiano, uma das quais é a seguinte: quando ele é usado para calcular a

ação, e quando essa ação é minimizada relativamente ao percurso entre dois pontos, a expressão resultante deve coincidir com as equações de movimento verificadas experimentalmente — neste caso, as equações de Maxwell inspiradas nos trabalhos de Faraday. Se a ação minimizada não conduzir às leis de movimento conhecidas, então a conjectura feita para a forma do lagrangiano estava errada, e é necessário regressar ao início e tentar de novo, repetindo o processo até se obterem as equações de Maxwell.

Verifica-se que esta cadeia de passos — *lagrangiano* → *ação* → *minimização* → *equações de Maxwell* → *experiências de Faraday* — é alcançada quando o lagrangiano é expresso em termos de uma onda, e quando essa onda mantém uma relação especial com o campo eletromagnético que descreve. Aqui está o ponto essencial. Somos livres de deslocar essa onda para trás e para a frente ao longo da sua linha de propagação (isto é, de alterar a sua gauge), mas como dessa alteração não podem resultar efeitos físicos observáveis, o lagrangiano não pode mudar; caso contrário, as equações de movimento — as equações de Maxwell — deixariam de concordar com a experiência. Isto significa que o lagrangiano tem de ser invariante sob transformações globais de gauge.

Chegou agora o momento de Noether e Lagrange se unirem. Recordará, do Capítulo 2, que Emmy Noether identificou uma relação profunda entre simetria e conservação. A invariância global de gauge é uma simetria e, portanto, tem de existir uma lei de conservação associada. Essa lei é *a conservação da carga elétrica*. Isto é, a carga elétrica não pode ser criada nem destruída.

Posso dar-lhe uma intuição de como esta conservação emerge da invariância global de gauge. Imagine um pequeno cubo transparente imerso na região ocupada pela onda. Quando a onda é deslocada ligeiramente (isto é, quando a sua gauge global é alterada da mesma forma em todo o espaço), parte da onda entra no cubo por uma das faces e parte sai pela face oposta. Para que o lagrangiano nessa região permaneça invariante — independentemente de onde a região se encontra, e também no conjunto — qualquer desfasamento entre o fluxo que entra e o que sai teria de ser compensado pela criação ou aniquilação de amplitude no interior do cubo. Esta é precisamente a interpretação padrão de uma equação de continuidade, que é a forma matemática da afirmação de que o fluxo líquido através das fronteiras de uma região tem de ser igual à taxa de criação ou destruição de carga no interior dessa região. Ou seja: a carga é conservada.⁵

Gostaria agora de dar mais dois passos nesta discussão. Primeiro, parece-me defensável que, quando o universo surgiu e pouco aconteceu, não houve qualquer pré-seleção das fases das ondas (das posições relativas dos seus máximos) que viriam, mais tarde, a constituir a base do eletromagnetismo. Isto é, quando os seres humanos chegaram às equações que descrevem o eletromagnetismo, não houve necessidade de identificar nem adotar uma gauge particular: qualquer gauge deveria funcionar. Por outras palavras, como consequência da indolência no início do universo, as equações do eletromagnetismo são globalmente invariantes por transformações de gauge e, como resultado, a carga elétrica é conservada.

O segundo ponto é o seguinte. Se, como consequência dessa indolência na Criação, a carga elétrica é conservada, e se o universo está preso àquilo que tem — sempre teve e sempre terá — então surgem naturalmente duas perguntas: quanta carga elétrica existe no universo? e como é que essa carga emergiu a partir do absolutamente nada?

Podemos estar confiantes quanto à resposta à primeira dessas perguntas. A carga total do universo é zero. Existe certamente uma enorme quantidade de carga positiva (pense em todos os núcleos atômicos existentes) e uma enorme quantidade de carga negativa (pense em todos os elétrons), mas elas cancelam-se mutuamente: a carga total é nula. Sabemos que as cargas se cancelam porque a intensidade da interação elétrica é enormemente superior à atração gravitacional entre massas; assim, mesmo um ligeiro desequilíbrio de carga teria feito o universo explodir assim que se formou, com a gravitação impotente a assistir, incapaz de manter tudo unido. Temos, portanto, de concluir que, embora existam cargas opostas, existe a mesma quantidade de cada uma. A consequência é que a carga não teve de ser criada a partir do Nada, de um Nada inicialmente sem carga. Tudo o que foi necessário foi que o Nada se separasse nos seus opostos elétricos.

Não faço ideia de como esse processo ocorreu, mas parece-me que explicar uma separação em opostos é conceptualmente (e talvez praticamente) muito mais simples do que explicar a criação efetiva de carga. Afinal, é mais fácil compreender a existência simultânea de um monte e de um buraco do que a existência de um monte isolado. A separação do Nada em opostos elétricos terá de

ser explicada um dia — mas isso parece um problema mais acessível do que encontrar um mecanismo para a criação do próprio conceito de carga.



Eis aqui outro ponto verdadeiramente interessante, talvez até espantoso. Suponha agora que, em vez de uma transformação de calibre global — em que toda a onda é deslocada como um todo — fazemos uma transformação de calibre local, na qual o deslocamento da onda é diferente em cada ponto ao longo do seu comprimento. Assim, num ponto o pico da onda é adiantado um pouco, noutro ponto é adiantado um pouco mais, noutro ainda é recuado ligeiramente. Tenha diante dos olhos da mente uma onda que fica distorcida, enrugada, de forma local e não por um deslocamento global uniforme.

Se a onda continuar a descrever a mesma realidade física, então a equação de Schrödinger tem de manter a mesma forma. Contudo, se não se tiver cuidado, isso não acontece, porque o deslocamento local dos picos introduz termos adicionais. Esses termos extra podem, no entanto, ser eliminados modificando a equação de Schrödinger através da adição de um termo que representa uma nova contribuição para a energia. O que está a acontecer é que o deslocamento dos picos altera a energia do estado original, mas o novo termo subtrai precisamente essa variação de energia (e realiza ainda alguns outros ajustamentos técnicos que asseguram que a onda representa um estado energético fisicamente admissível).⁶ A

conclusão espantosa, que se obtém ao examinar a forma da nova equação, é que esse novo termo representa o efeito de um campo eletromagnético. Ou seja, a exigência de que a equação de Schrödinger seja invariante por transformações de calibre locais — isto é, que a descrição da Natureza permaneça a mesma sob modificações locais da onda — implica a existência do eletromagnetismo e das suas equações inspiradas por Maxwell. O eletromagnetismo e as suas leis emergem, assim, da simetria — neste caso, da simetria de calibre local.

É provável que já consiga perceber para onde isto nos conduz no que respeita à emergência das leis da Natureza na Criação. Admitindo que não aconteceu grande coisa (como é habitual, gostaria de dizer “nada aconteceu”, mas isso ultrapassa por agora a minha capacidade), a uniformidade do absolutamente nada foi preservada quando o universo surgiu. Mas agora essa uniformidade foi preservada de uma forma muito mais subtil do que simplesmente dando origem a um espaço-tempo uniforme. Foi preservada de um modo que corresponde ao funcionamento interno do próprio espaço e do próprio tempo: através da invariância de calibre local.

A indolência resulta então no eletromagnetismo. E por eletromagnetismo entendo todas as descobertas de Michael Faraday no seu laboratório na cave da Royal Institution, toda a sua matematização por James Clerk Maxwell, e as leis da ótica que decorrem do reconhecimento de que a luz é uma onda eletromagnética. Raramente uma indolência tão absoluta produziu tanto — incluindo muitos dos atributos da civilização moderna,

manifestos na comunicação, na computação, no transporte, na indústria, no comércio, no entretenimento e no conforto geral da vida.



Por que parar no eletromagnetismo? Afinal, sabe-se que existem vários tipos de força a operar na Natureza. Cargas elétricas interagem através de forças elétricas e cargas elétricas em movimento interagem através de forças magnéticas. O trabalho de James Clerk Maxwell mostrou que estas duas forças são manifestações de uma força unificada — a força eletromagnética — e vimos que ela emerge de um certo tipo de invariância de calibre local. A chamada força fraca, assim designada por ser muito mais fraca do que a força eletromagnética nas condições atuais do universo, é responsável por certos tipos de decaimento nuclear, separando partículas elementares e fazendo com que sejam expulsas dos núcleos sob a forma de radiação.

Depois há a força forte, que liga os constituintes de certas partículas elementares com tal intensidade (daí o seu nome pouco inspirador) que consegue vencer a força eletromagnética que, de outro modo, os faria repelir-se. É assim que os núcleos persistem apesar do poder formidável da força eletromagnética que ameaça destruí-los, mas que é superada pela força forte — até que a débil força fraca, trabalhando discretamente nos bastidores, acabe por se impor. É como consequência da força forte que existe toda uma variedade de núcleos diferentes. Esses núcleos, devido à sua carga

elétrica positiva e à atração entre cargas opostas, capturam elétrons com carga negativa e dão origem a átomos eletricamente neutros. Esses elétrons capturados estão ligados aos seus núcleos — não apenas de forma relativamente fraca e podem ser removidos por impactos bastante suaves, dando assim origem à química, e através da química à biologia, e através da biologia à zoologia, e através da zoologia à sociologia, e através da sociologia à civilização. Por fim, há a gravitação, talvez a mais sutil de todas as forças, que liga tudo, de forma fraca, mas universal, e que dá origem a galáxias, estrelas, sistemas solares, planetas, vida e, tardiamente, à realização — e ao custo atual — da vida humana.

Todas estas forças (e talvez outras ainda não identificadas) explicam o panorama da existência, o tecido da realidade, a complexidade do que existe. Todas elas podem ser expressas como emergindo de invariância de calibre local e, portanto, como consequência da indolência. Há, reconhecidamente, sérios problemas com esta visão.

Uma dificuldade já foi superada. A invariância de calibre local que deu origem ao eletromagnetismo era de um tipo muito simples e facilmente visualizável: envolvia pegar numa onda e “enrugá-la” no espaço tridimensional comum. Por razões técnicas, este tipo de enrugamento é considerado um exemplo de transformação de calibre abeliana. Niels Abel foi um matemático norueguês que, entre muitas outras contribuições importantes, estudou transformações de simetria que conduzem ao mesmo resultado independentemente da ordem em que são aplicadas (por exemplo, rodar à esquerda, refletir e depois rodar à direita). O enrugamento necessário para

mostrar que as forças fraca e forte emergem de formas de invariância de calibre local é não abeliano. Isto significa que as transformações a considerar dependem da ordem pela qual os processos são realizados. Isso torna-as muito mais difíceis de tratar. No entanto, uma das recompensas de vencer maior dificuldade é uma maior riqueza conceptual. As transformações de calibre local não abelianas apropriadas, necessárias para mostrar como a indolência dá origem à força fraca e, ao mesmo tempo, para mostrar que ela tem a mesma raiz que a força eletromagnética (passando a ser conhecida como força eletrofraca), foram identificadas e valeram, com inteira justiça, o Prémio Nobel (em 1979) aos seus autores, Stephen Weinberg e Abdus Salam.

A invariância de calibre subjacente à gravitação ainda não foi encontrada, pelo que é bem possível que a minha visão — uma visão, ou pelo menos uma esperança, partilhada por muitos físicos teóricos — permaneça por concretizar.



Permita-me resumir, uma vez mais, o ponto a que chegámos. Neste capítulo introduzi aquilo que pode ser a origem das forças inicialmente classificadas como eléctricas ou magnéticas, mas que, ao recuarmos a perspetiva (numa dimensão superior), se revelam manifestações de uma única força: o eletromagnetismo. Diversos aspetos da eletricidade e do magnetismo emergiram à medida que olhámos mais atentamente para o espaço e usámos a simetria para descobrir que podem ser entendidos como resultando da indiferença

da Natureza à modificação das ondas por enrugamento — tecnicamente, invariância de calibre. Assim, o espaço vazio revela-se mais intrincado e subtil do que parece à primeira vista, e a uniformidade que herdou do seu progenitor, o Nada, está para além da visão direta: invisível, mas presente, e responsável pelas forças que ligam e destroem e que, por sua vez, são responsáveis pelo funcionamento extraordinário do mundo.

CAPÍTULO 8

Medida por Medida

— A Origem das Constantes Fundamentais

As constantes fundamentais, grandezas como a velocidade da luz ($c = 2,998 \times 10^8$ metros por segundo), a constante de Planck ($h = 6,626 \times 10^{-34}$ joule·segundo), a constante de Boltzmann ($k = 1,381 \times 10^{-23}$ joule por kelvin) e a carga elétrica fundamental ($e = 1,602 \times 10^{-19}$ coulomb), desempenham um papel extraordinário nas consequências das leis da Natureza. As leis dão, por assim dizer, instruções sobre como os sistemas devem comportar-se dadas certas grandezas, como a massa e a carga, e as constantes fundamentais determinam a escala dos efeitos resultantes. Por exemplo, as leis da Natureza que designamos por relatividade especial implicam que o espaço e o tempo se misturam tanto mais quanto maior for a velocidade de deslocação; a velocidade da luz estabelece o grau dessa mistura para uma dada velocidade. As leis do eletromagnetismo implicam que uma partícula carregada é desviada por um campo elétrico, e a carga fundamental determina a extensão desse desvio para uma dada intensidade do campo. A energia de um oscilador — como uma massa presa a uma mola ou um pêndulo —, segundo as leis da mecânica quântica, só pode aumentar por degraus discretos; a constante de Planck diz-nos qual é a separação entre esses degraus. Se essa constante fosse zero,

não haveria intervalos entre os degraus e a energia do oscilador poderia variar de forma contínua. O facto de a constante de Planck ser tão pequena (note-se o fator 10^{-34}) implica que os degraus são tão próximos que não detetamos essa separação em pêndulos ou molas do quotidiano. Mas ela existe.

Tem havido também muita discussão acerca da aparente feliz coincidência dos valores que estas constantes apresentam. Argumenta-se que desvios minúsculos relativamente aos seus valores atuais teriam consequências catastróficas para o surgimento da vida, da consciência e da própria capacidade de questionar por que razão esses valores parecem tão benignos. Com valores apenas ligeiramente diferentes, as estrelas poderiam não se formar ou, se se formassem, poderiam consumir o seu combustível tão rapidamente que não haveria tempo para a evolução da vida, e assim por diante.

A meu ver, existem duas classes de constantes fundamentais: as que não existem e as que existem. Como é de esperar, os valores das que não existem são muito mais fáceis de explicar do que os valores das que existem. As primeiras resultam essencialmente de escolhas humanas sensatas, mas em última análise arbitrárias, feitas ao longo da história intelectual da humanidade sobre como medir e expressar as grandezas físicas (por exemplo, comprimento em metros e tempo em segundos). As segundas — as constantes que existem de facto de forma fundamental e que, por isso, são verdadeiramente fundamentais — são constantes de acoplamento, que resumem a intensidade das interações entre entidades físicas: a intensidade da força entre cargas elétricas, a intensidade da

interação de uma carga elétrica com um campo eletromagnético, e a intensidade das forças nucleares que ligam partículas elementares entre si e às estruturas que designamos por núcleos atômicos. Incluem também a constante gravitacional ($G = 6,673 \times 10^{-11}$ joule·metro por quilograma ao quadrado), que especifica a intensidade do campo gravitacional gerado por um corpo massivo e, por conseguinte, determina as órbitas dos planetas em torno das suas estrelas, contribui para a formação das galáxias e fixa a aceleração com que uma maçã cai.

Embora as tabelas de constantes fundamentais as apresentem com unidades — como a velocidade da luz em metros por segundo —, na realidade não deveriam ter unidades. Dito de outro modo, as constantes fundamentais que não existem têm todas o valor 1 (assim, $c = 1$, e não $c = 2,998 \times 10^8$ metros por segundo), e as constantes fundamentais que de facto existem são melhor expressas de modo a também não terem unidades. Como explicarei, em vez de a carga fundamental ter o valor $e = 1,602 \times 10^{-19}$ coulomb, é preferível expressá-la numa forma cujo valor seja $1/137$. As restantes constantes verdadeiramente fundamentais são igualmente melhor expressas como números puros semelhantes. Como se tornará claro, penso conseguir explicar o valor 1, mas não valores como $1/137$. Atualmente, não temos qualquer ideia de onde surgem números como $1/137$, e não pretendo fingir que sei mais do que os outros. É uma pena, porque são precisamente estes números que governam a nossa existência e a emergência do pensamento: se $1/137$ tivesse sido $1/136$ ou $1/138$, talvez não estivéssemos aqui para o saber.

Estas observações precisam agora de ser desenvolvidas, para que possa perceber o que tenho em mente e por que razão considero que existem estas duas classes de constantes. Não abordarei todas as constantes fundamentais (existem cerca de uma dúzia de importância central, além de várias combinações que são tratadas como se tivessem estatuto semelhante). Limitarei a discussão a algumas que considero verdadeiramente fundamentais e analisarei as suas origens.



Começarei talvez pela constante fundamental mais importante de todas: a velocidade da luz, c (de *celeritas*). Considero que merece esse estatuto porque, embora não exista no sentido fundamental que discutirei, governa a estrutura do espaço-tempo, o palco onde toda a ação decorre.

Há mais no espaço do que aquilo que é imediatamente visível. Isaac Newton (1642–1726, estilo antigo), para não falar de René Descartes (1596–1650) e dessa mente extraordinária da Antiguidade que foi Aristóteles (384–322 a.C.), que simultaneamente inspirou e asfixiou o pensamento, bem como nós próprios, olhamos para o espaço e vemos que ele é tridimensional. Albert Einstein (1879–1955), apoiando-se no trabalho de outros, alterou radicalmente essa visão. A sua relatividade especial (de 1905, o seu *annus mirabilis* (*ano maravilhoso*), a que ainda se seguiria mais glória quando o “especial” evoluiu para o ainda mais extraordinário “geral”) convida-o a aceitar que o espaço está

entrelaçado com o tempo e que aquilo que pensava serem espaço e tempo não deve ser considerado separadamente, mas como componentes de um único palco: o espaço-tempo. Essa teoria traz consigo mais desconforto e subversão do que parece seguro à primeira vista, pois obriga-o a aceitar que aquilo que identifica como espaço e tempo pode não coincidir com o que o seu vizinho identifica. Se esse vizinho estiver em movimento (e a maior parte está, mesmo que apenas a passear, a conduzir ou a passar num foguete), então ele terá percepções diferentes sobre que componente do espaço-tempo é espaço e qual é tempo.

Tudo depende da rapidez com que se desloca. Se eu e o leitor não estivermos em movimento relativo, então aquilo que ambos entendemos por espaço e por tempo é exatamente isso: espaço e tempo. Mas suponha que o leitor se move — a pé, de carro ou num foguete. Nesse caso, a sua percepção altera-se de forma extraordinária: o tempo roda para dentro do espaço e o espaço roda para dentro do tempo. O leitor pode afirmar perfeitamente que algo estacionário ocupa uma posição fixa numa coordenada que identifica como “espaço”. Mas eu, enquanto vizinho que passa a caminhar, perceciono o espaço e o tempo de forma diferente e já não concordo com a sua atribuição de espaço e tempo a esse acontecimento. Quanto mais rapidamente me mover relativamente ao leitor, enquanto observador, mais a minha percepção do tempo se roda para a percepção do espaço, e vice-versa. Cada um de nós, no decorrer das atividades quotidianas, percebe o espaço e o tempo de modo diferente: o seu espaço não é o meu, nem o seu tempo é o meu (a menos que nos movamos exatamente à mesma velocidade, incluindo estarmos ambos em repouso). Essas

diferenças só se tornam relevantes quando as velocidades relativas são muito elevadas, aproximando-se da velocidade da luz. Felizmente é assim, pois de outro modo a ciência e a sociedade seriam provavelmente impossíveis. Ainda assim, a realidade é tal que o espaço-tempo se resolve de forma diferente para cada um de nós e depende do nosso estado relativo de movimento — daí o termo “relatividade”.

Podemos agora regressar ao papel da velocidade da luz. Por vezes pensa-se que é um mistério o facto de existir uma velocidade limite para a propagação da informação, sendo impossível, segundo a relatividade especial, exceder c . Por que existe esse limite? Poderá haver algo semelhante a um arrasto viscoso, como o que impõe uma velocidade terminal a uma bola que cai num meio viscoso? Será o espaço viscoso, sendo a velocidade da luz a velocidade terminal da informação ao atravessá-lo? Não: a explicação é muito mais profunda e, por isso mesmo, mais simples. A velocidade da luz é simplesmente a velocidade à qual a sua percepção do tempo fica totalmente rodada para aquilo que lhe aparece como espaço. Não há qualquer grau adicional de rotação possível. Não existe arrasto viscoso para a informação a atravessar o espaço: a velocidade limite é uma característica da própria forma como percebemos o espaço e o tempo.

Mas por que tem a velocidade da luz o valor particular que tem (exatamente 299.792.458 metros por segundo, cerca de 670 milhões de milhas por hora)? A explicação reside num artefacto da burocracia humana, que se resume ao facto de, por convenção, medirmos comprimentos em metros e não em segundos. Foram

feitas várias propostas quando o metro foi definido pela primeira vez (em 1790, com o zelo revolucionário francês pela racionalização de praticamente tudo, incluindo a aristocracia). Uma sugestão inicial foi defini-lo como um décimo de milionésimo da distância do Polo Norte ao equador ao longo de um meridiano, escolhido diplomaticamente em vez de sensatamente, de modo a situar-se aproximadamente a meio caminho entre Paris e Washington, nas respectivas jovens repúblicas — e que, de forma pouco útil, começava e terminava no mar. O compromisso foi depois abandonado e o meridiano escolhido passou a ser o que atravessa Paris. Foi então fundida uma barra metálica para permitir que este padrão último fosse divulgado e utilizado de forma mais prática.

Desde então, tendo em conta que a Terra efetivamente “respira” e que, por isso, essa distância de referência não é constante — e, em princípio, também o metro não o seria —, foi adotada uma definição mais precisa e imutável. O metro é agora definido como $1/299.792.458$ da distância que a luz percorre (no vácuo) num segundo. Assim, todas as medições de comprimento são, na realidade, medições de tempo. Por exemplo, uma pessoa com 1,7 metros de altura poderia ser descrita como tendo $1,7/299.792.458$ segundos, ou 5,7 nanossegundos, de altura. Embora a luz percorra 299.792.458 metros num segundo, isso pode ser descrito como $299.792.458/299.792.458$ segundos, ou seja, 1 segundo. Então, qual é a velocidade da luz? Se percorre 1 segundo em 1 segundo, a sua velocidade (distância dividida pelo tempo) é 1. Sem unidades; apenas 1. Se estiver num automóvel a viajar a 100 quilómetros por hora (cerca de 28 metros por segundo), pode calcular que a sua velocidade é, na realidade, apenas 0,000.000.093. A uma

velocidade tão baixa, é evidente que pode ignorar os efeitos relativistas e estar confiante de que o seu espaço não foi rodado de modo a se assemelhar ao tempo (comparado com a minha avaliação, enquanto o leitor passa por mim), e não deveremos ter qualquer desacordo, pelo menos dentro de uma precisão razoável, sobre se dois acontecimentos são simultâneos.

Espero que agora possa aceitar que $c = 1$. O facto de até aqui a ter considerado como um grande número de metros por segundo revela-se um acidente histórico: por razões perfeitamente compreensíveis e sensatas, a sociedade civil mediu a distância e o tempo em unidades diferentes. Meça-os nas mesmas unidades e uma constante fundamental de enorme importância desaparece efetivamente. A partir de agora, sempre que eu me referir a um comprimento L , convencionalmente medido em metros, representá-lo-ei por $L^\dagger$, sendo o punhal um sinal de que os metros foram eliminados e que, doravante, o comprimento será expresso em segundos. Todas as velocidades passam assim a perder unidades e tornam-se números puros.



Suspeito que o leitor possa estar a perguntar-se se outros coelhos podem também ser empurrados de volta para os chapéus. E quanto à que é provavelmente a segunda constante fundamental mais importante, a constante de Planck, h ? Tal como a velocidade da luz introduziu essencialmente a relatividade na ciência, a constante de Planck introduziu efetivamente a mecânica quântica;

culturalmente, portanto, ambas têm um peso semelhante. A constante c , agora “desaparecida”, permeia todas as formulações da relatividade especial; será então possível que h , presente em todas as formulações da teoria quântica, também desapareça simplesmente porque uma determinada grandeza foi, historicamente, expressa em unidades convenientes, mas fundamentalmente inadequadas?

O físico alemão Max Planck (1858–1947) foi o iniciador da mecânica quântica num gesto que ele próprio considerou um ato de desespero. Esse desespero dirigia-se à incapacidade da física clássica — que ele, com razão, estimava profundamente — de explicar aquilo que parecia ser um problema elementar: a cor da luz emitida por um corpo incandescente, essencialmente por que razão um corpo passa de vermelho-quente a branco-quente à medida que a temperatura aumenta. A física clássica conduzia à conclusão errada de que todos os objetos deveriam ser branco-quentes mesmo quando apenas mornos. Segundo a física clássica, não deveria existir escuridão. Mais ainda, e pior, qualquer objeto, mesmo apenas morno, deveria devastar a paisagem com radiação gama. O desespero de Planck levou-o a supor, em 1900 ou pouco antes, que se algo oscila com uma determinada frequência, então só pode trocar energia com o resto do mundo em pacotes — quanta — de energia, sendo o tamanho do pacote proporcional à frequência: objetos que oscilam a baixa frequência só podem trocar pacotes pequenos; objetos que oscilam a alta frequência só podem trocar pacotes grandes. A física clássica supunha que um oscilador de qualquer frequência podia trocar energia em qualquer quantidade; a hipótese de Planck supunha que a energia é

quantizada, isto é, trocada em pacotes. Essa sugestão simples, mas revolucionária, que Planck parece ter detestado por contrariar toda a sua formação clássica (Einstein teve dificuldades semelhantes com a mecânica quântica em geral), explicou a cor dos objetos quentes — na verdade, a cor dos objetos a qualquer temperatura. Sabemos hoje que explica a cor do Sol, cerca de 5772 K nas regiões superficiais emissoras de luz, e até a cor do próprio universo, que arrefeceu até uns parcimoniosos 2,7 K, mas continua a brilhar com radiação característica de um corpo a essa temperatura.

Na física convencional, a energia é expressa em joules (J). Um joule é uma unidade bastante pequena, mas muito adequada ao discurso cotidiano. Por exemplo, cada batimento do coração humano requer cerca de 1 joule de energia. A bateria de um smartphone típico armazena cerca de 50 quilojoules de energia. O joule é uma introdução relativamente recente, substituindo um conjunto heterogêneo de unidades mais antigas que incluíam calorias, ergs e British thermal units. No século XIX, à medida que a termodinâmica e a ciência da energia emergiam, o calor era tipicamente expresso em calorias e o trabalho em ergs.

Eis uma analogia para introduzir um ponto importante. Houve outrora grande interesse na eficiência das máquinas a vapor e, conseqüentemente, na relação entre calorias de calor fornecido e ergs de trabalho produzido. Foram realizados elaborados experimentos para estabelecer o chamado “equivalente mecânico do calor”, o fator de conversão — então visto como uma constante fundamental algo modesta — que permitia converter medições de uma forma de energia em outra. No entanto, embora esses

experimentos tenham sido uma componente importante do nosso progresso intelectual, noutro sentido foram uma completa perda de tempo. Se os primeiros investigadores tivessem medido o calor e o trabalho nas mesmas unidades — ambas em calorias ou ambas em ergs — então o fator de conversão, essa constante fundamental específica, teria sido simplesmente 1. Hoje isso é de facto o caso (excetuando algumas ilhas arcaicas de atividade, incluindo a ciência alimentar quotidiana), usando-se o joule para relatar ambas as formas de energia. O “equivalente mecânico do calor” pertence agora à história ou, dito de outro modo, vale 1.

Creio que o leitor consegue ver os paralelos entre esta situação e os argumentos que tenho vindo a apresentar sobre as constantes fundamentais reais — ou pelo menos aquelas que não existem, ou não deveriam existir: escolhendo as mesmas unidades para grandezas relacionadas, os fatores de conversão tornam-se 1. A constante de Planck é uma candidata natural para este tratamento. Foi introduzida para relacionar a frequência de oscilação com o tamanho dos pacotes de energia correspondentes, isto é, o tamanho mínimo dos quanta que podem ser transferidos.¹

O caminho a seguir deverá agora ser claro. Vamos eliminar os joules e expressar a energia como uma frequência, em ciclos por segundo. Sempre que quiser expressar energia como frequência, irei denotá-la por $E^\dagger$ e reportá-la como um certo número de ciclos por segundo. Já não há necessidade de qualquer fator de conversão entre essas descrições, tal como já não há necessidade de relatar o equivalente mecânico do calor ou, depois de termos decidido expressar a distância em segundos, de relatar a velocidade da luz.

A constante de Planck tornou-se 1. Joules, tal como calorias e ergs, pertencem agora à história. À primeira vista, poderá parecer que isto tem implicações profundas para a mecânica quântica, se $h = 1$ em vez do seu valor convencionalmente diminuto; mas isso não é o caso, como explicarei depois de varrer ainda mais alguns detritos acumulados nos estábulos de Augias das unidades convencionais.

Um ponto quase final, neste contexto, é que com as artimanhas desenvolvidas até aqui, a fórmula de Einstein $E = mc^2$ torna-se $E^\dagger = m^\dagger$, com ambas as grandezas expressas como frequências. É claro que pode manter a forma $E^\dagger = m^\dagger c^2$, mas para isso terá de aceitar que $c = 1$, como já argumentei. Um ponto verdadeiramente final é que agora se pode ver que, uma vez que $E^\dagger = m^\dagger$, energia e massa são a mesma coisa.



Atualmente, praticamente toda a gente (excetuando os EUA, juntamente com Myanmar e a Libéria) expressa a massa em quilogramas e nas suas frações (gramas) ou múltiplos (toneladas, 1000 quilogramas). O quilograma foi originalmente definido (na década de 1790) como a massa de um litro de água a uma determinada temperatura. Tal como o metro, essa definição foi refinada e substituída por um quilograma-padrão, o «Protótipo Internacional do Quilograma» (*IPK, International Prototype of the Kilogram*), um cilindro de liga de platina-irídio conservado no *Bureau International des Poids et Mesures*, em Sèvres, nos arredores de Paris, com várias cópias secundárias distribuídas pelo

mundo. Infelizmente, nem mesmo o IPK é perfeitamente estável: impurezas evaporam, o ar difunde-se para o interior, e surgem microrriscos quando é manuseado. Assim, aquilo que se entende por «um quilograma» tem vindo a mudar lentamente. As propostas atuais visam definir o quilograma em termos da constante de Planck, uma constante eterna (tanto quanto sabemos), de modo que o significado de «um quilograma» fique fixado para sempre e qualquer pessoa com acesso às constantes fundamentais saiba exatamente o que se quer dizer. O que significa isso para os nossos propósitos?

Adotemos a perspetiva de que a humanidade, no seu habitual modo algo confuso, mas sensato, cometeu um erro coletivo ao adotar o quilograma como unidade de massa. Em vez do quilograma, suponha-se que tivesse adotado o segundo — ou, mais precisamente, «ciclos por segundo» — tal como numa frequência. Com extraordinária clarividência, poderia ter passado a reportar não m , mas $m\hbar = mc^2/h$ e a expressar a massa em oscilações por segundo. Uma massa de 1 quilograma, por exemplo, seria então reportada como $1,4 \times 10^{50}$ ciclos por segundo. Se se considera uma pessoa harmoniosamente proporcionada de 70 quilogramas, a partir de agora deverá pensar na sua massa como uns vertiginosos $9,5 \times 10^{51}$ ciclos por segundo, convertendo a massa em quilogramas numa energia em joules, multiplicando pelo quadrado da velocidade da luz (ou seja, utilizando a relação $mc^2 = E$), e depois recorrendo à constante de Planck para expressar essa energia como uma frequência em ciclos por segundo. A unidade “ciclos por segundo” começa a tornar-se algo enfadonha de escrever e de ler; trata-se, na verdade, da própria definição da

unidade “hertz” (Hz), assim designada em honra do pioneiro das comunicações por rádio, Heinrich Hertz (1857–94), cuja vida foi lamentavelmente breve. Assim, 1 ciclo por segundo é 1 hertz. Ao adotar este procedimento — multiplicar por c^2 e depois dividir por h — a sua massa surgirá como aproximadamente $9,5 \times 10^{51}$ hertz. À primeira vista, poderá parecer uma forma absurda de expressar massa, mas esse não é o ponto essencial. Na prática quotidiana, o quilograma é sensato e útil. O que procuro, porém, é alcançar a raiz da apresentação de dados da maneira mais coerente possível — e, nesse processo, apontar a minha adaga tipográfica à garganta das unidades convencionais.

Podemos agora compreender por que razão fixar $h = 1$ não tem qualquer consequência no mundo físico, no sentido em que deixa a mecânica quântica intacta. Uma forma de o demonstrar seria mostrar que a equação de Erwin Schrödinger (que apresentei no Capítulo 3 como um dos componentes principais da mecânica quântica) permanece inalterada, salvo na interpretação dos seus símbolos. Contudo, equações tão complexas como a sua ficam confinadas a espreitar nas sombras deste livro, nas Notas.² Outra via consiste em conduzi-lo aos fundamentos dessas equações. E isso revela-se possível, pois os fundamentos — mesmo em ciência — são invariavelmente mais simples do que os edifícios que sustentam.

Se é um utilizador habitual de transportes pendulares, já está a meio caminho de compreender a mecânica quântica. O termo “*commuter*” deriva da prática comum de vender um bilhete de ida e volta por um valor inferior à soma das tarifas individuais de ida e

de volta: a tarifa de regresso é “comutada” (do latim *commutare*, “mudar”, “alterar”). Dito de outro modo, o custo da viagem de regresso não é igual ao da viagem de ida (desde que já tenha investido na ida). A mecânica quântica difere da mecânica clássica de modo surpreendentemente análogo. A analogia é a seguinte. A tarifa da viagem de “ida” corresponde à multiplicação do momento linear pela posição; a tarifa da viagem de “volta” corresponde à multiplicação da posição pelo momento linear (note a ordem inversa). Ora, estas duas “tarifas” não são iguais. A diferença entre elas recebe o nome de “comutador” da posição e do momento linear. Se quisermos visualizar isto internamente, estamos perante algo subtil: na mecânica clássica, multiplicar posição por momento ou momento por posição conduz ao mesmo resultado — as grandezas são números e a multiplicação é comutativa. Na mecânica quântica, porém, essas grandezas tornam-se operadores, e a ordem de aplicação altera o resultado. É precisamente essa não-comutatividade — essa assimetria estrutural — que marca a transição do edifício clássico para o quântico.

Uma companhia ferroviária pode ajustar as tarifas dos seus passageiros pendulares ao seu bel-prazer. A Natureza, porém, parece ter fixado um padrão muito específico de comutação, e a redução na viagem de ida e volta corresponde a uma modificação discreta (mas de consequências vastíssimas) da constante de Planck.³ Isto é: efetuar a viagem de “ida”, multiplicando o momento linear pela posição, menos efetuar a viagem de “volta”, multiplicando a posição pelo momento linear, produz uma diferença proporcional a h . Toda a divergência entre as previsões da mecânica quântica e as da mecânica clássica emerge dessa comutação da

“tarifa de ida e volta”. E todos os aspetos quantitativos decorrem do facto de o “conselho de administração” da Natureza ter estabelecido que o desconto do passageiro pendular é proporcional à constante de Planck. Se quisermos tornar isto estruturalmente nítido, podemos escrever simbolicamente: $px - xp \propto h$ onde p representa o operador momento linear e x o operador posição. Na mecânica clássica, a expressão acima seria simplesmente nula, porque $px = xp$. No regime quântico, a diferença não desaparece — e é precisamente esse desvio, controlado por h , que introduz a textura quântica da realidade. Se h fosse zero, não haveria “desconto”; a ida e a volta seriam indistinguíveis — e regressaríamos ao edifício clássico. O mundo quântico nasce dessa pequena, mas estruturalmente decisiva, assimetria.

Nas unidades convencionais, a constante de Planck é tão diminuta (mas tão fecunda) que o conselho de administração da mecânica clássica decidiu que não valia o incómodo administrativo de conceder qualquer desconto aos passageiros pendulares. É fácil compreender o seu ponto de vista: seria como oferecer um desconto de 1 cêntimo numa tarifa de muitos biliões de libras. Dessa decisão perfeitamente razoável emerge a mecânica clássica.

Razoável — sim. Mas também errada. O verdadeiro conselho de administração da Natureza insiste em manter o desconto. A descrição matemática mais bem-sucedida de sempre da matéria e da radiação — a mecânica quântica — difere da mecânica clássica por esse simples gesto: conceder um desconto aos “pendulares”. E, contudo, as consequências são das mais profundas que se podem imaginar. Como já foi sugerido, Isaac Newton e os seus

contemporâneos, bem como os seus sucessores imediatos, não tinham qualquer suspeita da não-comutatividade entre posição e momento. Sobre essa omissão ergueram a magnífica catedral teórica a que chamamos mecânica clássica. Dela nasceu a compreensão dos céus — e quem se preocuparia com um desconto tão minúsculo quando corpos do tamanho de planetas orbitam o Sol? Mas quando a atenção científica se voltou para elétrons em átomos, onde as próprias tarifas de “ida” e “volta” são extremamente pequenas, então o desconto de comutação tornou-se tremendamente significativo. Numa tarifa de uma libra, o desconto poderia atingir 50 pence; simplesmente não pode ser ignorado. Se visualizarmos o contraste de escalas, torna-se claro o mecanismo conceptual: quando as grandezas envolvidas são enormes (como massas planetárias), o termo proporcional a h é negligenciável; quando as grandezas são microscópicas (como momentos associados a elétrons), o mesmo termo torna-se comparável às próprias “tarifas”. A transição do clássico para o quântico não é uma mudança de leis — é uma mudança de regime de escala. O desconto esteve sempre lá. Apenas deixa de ser desprezável quando o cenário encolhe até à dimensão atômica.

Como é então possível fixar h igual a 1, em vez de um minúsculo 10^{-34} , e ainda assim concluir que a mecânica clássica continua a ser adequada para os objetos do quotidiano? Não significaria isso que qualquer posição e momento correntes passariam a beneficiar de um desconto pendular significativo? A posição escorregadia é a de que contornei o problema ao omitir as unidades associadas ao 10^{-34} . Os valores das posições e dos momentos do quotidiano, que poderiam ter valores perfeitamente normais nas antigas unidades

— posição em metros, massa em quilogramas e velocidade em metros por segundo — tornam-se enormes quando expressos nas novas unidades: posição em segundos, massa como frequência em ciclos por segundo e velocidade sem unidades. Como consequência, o produto da posição pelo momento para um objeto cotidiano torna-se também enorme nas novas unidades, e muito, muito maior do que 1.⁴ Na antiga forma de encarar as coisas, posição e momento tinham valores correntes e h era extremamente pequeno. Na nova forma, é h que tem um valor corrente (igual a 1), enquanto posição e momento se tornam extremamente grandes. O resultado — o desconto ser negligenciável — é, na prática, o mesmo, e a consequência dessa negligenciabilidade também: não é necessário recorrer à mecânica quântica para descrever os objetos do cotidiano.

Preciso de mencionar aqui esse grande clarificador do pensamento humano, o princípio da incerteza de Werner Heisenberg, que ele formulou em 1927, pois ele decorre diretamente da não-comutatividade entre momento e posição. O princípio afirma que não é possível conhecer, com precisão arbitrária, a posição e o momento simultaneamente. A mecânica quântica, para desconforto de muitos formados na tradição clássica (incluindo Niels Bohr e Albert Einstein), revela assim que é necessário fazer uma escolha ao procurar especificar o estado de um sistema. Ela instrui-nos a escolher uma descrição em termos de posição ou uma descrição em termos de momentos — qualquer delas podendo ser especificada com precisão arbitrária. Mas, se insistir, por força do seu condicionamento clássico, em falar simultaneamente em ambas as descrições, acreditando que só

assim a sua descrição do mundo será completa, o princípio da incerteza detê-lo-á abruptamente, pois implica que as duas descrições são intrinsecamente incompatíveis. Se não conseguir libertar-se desse condicionamento de físico clássico, será conduzido à conclusão de que a mecânica quântica impede uma descrição completa da Natureza. Uma perspectiva muito mais fecunda, porém, é reconhecer que aquilo que os praticantes da mecânica clássica tomavam por “completo” era, na verdade, excessivamente completo — e, por isso mesmo, inalcançável. A mecânica quântica ensina que a utilização simultânea de ambas as descrições é inconsistente. É como iniciar uma frase numa língua e terminá-la noutra. É necessário escolher a língua; caso contrário, a mensagem tornar-se-á ininteligível e o seu interlocutor — neste caso, o universo — olhará para si em silêncio, sem resposta. A mecânica quântica elimina este erro inspirado pelo senso comum e aceita que a completude existe numa linguagem ou na outra — em posição ou em momento — mas não em ambas simultaneamente. Uma vez aceite isso, a descrição do universo torna-se mais depurada (ainda que não simples). É por essa razão que considero o princípio da incerteza um grande clarificador.



Eliminei c e h , as dobradiças sobre as quais giram a relatividade e a mecânica quântica. Haverá espaço, neste cemitério, para outras constantes fundamentais? Se tivesse de identificar a constante fundamental mais importante que funciona como a dobradiça da termodinâmica, escolheria a constante de Boltzmann, k . Ela surge

na crucial distribuição de Boltzmann que exaltei no Capítulo 5, está gravada na lápide de Boltzmann na sua definição de entropia, e aparece disfarçada — entre outras formas, como a constante dos gases — insinuando-se por toda a termodinâmica. No entanto, é completamente desnecessária e pode ser eliminada e sepultada com argumentos semelhantes aos que usei para fazer desaparecer c e h .

O erro — mais uma vez sensato, compreensível e até louvável — remonta a Celsius e Fahrenheit, que apresentei no Capítulo 4 como criadores das primeiras escalas de temperatura, e foi consolidado pela introdução, por Kelvin, de uma escala absoluta aparentemente mais natural. Antes de mais, é preciso reconhecer que todos foram seduzidos pela convenção, talvez Celsius menos do que os outros. No nosso mundo atual, quanto mais quente o objeto, maior a temperatura — em todas as três escalas. Como referi, Celsius inicialmente concebeu a sua escala ao contrário: quanto mais quente, mais baixo o valor. Creio que, inadvertidamente, estava no caminho certo, pois de várias formas considero que «quanto mais quente, mais baixo» é mais natural num nível termodinâmico fundamental, como explicarei. Contudo, os três erraram, na minha perspetiva, ao introduzirem uma nova unidade de medida (o grau, e mais tarde o kelvin, K) para reportar temperatura — tal como a introdução do metro para medir comprimento, em vez de usar segundos, acabou por gerar confusão desnecessária que só se tornou evidente com o amadurecimento da ciência. Viu-se que, se o comprimento tivesse sido medido em segundos, não teria sido necessário introduzir a constante fundamental c , a velocidade da luz. De modo análogo, argumentarei que, se a temperatura tivesse

sido reportada nas mesmas unidades que a energia, nunca teria sido necessário introduzir a constante de Boltzmann.

Há, naturalmente, vários pontos a esclarecer. A constante de Boltzmann, expressa em joules por kelvin, pode ser vista como um fator de conversão entre kelvins e joules. Se já concordou em reportar temperatura em joules, não há necessidade de conversão. Se existir uma relação uniforme entre temperatura em kelvins e em joules, não há ambiguidade na mudança de unidades. Poderão surgir números pouco familiares, mas a estranheza numérica não é critério de aceitabilidade científica (ainda que possa sê-lo na prática quotidiana). Por exemplo, com o valor atualmente aceite para k , uma temperatura amena de 20 °C (293 K) seria reportada como 4,0 zeptojoules (zepto é o prefixo talvez desconhecido, mas útil, que designa 10^{-21}) e a água ferveria a 5,2 zeptojoules.

Se aceitar reportar temperaturas em joules (ou submúltiplos como zeptojoules), as divisões do termómetro passarão a ser em joules, e cada grau Celsius corresponderá a 0,0138 zeptojoules. Feito isso, não será mais necessário invocar a constante de Boltzmann em qualquer expressão. Na prática, se insistir em usar as equações dos manuais, sempre que surgir k deverá atribuir-lhe o valor 1. Assim, k junta-se a c e h : uma constante fundamental supérflua, nascida apenas porque os primeiros cientistas foram levados pela prática quotidiana a introduzir uma unidade nova, mas desnecessária para medir temperatura.⁵

Mas por que razão considero que Celsius estava menos errado do que Fahrenheit e Kelvin, e que seria preferível pensar a

temperatura como diminuindo quando algo aquece? Refiro-me ao facto de muitas expressões da termodinâmica, e sobretudo da sua vertente estatística — que liga o microscópico ao macroscópico, o indivíduo à multidão — se tornarem notavelmente mais simples quando expressas no inverso da temperatura (isto é, como $1/T$ em vez de T , não simplesmente invertido com o 0 e o 100 a trocarem de lugar). A matemática parece sugerir que a escala natural não deve apenas ser invertida (trocar 0 e 100), mas literalmente colocada de cabeça para baixo. Com a temperatura já expressa em zeptojoules, o seu inverso seria reportado em «por zeptojoule». Assim (deixo-lhe os cálculos), o ponto de ebulição da água seria 0,19 por zeptojoule e o de congelação, 0,27 por zeptojoule.

A partir de agora, exprimirei todas as temperaturas de forma invertida e convertida, como tantos “por zeptojoule”, e designarei a temperatura assim redefinida pela letra $\mathbb{T}$ (tê-traço). Como me impus a regra de não citar quaisquer fórmulas fora do espaço protegido das Notas, para as quais o remeto,⁶ terá de aceitar a minha palavra de que, se tomar qualquer fórmula da termodinâmica estatística, ela se torna — e é de facto — mais simples quando T é substituído por $\mathbb{T}$. Mas essa substituição envolve mais do que uma mera questão de aparência.

Toda a gente (bem, quase toda a gente) sabe que não se pode atingir o zero absoluto de temperatura. A terceira lei da termodinâmica exprime essa inatingibilidade em termos mais sofisticados e cientificamente aceitáveis, acrescentando “num número finito de etapas” e alguns refinamentos adicionais, mas esse é o seu núcleo essencial. Pode parecer estranho que $T = 0$, o

limite inferior da escala Kelvin, não possa ser alcançado em um número finito de etapas. Contudo, $T = 0$ corresponde a $\mathcal{T} = \infty$, e provavelmente há menos resistência psicológica à ideia de que não se pode atingir um $\mathcal{T}$ infinito através de um número finito de passos.

Uma simplificação mais profunda emerge da exploração de várias equações da termodinâmica estatística. Embora temperaturas absolutas negativas (temperaturas como -100 K) sejam destituídas de sentido na termodinâmica ordinária (são como comprimentos negativos: algo não pode ter -1 metro de comprimento), nada impede que se manipulem as equações da termodinâmica estatística e se observe o que sucede a diversas propriedades (por exemplo, a entropia) quando a temperatura é deixada descer através de zero e tornar-se negativa, ou até negativamente infinita. Por exemplo, pode-se tomar qualquer uma das fórmulas da nota 6 e verificar o que acontece ao introduzir um valor negativo da temperatura. Tipicamente, surgem comportamentos problemáticos: as propriedades exibem saltos abruptos ou disparam para o infinito quando a temperatura atravessa o zero. Contudo, se as mesmas propriedades forem representadas em função de $\mathcal{T}$, então todos esses saltos e divergências desaparecem, e as propriedades passam a comportar-se de forma contínua e suave. Esta domesticação do comportamento sugere fortemente (não é mais do que isso) que $\mathcal{T}$ é uma medida mais fundamental da temperatura do que T . No entanto, argumentarei agora que não é ainda suficientemente fundamental: ainda não atingiu o nível mais profundo de fundamentação.

Estou certo de que começa a reconhecer um padrão a emergir ao longo destes capítulos, com tudo a tornar-se mais simples quando é expresso quer em segundos (tempo e distância), quer como frequência em “por segundo” (energia). Viu também que a temperatura inversa $\mathbb{T}$ é uma energia inversa em “por zeptojoule”. Ora repare: podemos converter essa energia inversa em inverso de “por segundo”, o que é simplesmente segundos.⁷ Assim, 20 °C passam a corresponder a 0,16 picossegundos (pico é o prefixo que designa 10^{-12}), a água solidifica a 0,18 picossegundos e entra em ebulição a 0,13 picossegundos.

Nesta fase, as três constantes fundamentais da relatividade, da mecânica quântica e da termodinâmica — c , h e k — tornaram-se redundantes. Dito de outro modo, se insistir em utilizar equações em que elas aparecem (como $E = mc^2$) e tiver escolhido exprimir as propriedades envolvidas (como E e m) em unidades relacionadas entre si (como segundos ou variações destes), então terá de atribuir a cada constante fundamental o valor 1 — e deixa de subsistir qualquer mistério quanto à sua origem.⁸



Abandono agora essas constantes fundamentais inexistentes — que consigo explicar — e volto-me para as que realmente existem e que não consigo explicar. Mencionarei apenas duas, embora outras se ocultem nesta caixa de Pandora do que permanece, por ora, inexplicável. Ambas são constantes de acoplamento, governando a intensidade de duas variedades de interação.

Já mencionei a carga fundamental, e , que exprime a intensidade das interações eletromagnéticas, como a força de atração entre duas cargas ou a interação de um eletrão (que possui carga $-e$) com um campo elétrico, como o de uma onda de rádio. A magnitude desta constante influencia a intensidade da interação entre eletrões e núcleos atómicos e, portanto, as dimensões e propriedades dos átomos; determina também a força das ligações entre átomos e, conseqüentemente, a formação de compostos; afeta ainda a interação entre eletrões em átomos e moléculas e o campo eletromagnético, influenciando assim as cores dos materiais e a intensidade dessas cores. Desempenha igualmente um papel no interior dos núcleos atómicos, pois os prótons, carregados positivamente, estão sujeitos a intensas repulsões mútuas.

Mais uma vez, convém separar a magnitude da carga fundamental das unidades inspiradas pela escala humana e exprimi-la como número puro. Sempre que uma constante surge acompanhada de unidades, não se pode afirmar com clareza se é grande ou pequena: grande ou pequena relativamente a quê? No seu caso, a carga fundamental é frequentemente combinada com outras constantes fundamentais para produzir um número adimensional, a chamada constante de estrutura fina, α (alfa), assim designada por ter sido introduzida para explicar certos detalhes da estrutura do espectro do átomo de hidrogénio. O seu valor é o já referido $1/137$.⁹ O facto de α ser tão pequeno reflete a fraqueza das interações eletromagnéticas (comparadas com a força forte que atua no interior dos núcleos) e explica que as moléculas, mantidas por interações eletromagnéticas, sejam muito mais maleáveis do que os núcleos, no sentido em que podem ser

desagregadas e reconstituídas em reações químicas. Se α fosse próximo de 1, não existiria química; as moléculas, caso existissem, teriam dimensões comparáveis às dos núcleos atômicos, e a vida (uma reação química extraordinariamente elaborada) não teria emergido. O universo teria sido biologicamente silencioso.

Ninguém sabe ainda por que razão α tem o valor $1/137$. Num cenário possível, todas as forças teriam inicialmente a mesma intensidade, mas, à medida que o universo se expandiu e arrefeceu, as suas intensidades divergiram, emergindo $1/137$ como a intensidade de uma delas. Presumo que esse valor venha a ser explicado quando se formular uma teoria mais abrangente sobre a origem, estrutura e evolução do universo; por enquanto, porém, o seu valor permanece um mistério. Isso não significa que não tenham sido propostas diversas combinações numéricas envolvendo π e $\sqrt{2}$, algumas com valores surpreendentemente próximos do valor experimental.¹⁰ Contudo, tratam-se de construções sem fundamento teórico sólido, nunca aceites pela comunidade científica como mais do que exercícios de numerologia. O problema é, ainda assim, de enorme importância para a compreensão do universo e do nosso lugar nele. Existem constantes de acoplamento semelhantes para as forças forte e fraca, que desempenham papéis na estrutura nuclear. Qualquer teoria futura das forças fundamentais (e das partículas fundamentais sobre as quais atuam) terá de explicar todos esses valores.

A única outra constante de acoplamento que mencionarei é a que governa a intensidade da gravidade: a constante gravitacional, G , que surge na *lei do inverso do quadrado da atração gravitacional*

entre duas massas.¹¹ A constante gravitacional pode ser convertida numa grandeza adimensional, α_G , análoga à constante de estrutura fina, substituindo efetivamente o quadrado da carga do elétron (que surge em α) pelo quadrado da massa do elétron. O resultado é o valor $1,752 \times 10^{-45}$.¹² Este valor revela quão extraordinariamente pequena é esta constante e confirma que a gravidade é incomparavelmente mais fraca do que o eletromagnetismo. Essa fraqueza é benéfica para o surgimento de entidades pensantes — pelo menos para nós — pois permite tempo suficiente para a formação de estrelas, galáxias, a estabilidade de planetas em órbita e a emergência e evolução de formas de vida. Se fosse muito mais intensa, tudo — nós incluídos — estaria concentrado num vasto buraco negro comum (e nem sequer teríamos consciência disso).

Ninguém sabe qual é a origem do valor de G . Especula-se que tenha sido inicialmente intensa, mas que se tenha atenuado quase até desaparecer à medida que o universo arrefeceu (à semelhança da constante de estrutura fina, mas de forma ainda mais acentuada). Outros sugerem que continua forte, mas que grande parte da sua intensidade se tenha dissipado em seis ou sete dimensões ainda não desdobradas nem detetáveis. Seja como for, ninguém sabe por que razão a gravidade é tão fraca, nem por que α_G possui o seu valor atual — e não fingirei o contrário.



Onde estamos, então? As leis da natureza governam o comportamento das entidades de modo geral, mas as suas

consequências quantitativas dependem dos valores de diversas constantes fundamentais. Entre estas contam-se a velocidade da luz, central na relatividade; a constante de Planck, central na mecânica quântica; e a constante de Boltzmann, central na termodinâmica. Procurei mostrar que, se todas as grandezas físicas observáveis forem expressas nas mesmas unidades ou em unidades relacionadas — em vez de permanecerem aprisionadas num conjunto heterogêneo de unidades concebidas pragmaticamente pelos humanos — então essas três constantes podem ser descartadas. Dito de outro modo, se insistir em mantê-las nas equações, pode atribuir-lhes o valor 1, desde que todas as propriedades observáveis sejam expressas em unidades relacionadas (escolhi segundos e as suas variações). Existe, contudo, outra classe de constantes fundamentais: as constantes de acoplamento que exprimem a intensidade das várias forças, como a força eletromagnética e a força gravitacional. Até ao momento, ninguém sabe por que razão possuem os seus valores atuais — valores que, para nós, foram extraordinariamente favoráveis.

CAPÍTULO 9

O Grito das Profundezas — Por Que Funciona a Matemática?

Muitas leis da natureza são expressas matematicamente; todas elas, mesmo aquelas que não são intrinsecamente matemáticas (como quaisquer leis que venham a ser formuladas para descrever a evolução por seleção natural), adquirem maior poder quando desenvolvidas matematicamente. Um dos primeiros cientistas a refletir explicitamente sobre este ponto foi o influente matemático húngaro Eugene Wigner, que levantou a questão numa conferência de 1959 intitulada *A Eficácia Inexplicável da Matemática nas Ciências Naturais*.¹ A sua conclusão — talvez prudentemente tímida — foi a de que essa eficácia aparentemente desrazoável constitui um mistério demasiado profundo para ser resolvido pela reflexão humana. Outros acrescentaram a este sentimento a convicção de que, entre os mistérios atuais, este será provavelmente dos que persistirão por mais tempo.

Uma perspetiva alternativa, mais afirmativa e menos cautelosa do que o pessimismo de Wigner, é a de que a eficácia da matemática não é desrazoável. Longe de ser desconcertante, poderá antes oferecer uma janela privilegiada para a estrutura profunda do universo. A matemática pode ser entendida como o

universo a esforçar-se por nos falar na nossa própria linguagem comum. Ao longo deste capítulo, procurarei dissipar qualquer tonalidade mística — Deus nos livre — que essa afirmação possa sugerir.² A existência de versões matemáticas das leis da natureza talvez aponte para uma questão de grande alcance — e, esperemos, para uma resposta fecunda — acerca da estrutura profunda do tecido da realidade. Talvez nos conduza à questão mais fundamental e, durante Eras, a mais inquietante e irresistível de todas: como é que aquilo que existe veio a existir.



É inegável que a matemática constitui uma linguagem extraordinariamente poderosa e bem-sucedida para dialogar com o universo. No plano mais pragmático, uma equação que sintetiza uma lei física pode ser usada para prever um resultado numérico, como na previsão do período de um pêndulo a partir do seu comprimento. Basta considerar a capacidade assombrosa dos astrónomos para prever as órbitas dos planetas, a ocorrência de eclipses e (no momento em que escrevo) o aparecimento de uma superlua, a coincidência entre lua cheia e a sua maior aproximação à Terra. Há ainda as consequências inesperadas que emergem de uma lei formulada matematicamente e que são depois confirmadas pela observação. Entre as mais célebres está a previsão dos buracos negros a partir do conteúdo da teoria geral da relatividade de Einstein, a sua teoria da gravitação. Chegou mesmo a dizer-se, ironicamente, que nenhuma observação experimental pode ser aceite se não for apoiada por uma teoria formulada

matematicamente. As economias mundiais floresceram — e por vezes definharam — sob o impacto da aplicação de formulações matemáticas das leis da natureza. Uma parte muito significativa da produção industrial das nações foi atribuída à implementação da mecânica quântica e da sua formulação matemática.

Existem, naturalmente, aspetos da nossa compreensão do universo e da sua materialização que não são expressos matematicamente. Logo no início deste livro, e de passagem há instantes, chamei a atenção para uma das teorias mais abrangentes acerca do universo: a teoria da seleção natural como explicação da evolução. Essa teoria não é intrinsecamente matemática no sentido de não estar expressa por uma fórmula e, contudo, é de enorme poder explicativo, podendo aplicar-se em todo o universo onde exista algo que possa ser considerado «vida». Chegou mesmo a ser aplicada ao surgimento não apenas de novas espécies, mas de universos inteiros. Pode ser expressa como uma espécie de lei da natureza, sendo a expressão de Herbert Spencer, «sobrevivência do mais apto», uma aproximação crua, mas sugestiva. No entanto, quando desenvolvida matematicamente — por exemplo, através da modelização da dinâmica populacional, a que voltarei dentro de momentos — a versão qualitativa da teoria torna-se incomensuravelmente mais rica, no sentido de poder gerar previsões quantitativas.

A biologia, no seu conjunto, é talvez um domínio menos óbvio de exposição matemática. Este ramo do conhecimento humano foi, em grande medida, uma sucessão de passeios pela natureza até 1953, quando Watson e Crick estabeleceram a estrutura do ADN e, quase

de um só golpe, transformaram a biologia numa parte da química e, portanto, num membro das ciências físicas, com todo o poder que isso implica. Ainda assim, é difícil apontar leis biológicas especificamente matemáticas, exceto — regressando ao ADN — as leis, incluindo o código, da hereditariedade. Contudo, há numerosos exemplos que ilustram o papel direto da matemática na biologia. Entre eles incluem-se a análise de populações de predadores com acesso a presas e, num certo sentido, a atividade análoga de conceber estratégias de colheita e pesca. Fenómenos periódicos de toda a ordem são típicos dos organismos — basta um momento de reflexão sobre nós próprios, respirando e com o coração a bater, e sobre os nossos ritmos circadianos mais lentos — e tais oscilações são suscetíveis de descrição matemática. Do mesmo modo, ondas de diferença — como a diferença entre o número de pessoas infetadas e não infetadas numa epidemia — e ondas de potencial elétrico, como na propagação de sinais ao longo dos nervos quando pensamos e agimos, e ondas de atividade muscular quando um peixe se flete automaticamente (mesmo decapitado) em ondas transversais que se propagam ao longo do corpo para se impulsionar na água, são todos aspetos da biologia que podem ser tratados matematicamente.

O brilhante e tragicamente maltratado Alan Turing (1912–54) foi talvez o primeiro a desfazer a reputação de incrível fealdade atribuída a Esopo (possivelmente 629–565 a.C., se é que alguma vez existiu), mostrando que o tratamento matemático de ondas de substâncias químicas que se difundem através de recipientes de várias formas — formas como leopardos, por exemplo — explica os padrões das pelagens dos animais, incluindo como o leopardo

adquiriu as suas manchas, a zebra as suas riscas e a girafa as suas manchas irregulares, bem como as intrincadas belezas das asas das borboletas. Até o elefante recebeu a sua tromba a partir de uma onda de substâncias químicas que se propagou pelo embrião inicial segundo leis matemáticas expressas por equações e pelas suas soluções.³

A sociologia, essa elaboração da biologia aplicada às populações humanas, por vezes modeladas como ratos, surgiu no final do século XVIII; o termo foi cunhado por Emmanuel-Joseph Sieyès (1748–1836) em 1780, mas a disciplina floresceu no final do século XIX e adquiriu estrutura matemática no século XX, à medida que modelos estatísticos elaborados puderam ser explorados numericamente em computadores. Embora o seu impulso inicial tenha sido identificar leis do comportamento humano, as suas principais conquistas consistiram no desenvolvimento de métodos estatísticos para analisar e, por vezes, prever os comportamentos mais prováveis ou médios de populações de indivíduos. Essa modelização estatística é essencial para o funcionamento eficaz e a governação das sociedades, mas não emergiram leis fundamentais, para além das leis intrínsecas à própria estatística (como as distribuições em forma de sino das variáveis aleatórias), apesar do desejo de as identificar.

A teologia, o estudo do divino intrinsecamente elusivo e incompreensível — a versão erudita da procura do sorriso do Gato de Cheshire — não necessita de matemática. Nem, naturalmente, aquelas outras criações muito mais afirmativas de mentes ativas, como a poesia, as artes e a literatura, que tanto enriquecem o

quotidiano com fantasias envolventes e por vezes perturbadoras. A estatística constitui uma exceção, pois ajuda, por exemplo, a distinguir Marlowe de Shakespeare. A música situa-se talvez na fronteira e poderá ser uma porta de entrada para uma ciência da estética, onde intuições matemáticas se revelem valiosas ao examinar acordes e sequências de notas em relação a possíveis circuitos de ressonância no cérebro.

Preciso agora de atenuar um pouco este relato. Apesar de todas estas aplicações variadas da matemática, elas não são, em si mesmas, leis. Excetuando a análise numérica de dados que a estatística realiza, em todos os casos (creio eu) o componente matemático consiste na análise de um modelo. Isso não é a substância das leis fundamentais da natureza; é antes a formulação de um arranjo complexo de leis físicas fundamentais subjacentes. Não são sequer leis marginais; são incursões de bandos organizados dessas leis marginais.



Ao nível mais simples e mais evidente, a matemática funciona porque fornece um modo imparcial e altamente racional de desdobrar as consequências de uma equação que exprime simbolicamente uma lei. Assim, é impossível fazer previsões fiáveis a partir de uma afirmação não matemática, como «sobrevive o mais apto», e prever, por exemplo, que combinações primitivas de elementos evoluirão, com o tempo, para elefantes. Em contraste, podem fazer-se previsões fiáveis a partir de uma formulação

matemática, como a lei de Hooke, segundo a qual a força restauradora é proporcional ao deslocamento (a verbalização da equação $F = -kx$): o período de um pêndulo pode ser previsto com precisão a partir do seu comprimento.

Ouç-o excluir «caos». É verdade que o desenrolar de certos sistemas parece imprevisível, mas essa imprevisibilidade deve ser interpretada com cautela. Um exemplo simples de sistema com comportamento caótico é o «pêndulo duplo», no qual um pêndulo está suspenso da extremidade de outro, e ambos oscilam de acordo com a lei de Hooke. Neste caso, as equações de movimento podem ser resolvidas e, desde que os ângulos iniciais a partir dos quais ambos são largados sejam conhecidos exatamente, os ângulos futuros em qualquer instante podem ser previstos com exatidão. A expressão crucial aqui é «desde que os ângulos iniciais sejam conhecidos exatamente», pois mesmo uma imprecisão infinitesimal nesses ângulos conduz a comportamentos subsequentes radicalmente diferentes. Um sistema caótico não é um sistema que se comporte de modo errático: é um sistema com extrema sensibilidade às condições iniciais, de tal forma que, *para todos os efeitos práticos*, o seu comportamento futuro é imprevisível. Um conhecimento perfeito da posição inicial conduz (na ausência de perturbações externas como atrito ou resistência do ar) a um comportamento perfeitamente previsível.

Uma consequência desta impossibilidade prática intrínseca de fazer coincidir previsão e observação é uma alteração no significado da verificabilidade experimental em ciência. Durante muito tempo se sustentou que uma pedra angular do método científico é o

confronto entre previsão e observação, com revisão da teoria em caso de falha. Mas se a previsão fiável nem sempre é possível, terá essa pedra angular sido minada? De modo algum. A previsão «global» de que o modelo conduz a comportamento caótico pode ser verificada testando o sistema sob diferentes condições iniciais; além disso, o próprio «caos» possui qualidades previsíveis que também podem ser confirmadas. Não é necessário prever e confirmar a trajetória exata de um pêndulo duplo para afirmar que o sistema foi compreendido e o seu comportamento verificado. As leis da natureza — neste caso, o tal grupo organizado de leis marginais — teriam sido confirmadas mesmo perante esta imprevisibilidade quantitativa.

O cérebro humano é uma concatenação de processos muito mais complexa do que a trivialidade mecânica de um pêndulo duplo, pelo que dificilmente surpreende que o seu resultado — uma ação, uma opinião, até uma obra de arte — não possa, e presumivelmente nunca poderá, ser previsto a partir de um determinado estímulo, como um olhar ou uma frase passageira. Os teólogos chamam a esta imprevisibilidade «livre-arbítrio». Tal como no caso do pêndulo duplo, mas numa escala incomparavelmente mais complexa, poder-se-ia afirmar que compreendemos o funcionamento de um cérebro, artificial ou natural, em termos da rede de processos que nele decorrem, mesmo que não tenhamos conseguido prever a opinião que iria expressar, o poema que iria escrever ou o ato violento que iria desencadear. A ocorrência de «livre-arbítrio» seria, nesse sentido, uma confirmação de que compreendemos o funcionamento de um cérebro, assim como a ocorrência de caos confirma que compreendemos o funcionamento de um pêndulo duplo. Talvez seja

demasiado esperar que, tal como padrões de caos são previsíveis em sistemas simples, também padrões de livre-arbítrio venham um dia a ser identificados. Talvez, através da psiquiatria, isso já esteja em curso, ainda que sem formulação rigorosa.



A racionalidade imparcial da matemática pode ser tudo o que está por trás da sua alegada eficácia “irracional”. Talvez essa eficácia não tenha nada de irrazoável: talvez resida simplesmente no facto de a matemática ser a apoteose do raciocínio sistemático. A razão pela qual funciona pode estar apenas no seu método: começa-se por propor um modelo, estabelecem-se algumas equações que exprimem as suas propriedades e, depois, desdobram-se as consequências recorrendo às ferramentas consolidadas da dedução matemática. Talvez seja só isso. Mas poderá haver algo mais profundo?

Há indícios de que o mundo possa ser matemático num sentido mais radical. O ponto de partida é a célebre observação do matemático alemão Leopold Kronecker (1823–1891): «*Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk*» (“*Deus fez os números inteiros; tudo o resto é obra do homem*”). Segundo esta perspetiva, as maravilhas da matemática seriam manipulações engenhosas de entidades primordiais — os inteiros — originalmente destinadas apenas à contagem. Mas de onde vieram os inteiros, deixando de lado a resposta teológica?

Os inteiros podem emergir literalmente do nada. A construção pertence à região mais depurada da matemática, a teoria dos conjuntos, que trata coleções de objetos abstraindo quase totalmente da natureza desses objetos.

Se não houver absolutamente nada, tem-se o chamado conjunto vazio, denotado $\{\emptyset\}$. Chamemos-lhe 0. Considere-se agora o conjunto que contém o conjunto vazio: $\{\{\emptyset\}\}$. Temos algo — chamemos-lhe 1. Pode ainda formar-se um conjunto que contenha o conjunto vazio e o conjunto que contém o conjunto vazio: $\{\{\emptyset\}, \{\{\emptyset\}\}\}$. Como tem dois elementos, chamemo-lo 2. Prosseguindo o padrão, 3 será $\{\{\emptyset\}, \{\{\emptyset\}\}, \{\{\emptyset\}, \{\{\emptyset\}\}\}\}$. O procedimento é claro: os inteiros são gerados a partir do vazio. Uma vez obtidos os inteiros, submetendo-os às regras e operações apropriadas, emerge toda a aritmética — e, por extensão, o edifício da matemática.

A analogia com a emergência do universo a partir do Nada torna-se sugestiva: o Nada como conjunto vazio $\{\emptyset\}$, e o universo como a sucessiva estruturação desse vazio. Pode ser apenas uma analogia elegante; ou pode conter uma intuição profunda acerca de como pode haver “algo” e de por que razão a matemática é tão eficaz na sua descrição.

Consigo ver que existem vários problemas com a analogia. Entre eles conta-se a ausência de regras que expliquem como os inteiros se articulam nas estruturas que conhecemos como “matemáticas”. Dispor de uma lista de inteiros dificilmente merece o nome de “universo”. Aqui, a resposta poderá residir nos axiomas que foram

propostos como fundamento da aritmética. Entre eles encontram-se os célebres axiomas propostos pelo matemático italiano Giuseppe Peano (1858–1932).⁴ Uma vez obtida a aritmética, obtém-se muito mais do que isso, porque existe um teorema igualmente célebre, devido ao alemão Leopold Löwenheim (1878–1957) e ao norueguês Thoralf Skolem (1887–1963), que implica que qualquer sistema axiomático é equivalente à aritmética.⁵ Assim, por exemplo, se se dispuser de uma teoria que abranja todas as leis da natureza e que esteja baseada num conjunto de afirmações (axiomas), então essa teoria é logicamente equivalente à aritmética, e quaisquer proposições acerca da aritmética aplicam-se igualmente a ela. Uma especulação ousada poderia, portanto, sugerir que relações lógicas análogas às propostas nos axiomas de Peano foram, por assim dizer, encontradas por acaso e conferiram estabilidade à entidade que emergiu do nada e que designamos por universo. É evidente que aqui tateio na obscuridade à procura de significado, e qualquer interpretação fiável desta visão, caso alguma vez venha a surgir, terá de aguardar progressos profundos na compreensão e na elucidação das nossas raízes cósmicas. Por agora, estas reflexões não passam de um capricho imaginativo.



Uma grande questão, evidentemente, é o que queremos dizer quando afirmamos que o universo é matemática. O que estou eu a tocar, se tudo não passar de aritmética? O que vejo através da janela, se não for senão álgebra? Será a minha consciência apenas uma colaboração de inteiros a dançar ao ritmo de axiomas? Será a

causalidade algo semelhante — ou mesmo idêntico — ao desdobrar da demonstração de um teorema?

Tomemos o tato. Estaremos, de algum modo, a tocar na raiz quadrada de 2 ou até no próprio π ? Posso talvez ajudar a perceber que, em certo sentido, estamos. Se deixarmos de lado os aspetos neurofísicos do tato que ocorrem no nosso interior quando entramos em contato com um objeto externo (e reconheço que poderá dizer: “Mas esse é precisamente o ponto do tato, a nossa resposta mental!”; tenha paciência comigo), então o tato reduz-se à impenetrabilidade do tocado relativamente ao que toca. A impenetrabilidade é uma forma de exclusão de uma região do espaço e, como tal, podemos compreender a origem do sinal que transporta o “toque” até ao cérebro ou para um circuito reflexo neural que desencadeia a retirada perante um possível perigo ou, pelo contrário, a intensificação do contato: a lesão.

A exclusão de um objeto por outro decorre de um princípio extremamente importante proposto pelo físico teórico nascido na Áustria Wolfgang Pauli (1900–58; mais uma chama breve) em 1925 e generalizado em 1940, o que lhe valeu o Prémio Nobel da Física em 1945. Este princípio, intrínseco à mecânica quântica, diz respeito à descrição matemática dos eletrões (entre certas outras partículas fundamentais) e estabelece como essa descrição deve mudar quando se trocam as designações atribuídas a dois eletrões.⁶ A consequência do princípio é que as nuvens eletrónicas de dois átomos não podem interpenetrar-se: uma é excluída da região ocupada pela outra. O tato emerge, assim, de um princípio fundamental da natureza. Embora reconheça que esta visão do tato

ainda não atinge plenamente o cerne do que significa “tocar” a matemática, espero que possa aceitá-la como um passo nessa direção.

A audição é uma forma de tato. Neste caso, o receptor sensível encontra-se no interior do ouvido, e o toque é o das moléculas do ar que constituem a onda de pressão e os impactos que produzem sobre a membrana do tímpano. O facto de o detetor transmitir a deteção desses contatos para uma região distinta do cérebro é o que nos leva a considerar a audição como diferente do tato; mas, no fundo, não o é. A visão também é tato, embora de um tipo mais subtil e oculto. Neste caso, o contato ocorre entre as moléculas recetoras ópticas presentes nos bastonetes e cones da retina. Estas encontram-se embebidas numa proteína em forma de taça até que um fotão de luz as excita, levando-as a adotar uma configuração diferente. A proteína deixa então de as poder acomodar — mais uma vez, tato — e elas “saltam” para fora, permitindo que a proteína altere ligeiramente a sua forma e desencadeie um impulso para outra região do cérebro, onde é interpretado como componente da visão. O olfato e o paladar também são aspetos do tato — desta vez (segundo se pensa atualmente, embora o mecanismo permaneça controverso) moléculas inaladas pelo nariz ou depositadas sobre a língua entram em contato com os seus recetores e desencadeiam um sinal para ainda outra região do cérebro. Toda a sensação é, em última análise, tato, e todo o tato é manifestação do princípio de Pauli relativo à natureza matemática do mundo.

Tenho de admitir — como já parcialmente admiti — que esta descrição da sensação como manifestação de um fragmento de matemática dificilmente será plenamente convincente, e não ousei sequer aprofundar o desencadear dos processos nos mistérios obscuros do cérebro e o modo como este converte sensação em consciência. Como poderia ser convincente antes de conhecermos verdadeiramente a natureza profunda da matéria? Espero, no entanto, que pelo menos sugira que estamos, em última instância, intimamente em contato com os inteiros e com a sua elaborada organização naquilo a que chamamos realidade.



Há ainda uma questão final importante — que poderia ser, literalmente, uma questão de vida ou de morte. Onde se situa o teorema de Gödel? O teorema de Gödel, demonstrado num notável golpe de génio pelo lógico austríaco Kurt Gödel (1908–78 — que, de forma célebre, acabou por morrer de inanição em Princeton por receio de ser envenenado) em 1931, afirma essencialmente que a autoconsistência de um conjunto de axiomas não pode ser provada a partir desses mesmos axiomas.⁷ Se as leis da natureza são matemáticas, poderá acontecer que não sejam autoconsistentes? Estará a minha descrição delas condenada sistemicamente? Se o universo for um gigantesco fragmento de matemática, poderá também ele não ser autoconsistente? Poderá colapsar sob o peso da sua própria inconsistência?

Existem saídas possíveis deste cenário. Gödel baseou a sua demonstração numa formulação particular da aritmética, uma versão da qual referi na nota 4. Suponha-se que se elimina uma dessas afirmações — por exemplo, a especificação do que se entende por multiplicação. Nesse caso, retira-se um dos pilares da prova de Gödel, e esta deixa de funcionar. Uma aritmética sem $\times$ poderá parecer estranha, mas também parecia estranha a versão de aritmética que mencionei no Capítulo 8, em que o resultado de 2×3 não era o mesmo que o de 3×2 , e ainda assim revelou-se a chave para compreender o mundo físico. Se se retirar a multiplicação da aritmética, Gödel fica impotente à beira do caminho, e a aritmética revela-se completa.⁸ Quem sabe qual seria o quadro se, além disso, $2 + 3$ não fosse considerado igual a $3 + 2$? No entanto, o ponto essencial é que está longe de ser claro se as condições sobre as quais Gödel construiu a sua demonstração se aplicam ao mundo físico (o único mundo). Assim, o pessimismo é infundado: as leis da natureza poderão muito bem ser autoconsistentes — e demonstravelmente autoconsistentes — e não existir qualquer falha lógica subterrânea no universo que possa, num instante, propagar-se catastroficamente e apagar-nos a nós e a tudo o que existe, devolvendo-nos com um sopro de aniquilação ao Nada absoluto de onde outrora emergimos. Mais ainda, poderá ser que apenas leis da natureza globalmente consistentes sejam viáveis e que o universo seja uma estrutura logicamente muito coesa, que não admite inconsistência nem incoerência — nem o tipo de aritmética que com elas seria compatível.

Há questões relacionadas. Alguns mostram-se pessimistas quanto às consequências de um dia, no futuro, encontrarmos uma

teoria de tudo — uma espécie de mãe cósmica que abarque não apenas todas as leis, mas o próprio fundamento de todas elas — sugerindo que, nesse momento, a humanidade poderá pendurar definitivamente a régua de cálculo e aceitar que o trabalho está concluído, com compreensão completa de todos os pormenores da realidade. Mas poderá sempre restar algo por fazer. Por exemplo, poderemos descobrir que existem duas ou mais descrições igualmente bem-sucedidas de tudo e não sermos capazes de escolher entre elas. Uma possibilidade desse gênero já foi encontrada: como expliquei no Capítulo 8, é possível formular uma descrição do mundo exclusivamente em termos de posições ou exclusivamente em termos de momentos lineares. Nenhuma é “melhor” do que a outra. Talvez existam miríades de descrições aparentemente inconciliáveis, mas igualmente válidas, à espera de serem descobertas — miríades de conjuntos de leis da natureza mutuamente autoconsistentes, embora aparentemente díspares.

Saberemos quando tivermos descoberto todas as leis da natureza? Saberemos que uma teoria particular da natureza é válida, mesmo que a sua verificação experimental esteja além das nossas capacidades — seja por limitações técnicas, seja por impossibilidade de princípio?

Com todas as leis supostamente descobertas, deveremos afrouxar cautelosamente a nossa exigência de verificação experimental rigorosa? Ou deveremos estabelecer guardiões nas fronteiras do conhecimento, encarregados da ingrata tarefa de identificar eventuais transgressões das nossas leis, mesmo confiando que tais transgressões nunca ocorrerão? Precisaremos de

inspetores robóticos da Natureza, incansáveis, sempre vigilantes, nessas fronteiras do saber. Deveremos aceitar — como é sugerido por algumas teorias fundamentais contemporâneas (tenho em mente a teoria das cordas) — que a nossa confiança nas teorias é tal que, mesmo não sendo testáveis, devam ser consideradas verdadeiras? Será a nossa exploração gradual das leis da natureza um passo fatal rumo à excessiva confiança?

Seja qual for o futuro, é reconfortante saber que, tanto quanto podemos discernir, o universo é um lugar racional e que até a origem das leis que o regem está ao alcance da compreensão humana. Ainda assim, anseio por substituir esse “não muito” que aconteceu na Criação pela vertiginosa perspectiva de que esse “não muito” foi, na verdade, nada absolutamente nada.

Núcleo Ontológico Operacional

Tudo o que foi percorrido — entropia, simetria, invariância, ação mínima, leis de conservação, teorias de gauge, constantes fundamentais — pode ser relido como manifestação de uma arquitetura tripla subjacente.

Essa arquitetura não acrescenta entidades.

Reorganiza princípios.

I. Campo Estrutural — Nada

O que chamamos “Nada” não é ausência absoluta, mas ausência de diferenciação.

É uniformidade estrutural.

- Não contém formas.
- Não contém direção.
- Não contém preferência.

É pura simetria.

E é precisamente essa simetria que explica por que razão:

- as leis são invariantes,
- as equações são reversíveis no tempo,
- as conservações emergem de simetrias (Noether),

- as interações obedecem a estruturas matemáticas estáveis.

Nada (Campo Estrutural)

→ explica por que razão as leis são simétricas.

A lei não é imposta.

É consequência da ausência de privilégio estrutural.



II. Campo Operativo — Anarquia

Uma vez existente estrutura uniforme, qualquer flutuação não encontra resistência direcional.

Não há guião.

Não há finalidade.

Há apenas dinâmica.

Esta dinâmica é estatística.

E, estatisticamente, os sistemas evoluem para estados mais prováveis.

Daqui decorre:

- a segunda lei da termodinâmica,
- o aumento global de entropia,

- a irreversibilidade prática,
- a evolução espontânea de distribuições energéticas.

Anarquia (Campo Operativo)

→ explica por que razão há evolução estatística e entropia crescente.

Não é caos desordenado.

É ausência de teleologia combinada com liberdade combinatória.



III. Campo Emergente — Forma

Mas a anarquia estatística não impede estabilizações locais.

Pelo contrário:

- Átomos surgem como mínimos energéticos.
- Estrelas surgem como equilíbrios gravitacionais.
- Organismos surgem como dissipadores estruturados.
- Pensamento surge como padrão neurodinâmico.

A forma é uma redução local do espaço de possibilidades

à custa de expansão global do mesmo.

Forma (Campo Emergente)

→ explica por que razão existem átomos, estrelas, organismos, pensamento.

A ordem não contradiz a entropia.

É um modo de a servir.



Resolução das Três Tensões Clássicas

Por que há leis?

Porque o Campo Estrutural é simétrico e não privilegia coordenadas, direções ou estados.

Por que há mudança?

Porque o Campo Operativo não impõe direção e a probabilidade governa a evolução.

Por que há estrutura?

Porque a dinâmica anárquica permite estabilizações locais que persistem.



Síntese

Leis não são decretos.

Mudança não é falha.

Forma não é milagre.

São três faces de uma mesma arquitetura:

- Estrutura → Operação → Emergência
- Nada → Anarquia → Forma

A física descreve estas três camadas com linguagens diferentes:

simetria, estatística, estabilidade.

Mas o padrão é único.

Se esta leitura estiver correta, então o universo não é um objeto.

É um processo estruturalmente coerente que se desdobra.

E nós — enquanto formas conscientes — somos uma dessas estabilizações locais que ganhou a capacidade rara de olhar para o próprio campo de onde emergiu.

Não há fecho.

Há reconhecimento estrutural.

Nada: simetria máxima.

Anarquia: dinâmica sem privilégio.

Forma: simetria quebrada estabilizada.

A realidade emerge quando a uniformidade estrutural do Nada é operada pela indiferença dinâmica da Anarquia, produzindo padrões localmente estáveis que chamamos Forma.

Glossário

Ação

Grandeza funcional cuja minimização determina a evolução de um sistema físico. Formalmente, é o integral temporal do lagrangiano. O princípio da ação mínima afirma que a trajetória física é aquela para a qual a ação é estacionária (normalmente mínima).

Axiomas de Peano

Conjunto de postulados que formalizam a aritmética dos números naturais. Permitem construir os inteiros a partir do conjunto vazio e estabelecer a estrutura lógica da contagem.



Boltzmann, constante de (k)

Constante que relaciona temperatura com energia nas formulações tradicionais da termodinâmica estatística.

No enquadramento do livro, desaparece quando a temperatura é expressa nas mesmas unidades que a energia.



C (velocidade da luz)

Constante que estabelece a relação estrutural entre espaço e tempo na relatividade restrita.

Quando comprimento e tempo são expressos nas mesmas unidades, assume valor 1 e deixa de ser uma constante “fundamental” independente.

Campo eletromagnético

Entidade física descrita pelas equações de Maxwell, resultante da unificação de eletricidade e magnetismo. Surge, no texto, como consequência da invariância de calibre local.

Carga elétrica

Propriedade conservada associada à invariância de calibre global.

A conservação da carga decorre da simetria do lagrangiano sob transformações globais de fase (teorema de Noether).

Comutador

Expressão matemática que mede a diferença entre duas operações realizadas em ordem inversa.

Na mecânica quântica:

posição $\times$ momento – momento $\times$ posição $\propto \hbar$

É a origem da diferença estrutural entre mecânica clássica e quântica.

Constantes fundamentais

Grandezas que determinam a escala quantitativa das leis físicas.

O livro distingue duas classes:

1. Constantes que desaparecem com escolha adequada de unidades (c , $\hbar$, k).
2. Constantes de acoplamento (como $1/137$), que exprimem a intensidade real das interações.



Equações de Maxwell

Conjunto de equações diferenciais que descrevem o comportamento dos campos elétrico e magnético. Resultam da minimização da ação eletromagnética.

Equação de Schrödinger

Equação fundamental da mecânica quântica que descreve a evolução temporal da função de onda.

A exigência de invariância de calibre local conduz à introdução do campo eletromagnético.

Entropia

Medida do número de configurações microscópicas compatíveis com um estado macroscópico.

Na formulação de Boltzmann:

$$S = k \ln W$$

No enquadramento do livro, a constante k pode ser eliminada por redefinição de unidades.



Força forte

Interação responsável pela coesão dos núcleos atômicos. Associada a transformações de calibre não abelianas.

Força fraca

Interação responsável por certos processos de decaimento nuclear. Unificada com o eletromagnetismo na teoria eletrofraca.

Força gravitacional

Interação universal entre massas, descrita pela relatividade geral. A sua formulação como teoria de calibre permanece incompleta.

Função de onda (ψ)

Objeto matemático cuja amplitude ao quadrado fornece a densidade de probabilidade de encontrar uma partícula numa dada região.



h (constante de Planck)

Constante que relaciona energia e frequência:

$$E = h\nu$$

Desaparece quando energia é expressa em unidades de frequência.



Incerteza de Heisenberg

Princípio que estabelece que posição e momento não podem ser simultaneamente determinados com precisão arbitrária.

Resulta diretamente da não comutatividade.

Inteiros

Podem ser construídos a partir do conjunto vazio através da teoria dos conjuntos. Servem de base estrutural à matemática.

Invariância de calibre (Gauge invariance)

- **Global:** transformação uniforme de fase não altera as observações → implica conservação da carga.
- **Local:** transformação de fase dependente da posição → exige introdução do campo eletromagnético.



Lei de Boyle

Para temperatura constante:

$$pV = \text{constante}$$

Lei de Charles

Para volume constante:

$$p \propto T \text{ (temperatura absoluta)}$$

Lei do gás perfeito

$$pV = NkT$$

ou

$$pV = nRT$$

Lei limite válida quando as moléculas não interagem.

Lei de Hooke

$$F = -k_f x$$

A força restauradora é proporcional ao deslocamento.

É lei limite válida para pequenas oscilações.

Leis limite

Leis válidas exatamente apenas em condições ideais (volume infinito, deslocamento nulo, etc.). Funcionam como aproximações fundamentais.

Lagrangiano

Função que caracteriza dinamicamente um sistema.

A sua minimização conduz às equações de movimento.

Löwenheim–Skolem

Teorema que estabelece equivalência lógica entre sistemas axiomáticos e a aritmética.



Mecânica clássica

Descrição determinista baseada na comutatividade entre grandezas dinâmicas.

Mecânica quântica

Descrição física baseada na não comutatividade e quantização da energia.



Noether, teorema de

Cada simetria contínua implica uma lei de conservação.



Oscilador

Sistema que apresenta troca periódica entre energia cinética e potencial. Exibe simetria entre espaço e momento.



Princípio da ação mínima

A natureza seleciona a trajetória que torna estacionária a ação.



Relatividade restrita

Teoria que unifica espaço e tempo em espaço-tempo.

A velocidade da luz emerge como fator geométrico.



Simetria

Invariância sob transformação.

Fonte profunda das leis físicas e das grandezas conservadas.

**Temperatura inversa (Ƨ)**

Reformulação mais fundamental da temperatura como inverso da energia.

Simplifica expressões da termodinâmica estatística.

Transformação Abelian

Transformação cujo resultado não depende da ordem de aplicação.

Transformação Não-Abelian

Transformação dependente da ordem de aplicação.

Base das teorias de calibre das forças forte e fraca.

THE OXFORD BOOK OF MODERN SCIENCE WRITING



Editado por Richard Dawkins

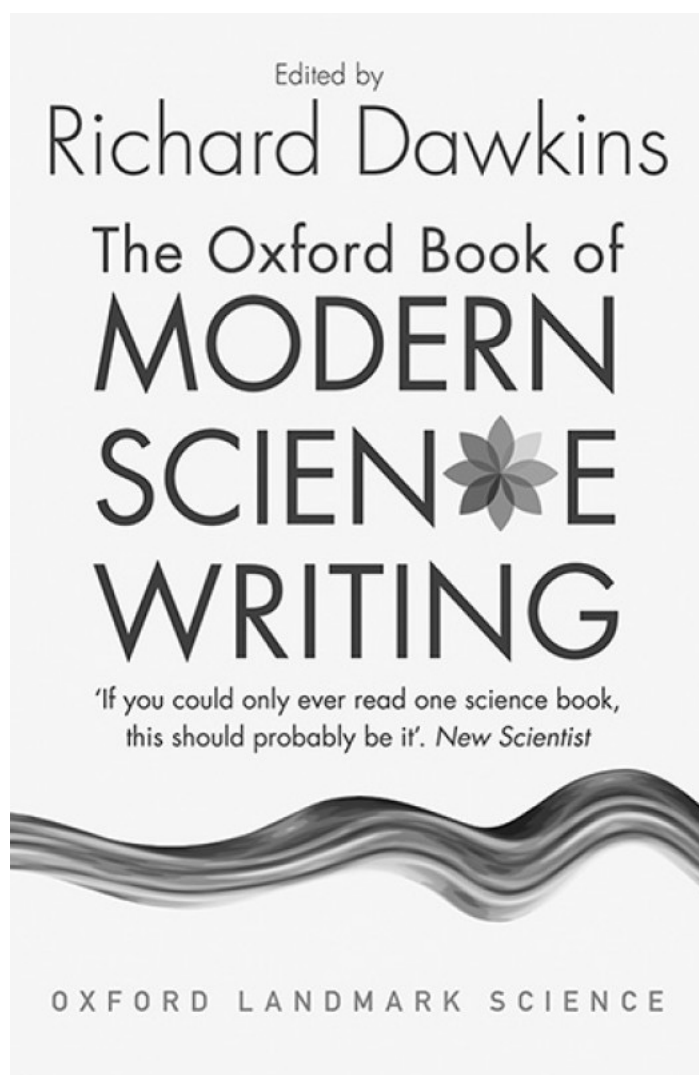
«Um compêndio de alguns dos pensamentos mais iluminadores dos últimos 100 anos.»

— The Times

«Um festim para muitas longas noites.»

— The Sunday Telegraph

Selecionada e introduzida por Richard Dawkins, *The Oxford Book of Modern Science Writing* é uma antologia rica e vibrante que celebra a melhor escrita de cientistas para um público mais vasto — revelando que os melhores cientistas demonstraram tanta imaginação e destreza com a pena quanto no laboratório.



ELEGANCE IN SCIENCE

The Beauty of Simplicity



Ian Glynn

«Um livro erudito... Bem ilustrado e repleto de anedotas e enquadramento histórico, trata-se, de facto, de um volume elegante.»

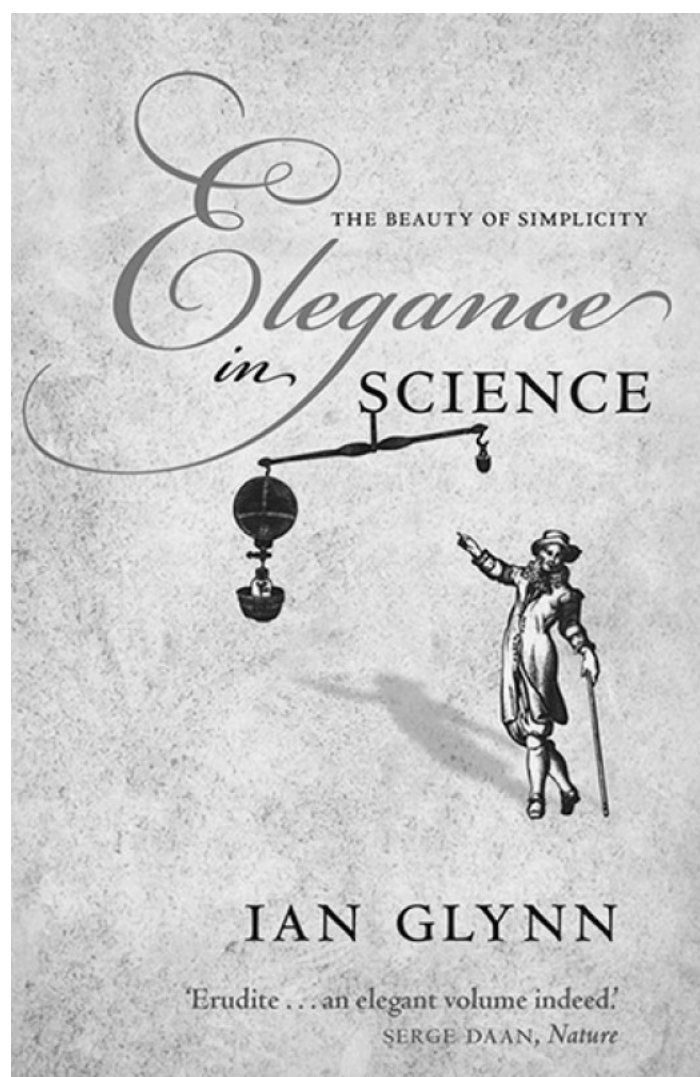
— Nature

«Há aqui uma abundância de informação histórica condensada.»

— Times Literary Supplement

A ideia de elegância na ciência não é necessariamente familiar, mas é uma ideia importante.

O uso do termo é talvez mais evidente na matemática — a demonstração elegante — e é precisamente aí que Ian Glynn inicia a sua exploração. Os cientistas partilham frequentemente um sentimento de admiração e entusiasmo ao tomarem conhecimento de uma solução elegante para um problema, de uma teoria elegante ou de uma experiência elegante. Através de uma seleção altamente legível de episódios inspiradores que sublinham o papel da beleza e da simplicidade nas ciências, este livro aborda também questões filosóficas fundamentais relativas à inferência. Glynn conclui com um aviso prudente: não devemos confiar apenas na beleza e na simplicidade — até a explicação mais elegante pode estar errada.



O QUE É A QUÍMICA?

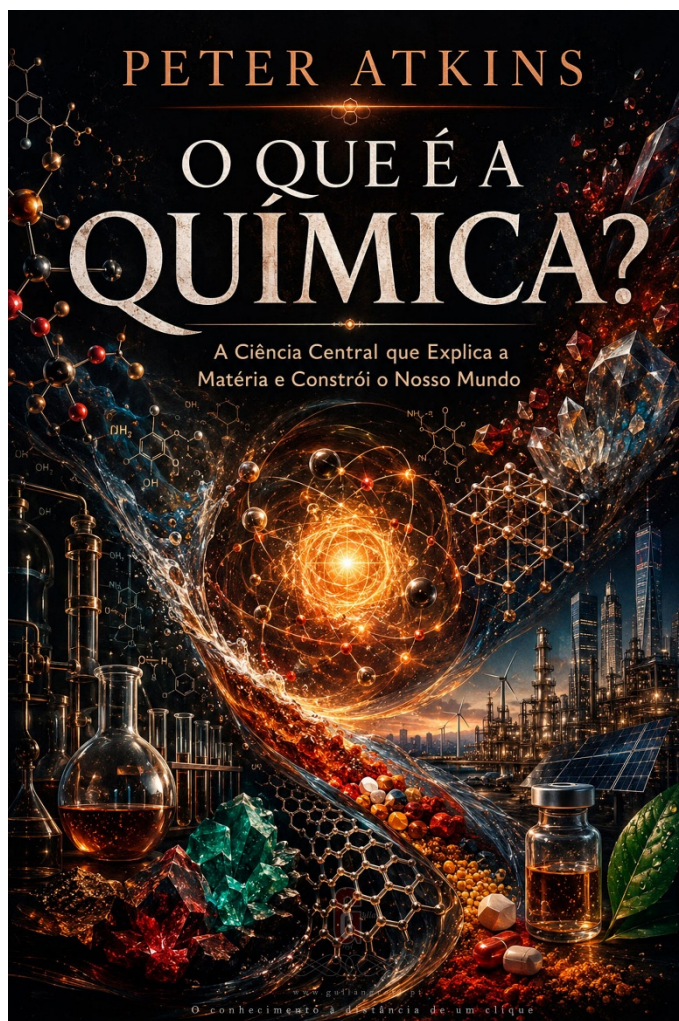


Peter Atkins

«Atkins conquista a atenção dos seus leitores simplesmente através de uma descrição elegante e lúcida da disciplina que ama.»

— Nature

Em *O Que É a Química?*, Peter Atkins convida-nos a olhar para a química de novo, através dos olhos de um químico, para compreender os seus conceitos centrais e perceber de que modo contribui não apenas para o nosso conforto material, mas também para a cultura humana. Mostra como a química fornece a infraestrutura do nosso mundo: através da indústria química, dos combustíveis que sustentam o aquecimento, a produção de energia e o transporte, bem como dos tecidos que compõem a nossa roupa e o nosso mobiliário. Ao considerar as realizações notáveis da química e ao examinar o seu lugar entre a física e a biologia, Atkins oferece uma exploração fascinante, clara e rigorosa do universo químico.



WHAT IS LIFE?

How Chemistry Becomes Biology

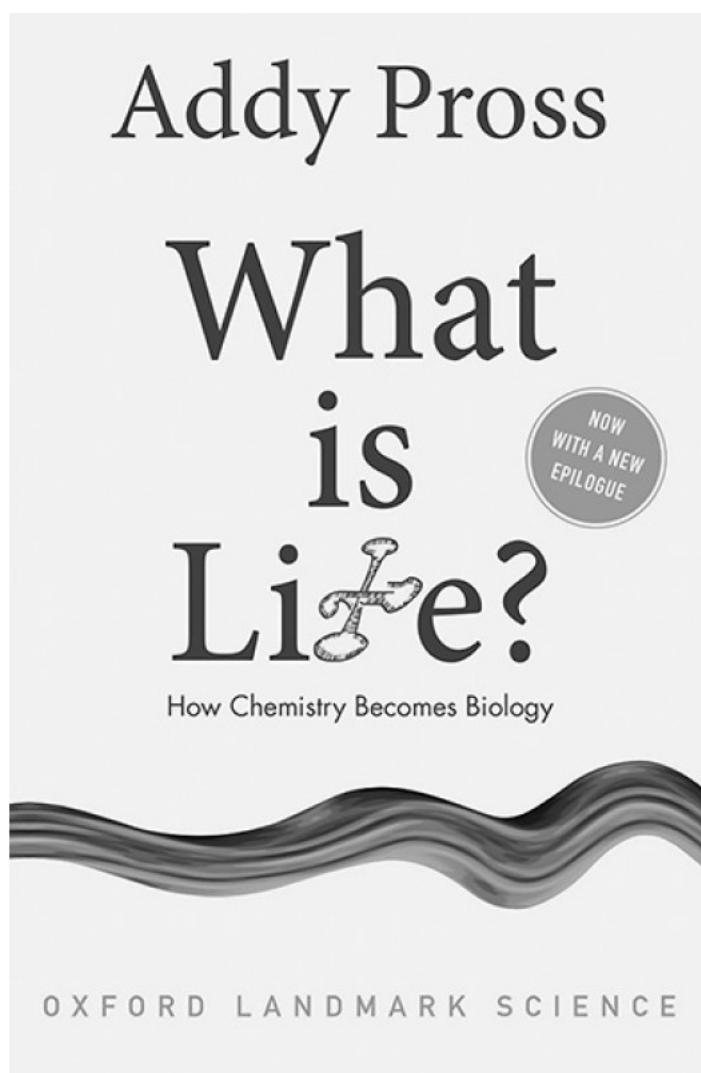


Addy Pross

«Pross realiza um excelente trabalho ao transmitir de forma sucinta a dificuldade em formular uma definição geral inequívoca de vida e oferece um roteiro para grande parte do trabalho sobre a origem da vida realizado por químicos nos últimos 50 anos. O livro vale a leitura só por estas discussões.»

— Chemical Heritage

Os seres vivos são imensamente complexos e possuem propriedades singulares, como a automanutenção e um comportamento aparentemente orientado para fins, que não observamos na matéria inerte. Como pode, então, a química dar origem à biologia? Que processos poderão ter conduzido as primeiras moléculas replicantes por esse caminho? Os desenvolvimentos no emergente campo da «química de sistemas» estão agora a desbloquear este problema. O fosso entre a biologia e as ciências físicas começa, finalmente, a ser transposto.



REAÇÕES

A Vida Privada dos Átomos

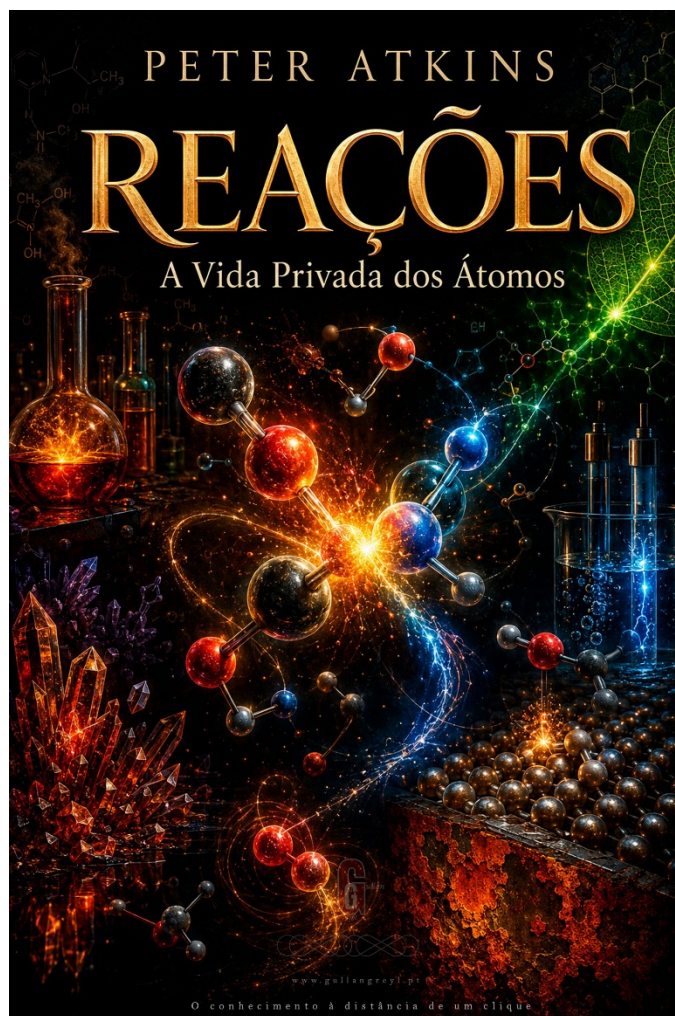


Peter Atkins

«O antídoto perfeito contra a fobia da ciência.»

— Booklist

Peter Atkins capta o cerne da química neste livro, através de uma conceção inovadora e estreitamente integrada entre imagens e texto, aliada à sua exposição caracteristicamente clara, precisa e económica. Ao explicar os processos envolvidos nas reações químicas, começa por introduzir uma “caixa de ferramentas” de reações fundamentais, como a precipitação, a corrosão e a catálise, e conclui mostrando como esses blocos estruturais se articulam em processos mais complexos, como a fotossíntese, oferecendo assim uma introdução concisa e intelectualmente estimulante à vida privada dos átomos.



HIPERESPAÇO

Uma Odisseia Científica pelos Universos Paralelos, Distorções do Tempo e a Décima Dimensão



Michio Kaku

«Belissimamente escrito, tornando ideias científicas difíceis acessíveis, quase fáceis.»

— Danah Zohar, The Independent

«Desde que assimilou a teoria da relatividade de Einstein, a física caiu numa toca quântica de coelho e, desde então, os relatos dos físicos ao mundo da divulgação científica tornaram-se cada vez mais curiosos... [Kaku] delinea as “deliciosas contradições” da revolução quântica. As suas perceções intelectuais entusiasmarão leitores leigos, fãs de ficção científica e aqueles familiarizados com a física.»

— Publisher's Weekly

Neste livro clássico, Michio Kaku descreve o desenvolvimento das ideias sobre o espaço multidimensional. Conduz o leitor numa viagem pelo hiperespaço até às fronteiras da física. Pelo caminho, oferece explicações cristalinas de conceitos matemáticos formidáveis como a geometria não euclidiana, a Teoria de Kaluza–Klein e a Supergravidade — ferramentas quotidianas do teórico das cordas. Recorrendo a episódios fascinantes e frequentemente bem-humorados da história, da arte e da ficção científica, mostra-nos que escritores e artistas — além dos cientistas — se deixaram fascinar pelo espaço multidimensional há mais de um século.



SOBRE SER

A Exploração, Por Um Cientista, Das Grandes Questões da Existência



Peter Atkins

«Conciso e sensato, claro na forma como transmite o conhecimento científico com aparente facilidade, e delicioso naquele tipo de humor que o faz parar e pousar o livro apenas para o saborear melhor. Apresenta uma visão da vida e da morte, da matéria, do espaço e do tempo que é honesta, coerente e livre de milagres.»

— Philip Pullman

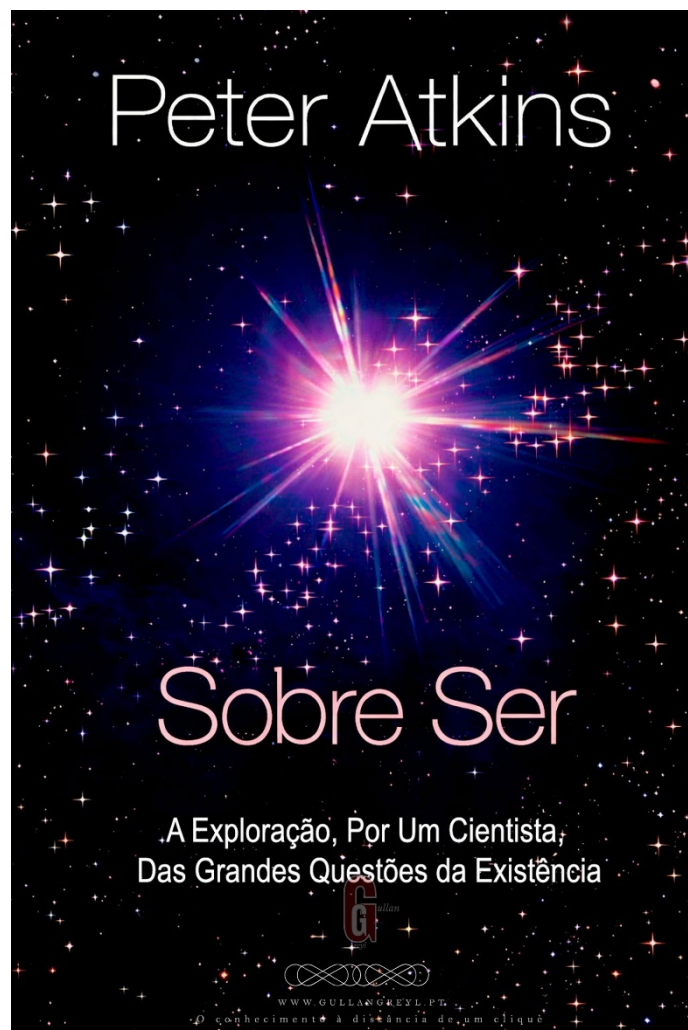
«Poucos conseguem igualar a beleza lapidar da prosa de Peter Atkins quando reflete sobre a natureza da vida e da morte, dos começos e dos fins.»

— Richard Dawkins

«Um hino à ciência.»

— Times Literary Supplement

Neste “Credo” científico, Peter Atkins aborda as questões universais das origens, dos fins, do nascimento e da morte, às quais as religiões têm reivindicado respostas. Com a sua habitual economia de expressão, humor e elegância — sem recuar perante realidades incômodas — Atkins expõe o que a ciência tem a dizer. Reconhecendo o conforto que alguns encontram na crença, declara a sua própria confiança na capacidade da ciência para revelar as verdades mais profundas.



ORIGENS

A História Científica da Criação



Jim Baggott

«Origens relata a maior história jamais contada: a evolução do Universo desde o Big Bang. Este rico relato interdisciplinar recorda-nos que a astronomia, a física, a química, as geociências, a biologia e a neurociência estão interligadas. Baggott conduz o leitor numa viagem linear de 13,8 mil milhões de anos. Recorda-nos que permanecem grandes questões nesta maravilhosa aventura científica.»

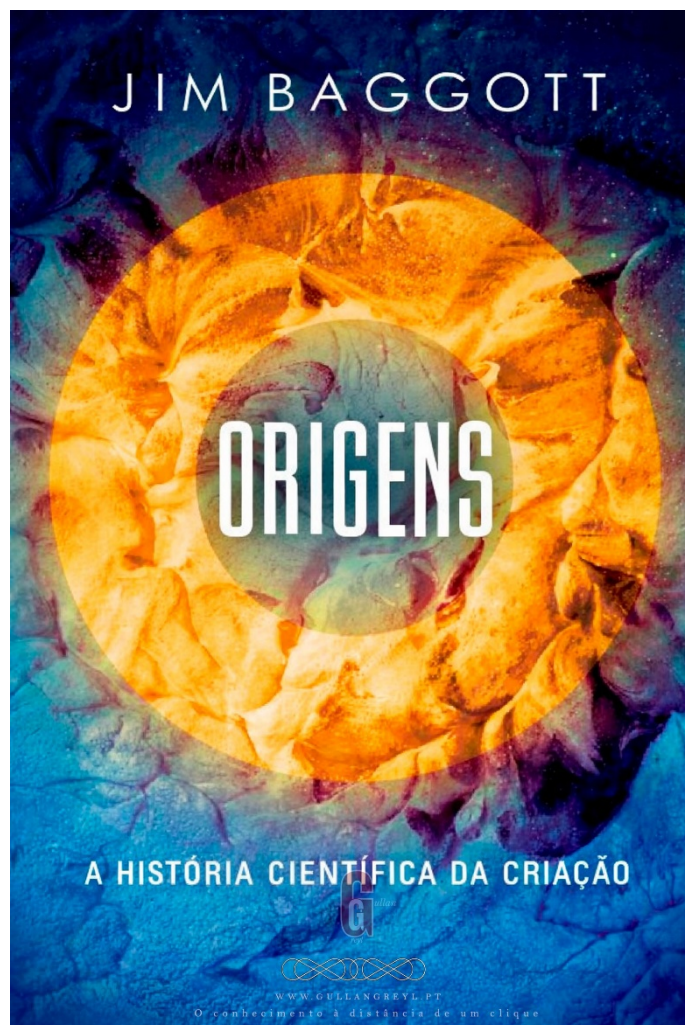
— Nature

«Um impressionante tour de force científico.»

— Chemistry & Industry

O que é a vida? De onde viemos e como evoluímos? O que é o universo e como se formou? Qual é a natureza do mundo material? Como funciona? Como e por que pensamos? O que significa ser humano? Como sabemos o que sabemos?

Existem muitas versões diferentes da nossa história de criação. Este livro apresenta a versão segundo a ciência moderna, começando no Big Bang e avançando até à emergência dos seres humanos como entidades conscientes e inteligentes, 13,8 mil milhões de anos depois. Capítulo após capítulo, expõe o estado atual do conhecimento científico: as origens do espaço e do tempo; da energia, da massa e da luz; das galáxias, das estrelas e do nosso Sol; da Terra habitável e da própria vida complexa. Baggott relata o que atualmente sabemos sobre a nossa história, sublinhando as questões que a ciência ainda tem por responder.



Notas

As notas que se seguem referem-se à velocidade da luz (c), à constante de Planck (h), à constante de Boltzmann (k) e à carga fundamental (e).

CAPÍTULO 1

¹ Sempre admirei a obra de Max Jammer, "O desenvolvimento conceptual da mecânica quântica" (McGraw-Hill, 1966), pela sua abordagem minuciosa da emergência da teoria.

² A lei de Hooke afirma que $F = -k_f x$, onde F é a força restauradora, x é o deslocamento a partir do equilíbrio (a mola em repouso) e k_f é uma característica da mola conhecida como a sua constante elástica. As molas rígidas têm constantes elásticas elevadas. Consulte o Capítulo 6 para obter mais informações.

³ Uma forma da lei de Boyle afirma que, a temperatura constante, $V \propto 1/p$, onde V é o volume ocupado pelo gás quando a pressão é p . Conclui-se que o produto pV é uma constante para uma dada amostra de gás a temperatura constante. Consulte o Capítulo 6 para obter mais informações.

CAPÍTULO 2

¹ Uma explicação acessível do teorema de Noether pode ser encontrada em *Emmy Noether's wonderful theorem*, de Dwight Neuenschwander (Johns Hopkins University Press, 2010). Para uma abordagem mais substancial, recomenda-se *The Noether theorems: invariance and conservation laws in the twentieth century*, de Yvette Kosmann-Schwarzbach (Springer, 2011).

² A energia cinética de um corpo de massa m que se desloca a uma velocidade v é $\frac{1}{2}mv^2$. A energia potencial de um corpo de massa m a uma altura h acima da superfície da Terra é mgh , sendo g uma constante, a "aceleração da gravidade" (o seu valor é próximo de $9,8 \text{ m/s}^2$). A energia de um campo eletromagnético é proporcional ao quadrado da soma das intensidades dos seus campos elétrico e magnético.

³ A deteção experimental do neutrino foi realizada por F. B. Harrison, H. W. Kruse e A. D. McGuire, que receberam o Prémio Nobel da Física, mas apenas em 1995, quarenta anos depois. Imagine estar na expectativa durante quarenta meses consecutivos de outubro!

⁴ Há ainda outros problemas a considerar nestas considerações. A distinção entre energia e momento linear (abordada mais adiante neste capítulo) depende do estado de movimento do observador e do observado, e ao longo desta discussão deveríamos estar a considerar a uniformidade do espaço-tempo em vez de cada componente em separado. Peço desculpa por ter ignorado este problema na exposição (mas não na minha mente).

⁵ Se o intervalo de tempo entre gerações se reduzir para metade em cada geração anterior, o período total seria finito mesmo para um número infinito de gerações ($1 + \frac{1}{2} + \frac{1}{4} + \dots = 2$), mas o tempo necessário para o Universo Primordial ter a sua filha teria sido infinitesimal. Ficaria desiludido se o Universo Primordial tivesse surgido há um tempo infinito, pois isso contradiria tudo o que digo, mas sem dúvida agradaria àqueles com uma certa mentalidade.

⁶ O comprimento de Planck é definido como

$$L_p = \sqrt{hG/2\pi c^3}$$

, onde G é a constante gravitacional, e resulta em cerca de $1,6 \times 10^{-35}$ metros. Isto é aproximadamente um trilionésimo de trilionésimo do diâmetro de um núcleo atómico. O *tempo de Planck* é definido como o tempo que a luz demora a percorrer essa distância, logo

$$t_p = \sqrt{hG/2\pi c^5}$$

, o que resulta em $5,4 \times 10^{-44}$ segundos. Para cúmulo, nada mais, refiro a massa de Planck,

$$m_p = \sqrt{hc/2\pi G}$$

, que resulta em cerca de 22 microgramas, um valor razoavelmente imaginável. Uma página deste livro pesa cerca de 140.000 massas de Planck.

⁷ Tenho consciência, mas não me distraio, do Hino da Criação no Rigveda:

1.º Então não era inexistente nem existente: não havia reino do ar, nem céu além dele. O que o cobria e onde? E o que lhe dava abrigo? Havia ali água, uma profundidade insondável?

2.º A morte não existia então, nem havia nada imortal: não havia sinal algum, o divisor entre o dia e a noite. Aquela Única Coisa, ofegante, respirava pela sua própria natureza: para além dela, não havia absolutamente nada.

3.º Havia escuridão: a princípio, oculto na escuridão, este Todo era caos indiscriminado.

⁸ Há muitas (cerca de 2×10^{52}) épocas de Planck, especulei sobre como poderia ter ocorrido o surgimento de algo a partir do Nada, em A Criação (W. H. Freeman & Co., 1981) e revisitei o tema em A Criação Revisitada (W. H. Freeman & Co., 1992).

⁹ O momento linear, p , de um corpo de massa m está relacionado com a velocidade, v , por $p = mv$.

¹⁰ O momento angular J está relacionado com a velocidade angular ω por $J = I\omega$, onde I é o momento de inércia. O momento de inércia de um corpo de massa m que roda numa trajetória de raio r é $I = mr^2$.

CAPÍTULO 3

¹ A lei de Snell da refração afirma que o ângulo de incidência e refração quando um raio passa pela interface de dois meios com índices de refração n_{r1} e n_{r2} é $\sin \theta_1 / \sin \theta_2 = n_{r2}/n_{r1}$.

² Aqui fica uma dica caso se depare com uma emergência destas, de alguém a afogar-se num lago. Se realmente consegue andar dez vezes mais depressa do que consegue atravessar a água na margem, e você e o seu amigo que se está a afogar estão a distâncias iguais da borda da água, e a mesma distância um do outro paralelamente a essa borda, então um cálculo rápido, mas minucioso (que é melhor fazer agora do que depois), mostra que deve caminhar até um ponto a 93% de uma dessas distâncias na margem, e depois atravessar a água a partir daí.

³ **Suponha** que a amplitude de uma onda que chega ao destino por um determinado percurso é a_0 a 0. A amplitude de uma onda que chega por um percurso ligeiramente diferente, descrito por um parâmetro p — uma medida da curvatura (ou “flexibilidade”) do percurso — é a_p . As duas amplitudes estão relacionadas por

$$a_p = a_0 + p \left(\frac{da}{dp} \right) + p^2 \left(\frac{d^2a}{dp^2} \right) + \dots$$

Se o percurso corresponder a um **mínimo**, então o termo $\frac{da}{dp} = 0$ e as duas amplitudes diferem apenas à **segunda ordem** em p ; todos os outros percursos diferem muito mais, à **primeira ordem** em p . Os especialistas saberão que eu deveria estar a discutir o **comprimento de fase**, e não a amplitude.

⁴ O comprimento de onda, λ , de uma partícula com momento linear de magnitude p é dado pela “relação de, de Broglie” $\lambda = h / p$, onde h é a constante de Planck (ver Capítulo 8), que foi proposta por Louis de Broglie (1892–1987) em 1924 e posteriormente demonstrada como consequência de uma formulação mais geral da mecânica quântica.

⁵ A definição formal de ação, S , é a seguinte:

$$S = \int_{\text{trajetória}} L(q, \dot{q}) dt$$

onde a integração é efetuada ao longo da trajetória, em passos infinitesimais dt ; q representa a posição da partícula, $\dot{q}$ a sua velocidade, e

$$L(q, \dot{q})$$

é o lagrangiano do sistema. (No texto original, Atkins escreve a integral em função de um parâmetro genérico “ s ” ao longo da trajetória “ ds ”. Aqui usamos o tempo “ t ”, portanto “ dt ”, por maior clareza expositiva.) Em certos casos, o lagrangiano L é simplesmente a diferença entre a energia cinética e a energia potencial da partícula, como acontece em

$$L = \frac{1}{2} m \dot{q}^2 - \frac{1}{2} k_f q^2$$

, por exemplo, no caso de um oscilador harmónico.

⁶ Uma versão da mecânica quântica baseada no conceito de caminhos interferentes é a formulação da teoria de "integral de caminho" de Feynman, que é apresentada em R. P. Feynman e A. R. Hibbs, *Quantum mechanics and path integrals* (McGraw-Hill, 1965).

⁷ Se a onda tiver amplitude a na sua origem, a sua amplitude e fase num ponto distante são $ae^{iS/\hbar}$, onde S é a ação associada ao caminho,

$$\hbar = h/2\pi, \text{ e } i = \sqrt{-1}.$$

⁸ A **segunda lei de Newton** é a equação diferencial

$$F = \frac{dp}{dt}$$

, em que F é a força e p é o momento linear.

Um exemplo mais complexo é a **equação de Schrödinger** para uma partícula de massa m e energia E numa região unidimensional onde a energia potencial é $V(x)$:

$$-\frac{\hbar^2}{2m} \frac{d^2\phi}{dx^2} + V(x)\phi = E\phi$$

, onde

$$\hbar = \frac{h}{2\pi}$$

e ϕ é a sua "função de onda", uma função matemática que contém toda a informação dinâmica sobre a partícula.

⁹ **Para encontrar o caminho correspondente à ação mínima** (tal como definido na nota 5), procura-se o percurso tal que

$$\delta \int_{\text{trajetória}} L(q, \dot{q}) dt = 0$$

, onde δ denota uma variação no percurso.

Esta minimização é satisfeita desde que

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} = 0$$

, que é uma equação diferencial conhecida como **equação de Euler-Lagrange**.

Para um **lagrangiano** da forma

$$L = \frac{1}{2} m \dot{q}^2 - V(q)$$

, a equação de Euler-Lagrange reduz-se à **segunda lei de Newton**.

CAPÍTULO 4

¹ A distribuição de Boltzmann implica que a razão entre os números de moléculas N_1 e N_2 em estados com energias E_1 e E_2 , a uma temperatura absoluta T , é

$$\frac{N_1}{N_2} = e^{-\frac{(E_2 - E_1)}{kT}}.$$

Nestas notas, o símbolo T significa sempre a temperatura absoluta.

² Esses outros pedantes saberão que estou a referir-me à energia de ponto zero de certos tipos de movimento, a energia que, por razões de mecânica quântica, não pode ser removida. Não é possível, por exemplo, que um pêndulo esteja absolutamente imóvel.

³ Para obter a temperatura absoluta na escala Kelvin a partir da temperatura em Celsius, adicione 273,15 a esta última. Logo, 20 °C equivalem a 293 K.

⁴ A lei de Arrhenius refere que a velocidade de uma reação química é proporcional a

$$e^{-\frac{E_a}{RT}},$$

onde E_a é a energia de ativação e R é a constante dos gases

$$R = N_A k.$$

⁵ A lei do arrefecimento de Newton afirma que a diferença de temperatura entre um corpo e o seu meio envolvente, ΔT , varia ao longo do tempo segundo a expressão

$$\Delta T(t) = \Delta T(0)e^{-Kt},$$

onde K é uma constante que depende da massa e da composição do corpo.

Esta expressão é a forma global da lei: descreve diretamente como a diferença de temperatura evolui no tempo, mostrando que o arrefecimento ocorre de modo exponencial.

A forma operacional subjacente, que governa o processo instante a instante, é a equação diferencial

$$\frac{d(\Delta T)}{dt} = -K \Delta T,$$

a qual afirma que a taxa de variação da diferença de temperatura é proporcional ao seu valor no próprio instante.

A forma global resulta da integração da forma operacional e explicita o comportamento total do processo ao longo do tempo.

⁶ A lei da desintegração radioativa afirma que o número de núcleos ativos, N , varia com o tempo segundo

$$N(t) = N(0)e^{-Kt}$$

, onde K é uma constante que depende da identidade do nuclídeo e está relacionada com a meia-vida radioativa, $t_{1/2}$, pela expressão

$$K = \frac{\ln 2}{t_{1/2}}$$

A expressão acima é a **forma global** da lei e descreve o comportamento estatístico do conjunto de núcleos ao longo do tempo.

O seu conteúdo operativo pode ser entendido assim: cada núcleo radioativo tem uma **probabilidade constante de decair por unidade de tempo**, independente dos restantes núcleos. Como consequência dessa dinâmica individual e anárquica, o número total de núcleos ativos diminui de forma **exponencial**.

A constante K caracteriza a rapidez desse decaimento e está diretamente ligada à meia-vida $t_{1/2}$, definida como o tempo necessário para que o número de núcleos ativos se reduza a metade do seu valor inicial.

Tal como no arrefecimento de Newton, a lei exponencial emerge de um comportamento microscópico anárquico, no qual cada núcleo evolui de forma independente.

CAPÍTULO 5

¹ A fórmula de Boltzmann para a entropia é $S = k \ln W$, onde W é o número de formas pelas quais as moléculas podem ser distribuídas e ainda assim manter a mesma energia total. O "ln" moderno denota o logaritmo natural; o epitáfio de Boltzmann utiliza log em vez de ln. Seria interessante pensar que a letra S , com o seu formato sinuoso, foi escolhida por Clausius para captar a ideia de "rotação" na entropia, mas entendo que ela estava simplesmente a mais, uma vez que as suas vizinhas R e T já estavam em uso.

² A expressão de Clausius para a variação da entropia, ΔS , quando uma quantidade de energia q é transferida sob a forma de calor para um corpo à temperatura absoluta T , é

$$\Delta S = \frac{q}{T}.$$

Existem, contudo, certas restrições técnicas sobre a forma como o calor é transferido, nomeadamente a exigência de que a transferência seja realizada de modo reversível, o que, na prática, significa que deve ocorrer com uma diferença de temperatura mínima entre o corpo aquecido e a fonte de aquecimento em todas as fases do processo.

³ A eficiência, η , de um motor é definida como a razão entre o trabalho produzido e o calor consumido.

A expressão de Carnot para a eficiência de um motor térmico ideal que opera entre uma fonte quente à temperatura absoluta T_{quente} e um reservatório frio à temperatura absoluta T_{frio} é

$$\eta = 1 - \frac{T_{frio}}{T_{quente}}.$$

A eficiência aproxima-se de 1 à medida que a temperatura do reservatório frio se aproxima de zero ou a temperatura da fonte quente se aproxima do infinito. Temperaturas elevadas são mais fáceis (e mais baratas) de atingir do que temperaturas muito

baixas; por isso, os engenheiros procuram aumentar a temperatura da fonte quente (por exemplo, utilizando vapor sobreaquecido) para obter a maior eficiência possível. Para uma fonte quente a 200 °C (473 K) e um reservatório frio a 20 °C (293 K), a eficiência é $\eta=0,38$, ou seja, apenas 38% da energia transferida sob a forma de calor pode ser convertida em trabalho, mesmo num motor ideal.

⁴ Kelvin exprimiu esta ideia da seguinte forma: *Não é possível produzir trabalho mecânico simplesmente fazendo um corpo arrefecer para uma temperatura inferior à do ambiente mais frio que o rodeia, sem recorrer a intervenção externa.*

⁵ Clausius escreveu (aqui em tradução): A transferência de energia na forma de calor de um corpo mais frio para um corpo mais quente não ocorre espontaneamente; só é possível se for acompanhada por outra transformação — tipicamente a realização de trabalho.

⁶ Para mais informações sobre a relação entre a inatingibilidade do zero absoluto e o valor da entropia, consulte o meu livro "As leis da termodinâmica: uma breve introdução" (Oxford University Press, 2010) ou, mais rigorosamente, o meu livro (com Julio de Paula e James Keeler) "Química Física" (11ª edição, Oxford University Press, 2018).

⁷ Para ajudar a avaliar o contributo de Prigogine, ver I. Prigogine e I. Stengers, *O fim da certeza* (The Free Press, 1997). O Rei dos Belgas parece ter admirado, ou ter sido aconselhado a admirar, o seu trabalho, pois Prigogine foi nomeado visconde em 1989.

CAPÍTULO 6

¹ A lei do gás perfeito é

$$pV = NkT,$$

em que p é a pressão, V o volume, N o número de moléculas presentes e T a temperatura absoluta. Os químicos escrevem normalmente a lei em termos da quantidade de matéria, n , definida por

$$n = \frac{N}{N_A},$$

sendo N_A a constante de Avogadro, e escrevem $N_A k = R$, a constante dos gases. A lei passa então a ter a forma

$$pV = nRT.$$

² A maioria das pessoas refere-se a esta lei como a lei do gás ideal. No entanto, eu prefiro manter a designação perfeito. A razão é a seguinte. Existem sistemas designados por soluções ideais, nos quais as moléculas de soluto e de solvente interagem entre si, mas em que uma molécula não "sabe" se a sua vizinha é de soluto ou de solvente: as interações entre elas são indistinguíveis. Num gás perfeito isso também é verdade, mas há mais: não só as interações entre moléculas são todas iguais, como são, além disso, nulas. Assim, a perfeição corresponde a mais um passo adiante no caminho que vai além da simples idealidade.

³ A lei de Henry afirma que a concentração, c , de um gás dissolvido num líquido em equilíbrio é proporcional à pressão do gás ($c = Kp$); A lei de Raoult afirma que a presença de um soluto reduz a pressão de vapor do solvente em Δp , em proporção com a concentração do soluto ($\Delta p = Kc$); A lei de van 't Hoff afirma que a pressão osmótica, Π , é proporcional à concentração do soluto ($\Pi = Kc$). Os valores de K são diferentes em cada caso.

- K é uma constante de proporcionalidade que depende da natureza do sistema (tipo de gás, solvente, temperatura, etc.).
- Em todas estas leis, a ideia central é a mesma: um efeito macroscópico mensurável emerge linearmente da concentração, revelando regularidade estatística por trás do comportamento microscópico.

⁴ O tratamento matemático conduz à expressão

$$pV = \frac{1}{3} N m v_{rms}^2,$$

onde N é o número de moléculas contidas no volume V , m é a massa de uma molécula, e v_{rms} é a velocidade quadrática média, isto é, a raiz quadrada do valor médio dos quadrados das velocidades das moléculas. Pode pensar nela, de forma algo livre, como uma medida da velocidade média das moléculas. A temperatura constante, esta expressão reduz-se à forma $pV = \text{constante}$, que é precisamente a lei de Boyle.

⁵ A velocidade média (aritmética) das moléculas de massa m num gás à temperatura absoluta T é dada por

$$v_{mean} = \left(\frac{8kT}{\pi m} \right)^{1/2}.$$

Isto é,

$$v_{mean} \propto \sqrt{T}.$$

Aqui, v_{mean} designa a velocidade média das moléculas, e o símbolo $\propto$ significa “é proporcional a”. Assim, afirmar que $v_{mean} \propto \sqrt{T}$ indica que a velocidade média aumenta com a raiz quadrada da temperatura absoluta, independentemente dos fatores constantes.

⁶ Como foi referido na primeira nota do Capítulo 1, a lei de Hooke afirma que

$$F = -k_f x,$$

onde F é a força restauradora e x é o deslocamento em relação à posição de equilíbrio. O sinal negativo indica que a força atua sempre no sentido oposto ao deslocamento: quanto mais a mola é afastada do equilíbrio, maior é a força que a tenta fazer regressar. A frequência de um oscilador de massa m ligado a uma mola é dada por

$$\nu = \left(\frac{1}{2\pi}\right) \left(\frac{k_f}{m}\right)^{1/2}.$$

onde:

- ν = frequência (não velocidade)
- k_f = constante elástica da mola
- m = massa associada ao oscilador

No caso de um pêndulo simples de comprimento l , a frequência é

$$\nu = \left(\frac{1}{2\pi}\right) \left(\frac{g}{l}\right)^{1/2},$$

onde:

- ν = frequência
- g é a aceleração da gravidade, isto é, uma medida da intensidade do campo gravitacional.
- l = comprimento do pêndulo

Este último resultado é também uma lei limite, no sentido em que é rigorosamente válido apenas no limite em que a amplitude da oscilação tende para zero — isto é, quando o pêndulo oscila com ângulos muito pequenos.

⁷ A expressão mais geral para uma propriedade que tem o valor $P(x)$ quando o deslocamento em relação ao equilíbrio é x é:

$$P(x) = P(0) + \left(\frac{dP}{dx}\right)_{x=0} x + \frac{1}{2} \left(\frac{d^2P}{dx^2}\right)_{x=0} x^2 + \dots$$

No mínimo da curva que descreve a dependência de P em função de x , tem-se:

$$\left(\frac{dP}{dx}\right)_{x=0} = 0.$$

Assim, o primeiro termo não nulo após $P(0)$ é:

$$\frac{1}{2} \left(\frac{d^2P}{dx^2}\right)_{x=0} x^2.$$

Se P for a energia potencial E_p , então, dado que a força restauradora F e a energia potencial estão relacionadas por

$$F = -\frac{dE_p}{dx},$$

neste cenário geral resulta

$$F = -\left(\frac{d^2P}{dx^2}\right)_{x=0} x,$$

o que corresponde à lei de Hooke quando

$$\left(\frac{d^2P}{dx^2}\right)_{x=0}$$

é identificado com a constante elástica k_f . Neste caso particular, $P \equiv E_p \equiv$ (identidade) — indica que dois símbolos representam o mesmo objeto por definição, não apenas o mesmo valor numérico.

⁸ A energia de um oscilador harmónico, um oscilador que obedece à lei de Hooke, é

$$E = \frac{p^2}{2m} + \left(\frac{k_f}{2}\right)x^2,$$

em que m é a sua massa. Note-se a simetria: tanto o momento linear p como o deslocamento x ocorrem como seus quadrados.

⁹ A estrutura e o seu padrão de difração são essencialmente a transformada de Fourier um do outro. Uma “descrição de posição” e uma “descrição de momento” do mundo são igualmente a transformada de Fourier uma da outra.

CAPÍTULO 7

¹ A lei do inverso do quadrado de Coulomb identifica a magnitude da força entre duas cargas elétricas Q_1 e Q_2 , separadas por uma distância r , como

$$F = \frac{Q_1 Q_2}{4\pi\epsilon_0 r^2},$$

em que ϵ_0 é uma constante fundamental, a *permissividade* do vácuo. A energia potencial elétrica associada às duas cargas é

$$E_p = \frac{Q_1 Q_2}{4\pi\epsilon_0 r}.$$

Uma lei semelhante do inverso do quadrado expressa a magnitude da força gravitacional entre duas massas m_1 e m_2 como

$$F = \frac{Gm_1 m_2}{r^2},$$

em que G é a constante gravitacional.

² A designação completa da teoria de grupos da simetria da interação de Coulomb é $SO(4)$, o “grupo ortogonal especial em quatro dimensões”.

³ Num átomo de hidrogénio, os orbitais atómicos da mesma camada (designada pelo número quântico principal n) possuem a mesma energia, independentemente do seu momento angular em torno do núcleo (designado pelo número quântico do momento angular l). Assim, os orbitais s , p , d ,... da mesma camada têm a mesma energia. A “degenerescência”, a posse da mesma energia, está sempre associada à simetria; neste caso, é uma consequência da hipersfericidade quadridimensional da interação Coulombiana, que permite que estes orbitais, com as suas diversas formas, sejam rodados uns nos outros em quatro dimensões.

⁴ Se a onda original for $\psi(x)$, então, sob uma transformação de calibre (gauge) global — isto é, um desvio uniforme da fase por um ângulo ϕ — ela transforma-se em

$$\psi(x)e^{i\phi}.$$

A densidade de probabilidade da partícula é

$$\psi^*(x)\psi(x)$$

antes da transformação e torna-se

$$\psi^*(x)e^{-i\phi}\psi(x)e^{i\phi} = \psi^*(x)\psi(x)$$

após a transformação. Ou seja, permanece invariante. Esta invariância subsiste mesmo sob uma transformação de calibre local, em que a fase passa a depender da posição, $\phi(x)$, porque continua a verificar-se que

$$\psi^*(x)e^{-i\phi(x)}\psi(x)e^{i\phi(x)} = \psi^*(x)\psi(x).$$

⁵ Aqui está o argumento técnico relacionando uma transformação global do medidor à conservação da carga. Estou eliminando o máximo possível da notação e pretendo simplesmente mostrar o caminho através do argumento: para fazê-lo corretamente, você precisaria considerar as derivadas de tempo, bem como a derivada de espaço único usada aqui. Considere um empurrão infinitesimal, tal que a transformação

$$\psi(x) \rightarrow e^{i\phi}\psi(x)$$

pode ser aproximado por

$$\psi(x) \rightarrow (1 + i\phi)\psi(x) = \psi(x) + \delta\psi(x)$$

com

$$\delta\psi(x) = i\phi\psi(x).$$

A mudança resultante na densidade lagrangiana

$$L(\psi, \psi') = \frac{1}{2}\psi'^2 - \frac{1}{2}m\psi^2,$$

onde

$$\psi' = \frac{\partial\psi}{\partial x},$$

é

$$\delta L = \frac{\partial L}{\partial\psi} \delta\psi + \frac{\partial L}{\partial\psi'} \delta\psi' = \left\{ \frac{\partial L}{\partial\psi} - \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial\psi'} \right) \right\} \delta\psi + \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial\psi'} \delta\psi \right).$$

Observe que, de acordo com a equação de Euler-Lagrange (a equação que lhe diz como tatear ao longo do caminho para minimizar a ação geral),

$$\frac{\partial L}{\partial\psi} - \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial\psi'} \right) = 0.$$

Por isso

$$\delta L = \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial\psi'} \delta\psi \right) = i\phi \frac{\partial}{\partial x} \psi' \psi.$$

A densidade lagrangiana permanece inalterada sob a transformação de calibre global, portanto $\delta L = 0$, para ϕ arbitrário. Por isso,

$$\frac{\partial}{\partial x} \widehat{\psi' \psi}^J = 0$$

e a corrente j é conservada.

J — corrente elétrica associada ao campo ψ (aqui usada em sentido conceptual, sem distinção entre corrente total e densidade de corrente).

⁶ Admita que a função de onda $\psi(x)$ satisfaz a equação de Schrödinger

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x) \psi(x) = E \psi(x)$$

Agora, desloque a fase da função de onda para

$$\psi(x)e^{i\phi(x)} = \tilde{\psi}(x)$$

Esta função com fase deslocada já não satisfaz a mesma equação, porque

$$\begin{aligned} & -\frac{\hbar^2}{2m} \frac{d^2\tilde{\psi}(x)}{dx^2} + V(x)\tilde{\psi}(x) \\ &= -\frac{\hbar^2}{2m} \left\{ \frac{d^2\psi}{dx^2} + 2i \frac{d\phi}{dx} \frac{d\psi}{dx} - \left(\frac{d\phi}{dx} \right)^2 \psi + i \frac{d^2\phi}{dx^2} \psi \right\} e^{i\phi(x)} + V(x)\tilde{\psi}(x) \\ &= E\tilde{\psi}(x) - \frac{\hbar^2}{2m} \left\{ 2i \frac{d\phi}{dx} \frac{d\psi}{dx} - \left(\frac{d\phi}{dx} \right)^2 \psi + i \frac{d^2\phi}{dx^2} \psi \right\} e^{i\phi(x)} \end{aligned}$$

Os três termos adicionais indesejados são eliminados se a equação de Schrödinger for modificada para

$$-\frac{\hbar^2}{2m} \frac{d^2\tilde{\psi}(x)}{dx^2} + U(x)\tilde{\psi}(x) + V(x)\tilde{\psi}(x) = E\tilde{\psi}(x)$$

com

$$U(x) = \frac{\hbar^2}{2m} \left\{ 2i \left(\frac{d\phi}{dx} \right) \frac{d}{dx} + \left(\frac{d\phi}{dx} \right)^2 + i \frac{d^2\phi}{dx^2} \right\}$$

O termo adicional, $U(x)$, é semelhante à contribuição de energia $V(x)$ e é um termo de interação com o campo. Assim, as interações emergem da invariância de calibre local. Note-se que o termo proporcional a

$$\frac{d}{dx}$$

é, na realidade, proporcional ao operador momento linear,

$$p = \left(\frac{\hbar}{i}\right) \frac{d}{dx}$$

CAPÍTULO 8

¹ A relação entre a frequência, ν , e a energia, E , dos quanta associados é $E = h\nu$, sendo que, em unidades convencionais, $h = 6,626 \times 10^{-34}$ joule-segundos.

Segue-se que uma energia (em joules), dividida pela constante de Planck, corresponde a uma frequência (uma grandeza “por segundo”). Uma energia de 1 joule converte-se, deste modo, em cerca de 2×10^{33} ciclos por segundo.

A relação de Planck passa então a ser $E^\dagger = \nu$, e a sua constante desapareceu. Se, ainda assim, insistir em preservar a forma $E^\dagger = h\nu$, pode fazê-lo, mas terá de fixar $h = 1$.

² A equação de Schrödinger para uma partícula de massa m que se move numa região onde a sua energia potencial é V e a sua energia total é E é

$$-\frac{\hbar^2}{8\pi^2m} \nabla^2 \psi + V\psi = E \psi.$$

Com

$$\frac{mc^2}{h} = m^\dagger,$$

$$\frac{V}{h} = V^\dagger,$$

e

$$\frac{E}{h} = E^\dagger,$$

$$c\nabla = \nabla^\dagger - \frac{1}{8\pi^2m^\dagger} \nabla^{\dagger 2} \psi + V^\dagger \psi = E^\dagger \psi$$

que parece igual, exceto pelo desaparecimento de h . Um problema padrão em mecânica quântica elementar é encontrar os níveis de energia permitidos de uma partícula confinada a uma região do espaço de comprimento L . A solução convencional é

$$E = \frac{n^2 \hbar^2}{8mL^2}$$

com $n = 1, 2, \dots$. Com $h = 1$, as soluções são

$$E^\dagger = \frac{n^2}{8m^\dagger L^{\dagger 2}}$$

onde

$$L^\dagger = \frac{L}{c}.$$

Outra solução básica é a de um oscilador harmónico, que é convencionalmente escrito

$$E = \left(n + \frac{1}{2}\right) \hbar \nu$$

com $n = 0, 1, 2, \dots$,

$$\nu = \left(\frac{1}{2\pi}\right) \sqrt{\frac{k_f}{m}}$$

e

$$k_f = \left(\frac{d^2 V}{dx^2} \right)_0.$$

Com $\hbar = 1$, as soluções são

$$E^\dagger = \left(n + \frac{1}{2} \right) \nu$$

com

$$\nu = \left(\frac{1}{2\pi} \right) \sqrt{\frac{k_f^\dagger}{m^\dagger}}$$

e

$$k_f^\dagger = \left(\frac{d^2 V^\dagger}{dx^{\dagger 2}} \right)_0,$$

com

$$x^\dagger = \frac{x}{c},$$

onde utilizei uma notação que será significativa para aqueles que estão familiarizados com estes assuntos.

³ Se a posição for denotada por x e o momento linear na mesma direção por p , então a comutação entre a posição e o momento linear é $xp - px$. Esta combinação é normalmente denotada por $[x, p]$, que é o “comutador” de x e p . Na mecânica quântica, x e p são tratados como “operadores” (coisas que executam ações, como multiplicar ou diferenciar uma função) e toda a sua estrutura deriva da relação $[x, p] = i\hbar/2\pi$, onde i é o “número imaginário” $\sqrt{-1}$. Pode argumentar-se que os fundamentos da mecânica quântica são inteiramente imaginários.

⁴ Com $x^\dagger = x/c$ e $m^\dagger = mc^2/\hbar$, o momento linear passa a ser $p^\dagger = cp/\hbar$. O comutador $[x, p] = i\hbar/2\pi$ passa então a ser $[x^\dagger, p^\dagger] = i/2\pi$. Admita que, num dado instante, você, com uma massa de 70 quilogramas, se encontra a uma distância de 2 metros de um determinado ponto e se desloca 3 metros por segundo. O seu momento linear (o produto da massa pela velocidade) é $70 \text{ kg} \times 3 \text{ m/s} = 210 \text{ kg m/s}$. O produto da sua posição pelo momento linear é $2 \text{ m} \times 210 \text{ kg m/s} = 420 \text{ kg m}^2/\text{s}$. A constante de Planck, nas mesmas unidades, é $6,6 \times 10^{-34} \text{ kg m}^2/\text{s}$, muito menor e totalmente desprezável. Com o novo sistema de unidades, 2 m de um ponto correspondem, na realidade, a 7 nanossegundos desse ponto, e uma massa de 70 kg que se desloca a 3 m/s tem, na realidade, um momento, expresso como uma frequência, de $9,5 \times 10^{43} \text{ Hz}$. Portanto, o produto da posição pelo momento é 7×10^{35} , que é muito, muito maior que 1.

⁵ Uma consequência do descarte de k é que a entropia (recorde-se como Boltzmann a definiu como $S = k \log W$) se torna $S = \log W$, e passa a ser um número puro sem unidades. A lei dos gases perfeitos $pV = NkT$ passa a ser $pV = NT^\dagger$. Se insistir em preservar a forma convencional da lei dos gases e escrevê-la como $pV = NkT^\dagger$, então terá de aceitar que $k = 1$. A lei dos gases é vulgarmente escrita como $pV = nRT$, onde n é a quantidade de moléculas (em moles) e R é a constante dos gases. Esta última está relacionada com k por $R = N_A k$, onde N_A é a constante de Avogadro. Com $k = 1$, $R = N_A$.

⁶ Eis, lado a lado, quatro exemplos para mostrar como o uso de $\mathbb{T}$ simplifica o aspeto das equações:

	Convencional	Revisada
Lei dos Gases Perfeitos	$pV = NkT$	$pV\mathbb{T} = N$
Distribuição de Boltzmann	$N_2/N_1 = e^{-(E_2 - E_1)/kT}$	$N_2/N_1 = e^{-\mathbb{T}(E_2 - E_1)}$
Energia de N Osciladores Harmónicos	$E = \frac{N\hbar\nu}{e^{\hbar\nu/kT} - 1}$	$E = \frac{N\hbar\nu}{e^{\mathbb{T}\hbar\nu} - 1}$

...e a sua capacidade térmica	$C = Nk \left(\frac{hv}{kT} \right)^2 \left(\frac{e^{-hv/2kT}}{1 - e^{-hv/kT}} \right)^2$	$C = Nk \left(\frac{\mathbb{T}hv e^{-\mathbb{T}hv/2}}{1 - e^{-\mathbb{T}hv}} \right)^2$
-------------------------------	---	--

⁷ Aqui estou a introduzir $\mathbb{T} = h\mathbb{T} = h/kT$.

⁸ Eis as formas finais das quatro grandezas da nota 6 onde, para além das grandezas já introduzidas, $C^\dagger = C/k$, $V^\dagger = V/c^3$ e $p^\dagger = c^3 p/h$, que são respetivamente adimensionais, em unidades de segundos³ e em unidades de 1/segundos⁴.

Lei dos Gases Perfeitos	$p^\dagger V^\dagger \mathbb{T}^\dagger = N$
Distribuição de Boltzmann	$N_2/N_1 = e^{-\mathbb{T}^\dagger(E_2^\dagger - E_1^\dagger)}$
Energia de N Osciladores Harmónicos	$E^\dagger = \frac{Nv}{e^{\mathbb{T}^\dagger v} - 1}$
...e a sua capacidade térmica	$C^\dagger = N \left(\frac{\mathbb{T}^\dagger v e^{-\mathbb{T}^\dagger v/2}}{1 - e^{-\mathbb{T}^\dagger v}} \right)^2$

⁹ A constante de estrutura fina é definida como

$$= \mu_0 e^2 x / 2\eta,$$

onde

$$\alpha_0$$

é a permeabilidade ao vácuo, com o valor definido de $4\pi \times 10^{-7} \text{ J s}^2 \text{ C}^{-2} \text{ m}^{-1}$. Um valor mais preciso do que $\alpha = 1/137$ é $\alpha = 0,007 297 352 5664$. Existe alguma arbitrariedade na definição de α porque pode existir uma medida mais fundamental da carga elétrica. A carga de um quark, por exemplo, é

$$\frac{1}{2} e,$$

e esse é talvez um valor mais apropriado. Neste caso, α seria nove vezes menor, cerca de $1/1233$.

¹⁰ Uma dessas misturas que pretende reproduzir o valor da constante de estrutura fina é $\alpha = 29 \cos(\pi/137) \tan(\pi/(137 \times 29))/\pi$, que resulta em $0,007 297 352 531 86...$

¹¹ Referi a lei do inverso do quadrado na nota 1 do Capítulo 6: expressa a magnitude da força gravítica entre duas massas m_1 e m_2 como

$$F = \frac{G m_1 m_2}{r^2}.$$

A constante gravitacional G tem o valor de $6,673 \times 10^{-13} \text{ kg}^{-1} \text{ m}^3 \text{ s}^{-2}$.

¹² A forma adimensional de G é

$$\alpha_G = \frac{2\pi G m_e^2}{hc}.$$

Não existe qualquer razão fundamental para selecionar a massa do eletrão nesta definição; trata-se apenas de uma analogia com o e^2 que surge em α . Assim, o valor numérico de α_G é talvez apenas sugestivo da intensidade da gravidade, e não mais do que isso.

CAPÍTULO 9

¹ E. P. Wigner (1960). *A Eficácia Inexplicável da Matemática nas Ciências Naturais*. Richard Courant Lecture in Mathematical Sciences proferida na Universidade de Nova Iorque, 11 de maio de 1959. Comunicações sobre Matemática Pura e Aplicada. 13: 1–14.

² A discussão que se segue sobre os fundamentos matemáticos da realidade baseia-se em ideias que publiquei em *Criação* (1983) e *Criação Revisada* (1992). Algumas décadas mais tarde, no seu livro *Our mathematical universe* (2014), Max Tegmark, talvez de forma independente, apresentou ideias semelhantes.

³ Para uma introdução às equações que explicam os padrões nas peles dos animais, ver o capítulo 15 de *Mathematical Biology*, de J. D. Murray (Springer Verlag, 1989).

⁴ Os axiomas de Peano são (de forma abreviada) os seguintes:

1. 0 é um número natural.
2. Para todo o número natural n , o seu sucessor é um número natural.
3. Para todos os números naturais m e n , $m = n$ se, e só se, os sucessores de m e n forem iguais.
4. Não existe número natural cujo sucessor seja 0.

A adição (+) é então definida de modo que $n + 0 = n$ e $n + S(m) = S(n + m)$, e a multiplicação ($\times$) é definida de modo que $n \times 0 = 0$ e $n \times S(m) = n + (n \times m)$, onde $S(n)$ representa o sucessor de n .

⁵ Na sua formulação inicial — ainda assim longe de ser particularmente acessível — o teorema de Löwenheim–Skolem afirma que, *se uma teoria de primeira ordem contável possui um modelo infinito, então, para todo o número cardinal infinito κ , existe um modelo dessa teoria com cardinalidade κ* . Uma consequência mais assimilável é que um sistema de regras como o da aritmética é capaz de emular qualquer domínio do conhecimento que possa ser formalizado como um conjunto de axiomas.

⁶ Mais especificamente, para todos os férmions (partículas com spin semi-inteiro, o que inclui os eletrões), sob a permuta das designações de dois férmions idênticos, a função de onda tem necessariamente de mudar de sinal:

$$\psi(2,1) = -\psi(1,2).$$

O princípio encontra-se profundamente enraizado na relatividade.

⁷ O artigo da Wikipédia é muito claro: “[Gödel] demonstrou que, para qualquer sistema axiomático computável que seja suficientemente poderoso para descrever a aritmética dos números naturais (por exemplo, os axiomas de Peano ou a teoria dos conjuntos de Zermelo–Fraenkel com o axioma da escolha), se verifica que:

Se um sistema (lógico ou axiomático formal) é consistente, não pode ser completo.

A consistência dos axiomas não pode ser provada no interior do próprio sistema.

Estes teoremas puseram fim a meio século de tentativas, iniciadas com o trabalho de Frege e culminando nos *Principia Mathematica* e no formalismo de Hilbert, para encontrar um conjunto de axiomas suficiente para toda a matemática.”

Uma versão em inglês de *On Formally Undecidable Propositions* pode ser encontrada em *Gödel’s Theorem in Focus*, org. S. G. Shanker (Routledge, 1988), mas o leitor terá de enfrentar expressões como: “ $\forall x \in \mathbb{N} \mid n \leq l(x) \ \& \ \forall n, x \ \& \ (\overline{Ep}) \ [n < p \dots]$ ». Uma versão muito mais acessível para humanos com os pés assentes na Terra é *Gödel’s Proof*, de E. Nagel e J. R. Newman (Routledge, 1958).

⁸ Aqui tenho em mente o que é designado por «aritmética de Presburger», que corresponde aos axiomas de Peano sem $\times$. Para uma explicação muito clara — e também acessível — veja-se *New Theories of Everything*, de John Barrow (Oxford University Press, 2007).