

BREVE HISTÓRIA DO TEMPO



Digitalizado, Corrigido e Adaptado por

Gullan Greyl

<http://www.gullangreyl.pt>

1ª Edição, 1988

18-03-2024

SINTESE

Escrito por uma das mentes mais brilhantes do nosso tempo, Breve História do Tempo explora as questões profundas que todos já nos colocámos:

Como teve início o Universo e o que tornou esse início possível? O tempo segue sempre em frente? O Universo é mesmo infinito, ou há um limite para a sua expansão? Existem outras dimensões no espaço? O que acontecerá quando tudo acabar?

Numa linguagem que todos conseguem compreender, Breve História do Tempo mergulha nos reinos exóticos dos buracos negros e dos quarks, da antimatéria e das «setas do tempo», do *Big Bang* e de Deus – reinos onde as possibilidades são assombrosas e surpreendentes.

Recorrendo a comparações inesperadas e fazendo uso da sua prodigiosa imaginação, Stephen Hawking põe-nos frente a frente com os derradeiros segredos da criação.

Índice

PREFÁCIO	1
CAPÍTULO 1 – A NOSSA REPRESENTAÇÃO DO UNIVERSO –	3
CAPÍTULO 2 – ESPAÇO E TEMPO –.....	13
CAPÍTULO 3 – O UNIVERSO EM EXPANSÃO –.....	29
CAPÍTULO 4 – O PRINCÍPIO DA INCERTEZA –.....	42
CAPÍTULO 5 – AS PARTÍCULAS ELEMENTARES E AS FORÇAS DA NATUREZA –.....	48
CAPÍTULO 6 – BURACOS NEGROS –.....	61
CAPÍTULO 7 – OS BURACOS NEGROS NÃO SÃO ASSIM TÃO NEGROS –.....	75
CAPÍTULO 8 – A ORIGEM E O DESTINO DO UNIVERSO –.....	86
CAPÍTULO 9 – A SETA DO TEMPO –.....	106
CAPÍTULO 10 – BURACOS DE VERME E VIAGENS NO TEMPO –	113
CAPÍTULO 11 – A UNIFICAÇÃO DA FÍSICA –.....	121
CAPÍTULO 12 – CONCLUSÃO –.....	133
ALBERT EINSTEIN	137
GALILEU GALILEI	139
ISAAC NEWTON	141
GLOSSÁRIO	142
AGRADECIMENTOS.....	148

PREFÁCIO

Não escrevi a introdução da edição original de *Breve História do Tempo*; foi escrita por Carl Sagan. Em vez disso, escrevi um pequeno texto intitulado «Agradecimentos», onde fui aconselhado a agradecer a toda a gente. Algumas das fundações que me apoiaram não ficaram, contudo, muito satisfeitas por terem sido mencionadas, pois tal levou a um grande aumento de solicitações.

Penso que ninguém, nem os editores, nem o meu agente, nem eu próprio, esperava o que aconteceu. O livro foi *best-seller* em Londres, na lista do *Sunday Times*, durante 237 semanas, mais tempo do que qualquer outro livro (aparentemente, a Bíblia e Shakespeare não são contados). Foi traduzido para quarenta línguas e foi vendido um exemplar para cada 750 homens, mulheres e crianças do mundo. Como afirmou Nathan Myhrvold, da Microsoft (um aluno meu de pós-doutoramento): vendi mais livros sobre física do que a Madonna sobre sexo.

O sucesso de *Breve História do Tempo* mostra que há um interesse generalizado pelas grandes questões, como: donde viemos? Por que é o universo como é?

Tive oportunidade de atualizar o livro e incluir novos resultados teóricos e experimentais, obtidos desde que o livro foi publicado pela primeira vez (no dia 1 de Abril de 1988). Incluí um novo capítulo sobre buracos de verme e viagens no tempo. A teoria geral da relatividade de Einstein parece oferecer a possibilidade de criarmos e mantermos buracos de verme, pequenos tubos que ligam diferentes regiões do espaço-tempo. Se assim for, podemos usá-los para viajar rapidamente à volta da galáxia ou retroceder no tempo. Evidentemente, ainda não vimos ninguém provindo do futuro (ou vimos?), mas encontro para isso uma explicação satisfatória.

Também descrevo o recente progresso da descoberta de «dualidades» ou correspondências entre teorias da física aparentemente diferentes. Estas correspondências constituem um indício seguro de que há uma teoria da física unificada completa, mas também sugerem que não é possível exprimir esta teoria numa única formulação fundamental. Em vez disso, temos de usar diferentes reflexões da teoria subjacente em situações diversas. É como se fôssemos incapazes de representar a superfície da Terra num só mapa e tivéssemos de usar mapas diferentes em diferentes regiões. Isto constituiria uma revolução na nossa visão da unificação das leis da física, mas não alteraria a questão mais importante: a de que o universo é governado por um conjunto de leis racionais que podemos descobrir e compreender.

No que se refere às observações, o desenvolvimento mais importante foi a medição das flutuações da radiação de micro-ondas do fundo cósmico pelo COBE (satélite de exploração do fundo cósmico) e outras contribuições. As flutuações são as impressões digitais da criação, pequenas irregularidades iniciais no universo primitivo, regular e uniforme, que mais tarde evoluíram para darem origem a galáxias, estrelas e a todas as estruturas que vemos à nossa volta. A sua forma concorda com as previsões da hipótese de o universo não ter fronteiras ou limites na direção do tempo imaginário; mas serão necessárias observações ulteriores para distinguir as consequências desta hipótese de

outras explicações possíveis para as flutuações do fundo. No entanto, dentro de poucos anos saberemos se vivemos num universo delimitado ou num universo sem princípio nem fim.

STEPHEN HAWKING

CAPÍTULO 1

- A NOSSA REPRESENTAÇÃO DO UNIVERSO -

Um conhecido cientista (segundo as más línguas, Bertrand Russell) deu uma vez uma conferência sobre astronomia. Descreveu como a Terra orbita em volta do Sol e como o Sol, por sua vez, orbita em redor do centro de um vasto conjunto de estrelas que constitui a nossa galáxia¹. No fim da conferência, uma velhinha, no fundo da sala, levantou-se e disse: «O que o senhor nos disse é um disparate. O mundo não passa de um prato raso equilibrado nas costas de uma tartaruga gigante.» O cientista sorriu com ar superior e retorquiu com outra pergunta: «E onde se apoia a tartaruga?» A velhinha então exclamou: «Você é um jovem muito inteligente, mas são tudo tartarugas por aí a baixo!»

A maior parte das pessoas acharia bastante ridícula a imagem do universo como uma torre infinita de tartarugas. Mas o que nos leva a concluir que sabemos que não é assim? Que sabemos ao certo sobre o universo e como atingimos esse conhecimento? Onde provém e qual o seu destino? Teve um princípio e, nesse caso, que aconteceu *antes* dele? Qual é a natureza do tempo? Acabará alguma vez? Poderemos retroceder no tempo? Descobertas recentes em física, tornadas possíveis em parte pela fantástica tecnologia atual, sugerem respostas e algumas destas perguntas antigas. Um dia essas respostas poderão parecer tão óbvias para nós como o facto de a Terra girar em volta do Sol; ou talvez tão ridículas como uma torre de tartarugas. Só o tempo (seja ele o que for) o dirá.

Já no ano 340 a. C. o filósofo grego Aristóteles, no seu livro *Sobre os Céus*, foi capaz de apresentar dois bons argumentos para se acreditar que a Terra era uma esfera, e não um prato raso. Primeiro compreendeu que os eclipses da Lua eram causados pelo facto de a Terra se interpor entre o Sol e a Lua. A sombra da Terra projetada na Lua era sempre redonda, o que só poderia acontecer se a Terra fosse esférica. Se esta fosse um disco achatado, a sombra seria alongada e elíptica, a não ser que o eclipse ocorresse sempre numa altura em que o Sol estivesse diretamente por baixo do centro do disco. Em segundo lugar, os Gregos sabiam, das suas viagens, que a Estrela Polar surgia menos alta no céu quando era observada mais a sul das regiões onde se encontra mais alta. (Uma vez que a Estrela Polar se encontra no zénite do Polo Norte, parece estar diretamente por cima de um observador no polo boreal, mas para um observador no equador ela encontra-se na direção do horizonte.)

A partir da diferença da posição aparente da Estrela Polar no Egito e na Grécia, Aristóteles estimou o perímetro da Terra em 400.000 estádios. Não se sabe exatamente o valor da medida de comprimento que os Gregos designavam por estádio, mas pensa-se que seria de 180 metros, o que equivale a dizer que Aristóteles calculou cerca de duas vezes o valor real do perímetro da Terra. Os Gregos encontraram ainda um terceiro argumento em prol da esfericidade da Terra: por que motivo se vislumbram primeiro as velas de um navio que surge no horizonte e somente depois o casco?

Aristóteles pensava que a Terra se encontrava imóvel e que o Sol, a Lua, os planetas e as estrelas se moviam em órbitas circulares em volta dela. Pensava assim porque sentia, por razões místicas, que a Terra era o centro do universo e que o movimento circular era o mais perfeito. Esta ideia foi depois sintetizada por Ptolemeu, no século da Era cristã, num modelo cosmológico acabado. A Terra ocupava o centro, rodeada por oito esferas relativas à Lua, ao Sol, às estrelas e aos cinco planetas então conhecidos: Mercúrio, Vénus, Marte, Júpiter e Saturno (fig. 1.1).

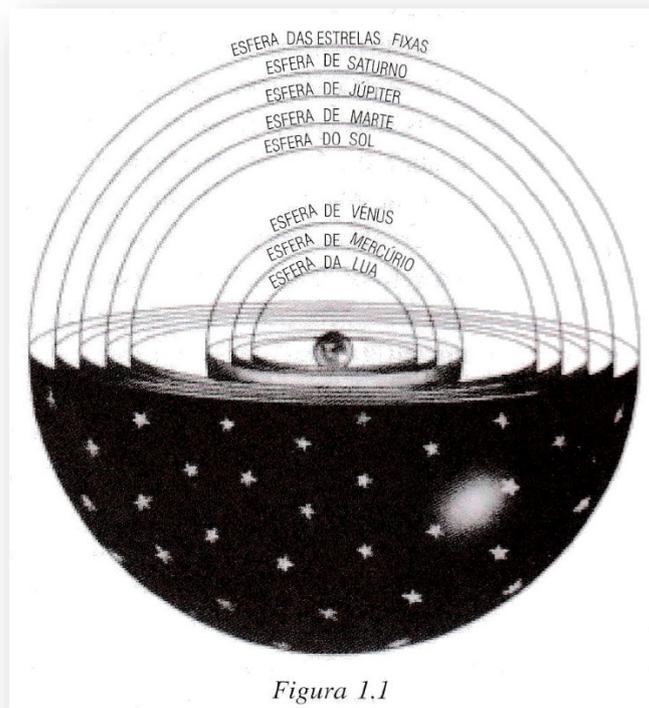


Figura 1.1

Os planetas moviam-se em círculos menores ligados às esferas respetivas, o que explicava as muito complicadas trajetórias observadas no céu. A esfera mais afastada do centro continha as chamadas «estrelas fixas», que ocupam sempre as mesmas posições relativamente umas às outras, mas que têm um movimento de rotação conjunto no céu. O que ficava para além da última esfera nunca foi bem esclarecido, mas não era certamente parte do que podia ser observado pela humanidade².

O modelo de Ptolemeu forneceu um sistema razoavelmente preciso para prever as posições dos corpos celestes no céu. Mas para prever estas posições corretamente teve de partir do princípio de que a Lua seguia uma trajetória tal que, por vezes, se encontrava duas vezes mais próxima do que noutras. Por consequência, haveria ocasiões em que a Lua pareceria duas vezes maior do que noutras. Ptolemeu reconheceu esta falha, o que não impediu que o seu modelo fosse quase universalmente aceite. Foi adotado pela igreja cristã como modelo do universo, de acordo com a Bíblia, pois tinha a grande vantagem de deixar imenso espaço, fora da esfera das estrelas fixas, para o Céu e o Inferno.

Contudo, em 1514, um modelo mais simples foi proposto por um cônego polaco de nome Nicolau Copérnico. (Ao princípio, talvez com medo de ser classificado de herege pela Igreja, Copérnico apresentou o seu modelo anonimamente.) A sua ideia consistia em que o Sol se encontrava imóvel no centro e os planetas se moviam em órbitas circulares à sua volta.

Foi necessário cerca de um século para esta ideia ser levada a sério. Então, dois astrónomos, o alemão Johannes Kepler e o italiano Galileu Galilei, defenderam publicamente a teoria de Copérnico, apesar do facto de as órbitas que predizia não coincidirem completamente com as que eram observadas. O golpe mortal para a teoria de Aristóteles e de Ptolemeu chegou em 1609. Nesse ano, Galileu começou a observar o céu à noite com um telescópio, que acabara de ser inventado. Quando olhou para o planeta Júpiter, descobriu que se encontrava acompanhado de vários pequenos satélites, ou luas, que orbitavam à sua volta. Isto implicava que *nem tudo* tinha de ter uma órbita à volta da Terra, como pensavam Aristóteles e Ptolemeu. (Claro que ainda era possível pensar que a Terra estava imóvel no centro do universo e que as luas de Júpiter se moviam em trajetórias extremamente complicadas em volta da Terra, *aparentando* girarem em volta de Júpiter. No entanto, a teoria de Copérnico era muito mais simples.) Kepler modificou a teoria, sugerindo que os planetas se moviam, não em círculos, mas sim em elipses («círculos» oblongos). As previsões, finalmente, condiziam com as observações.

Para Kepler, as órbitas elípticas eram apenas uma hipótese *ad hoc*, e até bastante inaceitável, porque as elipses eram claramente menos perfeitas do que os círculos. Tendo descoberto quase por acaso que as órbitas elípticas condiziam com as observações, não conseguiu reconciliá-las com a sua ideia de que os planetas giravam em volta do Sol devido a forças magnéticas. Só muito mais tarde, em 1687, surgiu uma explicação, quando Sir Isaac Newton publicou a sua obra *Philosophiae Naturalis Principia Mathematica*, provavelmente o mais importante livro de física alguma vez publicado. Nele, Newton não só apresentou uma teoria sobre o movimento dos corpos, como desenvolveu o aparato matemático necessário para análise do movimento. Além disso, Newton postulou uma lei universal segundo a qual quaisquer dois corpos do universo se atraíam com uma força tanto mais intensa quanto maiores fossem as respetivas massas e maior a sua proximidade. Era esta mesma força que atraía os corpos para o chão. (A história de que Newton se inspirou numa maçã que lhe caiu na cabeça é quase de certeza apócrifa. Ele

ter-se-á limitado a dizer que a ideia da gravidade lhe tinha ocorrido quando estava sentado «imerso nos seus pensamentos» e lhe «tinha sido sugerida pela queda de uma maçã».) Newton mostrou ainda que, segundo a sua lei, a gravidade faz com que a Lua se mova numa órbita elíptica à volta da Terra e a Terra e os outros planetas sigam trajetórias elípticas em volta do Sol.

O modelo de Copérnico fez desaparecer as esferas celestes de Ptolemeu³ e, com elas, a ideia de que o universo apresentava uma fronteira natural. Uma vez que as «fixas» não pareciam alterar a sua posição, excetuando o movimento aparente de rotação, que tem origem no movimento da Terra em torno do seu eixo em sentido contrário, tornou-se natural supor que as estrelas, assimiláveis ao nosso Sol, se encontravam muito mais longe.

Newton compreendeu que, segundo a sua teoria da gravitação, as estrelas deviam atrair-se umas às outras, de modo que parecia não poderem permanecer em repouso. Não colapsariam todas a um tempo em algum ponto? Numa carta escrita em 1691 a Richard Bentley, outro importante pensador desse tempo, Newton argumentava que isso aconteceria realmente se houvesse um número finito de estrelas distribuídas numa região finita de espaço. Mas afirmava também que, se, por outro lado, houvesse um número infinito de estrelas, distribuídas mais ou menos uniformemente num espaço infinito, tal não aconteceria, porque careceriam de um ponto privilegiado para o colapso.

Este raciocínio é um exemplo das rasteiras que podem ser encontradas ao falar-se acerca do infinito. Num universo infinito, cada ponto pode ser eleito o centro, porque em cada direção que cruza o ponto podem contar-se estrelas infinitas. A maneira correta de pensar o assunto, compreendeu-se muito mais tarde, consiste em considerar a situação numa região finita onde as estrelas caem todas umas sobre as outras e, depois, perguntar se uma distribuição mais ou menos uniforme de estrelas fora daquela região alteraria alguma coisa. Segundo a lei de Newton, as estrelas exteriores não introduziriam, em média, a menor diferença na situação das já existentes, de maneira que estas cairiam com a mesma rapidez. Podemos acrescentar as estrelas que quisermos, que continuarão a cair sobre si mesmas. Sabemos agora que é impossível conceber um modelo estático de um universo infinito em que a gravidade seja sempre atrativa.

É interessante refletir acerca de ideias gerais sobre o universo antes do século XX, quando ainda não tinha sido sugerido que o universo estivesse a expandir-se ou a contrair-se. Era geralmente aceite que o universo tinha permanecido imutável através dos tempos, ou que tinha sido criado num certo instante do passado, mais ou menos como o observamos hoje. Em parte, isto pode dever-se à tendência das pessoas para acreditarem em verdades eternas, bem como ao conforto que lhes dá o pensamento de que, embora possam envelhecer e morrer, o universo é eterno e imutável.

Até aqueles que compreenderam que a teoria da gravidade de Newton mostrava que o universo não podia ser estático não pensaram sugerir que

podia estar a expandir-se. Em vez disso, procuraram modificar a teoria, tornando a força gravitacional repulsiva a distâncias muito grandes. Isto não afetou significativamente as suas previsões dos movimentos dos planetas, mas permitiu entender que uma distribuição infinita de estrelas pode permanecer em equilíbrio, opondo às forças atrativas entre estrelas próximas as forças repulsivas das mais afastadas. Contudo, acreditamos agora que esse equilíbrio seria instável: se as estrelas numa região se aproximassem, ainda que ligeiramente, umas das outras, as forças atrativas mútuas tornar-se-iam mais intensas e dominariam as forças repulsivas, de modo que as estrelas continuariam a aproximar-se umas das outras. Por outro lado, se as estrelas se afastassem um pouco umas das outras, as forças repulsivas tornar-se-iam dominantes e afastá-las-iam ainda mais umas das outras.

Outra objeção a um universo estático infinito é normalmente atribuída ao filósofo alemão Heinrich Olbers, que escreveu sobre esta teoria em 1823. De facto, vários contemporâneos de Newton⁴ tinham levantado o problema e o artigo de Olbers nem sequer foi o primeiro a apresentar argumentos relevantes. Foi, no entanto, o primeiro a ser largamente divulgado. A dificuldade reside em que, num universo infinito⁵ estático, quase todas as direções do olhar iriam culminar na superfície de uma estrela.

Assim, esperar-se-ia que o céu fosse tão brilhante como o Sol, mesmo à noite. A proposta de Olbers para resolver este problema era a de que a luz das estrelas distantes seria absorvida pela matéria interestelar. No entanto, se isso acontecesse, a matéria interveniente aqueceria eventualmente até brilhar como as estrelas. A única maneira de evitar a conclusão de que todo o céu noturno seria tão brilhante como a superfície do Sol seria admitir que as estrelas não tinham estado sempre a brilhar, mas haviam iniciado as suas carreiras há um tempo finito. Nesse caso, a matéria absorvente podia não ter ainda aquecido ou a luz das estrelas distantes não ter ainda chegado até nós. E isto leva-nos à questão de qual poderia ter sido a causa de as estrelas se terem acendido.

O começo do universo tinha, evidentemente, sido discutido antes. Segundo algumas das mais antigas cosmologias e a tradição judaico-cristã-muçulmana, o universo teve origem há um tempo finito e não muito distante no passado. Um dos argumentos a favor desta teoria era a sensação de ser necessária a «causa primeira» para explicar a existência do universo. (Dentro deste sempre se explicou um acontecimento como causado por outro anterior, mas a existência do próprio universo só podia ser explicada desta maneira se tivesse tido um começo.) Outro argumento foi exposto por Santo Agostinho no seu livro *A Cidade de Deus*. Chamou a atenção para o facto de a civilização estar a progredir e de nos lembrarmos daqueles que realizaram feitos heroicos e dos que inventaram novas técnicas. Portanto, o homem, e talvez também o universo, não podiam existir há tanto tempo. Santo Agostinho aceitou a data de 5000 anos antes de Cristo para a criação do universo, segundo o livro do *Génese*. (É interessante verificar que esta data não está muito longe do fim do último período glaciário, cerca de 10.000 anos antes de Cristo, data a que os arqueólogos fazem remontar o início da civilização.)



Aristóteles, bem como a maioria dos filósofos gregos, por outro lado, não se afeiçoaram à ideia da criação porque tinha demasiado sabor a intervenção divina. Acreditavam que a raça humana e o mundo à sua volta sempre tinham existido e existiriam para sempre. Os antigos haviam tido em conta o argumento acima referido acerca da evolução e explicavam-no recorrendo a dilúvios cíclicos e outros desastres que periodicamente tinham reconduzido a raça humana ao começo da civilização.

As questões de o universo haver tido ou não um começo no tempo e de ser ou não limitado no espaço foram mais tarde examinadas em pormenor pelo filósofo Immanuel Kant na sua monumental e muito obscura obra *Crítica da Razão Pura*, publicada em 1781. Chamou a essas questões antinomias (ou seja, contradições) da razão pura, porque achava que eram argumentos igualmente atraentes para se acreditar na tese de o universo ter tido um começo e na antítese de haver existido sempre. O seu argumento em defesa da tese era o de que, se o universo não houvesse tido um começo, teria havido um período infinito de tempo antes de qualquer acontecimento, o que considerava absurdo. O argumento antitético consistia em que, se o universo havia tido um princípio, teria havido um período de tempo infinito antes da sua origem: então por que tinha o universo começado num momento especial? De facto, os argumentos que apresenta tanto para a tese como para a antítese são realmente os mesmos. Baseiam-se ambos na suposição não expressa de o tempo continuar indefinidamente para trás, quer o universo tenha ou não existido sempre. Como veremos, o conceito de tempo não tem qualquer significado *antes*⁶ do começo do universo. Este facto foi apontado por Santo Agostinho. Quando lhe perguntaram: «Que fazia Deus antes de criar o mundo?», Agostinho não respondeu: «Andava a preparar o Inferno para todos os que fazem perguntas.» Em vez disso, respondeu que o tempo era uma propriedade do universo que Deus tinha criado e que não existia *antes*⁷ do começo do universo.

Quando a maior parte das pessoas acreditava num universo essencialmente estático e imutável, a questão de saber se havia ou não tido um começo era na verdade do domínio da metafísica ou da teologia. Podia explicar-se o que se observava tanto segundo a teoria de que o universo existira sempre como segundo a de que tinha sido acionado há um tempo finito, mas de tal modo que parecesse ter existido sempre⁸. Mas, em 1929, Edwin Hubble apresentou factos da observação que iniciaram uma nova Era: seja para onde for que se olhe, as galáxias distantes afastam-se velozmente. Por outras palavras, o universo está em expansão, o que significa que nos primeiros tempos os corpos celestes se encontrariam mais perto uns dos outros. De facto, parece ter havido um tempo, há cerca de 10.000 ou 20.000 milhões de anos, em que os objetos estavam todos exactamente no mesmo lugar e em que, portanto, a densidade do universo era infinita. Esta descoberta trouxe, finalmente, a questão das origens para o domínio da ciência.

As observações de Hubble sugeriam que tinha havido um tempo para uma grande explosão (um *big bang*)⁹, em que o universo era infinitamente pequeno e denso. Nessas condições, todas as leis da física e, portanto, toda a

possibilidade de prever o futuro cairiam por terra. Se houve acontecimentos antes desse tempo, não podem afetar o que acontece no presente. A sua existência pode ser ignorada, por não ter consequências observáveis. Pode dizer-se que o tempo começou com o *big bang*, no sentido de que os primeiros momentos não podiam ser definidos. Deve sublinhar-se que este começo no tempo é muito diferente dos que tinham sido considerados previamente. Num universo imutável, um começo no tempo é uma coisa que tem de ser imposta por algum ser exterior ao universo; não há necessidade física de um começo. Pode imaginar-se que Deus criou o universo em qualquer momento do passado. Por outro lado, se o universo está em expansão, pode haver razões de natureza física para um começo. Poderia continuar a imaginar-se que Deus criou o universo no instante do *big bang*, ou mesmo depois, de tal modo que o *big bang* parecesse ter ocorrido, mas não teria qualquer significado supor que tinha sido criado *antes* do *big bang*. Um universo em expansão não exclui um criador, mas impõe limitações ao momento do desempenho da criação!

Para se falar da natureza do universo e discutir assuntos como o princípio e o fim é necessário ser claro acerca do que é uma teoria científica. Vou partir do princípio simplista de que uma teoria não é mais do que um modelo do universo, ou de uma parte restrita deste, bem como um conjunto de regras que relacionam as quantidades do modelo com as observações. Existe apenas na nossa mente e não tem qualquer outra realidade, seja o que for que signifique. Uma teoria é boa quando satisfaz dois requisitos: deve descrever com precisão um grande número de observações que estão na base do modelo, que pode conter um pequeno número de elementos arbitrários, e deve permitir prever resultados de observações futuras. Por exemplo, Aristóteles acreditava na teoria, que é de Empédocles, segundo a qual todas as coisas eram compostas por quatro elementos — terra, ar, fogo e água. Era suficientemente simples para valer como tal, embora apresentasse um conteúdo previsível pobre. Por outro lado, a teoria da gravitação de Newton baseava-se num modelo ainda mais simples, em que os corpos se atraíam uns aos outros com uma força proporcional às suas massas e inversamente proporcional ao quadrado da distância entre eles. No entanto, prevê os movimentos do Sol, da Lua e dos planetas com elevado grau de precisão.

Qualquer teoria física é sempre provisória, porquanto não passa de uma hipótese: nunca consegue provar-se. Por muitas vezes que os resultados da experiência estejam de acordo com alguma teoria, nunca pode ter-se a certeza de que na vez seguinte o resultado não a contrarie. Por outro lado, pode refutar-se uma teoria descobrindo uma única observação em desacordo com as suas previsões. Como o filósofo da ciência Karl Popper realçou, uma boa teoria caracteriza-se pelo facto de fazer previsões que podem, em princípio, ser contestadas ou falseadas pela observação. Sempre que novas experiências concordam com as previsões, a teoria sobrevive e a nossa confiança nela aumenta, mas, se uma nova observação surge em desacordo, abandona-se ou modifica-se a teoria. Pelo menos é o que se supõe acontecer, mas pode sempre pôr-se em dúvida a competência de quem efetuou a observação.

Na prática, o que acontece muitas vezes é surgir uma teoria que mais não é do que uma extensão de outra. Por exemplo, observações muito precisas do planeta Mercúrio revelaram uma pequena discrepância entre o movimento observado e o movimento previsto pela teoria da gravitação de Newton. A teoria da relatividade geral de Einstein previa um movimento ligeiramente diferente do da teoria de Newton. O facto de as previsões de Einstein condizerem com as observações, o que não acontece com as de Newton, foi uma das confirmações cruciais da nova teoria. Contudo, ainda se usa a teoria de Newton para fins práticos, porque as diferenças entre as previsões de uma e de outra são muito pequenas nas situações que normalmente se nos deparam. (A teoria de Newton tem também a grande vantagem de ser muito mais operacional do que a de Einstein!)

O objetivo final da ciência é fornecer uma única teoria que descreva todo o universo. No entanto, o caminho seguido pela maior parte dos cientistas consiste em separar o problema em duas partes. Primeiro há as leis que nos dizem como o universo evolui. (Se conhecermos o estado do universo num dado momento, essas leis permitem inferir o estado do universo em qualquer momento futuro.) Segundo, há a questão do estado inicial do universo. Há quem pense que a ciência devia preocupar-se apenas com a primeira parte: a questão do estado inicial é remetida para a metafísica ou para a religião. Pensam que Deus, sendo onipotente, poderia ter criado o universo como quisesse. Talvez seja verdade, mas nesse caso também podia tê-lo feito desenvolver-se de um modo completamente arbitrário. Contudo, parece que decidiu fazê-lo evoluir de um modo muito regular e segundo certas leis. Parece, portanto, igualmente razoável supor que também há leis que regem o estado inicial.

Acaba por ser muito difícil construir uma teoria que descreva o universo de uma vez só. Em vez disso, dividimos o problema em partes e inventamos teorias parciais. Cada uma destas descreve e prevê um certo conjunto limitado de observações, desprezando o efeito de outras quantidades ou representando-as por simples conjuntos de números, procedimento que pode estar completamente errado. Se tudo no universo depende de tudo o mais de uma maneira fundamental, pode ser impossível aproximarmo-nos de uma solução completa investigando isoladamente as partes do problema. Contudo, foi certamente este o processo com que lográmos progressos no passado. O exemplo clássico é mais uma vez a teoria de Newton da gravitação, que nos diz que a força atrativa entre dois corpos (a uma certa distância) depende apenas de um número associado a cada corpo, a sua massa, mas é independente da matéria de que os corpos são feitos. Desde modo, não é necessária uma teoria de estrutura e constituição do Sol e dos planetas para calcular as suas órbitas.

Os cientistas de hoje descrevem o universo em termos de duas teorias parciais fundamentais: a teoria da relatividade geral e a mecânica quântica. São estes os grandes feitos intelectuais da primeira metade do século. A teoria da relatividade geral descreve a força da gravidade e a estrutura em macroescala do universo, ou seja, a estrutura em escalas que vão de apenas

alguns quilómetros a alguns milhões de milhões de milhões de milhões (1.000.000.000.000.000.000.000.000) de quilómetros: as dimensões do universo observável. A mecânica quântica, por seu lado, lida com fenómenos que ocorrem em escalas extremamente reduzidas, como um milionésimo de milionésimo de centímetro. Infelizmente, contudo, estas duas teorias são incompatíveis: não podem estar ambas corretas. Uma das maiores demandas da física atual, e o assunto principal deste livro, é uma nova teoria que concilie as duas: uma teoria quântica da gravidade. Ainda não a encontramos e podemos estar muito longe dela, mas já conhecemos muitas das propriedades que devem caracterizá-la. E veremos em capítulos ulteriores que já sabemos muito sobre as previsões a que essa teoria irá conduzir-nos.

Portanto, se acreditamos que o universo não é arbitrário, mas sim governado por leis definidas, teremos, finalmente, de combinar as teorias parciais numa teoria unificada completa que descreva todo o universo. Mas existe um paradoxo fundamental na busca dessa teoria. As ideias sobre as teorias científicas que foram aqui delineadas presumem que somos seres racionais livres para observar o universo como quisermos e tirarmos conclusões lógicas daquilo que observamos. Num esquema como este é razoável supor que seremos capazes de progredir cada vez mais em direção às leis que governam o universo. Contudo, se houver realmente uma teoria unificada, ela também determinará presumivelmente as nossas ações. E, assim, a própria teoria determinaria o resultado da nossa busca. E por que motivo determinaria que chegássemos às conclusões certas a partir da evidência? Não poderia também determinar que chegássemos à conclusão errada? Ou a nenhuma conclusão?

A única resposta que posso dar tem por base o princípio da seleção natural de Darwin. A ideia consiste em que em qualquer população de organismos autorreprodutores haverá flutuações genéticas de geração para geração e na mesma geração. Estas variações significam também que alguns indivíduos são mais capazes do que outros para tirarem as conclusões certas sobre o mundo que os rodeia e para agirem de acordo com elas. Estes indivíduos terão mais hipóteses de sobreviver e de se reproduzirem; deste modo, o seu padrão de comportamento e pensamento virá a ser dominante. Tem-se verificado no passado que aquilo a que chamamos inteligência e descobertas científicas têm acarretado vantagens de sobrevivência. Já não é tão claro que este estado de coisas se mantenha: as nossas descobertas científicas podem perfeitamente acabar por destruir-nos a todos e, mesmo que tal não aconteça, uma teoria unificada pode não fazer grande diferença quanto às nossas hipóteses de sobrevivência. Contudo, desde que o universo tenha evoluído de modo regular, podemos esperar que a capacidade de raciocínio que nos foi dada pela seleção natural seja válida também na nossa busca de uma teoria unificada, não nos conduzindo a conclusões erradas.

Como as teorias parciais que já temos são suficientes para fazer previsões exatas em todas as situações, exceto nas mais extremas, a busca da teoria definitiva do universo parece de difícil justificação em termos práticos. (De nada vale, no entanto, que argumentos semelhantes possam ter sido utilizados

quer contra a relatividade, quer contra a mecânica quântica, e estas teorias deram-nos a energia nuclear e a revolução da microeletrónica!) A descoberta de uma teoria unificada, portanto, pode não ajudar a sobrevivência da espécie. Pode mesmo nem afetar a nossa maneira de viver. Mas, desde a alvorada da civilização, as pessoas não se contentam com verem acontecimentos desligados e sem explicação. Têm ansiado por um entendimento da ordem subjacente no mundo. Ainda hoje sentimos a mesma ânsia de saber por que estamos aqui e donde viemos. O mais profundo desejo de conhecimento da humanidade é justificação suficiente para a nossa procura contínua. E o nosso objetivo é nada mais nada menos do que uma descrição completa do universo em que vivemos.

¹ A nossa galáxia ou, mais familiarmente, a Galáxia. (N. do R. C.)

² Esta descrição encaixa fielmente na cosmologia física de Aristóteles, que subsiste paralelamente à astronomia matemática de Ptolemeu. (N. do R. C.)

³ De Aristóteles (N. do R. C.)

⁴ O próprio Kepler, numa das suas obras mais divulgadas, *Dissertatio cum Nuncio Sidereo*, usa um argumento similar para concluir da finitude do mundo. (N. do R. C.)

⁵ O argumento é válido num universo estático infinito no espaço e no tempo. (N. do R. C.)

⁶ O *itálico* é do revisor científico. Repare na incapacidade e ambiguidade da linguagem comum quando se exprime a temática: *antes* não faz sentido, em rigor, pois o tempo surge com a criação. (N. do R. C.)

⁷ Id. (N. do R. C.)

⁸ Tudo depende de admitirmos que o estado inicial do universo era mais ou menos complexo. (N. do R. C.)

⁹ Os comentários parentéticos são da responsabilidade do revisor científico. (N. do R. C.)

CAPÍTULO 2

- ESPAÇO E TEMPO -

As nossas ideias atuais sobre o movimento dos corpos vêm dos tempos de Galileu e de Newton. Antes deles, as pessoas acreditavam em Aristóteles, que afirmou que o estado natural de um corpo era o estado de repouso; mover-se ia apenas enquanto sobre ele atuasse uma força ou impulso. Assim, um corpo pesado cairia mais depressa do que um corpo leve porque sofreria um impulso maior em direção à Terra.

A tradição aristotélica também afirmava que era possível descobrir todas as leis que governam o universo só por puro pensamento, sem necessidade de confirmação observacional. Deste modo, até Galileu, ninguém se preocupou em ver se corpos de pesos diferentes caíam de facto com velocidades diferentes. Diz-se que Galileu demonstrou que a crença de Aristóteles era falsa deixando cair pesos da torre inclinada de Pisa. A história é certamente falsa, mas Galileu fez uma coisa equivalente: fez rolar bolas de pesos diferentes pelo suave declive de um plano inclinado. A situação é semelhante à de corpos pesados que caem verticalmente, mas mais fácil de observar, porque se movimentam com velocidades menores. As medições de Galileu indicavam que a velocidade de cada corpo aumentava na mesma proporção, qualquer que fosse o seu peso. Por exemplo, se deixarmos rolar uma bola por uma encosta que desce 1 metro por cada 10 metros de caminho andado, veremos que a bola se move com uma velocidade de cerca de 1 metro por segundo após um segundo, 2 metros por segundo após dois segundos, etc., por mais pesada que seja. É evidente que um peso de chumbo cairá mais depressa do que uma pena, mas tal sucede apenas porque a pena é retardada pela resistência do ar. Se deixarmos cair dois corpos que sofram pouca resistência por parte do ar, por exemplo dois pesos de chumbo diferentes, o tempo de queda será o mesmo. Na Lua, onde não há ar que oponha resistência aos corpos que caem, o astronauta David R. Scott fez experiências com penas e pesos de chumbo e verificou que estes atingiam o solo ao mesmo tempo.

As medições de Galileu foram utilizadas por Newton como base para as suas leis do movimento. Nas experiências de Galileu, quando um corpo rolava por um plano inclinado, exercia-se sobre ele sempre a mesma força (o seu peso), consistindo o efeito em fazer aumentar constantemente a velocidade. Isto mostrou que o verdadeiro efeito da força é modificar sempre a velocidade de um corpo, e não só imprimir-lhe o movimento, como se pensara antes. Também significava que, quando um corpo não sofre o efeito de qualquer força, se manterá em movimento retilíneo com velocidade constante. Esta ideia foi explicitada pela primeira vez na obra de Newton *Principia Mathematica*, publicada em 1687, e é conhecida por primeira lei de Newton. O

que acontece a um corpo quando uma força atua sobre ele é explicado pela segunda lei de Newton, que afirma que o corpo acelerará ou modificará a sua velocidade proporcionalmente à força. (Por exemplo, a aceleração será duas vezes maior se a força for duas vezes maior.) A aceleração também será menor quanto maior for a massa (ou quantidade de matéria) do corpo. (A mesma força atuando sobre um corpo com o dobro da massa produzirá metade da aceleração.) O automóvel é um exemplo familiar: quanto mais potente for o motor, maior será a aceleração, mas quanto mais pesado for o carro, menor será a aceleração para o mesmo motor. Além das leis do movimento, Newton descobriu uma lei para descrever a força da gravidade, que afirma que um corpo atrai outro corpo com uma força proporcional à massa de cada um deles. Assim, a força entre dois corpos será duas vezes mais intensa se um dos corpos (por exemplo, o corpo *A*) tiver o dobro da massa. É o que pode esperar-se, pois pode pensar-se que o novo corpo *A* é constituído por dois corpos com a massa original. Cada um atrairá o corpo *B* com a força original. Assim, a força total entre *A* e *B* será duas vezes a força original. E, se, por exemplo, um dos corpos tiver duas vezes a massa e o outro três vezes, então a força será seis vezes mais intensa. Vê-se, assim, por que razão todos os corpos caem com a mesma velocidade relativa; um corpo com o dobro do peso terá duas vezes a força de gravidade a puxá-lo para baixo, mas terá também duas vezes a massa original. De acordo com a segunda lei de Newton, estes dois efeitos anulam-se exatamente um ao outro, de modo que a aceleração será a mesma em todos os casos.

A lei da gravitação de Newton também nos diz que, quanto mais separados estiverem os corpos, menor será a força. E também nos diz que a atração gravitacional de uma estrela é exatamente um quarto da de uma estrela semelhante a metade da distância. Esta lei prevê as órbitas da Terra, da Lua e dos outros planetas com grande precisão. Se a lei estipulasse que a atração gravitacional de uma estrela diminuía mais depressa com a distância, as órbitas dos planetas não seriam elipses: os planetas mover-se-iam em espiral até colidirem com o Sol. Se diminuísse mais devagar, as forças gravitacionais das estrelas distantes dominariam a que o Sol exerce sobre a Terra.

A grande diferença entre as ideias de Aristóteles e as de Galileu e de Newton consiste em que Aristóteles acreditava num estado preferido de repouso, que qualquer corpo tomaria se não fosse atuado por qualquer força ou impulso. Pensava particularmente que a Terra estava em repouso. Mas das leis de Newton decorre que não existe um padrão único de repouso. Poder-se-ia igualmente dizer que o corpo *A* está em repouso e o corpo *B* em movimento com velocidade constante em relação ao corpo *A* ou que o corpo *B* está em repouso e o corpo *A* em movimento. Por exemplo, se por instantes pusermos de lado a rotação da Terra e a sua órbita em torno do Sol, podemos dizer que a Terra está em repouso e que um comboio se desloca para Este a 48 quilómetros por hora ou que o comboio está em repouso e a Terra se move para Oeste a 48 quilómetros por hora. Se efetuássemos experiências com corpos em movimento no comboio, todas as leis de Newton continuariam válidas. Por exemplo, jogando ténis de mesa no comboio, verificar-se-ia que a

bola obedecia às leis de Newton, tal como a bola numa mesa colocada junto à linha. Portanto, não existe maneira de dizer se é o comboio ou a Terra que está em movimento¹.

A falta de um padrão absoluto significava que não era possível determinar se dois acontecimentos que ocorriam em momentos diferentes ocorriam na mesma posição no espaço. Suponhamos, por exemplo, que a bola de ténis de mesa no comboio saltita verticalmente, para cima e para baixo, atingindo a mesa duas vezes no mesmo sítio com um segundo de intervalo. Para alguém na linha, os dois saltos pareceriam ocorrer a cerca de 13 metros um do outro, porque o comboio teria percorrido essa distância entre os dois saltos.

A não existência de repouso absoluto significava, portanto, que não poderia atribuir-se uma posição absoluta a um acontecimento no espaço, como Aristóteles acreditava. Os lugares dos acontecimentos e as distâncias entre eles seriam diferentes para uma pessoa no comboio e para outra na linha e não haveria motivo para dar preferência a qualquer delas.

Newton preocupou-se muito com esta falta de lugar absoluto, ou espaço absoluto, como se chamava, por não estar de acordo com a sua ideia de um deus absoluto. De facto, recusou-se a aceitar que o espaço não fosse absoluto, embora as suas leis o sugerissem. Muitas pessoas criticaram severamente a sua crença irracional, particularmente o bispo Berkeley, filósofo que acreditava que todos os objetos materiais e o espaço e o tempo não passavam de uma ilusão. Quando o famoso Dr. Johnson ouviu a opinião de Berkeley, gritou: «Refuto-a assim!», e deu um pontapé numa pedra.

Tanto Aristóteles como Newton acreditavam no tempo absoluto: ou seja, acreditavam que podiam medir sem ambiguidade o intervalo de tempo entre dois acontecimentos e que esse tempo seria o mesmo para quem o medisse, desde que utilizasse um bom relógio. O tempo era completamente separado e independente do espaço. Isto é o que a maior parte das pessoas acharia ser uma opinião de senso comum. Contudo, fomos obrigados a mudar de ideias quanto ao espaço e ao tempo. Embora estas noções de aparente senso comum funcionem perfeitamente quando lidamos com coisas como maçãs ou planetas, que se movem relativamente devagar, já não funcionam à velocidade da luz ou perto dela.

O facto de a luz se deslocar com uma velocidade finita, mas muito elevada, foi descoberto em 1676 pelo astrónomo dinamarquês Ole Christensen Roemer. Este observou que os períodos em que as luas de Júpiter pareciam passar por trás do planeta não eram regulares, como se esperaria se girassem à volta do planeta com uma velocidade constante. Como a Terra e Júpiter orbitam em volta do Sol, a distância entre eles varia. Roemer reparou que os eclipses das luas de Júpiter ocorriam tanto mais tarde quanto mais longe estivéssemos do planeta. Argumentou que isto acontecia porque a luz das luas levava mais tempo a chegar até nós quando estávamos mais longe. As suas medições das variações de distância entre a Terra e Júpiter não eram, contudo, muito precisas, sendo, assim, o valor da velocidade da luz de

225.000 quilômetros por segundo, aquém do valor real de 300.000 quilômetros por segundo. No entanto, a proeza de Roemer, não só ao provar que a luz se propaga a uma velocidade finita, mas também ao medi-la, foi notável, conseguida, como foi, onze anos antes da publicação dos *Principia Mathematica* de Newton.

A primeira teoria correta da propagação da luz surgiu em 1865, quando o físico britânico James Clerk Maxwell conseguiu unificar as teorias parciais utilizadas até então para descrever as forças da eletricidade e do magnetismo. As equações de Maxwell previam que podia haver perturbações de tipo ondulatório no campo eletromagnético e que elas se propagariam com uma velocidade determinada, como pequenas ondulações num tanque. Se o comprimento de onda destas ondas (a distância entre uma crista de onda e a seguinte) for de 1 metro ou mais, trata-se das hoje chamadas ondas de rádio. De comprimentos de onda mais curtos são as chamadas micro-ondas (alguns centímetros) ou ondas infravermelhas (um pouco mais de dez milésimos de centímetro). A luz visível tem um comprimento de onda compreendido entre quarenta e oitenta milionésimos de centímetro. São conhecidos comprimentos de onda mais curtos relativos a ondas ultravioletas, de raios X e de raios gama.

A teoria de Maxwell previa que as ondas de rádio ou de luz deviam propagar-se a uma velocidade determinada. Mas a teoria de Newton tinha acabado com a ideia do repouso absoluto, de maneira que, supondo que a luz se devia propagar a uma velocidade finita, era preciso dizer em relação a quê devia essa velocidade ser medida. Foi ainda sugerido que havia uma substância, chamada «éter», presente em todo o lado, mesmo no espaço «vazio». As ondas de luz propagar-se-iam através do éter como as ondas sonoras se propagam através do ar, sendo, assim, a sua velocidade relativa ao éter. Observadores diferentes que se movessem em relação ao éter veriam a luz propagar-se na sua direção com velocidades diferentes, mas a velocidade da luz em relação ao éter manter-se-ia fixa. Em particular, como a Terra se moveria no seio do éter, na sua órbita em torno do Sol, a velocidade da luz medida na direção do movimento da Terra através do éter (quando nos movemos em direção à fonte de luz) deveria ser mais elevada do que a velocidade da luz na direção perpendicular a esse movimento (quando não nos dirigimos para a fonte). Em 1878, Albert Michelson (que mais tarde veio a ser o primeiro americano galardoado com o Prémio Nobel da Física) e Edward Morley realizaram uma experiência cuidadosa na Case School de Ciências Aplicadas em Cleveland. Compararam a velocidade da luz na direção do movimento da Terra com a velocidade medida na direção perpendicular a esse movimento. Para sua grande surpresa, descobriram que os respetivos valores eram exatamente os mesmos!

Entre 1887 e 1905 houve várias tentativas, sobretudo as do físico holandês Hendrik Lorentz, para explicar o resultado da experiência de Michelson e Morley, em termos de contração de objetos e de atrasos nos relógios quando se moviam no éter. Contudo, num famoso trabalho de 1905, um funcionário, até então desconhecido, do gabinete suíço de patentes, Albert

Einstein, mostrou que a ideia do éter seria desnecessária desde que se abandonasse a ideia do tempo absoluto. Algumas semanas mais tarde, um importante matemático francês, Henri Poincaré, demonstrou a mesma coisa. Os argumentos de Einstein estavam mais próximos da física do que os de Poincaré, que encarava o problema do ponto de vista matemático. Geralmente, o crédito da nova teoria cabe a Einstein, mas o nome de Poincaré é lembrado por estar ligado a uma importante parte dela.

O postulado fundamental da teoria da relatividade, como era chamada, foi o de que as leis da física² deviam ser as mesmas para todos os observadores que se movessem livremente, qualquer que fosse a sua velocidade. Isto era verdadeiro para as leis do movimento de Newton, mas agora a ideia alargava-se para incluir a teoria de Maxwell e a velocidade da luz: todos os observadores deviam medir a mesma velocidade da luz, independentemente da velocidade do seu movimento. Esta ideia simples teve algumas consequências notáveis. Talvez as mais conhecidas sejam a equivalência da massa e da energia, denotada pela famosa equação de Einstein $E = mc^2$ (em que E representa a energia, m a massa e c a velocidade da luz), e a lei de que nada³ pode deslocar-se mais depressa do que a luz. Devido à equivalência entre massa e energia, a energia de um objeto devida ao seu movimento adicionar-se-á à sua massa. Por outras palavras, será mais difícil aumentar a sua velocidade. Este efeito só é realmente significativo para objetos que se movam a velocidades próximas da, da luz. Por exemplo, a 10% da velocidade da luz, a massa de um objeto é apenas 0,5% superior ao valor normal, ao passo que a 90% da velocidade da luz excederia o dobro do valor normal. Quando a velocidade de um objeto se aproxima da velocidade da luz, a sua massa aumenta ainda mais, pelo que é necessária cada vez mais energia para lhe aumentar a velocidade. De facto, nunca pode atingir a velocidade da luz porque nessa altura a sua massa se tornaria infinita e, pela equivalência entre massa e energia, seria necessária uma quantidade infinita de energia para incrementar indefinidamente a massa. Por este motivo, qualquer objeto normal está para sempre condicionado pela relatividade a mover-se a velocidades inferiores à da luz. Só esta ou as outras ondas que não possuam massa intrínseca podem mover-se à velocidade da luz.

Uma consequência igualmente notável da relatividade é a maneira como revolucionou as nossas concepções de espaço e de tempo. Na teoria de Newton, se um sinal luminoso for enviado de um local para outro, diferentes observadores estarão de acordo quanto ao tempo que essa viagem demorou, mas não quanto à distância que a luz percorreu. Como a velocidade da luz é exatamente o quociente da distância percorrida pelo tempo gasto, diferentes observadores mediriam diferentes velocidades para a luz. Em relatividade, por outro lado, todos os observadores têm de concordar quanto à velocidade de propagação da luz. Continuam ainda, no entanto, a não concordar quanto à distância que a luz percorreu, pelo que têm também de discordar quanto ao tempo que demorou. O tempo gasto é apenas a distância percorrida pela luz — com o que os observadores não concordam — dividida pela velocidade — valor comum aos observadores. Por outras palavras, a teoria da relatividade acabou com a ideia do tempo absoluto! Parecia que cada observador obtinha a sua

própria medida do tempo, registada pelo relógio que utilizava, e que relógios idênticos utilizados por observadores diferentes nem sempre coincidiam.

Cada observador pode usar o radar para dizer onde e quando um acontecimento ocorre, enviando um sinal de luz ou de rádio. Parte do sinal é refletido no momento do acontecimento e o observador mede o instante decorrido quando recebe o eco. Diz-se então que o tempo do acontecimento é o que está a meio entre o envio do impulso e a receção do eco; a distância do acontecimento é metade do tempo da viagem de ida e volta multiplicada pela velocidade da luz. (Um acontecimento, neste sentido, é qualquer coisa que ocorre num ponto do espaço e num momento específico do tempo.) Esta ideia está exemplificada na fig. 2.1, que representa um diagrama espaço-temporal.

Utilizando este procedimento, os observadores que se movem em relação uns aos outros podem atribuir ao mesmo acontecimento tempos e posições diferentes. Nenhuma medição de um observador em particular são mais corretas do que as de outro, mas estão todas relacionadas. Qualquer observador pode calcular com precisão o tempo e a posição que outro observador atribuirá a um acontecimento, desde que conheça a velocidade relativa desse outro observador.

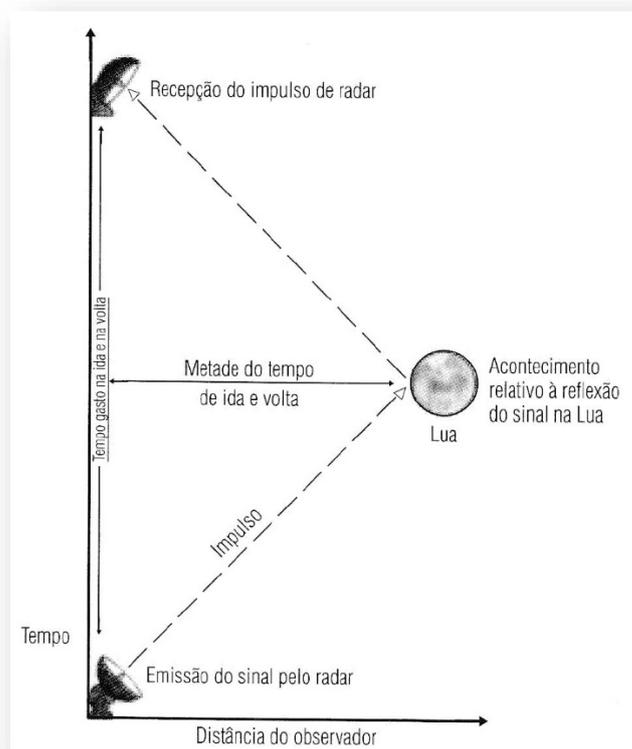


Figura 2.1 – O tempo é medido verticalmente e a distância a partir do observador é medida horizontalmente. O caminho percorrido e o tempo gasto pelo observador são representados pela linha vertical à esquerda. Os percursos dos raios de luz para e a partir do acontecimento são representados pelas linhas diagonais.

Hoje em dia utilizamos este método para medir com rigor distâncias, porque podemos medir o tempo com maior precisão do que as distâncias. Com efeito, o metro é definido como a distância percorrida pela luz em 0,00000003335640952 segundos, medidos por um relógio de césio. (A razão para este número em particular é o facto de corresponder à definição histórica do metro — em termos de distância entre duas marcas numa barra de platina guardada em Paris.) Do mesmo modo, pode usar-se uma nova e mais conveniente unidade de comprimento chamada segundo-luz. Este é simplesmente definido como a distância percorrida pela luz num segundo. Na teoria da relatividade define-se agora a distância em termos de tempo e de velocidade da luz, pelo que cada observador medirá a luz com a mesma velocidade (por definição, um metro por 0,00000003335640952 segundos). Não há necessidade de introduzir a ideia de um éter, cuja presença, aliás, não pode ser detetada, como mostrou a experiência de Michelson e Morley. A teoria da relatividade obriga-nos, contudo, a modificar fundamentalmente as nossas concepções de espaço e de tempo. Temos de aceitar que o tempo não é independente do espaço, mas se combina com ele para formar um objeto chamado espaço-tempo.

É um dado da experiência comum podermos descrever a posição de um ponto no espaço através de três números ou coordenadas. Por exemplo, pode dizer-se que um ponto numa sala está a 2 metros de uma parede, a 90 centímetros de outra e a 1,5 metros acima do chão. Ou podemos especificar que um ponto está numa determinada latitude e longitude e a determinada altitude acima do nível do mar. É-se livre de utilizar quaisquer coordenadas, embora a sua validade seja limitada. Não é possível especificar a posição da Lua em termos de quilómetros a norte e quilómetros a oeste de Piccadilly Circus e metros acima do nível do mar. Em vez disso, podemos descrevê-la em termos de distância ao Sol, distância ao plano das órbitas dos planetas e do ângulo entre a linha que une a Lua ao Sol e a linha que une o Sol a uma estrela próxima, como a Alfa do Centauro. Mesmo estas coordenadas não teriam grande utilidade para descrever a posição do Sol na Galáxia ou a posição da Galáxia no grupo local de galáxias. De facto, é possível descrever o universo em termos de um conjunto de partes justapostas. Em cada uma das partes pode utilizar-se um conjunto diferente de três coordenadas para especificar a posição de um ponto.

Um acontecimento é qualquer coisa que ocorre num determinado ponto no espaço e num determinado momento. Pode, portanto, ser especificado por quatro números de coordenadas. Mais uma vez, a escolha das coordenadas é arbitrária; podem ser usadas quaisquer três coordenadas espaciais bem definidas e qualquer medida de tempo. Em relatividade não há verdadeira distinção entre as coordenadas de espaço e de tempo, tal como não existe diferença real entre quaisquer duas coordenadas espaciais. Pode escolher-se um novo conjunto de coordenadas em que, digamos, a primeira coordenada de espaço seja uma combinação das antigas primeira e segunda coordenadas de espaço. Por exemplo, em vez de medirmos a posição de um ponto na Terra em quilómetros a norte de Piccadilly e quilómetros a oeste de Piccadilly, podemos usar quilómetros a nordeste de Piccadilly e a noroeste de Piccadilly. Do mesmo

modo, em relatividade podemos utilizar uma nova coordenada de tempo que é o tempo antigo em segundos mais a distância (em segundos-luz) a norte de Piccadilly.

Muitas vezes é útil pensar nas quatro coordenadas de um acontecimento para especificar a sua posição num espaço quadridimensional chamado «espaço-tempo». É impossível imaginar um espaço quadridimensional. Eu próprio já acho suficientemente difícil visualizar um espaço tridimensional! Contudo, é fácil desenhar diagramas de espaços bidimensionais, como a superfície da Terra. (A superfície da Terra é bidimensional porque a posição de um ponto pode ser especificada por duas coordenadas: a latitude e a longitude.) Usarei geralmente diagramas em que o tempo aumenta no sentido ascendente vertical e uma das dimensões espaciais é indicada horizontalmente. As outras duas dimensões espaciais ou são ignoradas ou, por vezes, uma delas é indicada em perspetiva. (São os diagramas de espaço-tempo como o da fig. 2.1.) Por exemplo, na fig. 2.2, o tempo é medido no sentido vertical ascendente em anos e a distância do Sol a Alfa do Centauro é medida horizontalmente em quilómetros. As trajetórias do Sol e de Alfa do Centauro através do espaço-tempo são representadas pelas linhas verticais à esquerda e à direita do diagrama. Um raio de luz do Sol segue a linha diagonal e leva quatro anos a chegar a Alfa do Centauro.

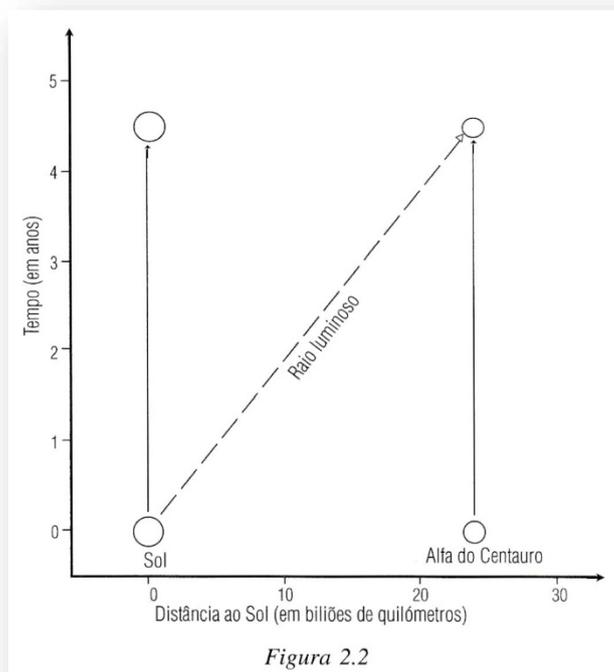
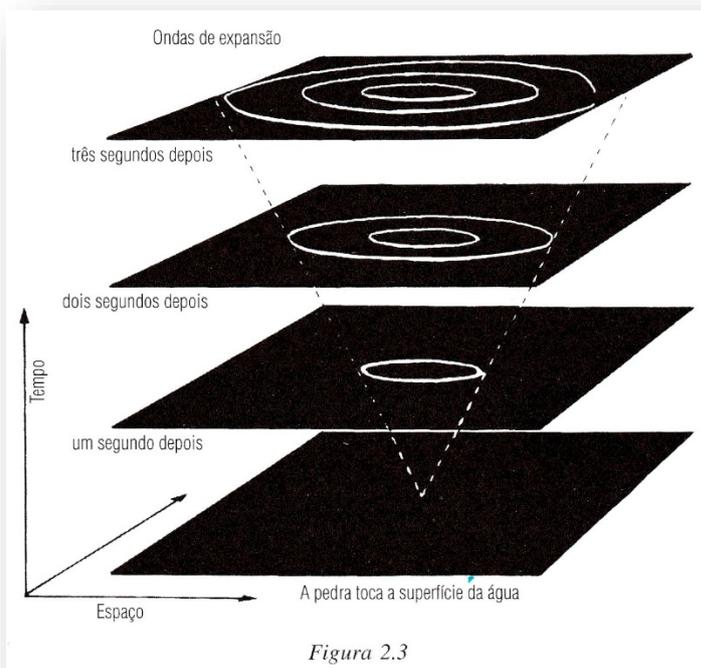


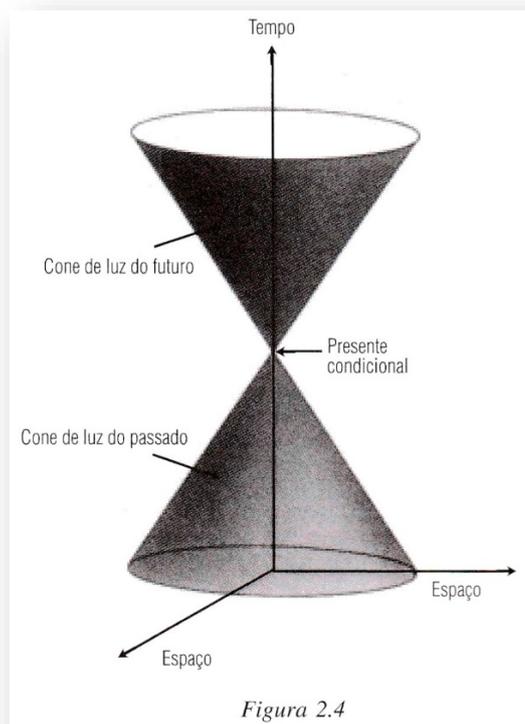
Figura 2.2

Como vimos, as equações de Maxwell previam que a velocidade da luz devia ser a mesma, qualquer que fosse a velocidade da sua fonte, o que foi confirmado por medições rigorosas. Daí que, se um impulso de luz é emitido em determinado momento e em dado ponto do espaço, à medida que o tempo passa, espalhar-se-á como uma esfera de luz cujos tamanho e posição são

independentes da velocidade da fonte. Um milionésimo de segundo depois a luz ter-se-á difundido para formar uma esfera com um raio de 300 metros; dois milionésimos de segundo depois o raio será de 600 metros, etc. Será como a ondulação que se propaga na superfície de um tanque quando se lhe atira uma pedra. A ondulação propaga-se como um círculo que aumenta à medida que o tempo passa. Se sobrepusermos películas fotográficas das ondulações em instantes diferentes, o círculo de ondulação que se expande será representado por um cone cujo vértice está no local e no instante em que a pedra atingiu a água (fig. 2.3).



Da mesma maneira, a luz que se propaga a partir de um acontecimento forma um cone (tridimensional) no espaço-tempo (quadrídimensional). Este cone designa-se por «cone de luz do futuro do acontecimento». Podemos, do mesmo modo, desenhar outro cone, chamado «cone de luz do passado», que reúne o conjunto dos pontos a partir dos quais um impulso de luz pode alcançar o acontecimento dado (fig. 2.4).



Dado um acontecimento P , pode classificar-se qualquer outro acontecimento no universo em três classes. Os acontecimentos ditos no futuro de P podem ser alcançados a partir de P por intermédio de uma partícula ou onda que se propague a uma velocidade igual ou inferior à da luz. Encontram-se no interior ou na superfície da esfera de luz emitida pelo acontecimento P . Deste modo, encontrar-se-ão no interior ou sobre o cone de luz do futuro de P no diagrama espaço-temporal. Só os acontecimentos no futuro de P podem ser afetados pelo que acontece em P porque nada pode se deslocar com velocidade superior à da luz.

Do mesmo modo, o passado de P pode definir-se como o conjunto de todos os acontecimentos a partir dos quais é possível alcançar o acontecimento P através de partículas ou ondas que se propaguem a uma velocidade inferior ou igual à da luz. Assim, é o conjunto de acontecimentos que pode afetar o que se passa em P . Os acontecimentos que não se encontram nem no futuro nem no passado de P , ditos no presente condicional de P , não podem causar nem ser causados por P (fig. 2.5).

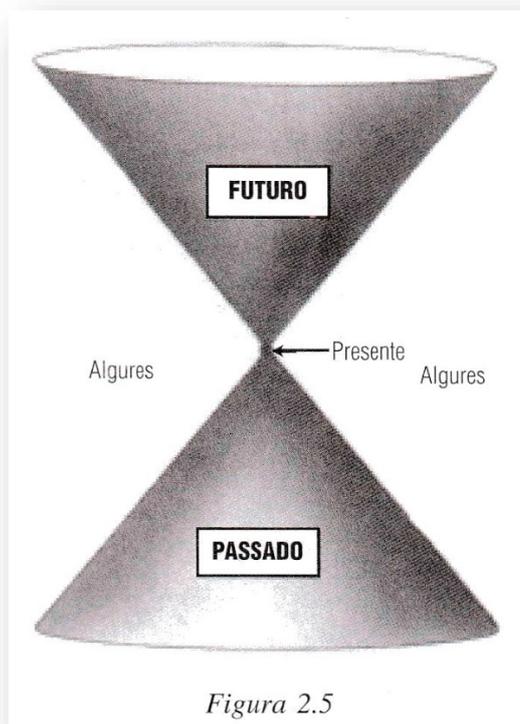


Figura 2.5

Por exemplo, se o Sol deixasse de brilhar neste preciso momento, não afetaria os acontecimentos atuais na Terra porque eles se situariam na região do presente condicional do acontecimento em que o Sol deixou de brilhar (fig. 2.6).

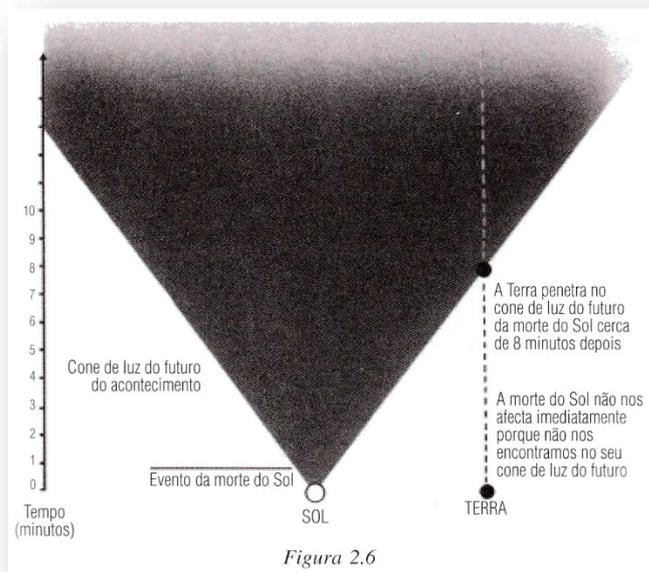
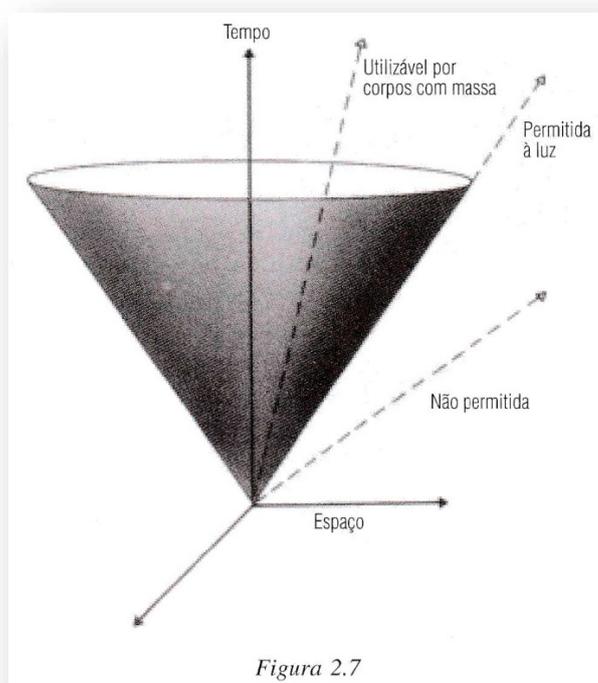


Figura 2.6

Só saberíamos o que se tinha passado decorridos 8 minutos, o tempo que a luz do Sol leva a alcançar-nos. Só nessa altura os acontecimentos na Terra ficariam no cone da luz do futuro do acontecimento da morte do Sol. Do mesmo modo, não sabemos o que se está a passar neste momento mais longe

no Universo: a luz que nos chega provinda de galáxias distantes deixou-as há milhões de anos; a luz do objeto mais longínquo que conseguimos avistar deixou-o há já cerca de 8000 milhões de anos. Assim, quando observamos o universo, vemo-lo como ele era no passado.

Se desprezarmos os efeitos da gravitação, como Einstein e Poincaré fizeram em 1905, obteremos a chamada «teoria da relatividade restrita». Para cada acontecimento no espaço-tempo podemos construir um cone de luz (conjunto de todas as trajetórias possíveis de luz, no espaço-tempo, emitida nesse acontecimento) e, uma vez que a velocidade da luz é a mesma para todos os acontecimentos e em todas as direções, todos os cones de luz serão idênticos e orientados na mesma direção. A teoria também nos diz que nada pode mover-se com velocidade superior à da luz⁴. Isto significa que a trajetória de qualquer objeto através do espaço e do tempo tem de ser representada por uma linha que fique dentro do cone de luz respetivo (fig. 2.7).



A teoria da relatividade restrita obteve grande êxito na explicação de que a velocidade da luz parece a mesma relativamente a todos os observadores (como a experiência de Michelson e Morley demonstrou) e na descrição do que acontece quando os objetos se movem a velocidades próximas da velocidade da luz. Contudo, era inconsistente com a teoria da gravitação de Newton, que afirmava que os objetos se atraíam uns aos outros com uma força que dependia da distância que os separava. Isto significava que, se se deslocasse um dos objetos, a força exercida sobre o outro mudaria instantaneamente. Por outras palavras, os efeitos gravitacionais deslocar-se-iam com velocidade infinita, e não à velocidade da luz ou abaixo dela, como a teoria da relatividade restrita exigia. Einstein tentou várias vezes, sem êxito, entre 1908 e 1914,

descobrir uma teoria da gravidade que fosse consistente com a relatividade restrita. Finalmente, em 1915 propôs a chamada «teoria da relatividade geral».

Einstein apresentou a sugestão revolucionária de a gravidade não ser uma força idêntica às outras, mas sim uma consequência do facto de o espaço-tempo não ser plano, como se pensara: é curvo ou «deformado» pela distribuição de massa e de energia. Corpos como a Terra não são feitos para se moverem em órbitas curvas por ação de uma força chamada «gravidade»; em vez disso, seguem o que mais se parece com uma trajetória retilínea num espaço curvo, chamada «geodésica». Uma geodésica é o caminho mais curto (ou mais longo) entre dois pontos próximos. Por exemplo, a superfície da Terra é um espaço curvo bidimensional. Uma geodésica na Terra chama-se «círculo máximo» e é o caminho mais curto entre dois pontos (fig. 2.8).

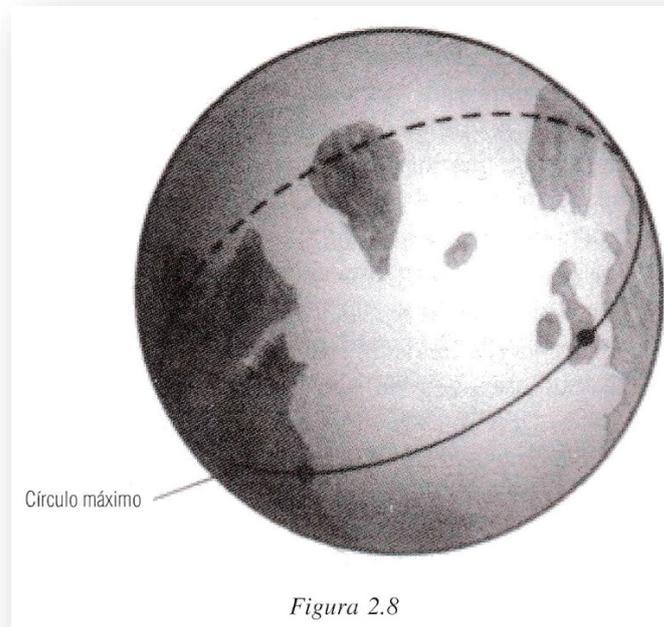


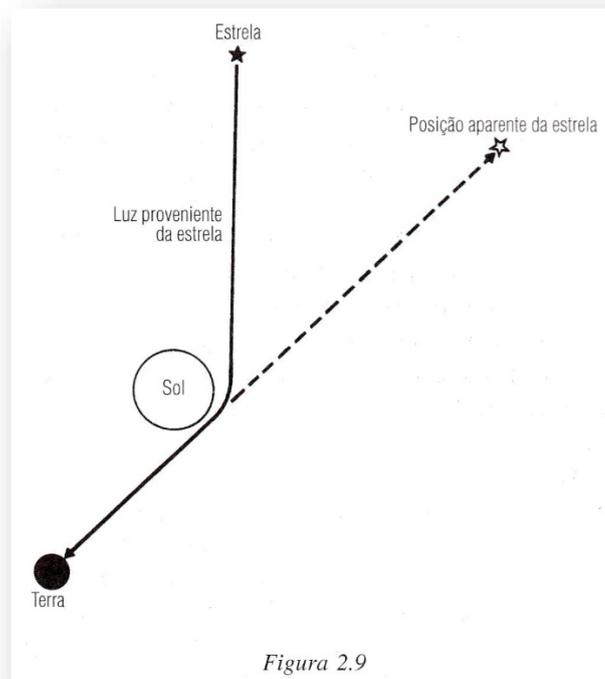
Figura 2.8

Como a geodésica é o caminho mais curto entre quaisquer dois aeroportos, é essa a rota que um navegador aeronáutico indicará ao piloto. Na relatividade geral, os corpos seguem sempre linhas retas no espaço-tempo quadridimensional, mas, aos nossos olhos, parecem continuar a mover-se ao longo de trajetórias curvas no espaço tridimensional. (Um bom exemplo é a observação de um voo de avião sobre colinas. Embora siga uma linha reta no espaço tridimensional, a sua sombra segue uma trajetória curva no espaço bidimensional.)

A massa do Sol encurva o espaço-tempo de tal modo que, embora a Terra siga uma trajetória retilínea no espaço-tempo quadridimensional, nos parece mover-se ao longo de uma órbita circular no espaço tridimensional. De facto, as órbitas dos planetas previstas pela relatividade geral são quase exatamente as mesmas que as previstas pela teoria da gravitação de Newton. Contudo, no caso de Mercúrio, que, sendo o planeta mais próximo do Sol, sofre os efeitos

gravitacionais mais fortes e tem uma órbita bastante alongada, a relatividade geral prevê que o eixo maior da elipse deve girar em volta do Sol à razão de cerca de um grau em 10.000 anos. Embora este efeito seja pequeno, foi anunciado antes de 1915 e foi uma das primeiras confirmações da teoria de Einstein. Em anos recentes, os desvios ainda mais pequenos das órbitas dos outros planetas, relativamente às previsões de Newton, têm sido medidos por radar, concordando com as previsões da relatividade geral.

Também os raios luminosos têm de seguir geodésicas no espaço-tempo. Mais uma vez, o facto de o espaço ser curvo significa que a luz já não parece propagar-se no espaço em linhas retas. Portanto, a relatividade geral prevê que a luz deve ser encurvada por campos gravitacionais. Por exemplo, a teoria prevê que os cones de luz de pontos perto do Sol serão ligeiramente encurvados para o interior devido à massa do Sol. Isto significa que a luz de uma estrela distante que passa perto do Sol é defletida de um pequeno ângulo, fazendo que a estrela pareça estar numa posição diferente relativamente a um observador na Terra (fig. 2.9).



É evidente que, se a luz da estrela passasse sempre perto do Sol, não poderíamos dizer se a luz estava a ser defletida ou se, em vez disso, a estrela estava realmente onde a víamos. No entanto, como a Terra orbita em volta do Sol, estrelas diferentes passam por trás deste, tendo, conseqüentemente, a sua luz defletida. Mudam, portanto, de posição relativamente às outras estrelas.

Normalmente, é muito difícil observar este efeito, porque a luz do Sol torna impossível a observação de estrelas que aparecem perto do Sol. Contudo, é possível fazê-lo durante um eclipse quando a luz do Sol é

bloqueada pela Lua. A previsão de Einstein da deflexão da luz não pôde ser testada imediatamente, em 1915, porque se estava em plena primeira guerra mundial; foi só em 1919 que uma expedição britânica, ao observar um eclipse na África ocidental, mostrou que a luz era realmente defletida pelo Sol, tal como havia sido previsto pela teoria. Esta comprovação de uma teoria alemã por cientistas britânicos foi louvada depois da guerra como um grande ato de reconciliação entre os dois países. É, portanto, irónico que o ulterior exame das fotografias tiradas durante essa expedição mostrasse que os erros eram tão grandes como o efeito que tentavam medir. As medidas tinham sido obtidas por mera sorte ou resultavam do conhecimento prévio do que se pretendia obter, o que não é assim tão invulgar em ciência. A deflexão da luz tem, contudo, sido confirmada com precisão por numerosas observações mais recentes.

Outra previsão da relatividade geral é a de que o tempo deve parecer decorrer mais lentamente perto de um corpo maciço, como a Terra. E isto porque há uma relação entre a energia da luz e a sua frequência (ou seja, o número de ondas luminosas por segundo): quanto maior for a energia, mais alta será a frequência. Quando a luz se propaga no sentido ascendente no campo gravitacional da Terra, perde energia e a sua frequência baixa. (Isto significa que o tempo decorrido entre uma crista de onda e a seguinte aumenta.) A um observador situado num ponto muito alto parecerá que tudo o que fica por baixo leva mais tempo a acontecer. Esta previsão foi testada em 1962 com dois relógios muito precisos, instalados no topo e na base de uma torre de água. Verificou-se que o relógio colocado na parte de baixo, que estava mais perto da terra, andava mais lentamente, em acordo absoluto com a relatividade geral. A diferença de velocidade dos relógios a alturas diferentes acima do globo é agora de considerável importância prática, com o advento de sistemas de navegação muito precisos, baseados em sinais emitidos por satélites. Se se ignorassem as previsões da relatividade geral, a posição calculada teria um erro de vários quilómetros!

As leis do movimento de Newton acabaram com a ideia da posição absoluta no espaço. A teoria da relatividade acaba de vez com o tempo absoluto. Consideremos dois gémeos: suponhamos que um deles vai viver para o cimo de uma montanha e que o outro fica ao nível do mar. O primeiro gémeo envelheceria mais depressa do que o segundo. Assim, se voltassem a encontrar-se, um seria mais velho do que o outro. Neste caso, a diferença de idades seria muito pequena, mas podia ser muito maior se um dos gémeos fosse fazer uma longa viagem numa nave espacial a uma velocidade próxima da velocidade da luz. Quando voltasse, seria muito mais novo do que o que tivesse ficado na Terra. Isto é conhecido por «paradoxo dos gémeos», mas só é um paradoxo se tivermos em mente a ideia de tempo absoluto. Na teoria da relatividade não existe qualquer tempo absoluto: cada indivíduo tem a sua medida pessoal de tempo, que depende do local onde está e da maneira como está a mover-se.

Até 1915 pensava-se que o espaço e o tempo eram um palco onde os acontecimentos ocorriam, mas que não era afetado por eles. Isto era verdade

mesmo para a teoria da relatividade restrita. Os corpos moviam-se, atraídos e repelidos por forças, mas o espaço e o tempo continuavam sem serem afetados. Era natural pensar que o espaço e o tempo continuassem para sempre.

A situação, no entanto, é completamente diferente na teoria da relatividade geral. O espaço e o tempo são agora quantidades dinâmicas: quando um corpo se move, ou uma força atua, a curvatura do espaço-tempo é afetada e, por seu lado, a estrutura do espaço-tempo afeta o movimento dos corpos e o efeito das forças. O espaço e o tempo não só afetam, como são afetados por tudo o que acontece no universo. Tal como não podemos falar de acontecimentos no universo sem as noções de espaço e de tempo, também na relatividade geral deixou de ter sentido falar sobre o espaço e o tempo fora dos limites do universo.

Nas décadas seguintes esta nova compreensão do espaço e do tempo iria revolucionar a nossa conceção do Universo. A velha ideia de um universo essencialmente imutável, que podia ter existido e podia continuar a existir para sempre, foi substituída pela noção de um universo dinâmico e em expansão, que parecia ter tido um início há um tempo finito e que podia acabar daqui a um tempo finito. Essa revolução constitui o assunto do próximo capítulo. E, anos mais tarde, foi também o ponto de partida para o meu trabalho em física teórica. Roger Penrose e eu mostrámos que a teoria da relatividade geral de Einstein implicava que o universo tinha de ter um princípio e, possivelmente, um fim.

¹ Quer dizer: não há processo mecânico de distinguir o estado de repouso do estado de movimento uniforme e retilíneo. Esta afirmação constitui o enunciado do princípio da relatividade galilaica. (N. do R. C.)

² As leis da física, e não somente as leis da mecânica (cf. princípio da relatividade galilaica). O conteúdo fundamental deste postulado passa muitas vezes despercebido ao leitor, que se deixa impressionar mais facilmente pelo postulado da constância da velocidade da luz: este é, de certa forma, implicado por aquele! (N. do R. C.)

³ Este nada refere-se a algo material. É claro que podem conceber-se velocidades meramente geométricas (v. g., expansão do espaço) tão grandes quanto se queira. (N. do R. C.)

⁴ O que está verdadeiramente em causa é a velocidade da luz, não a luz. Acidentalmente, a luz propaga-se à velocidade da luz, que, tanto quanto se sabe, também podia chamar-se «velocidade dos neutrinos»! (N. do R. C.)

CAPÍTULO 3

- O UNIVERSO EM EXPANSÃO -

Se olharmos para o céu numa noite de céu limpo e sem luar, os objetos mais brilhantes que podemos ver serão, possivelmente, os planetas Vénus, Marte, Júpiter e Saturno. Haverá também um grande número de estrelas, que são exatamente como o nosso Sol, mas encontram-se mais distantes de nós. Algumas destas estrelas fixas parecem de facto mudar muito ligeiramente de posição umas em relação às outras, enquanto a Terra gira em volta do Sol: não estão absolutamente nada fixas! Isto acontece por estarem comparativamente perto de nós. Como a Terra gira em volta do Sol, vemo-las de diferentes posições no pano de fundo das estrelas mais distantes. É uma sorte, porque nos permite medir diretamente a distância a que essas estrelas estão de nós: quanto mais próximas, mais parecem mover-se. A estrela que está mais perto de nós chama-se Próxima do Centauro e está, afinal, a cerca de 4 anos-luz de distância (a sua luz leva cerca de quatro anos a alcançar a Terra) ou a cerca de 37 milhões de milhões de quilómetros. A maior parte das outras estrelas visíveis a olho nu está a algumas centenas de anos-luz de nós. O nosso Sol, em comparação, está a uns meros 8 minutos-luz de distância! As estrelas visíveis aparecem espalhadas por todo o céu noturno, mas concentram-se particularmente numa faixa a que damos o nome de Via Láctea. Por volta de 1750, alguns astrónomos, entre os quais Thomas Wright, sugeriram que o aspeto da Via Láctea podia ser explicado por a maior das estrelas visíveis estar distribuída numa configuração de disco, como aquilo a que agora chamamos galáxia espiral. Só algumas décadas mais tarde outro astrónomo, Sir William Herschel, confirmou a ideia de Wright, catalogando pacientemente as posições e distâncias de um grande número de estrelas. Mesmo assim, a ideia só obteve aceitação completa no princípio deste século.

A representação moderna do universo data apenas de 1924, quando o astrónomo americano Edwin Hubble demonstrou que a Galáxia não era a única. Havia, na realidade, muitas outras, com vastidões de espaço vazio entre elas. Para o provar precisava de determinar as distâncias a que se encontravam essas outras galáxias, que estão tão longe que, ao contrário das estrelas próximas, parecem realmente fixas. Hubble teve de utilizar métodos indiretos para medir as distâncias. O brilho aparente de uma estrela depende de dois fatores: da quantidade de luz que radia (a sua luminosidade) e da distância a que se encontra de nós. Relativamente às estrelas próximas, podemos medir o seu brilho aparente e a distância a que se encontram e, assim, determinar a sua luminosidade. Pelo contrário, se conhecermos a luminosidade de estrelas de outras galáxias, podemos calcular a sua distância medindo o seu brilho aparente. Hubble notou que certos tipos de estrelas¹ têm sempre a mesma luminosidade quando estão suficientemente perto de nós

para que possamos medi-la; portanto, argumentou que, se encontrássemos estrelas dessas noutra galáxia, podíamos admitir que teriam a mesma luminosidade e, assim, calcular a distância dessa galáxia. Se pudéssemos consegui-lo com várias estrelas da mesma galáxia e os nossos cálculos indicassem sempre a mesma distância, podíamos confiar razoavelmente neles.

Deste modo, Edwin Hubble calculou as distâncias de nove galáxias diferentes. Sabemos agora que a Galáxia é apenas uma entre centenas de milhares de milhões que podem ser observadas com os telescópios modernos e que cada galáxia contém algumas centenas de milhares de milhões de estrelas. A figura 3.1 mostra uma galáxia espiral semelhante ao que pensamos que seja o aspeto da nossa galáxia para alguém que viva noutra. Vivemos numa galáxia que tem cerca de uma centena de milhares de anos-luz de diâmetro e roda vagarosamente; as estrelas, nos seus braços em espiral, orbitam em redor do centro cerca de uma vez em cada várias centenas de milhões de anos. O nosso Sol não passa de uma estrela amarela normal, de tamanho médio, perto do limiar de um dos braços em espiral. Percorremos realmente um longo caminho desde Aristóteles e Ptolemeu, quando se pensava que a Terra era o centro do universo!

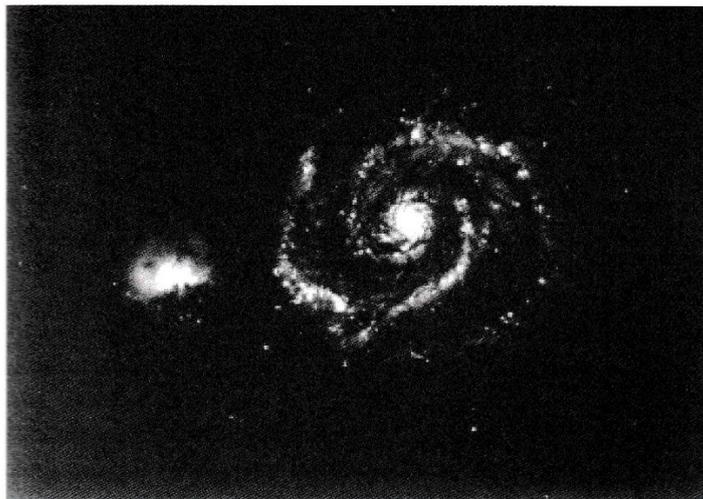


Figura 3.1

As estrelas estão tão distantes que nos parecem meros pontinhos de luz. Não podemos ver o seu tamanho nem a sua forma. Então como podemos distinguir diferentes tipos de estrelas? Na grande maioria das estrelas há apenas uma característica que podemos observar: a cor da sua luz. Newton descobriu que, se a luz do Sol passa através de um pedaço triangular de vidro, chamado «prisma», se decompõe nas cores componentes (o seu espectro), como num arco-íris. Focando uma estrela ou uma galáxia com um telescópio, podemos observar do mesmo modo o espectro da luz dessa estrela ou galáxia. Estrelas diferentes têm espectros diferentes, mas o brilho relativo das diferentes cores é sempre exatamente o que se esperaria encontrar na luz

emitida por um objeto incandescente. (Na realidade, a luz emitida por um objeto opaco ao rubro apresenta um espectro característico que depende apenas da sua temperatura — um espectro térmico. Isto significa que podemos medir a temperatura a partir do espectro da sua luz.) Além disso, sabe-se que algumas cores muito específicas estão ausentes dos espectros das estrelas e estas cores que faltam podem variar de estrela para estrela. Como sabemos que cada elemento químico absorve um conjunto característico de cores muito específicas, comparando-as com as que faltam no espectro de uma estrela, podemos determinar exatamente quais são os elementos presentes na atmosfera da estrela.

Nos anos 20, quando os astrónomos começaram a observar os espectros de estrelas de outras galáxias, descobriram algo muito estranho: faltavam as mesmas cores encontradas nos espectros das estrelas da Galáxia, porque eram todas desviadas na mesma proporção para o extremo vermelho do espectro. Para compreendermos as implicações deste fenómeno temos de entender primeiro o efeito de Doppler. Como vimos, a luz visível consiste em flutuações ou ondas no campo eletromagnético. A frequência (ou número de ondas por segundo) da luz é extremamente alta, indo de 400 a 700 milhões de milhões de ondas por segundo. As diferentes frequências da luz são o que o olho humano vê como cores diferentes, com as frequências mais baixas junto do extremo vermelho do espectro e as mais altas no extremo azul. Imaginemos agora uma fonte luminosa a uma certa distância de nós, tal como uma estrela, emitindo ondas luminosas com uma frequência constante. É óbvio que a frequência das ondas que recebemos será a mesma a que são emitidas (o campo gravitacional da Galáxia não tem um efeito significativo). Suponhamos agora que a fonte começa a mover-se na nossa direção. Quando a fonte emitir a crista de onda seguinte, estará mais perto de nós; por isso, o tempo que esta crista leva a chegar até nós será menor do que quando a estrela estava em repouso relativo. Isto significa que o tempo entre duas cristas de onda que chegam até nós é menor e, portanto, o número de ondas que recebemos por segundo (ou seja, a frequência) é maior do que quando a estrela está em repouso relativo. Da mesma maneira, se a fonte se afastar de nós, a frequência das ondas que recebemos será mais baixa. No caso da luz, isto significa que estrelas que se afastam de nós terão os seus espectros desviados para o extremo vermelho (desvio para o vermelho) e que as que se aproximam de nós terão os seus espectros deslocados para o azul. Esta relação entre a frequência e a velocidade, a que se chama «efeito de Doppler», faz parte da experiência de todos os dias. Basta escutar o ruído de um carro que passa na estrada: à medida que se aproxima, o motor soa mais alto (o que corresponde a uma frequência mais alta das ondas sonoras) e, quando passa e se afasta, o som é mais baixo. O comportamento das ondas de luz ou de rádio é semelhante. Na verdade, o efeito de Doppler é utilizado para medir a velocidade dos automóveis, medindo a frequência de ondas de rádio por eles refletidas.

Nos anos que se seguiram à demonstração da existência de outras galáxias, Hubble passou o tempo a catalogar as distâncias entre elas e a observar os seus espectros. Nessa altura, a maior parte das pessoas julgava

que as galáxias se moviam completamente ao acaso e, portanto, esperava encontrar tantos espectros desviados para o azul como para o vermelho. Constituiu, assim, uma surpresa descobrir que as riscas do espectro da maioria das galáxias surgiam desviadas para o vermelho; quase todas se afastavam de nós! Mais surpreendente ainda foi a descoberta que Hubble publicou em 1929: o valor do desvio para o vermelho de uma galáxia não é casual, mas sim diretamente proporcional à distância a que a galáxia está de nós. Ou, por outras palavras, quanto mais longe ela se encontra, mais depressa está a afastar-se! E isso significava que o universo não podia ser estático, como toda a gente tinha pensado antes, mas que está, de facto, em expansão; a distância entre as diferentes galáxias aumenta constantemente.

A descoberta de que o universo está em expansão foi uma das grandes revoluções intelectuais do século XX. Perante este facto, é fácil perguntarmos por que motivo ninguém tinha pensado nisso antes. Newton e outros deviam ter compreendido que um universo estático depressa começaria a contrair-se sob influência da gravidade. Mas pensemos, ao invés, num universo em expansão. Se se expandisse bastante devagar, a força da gravidade acabaria por travar a expansão, seguindo-se-lhe inevitavelmente a contração. Contudo, se estivesse a expandir-se acima de uma certa razão crítica, a gravidade nunca teria força suficiente para travar a expansão e o universo continuaria a expandir-se para sempre. É um pouco como o que acontece quando se dispara um foguetão para o espaço. Se tiver uma velocidade pequena, a gravidade acabará por detê-lo e ele cairá. Por outro lado, ultrapassada uma certa velocidade crítica (cerca de 11 quilómetros por segundo), a gravidade não terá força suficiente para o aprisionar, de maneira que continuará a afastar-se da Terra para sempre. Este comportamento do universo podia ter sido previsto a partir da teoria da gravidade de Newton em qualquer altura nos séculos XIX ou XVIII ou até no fim do século XVII. Mas a crença num universo estático era tão forte que prevaleceu até ao século XX. Até Einstein, quando formulou a teoria da relatividade geral, em 1915, estava tão certo de que o universo era estático que modificou a sua teoria para o tornar possível, introduzindo nas suas equações a chamada «constante cosmológica». Einstein introduziu uma nova força, a «antigravitação» que, ao contrário das outras forças, não provinha de qualquer origem especial, mas era construída na própria estrutura do espaço-tempo. Afirmava ele que o espaço-tempo tinha uma tendência intrínseca para se expandir, o que poderia levar a equilibrar exatamente a atração de toda a matéria no universo, de modo a daí resultar um universo estático. Só um homem, segundo parece, estava disposto a tomar a relatividade geral pelo que era e, enquanto Einstein e outros físicos procuravam maneiras de evitar, no contexto da relatividade geral, soluções não estáticas, o físico e matemático russo Alexander Friedmann dedicou-se a explicá-las.

Friedmann tirou duas conclusões muito simples sobre o universo: que este parece idêntico seja em que direcção se olhe e que tal também seria verdade se observássemos o universo de qualquer outro lugar. Apenas com estas duas ideias², Friedmann mostrou que não deveríamos esperar que o universo fosse

estático. De facto, em 1922, vários anos antes da descoberta de Edwin Hubble, Friedmann previu exatamente o que aquele veio a descobrir!

Evidentemente, a suposição de que o universo tem o mesmo aspeto em todas as direções não é, na realidade, verdadeira. Por exemplo, como já vimos, as outras estrelas da Galáxia formam uma faixa de luz distinta no céu noturno, chamada Via Láctea. Mas, se olharmos para galáxias distantes, parece haver mais ou menos o mesmo número delas qualquer que seja a direção em que olhemos. Portanto, o universo, na realidade, parece ser praticamente idêntico em todas as direções, desde que o observemos numa grande escala em comparação com a distância entre as galáxias e ignoremos as diferenças em pequenas escalas. Durante muito tempo isto constituiu justificação suficiente para a suposição de Friedmann: uma grosseira aproximação do universo. Mas, mais recentemente, por um acidente feliz, descobriu-se que a suposição de Friedmann é realmente uma notável e precisa descrição do universo.

Em 1965, dois físicos americanos dos Bell Telephone Laboratories de Nova Jérquia, Arno Penzias e Robert Wilson, efetuavam experiências com um detetor de micro-ondas muito sensível. (As micro-ondas são exatamente como as ondas luminosas, mas com uma frequência de cerca de 10.000 milhões de ondas por segundo.) Penzias e Wilson ficaram preocupados quando descobriram que o seu detetor captava mais ruídos do que devia. Os ruídos não pareciam provir de uma direção em particular. Primeiro, descobriram excremento de aves no detetor e procuraram outros defeitos possíveis, mas depressa abandonaram essa hipótese. Sabiam que qualquer ruído proveniente do interior da atmosfera seria mais forte quando o detetor não estivesse apontado verticalmente porque os raios de luz percorrem maior distância na atmosfera quando são recebidos perto do horizonte do que quando são recebidos diretamente de cima. Os ruídos extra eram os mesmos qualquer que fosse a direção para que estivesse apontado o detetor; portanto, deviam vir *de fora* da atmosfera. Também eram iguais de dia e de noite e durante todo o ano, embora a Terra rodasse sobre o seu eixo e orbitasse em volta do Sol. Isto mostrava que a radiação devia vir de fora do sistema solar e até de fora da Galáxia, porque, se assim não fosse, variaria quando o movimento da Terra apontasse o detetor para direções diferentes. De facto, sabemos que a radiação deve ter viajado até nós através da maior parte do universo observável e, uma vez que parece ser a mesma em direções diferentes, o universo também deve ser o mesmo em todas as direções, apenas a uma escala maior. Sabemos agora que, em qualquer direção que olhemos, estes ruídos nunca variam mais do que uma pequena fração: Penzias e Wilson tinham tropeçado por acaso numa confirmação incrivelmente precisa da primeira suposição de Friedmann. No entanto, como o universo não é exatamente o mesmo em qualquer direção, nas apenas em média em larga escala, as micro-ondas não podem ser exatamente as mesmas em todas as direções. Deve haver variações ligeiras de direção para direção. Este facto foi detetado pela primeira vez em 1992 pelo satélite *Cosmic Background Explorer*, ou COBE, a um nível de cerca de uma parte em cem mil. Por mais pequenas

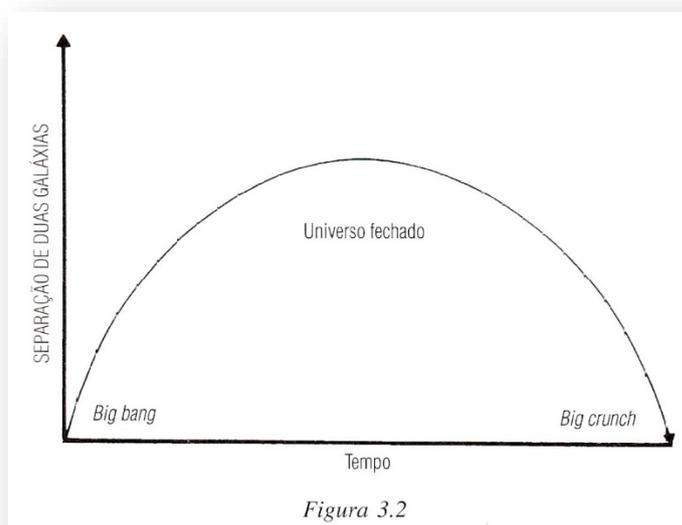
que estas variações sejam, elas são muito importantes, como explicaremos no capítulo 8.

Mais ou menos ao mesmo tempo em que Penzias e Wilson investigavam os ruídos no detetor, dois físicos americanos da Universidade de Princeton, ali perto, Bob Dicke e Jim Peebles, também se interessavam pelas micro-ondas. Estavam a trabalhar a sugestão de George Gamow (que tinha sido aluno de Friedmann) de que o universo primordial devia ter sido muito quente e denso, com brilho rubro-branco. Dicke e Peebles achavam que ainda devíamos poder ver esse brilho do universo primitivo porque a luz proveniente de partes muito distantes do universo primitivo devia estar agora a chegar até nós. Contudo, a expansão do universo significava que essa luz devia ser de tal maneira desviada para o vermelho que só podia aparecer-nos agora como uma radiação de micro-ondas. Dicke e Peebles preparavam-se para procurar esta radiação quando Penzias e Wilson ouviram falar do seu trabalho e compreenderam que já a tinham encontrado. Assim, Penzias e Wilson receberam o Prémio Nobel em 1978 (o que parece um pouco duro para Dicke e Peebles, para não falar de Gamow!).

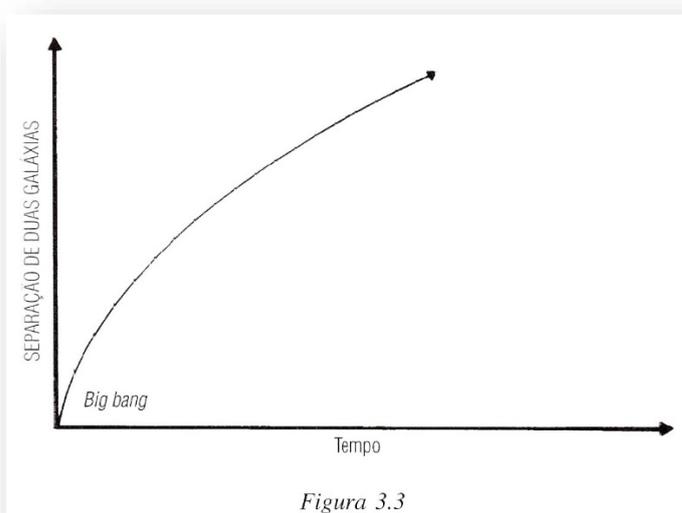
Ora, à primeira vista, todas estas provas de que o universo tem o mesmo aspeto, seja qual for a direção, podem parecer sugerir que existe algo de especial quanto ao nosso lugar no universo. Em particular, pode parecer que, se observamos a recessão de todas as outras galáxias, devemos estar no centro do universo. Há, no entanto, uma explicação alternativa: o universo pode ter o mesmo aspeto em todas as direções se for visto também de outra galáxia. Esta foi, como vimos, a segunda suposição de Friedmann. Não temos qualquer prova científica a favor ou contra ela. Acreditamos nela apenas por modéstia: seria absolutamente espantoso se o universo tivesse o mesmo aspeto em toda a nossa volta e não à volta de outros pontos! No modelo de Friedmann todas as galáxias se afastam diretamente umas das outras. A situação parece-se muito com a de um balão com várias manchas pintadas a ser enchido sem parar³. À medida que o balão se expande, a distância entre quaisquer duas manchas aumenta, mas não pode dizer-se que alguma mancha seja o centro da expansão. Além disso, quanto mais afastadas estiverem as manchas, mais depressa se afastam. Do mesmo modo, no modelo de Friedmann, a velocidade a que duas galáxias quaisquer se afastam uma da outra é proporcional à distância entre elas. Portanto, previa que o desvio para o vermelho de uma galáxia devia ser diretamente proporcional à distância a que se encontra de nós, exatamente como Hubble descobriu. Apesar do êxito deste modelo e da sua previsão das observações de Hubble, o trabalho de Friedmann permaneceu muito tempo desconhecido no Ocidente, até serem descobertos modelos semelhantes em 1935 pelo físico americano Howard Robertson e pelo matemático britânico Arthur Walker, em resposta à descoberta de Hubble da expansão uniforme do universo.

Embora Friedmann tenha descoberto apenas um, há de facto três modelos diferentes que obedecem às suas duas suposições fundamentais. O primeiro é um universo que se expande suficientemente devagar para que a atração gravitacional entre as diferentes galáxias provoque abrandamento e,

provavelmente, paragem da expansão. As galáxias começam então a mover-se umas em direção às outras e o universo contrai-se. A fig. 3.2 mostra como a distância entre duas galáxias vizinhas se modifica à medida que o tempo aumenta. Começa em zero, aumenta até um máximo e depois diminui novamente até zero.

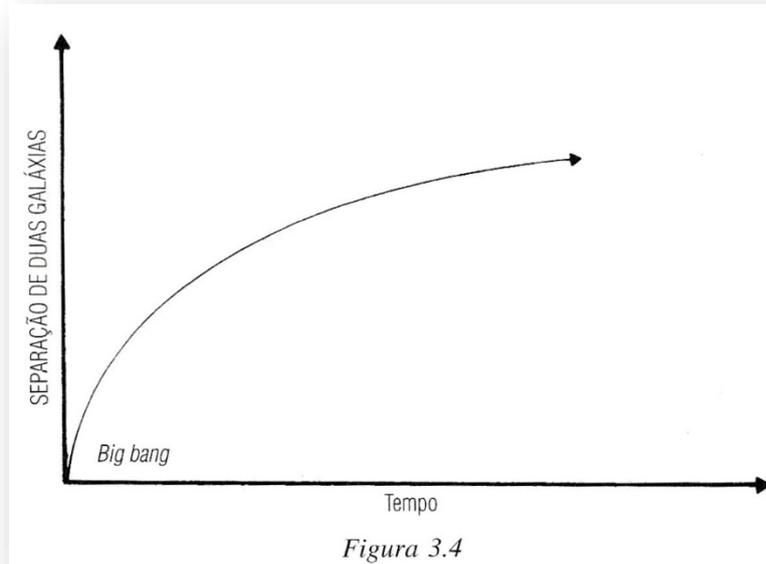


No segundo modelo, o universo expande-se tão rapidamente que a atração gravitacional nunca pode deter a expansão, embora faça que abrande um pouco. A fig. 3.3 mostra a separação entre galáxias vizinhas neste modelo. Começa em zero e depois as galáxias acabam por se afastar a uma velocidade constante.



Finalmente, existe uma terceira espécie de solução, na qual o universo se expande apenas à velocidade suficiente para evitar o colapso. Neste caso, a separação, ilustrada na fig. 3.4, também começa em zero e vai sempre

aumentando. Contudo, a velocidade a que as galáxias se afastam umas das outras torna-se cada vez menor, embora nunca se anule.



Uma característica notável do primeiro modelo de Friedmann é o facto de descrever um universo finito, mas ilimitado. A gravidade é tão forte que o espaço se curva sobre si próprio, o que o torna bastante semelhante à superfície da Terra. Se uma pessoa viajar continuamente em determinada direção na superfície da Terra, nunca chega a uma barreira intransponível nem cai da extremidade; acaba, sim, por voltar ao ponto de partida. No primeiro modelo de Friedmann, o espaço é exatamente assim, mas com três dimensões, em vez das duas à superfície da Terra. A quarta dimensão, o tempo, também é finito em extensão, mas é como uma linha com duas extremidades ou fronteiras, um começo e um fim. Veremos mais tarde que, quando se combina a relatividade geral com o princípio da incerteza da mecânica quântica, é possível que tanto o espaço como o tempo sejam finitos sem quaisquer extremidades ou fronteiras.

A ideia de que é possível andar à volta do universo e voltar ao ponto de partida originou boa ficção científica, mas não tem grande significado prático, porque pode demonstrar-se que o universo colapsaria antes de se conseguir dar a volta. Seria preciso viajar mais depressa do que a luz para voltar ao ponto de partida antes do fim do universo, o que não é possível!

No primeiro modelo de Friedmann, que se expande e depois colapsa, o espaço curva-se sobre si próprio, como a superfície da Terra. É, portanto, finito em extensão. No segundo modelo, que se expande para sempre, o espaço curva-se ao contrário, como a superfície de uma sela. Portanto, nesse caso, o espaço é infinito. Finalmente, no terceiro modelo de Friedmann, que se expande à razão crítica, o espaço é plano (e, portanto, também infinito).

Mas qual é o modelo de Friedmann que melhor descreve o universo? Será que este vai alguma vez parar de expandir e começar a contrair-se ou expandir-se-á para sempre? Para respondermos a esta pergunta precisamos de saber qual o ritmo de expansão do Universo e a sua densidade média. Se a densidade for menor do que um determinado valor crítico, fixado pela taxa de expansão, a atração gravitacional será demasiado fraca para deter a expansão. Se a densidade for maior do que o valor crítico, a gravidade deterá a expansão e conduzirá o universo ao colapso.

Podemos determinar o ritmo atual da expansão, determinando as velocidades a que as galáxias se afastam, através do efeito de Doppler. Isto pode ser conseguido com muita precisão. Contudo, como não conhecemos muito bem as distâncias das galáxias, só podemos medi-las indiretamente. Portanto, apenas sabemos que o universo está a expandir-se à razão de 5% a 10% em cada 1000 milhões de anos. No entanto, a nossa incerteza quanto à densidade média do universo ainda é maior. Se acrescentarmos as massas de todas as estrelas que podemos ver na Galáxia e noutras galáxias, o total será inferior a um centésimo da quantidade necessária para fazer deter a expansão do universo, mesmo para o cálculo mais baixo da taxa de expansão. A nossa e as outras galáxias devem, porém, conter uma grande quantidade de «matéria escura», que não podemos ver diretamente, mas que sabemos dever existir, por causa da influência da sua atração gravitacional nas órbitas das estrelas. Além disso, como a maioria das galáxias se encontra em aglomerados, podemos, assim, concluir que existe mais matéria escura nestes aglomerados pelo seu efeito no movimento das galáxias. Quando somamos toda esta matéria escura, continuamos a não obter mais do que um décimo da quantidade necessária para deter a expansão. Não devemos, porém, excluir a possibilidade da existência de outra forma de matéria, distribuída quase uniformemente pelo universo, que ainda não detetámos e que pode aumentar a densidade média do universo até ao valor crítico necessário para deter a expansão. A evidência sugere, portanto, que o universo, provavelmente, se expandirá para sempre, mas apenas podemos ter a certeza de que, mesmo que venha a colapsar, tal não acontecerá pelo menos durante os próximos 10.000 milhões de anos, uma vez que tem estado a expandir-se pelo menos desde há outro tanto tempo. O facto não deve preocupar-nos muito: entretanto, a não ser que tenhamos colonizado outros sóis, a humanidade ter-se-á extinguido há muito conjuntamente com o Sol!

Todas as soluções de Friedmann têm a característica de num dado momento do passado (entre 10.000 e 20.000 milhões de anos) a distância entre galáxias vizinhas deve ter sido zero. Nesse momento, a que chamamos *big bang*, a densidade do universo e a curvatura do espaço-tempo teriam sido infinitas. Como a matemática não pode realmente lidar com números infinitos, isto significa que a teoria da relatividade geral (em que se baseiam as soluções de Friedmann) prevê que há um ponto do universo onde a própria teoria não se aplica. Esse ponto é um exemplo daquilo a que os matemáticos chamam uma singularidade. De facto, todas as nossas teorias científicas são formuladas na suposição de que o espaço-tempo é quase plano, de modo que não se aplicam a singularidades, como o *big bang*, onde a curvatura do espaço-tempo

é infinita. Isto significa que, mesmo que tivesse havido acontecimentos anteriores ao *big bang*, não poderíamos utilizá-los para determinar o que veio a acontecer depois, porque tudo o que se previsse perde a relevância no momento do *big bang*. Do mesmo modo, se, como é o caso, sabemos apenas o que aconteceu a partir do *big bang*, não podemos determinar o que aconteceu antes. Tanto quanto sabemos, os acontecimentos anteriores ao *big bang* não podem ter quaisquer consequências, pelo que não devem fazer parte de um modelo científico do universo. Devemos, portanto, excluí-los do modelo e dizer que o tempo começou com o *big bang*.

Muitas pessoas não gostam da ideia de o tempo ter um começo, provavelmente porque isso cheira muito a intervenção divina. (A igreja católica, por seu lado, agarrou-se ao modelo do *big bang* e em 1951 afirmou oficialmente que estava de acordo com a Bíblia.) Houve, por isso, algumas tentativas para evitar a conclusão de que tinha havido um *big bang*. A proposta que obteve mais adeptos foi a teoria do estado estacionário. Foi sugerida em 1948 por dois refugiados da Áustria, ocupada pelos nazis, Hermann Bondi e Thomas Gold, juntamente com um inglês, Fred Hoyle, que tinha trabalhado com eles no desenvolvimento do radar durante a guerra. A ideia advogava que, enquanto as galáxias se afastam umas das outras, novas galáxias estão constantemente a formar-se nos intervalos a partir de nova matéria em criação contínua.

O universo, portanto, pareceria mais ou menos sempre igual em todas as épocas e em todos os lugares. A teoria do estado estacionário exigia uma modificação da relatividade geral que permitisse a criação contínua de matéria, mas a taxa de criação era tão baixa (cerca de uma partícula por quilómetro cúbico por ano) que não entrava em conflito com a observação. A teoria era cientificamente boa, no sentido descrito no capítulo 1: era simples e permitia previsões concretas que podiam ser testadas através da observação. Uma dessas previsões consistia em que o número de galáxias, ou objetos semelhantes, num dado volume do espaço seria o mesmo donde e quando quer que se olhasse. No fim dos anos 50 e princípio dos anos 60 foi feito um levantamento das fontes de ondas de rádio do espaço exterior, em Cambridge, por um grupo de astrónomos dirigidos por Martin Ryle (que também tinha trabalhado com Bondi, Gold e Hoyle no radar durante a guerra). O grupo de Cambridge mostrou que a maior parte das fontes de rádio se situava fora da Galáxia (na realidade, muitas podiam ser identificadas com outras galáxias) e também que as fontes fracas eram em muito maior número do que as fortes, interpretando as fontes fracas como sendo as mais distantes e as fortes como as mais próximas. Além disso, parecia haver menos fontes por unidade de volume no caso das fontes próximas do que no das distantes. Isto podia significar que estávamos no centro de uma grande região do universo em que as fontes são em menor número do que em outra parte qualquer. Ou, alternativamente, podia significar que as fontes tinham sido mais numerosas no passado, na época em que as ondas de rádio foram emitidas na nossa direção, do que são agora. Qualquer das explicações contradizia as previsões da teoria do estado estacionário. Além disso, a descoberta da radiação de micro-ondas por Penzias e Wilson, em 1965, também indicava que o universo

devia ter sido muito mais denso no passado. A teoria do estado estacionário teve, portanto, de ser abandonada.

Outra tentativa para evitar a conclusão da existência do *big bang*, portanto, de um princípio, foi realizada por dois cientistas russos, Evgenii Lifshitz e Isaac Khalatnikov, em 1963. Sugeriram que o *big bang* podia ser apenas uma peculiaridade dos modelos de Friedmann que, afinal, não passavam de modelos aproximados. De todos os modelos mais ou menos parecidos com o universo, talvez só o de Friedmann contivesse a singularidade do *big bang*. Nos modelos de Friedmann, as galáxias movem-se, afastando-se todas diretamente umas das outras, pelo que não admira que no passado tenham estado todas no mesmo lugar. Contudo, no universo, as galáxias não se afastam apenas diretamente umas das outras, exibem também pequenas velocidades laterais. Deste modo, na realidade, não precisavam de ter estado todas exatamente no mesmo local, mas apenas muito perto umas das outras. Então talvez o universo em expansão seja o resultado, não de uma singularidade, mas de uma fase anterior de contração; quando o universo colapsou, as partículas que o constituíam não colidiram todas, mas passaram ao lado para depois se afastarem umas das outras, produzindo a atual expansão. Em que base podemos então afirmar que o universo começou com o *big bang*? O que Lifshitz e Khalatnikov fizeram foi estudar universos que eram mais ou menos parecidos com os de Friedmann, mas consideraram as irregularidades e as velocidades aleatórias das galáxias no universo. Mostraram que havia modelos que começavam com o *big bang*, embora as galáxias já não se afastassem diretamente umas das outras; afirmaram então que esta propriedade só se verificava em modelos excepcionais. Argumentaram que, uma vez que parecia haver infinitamente mais modelos como o de Friedmann sem a singularidade do *big bang* do que com ela, devíamos concluir que, na realidade, não tinha havido *big bang*. Mais tarde, contudo, compreenderam que havia também muitos mais modelos como o de Friedmann, com singularidades, em que as galáxias não precisavam de se mover de uma maneira especial. Por isso, em 1970 acabaram por retirar as suas afirmações.

O trabalho de Lifshitz e Khalatnikov foi válido porque mostrou que o universo *podia*⁴ ter tido uma singularidade, um *big bang*, se a teoria da relatividade geral estivesse correta. Contudo, não resolvia a pergunta crucial: a relatividade geral encerra a inevitabilidade do *big bang*, um início dos tempos? A resposta surgiu de uma abordagem completamente diferente do problema, apresentada por um matemático e físico britânico, Roger Penrose, em 1965. Utilizando a maneira como os cones de luz se comportam na relatividade geral, juntamente com o facto de a gravidade ser sempre atrativa, mostrou que uma estrela que entra em colapso devido à própria gravidade fica presa numa região cuja superfície acaba eventualmente por reduzir-se a zero. E, como a superfície se reduz a zero, o mesmo se deve passar com o volume que ela encerra. Toda a matéria existente na estrela é comprimida numa região de volume nulo, de modo que a densidade da matéria e a curvatura do espaço-tempo se tornam infinitas. Por outras palavras, obtém-se uma

singularidade contida numa região de espaço-tempo conhecida por buraco negro.

À primeira vista, o resultado de Penrose aplicava-se apenas às estrelas; nada tinha a ver com a questão de saber se o universo teve ou não uma singularidade. Contudo, na altura em que Penrose apresentou o seu teorema era eu um estudante de investigação que procurava desesperadamente um problema para completar a minha tese de doutoramento. Dois anos antes tinham-me diagnosticado ALS, vulgarmente conhecida por doença de Gehrig, ou neuropatia motora, e tinham-me dado a entender que só tinha mais um ou dois anos de vida. Nessas circunstâncias, não parecia valer muito a pena trabalhar para a tese de doutoramento, pois não esperava viver o tempo suficiente. Contudo, passados dois anos, não tinha piorado muito. Na realidade, as coisas até me corriam bastante bem e tinha ficado noivo de uma excelente rapariga, Jane Wilde. Mas para poder casar tinha de arranjar emprego e para arranjar emprego precisava do doutoramento.

Em 1965 tomei conhecimento do teorema de Penrose de que um corpo em colapso gravitacional origina eventualmente uma singularidade. Depressa compreendi que, se invertesse o sentido do tempo no teorema de Penrose, de modo a transformar o colapso numa expansão, as condições do teorema se manteriam desde que o universo se comportasse, a grande escala e no presente, mais ou menos como no modelo de Friedmann. O teorema de Penrose mostrava que qualquer estrela em colapso devia culminar numa singularidade; o argumento com o tempo ao contrário mostrava que qualquer universo em expansão semelhante ao de Friedmann devia ter começado com uma singularidade. Por razões técnicas, o teorema de Penrose requeria que o universo fosse infinito no espaço. Assim, utilizei-o, de facto, para provar que a singularidade inicial é inevitável no caso de o universo se encontrar numa expansão suficientemente rápida para que não venha a colapsar (uma vez que só estes modelos de Friedmann são infinitos no espaço).

Durante os anos seguintes desenvolvi novas técnicas matemáticas para remover esta e outras condições técnicas dos teoremas que provavam que as singularidades eram inevitáveis. O resultado final foi um trabalho produzido em conjunto por Penrose e por mim, em 1970, o qual provou, por fim, que deve ter havido uma singularidade, contanto que a teoria da relatividade geral esteja correta e o universo não contenha muito mais matéria do que a que observamos. Houve grande oposição ao nosso trabalho por parte dos Soviéticos, por causa da sua fé marxista no determinismo científico, e por parte de pessoas que achavam que a própria ideia de singularidade era repugnante e estragava a beleza da teoria de Einstein. No entanto, não pode discutir-se realmente com um teorema matemático. Deste modo, no fim, o nosso trabalho foi geralmente aceite e hoje em dia quase toda a gente admite que o universo começou com a singularidade do *big bang*. Talvez seja irónico que, tendo mudado de ideias, me encontre agora a tentar convencer outros físicos de que não houve na realidade qualquer singularidade no começo do universo; como veremos mais tarde, a singularidade pode desaparecer quando levamos em conta os efeitos quânticos.

Vimos neste capítulo como, em menos de metade de um século, se transformou a ideia que o homem fazia do universo, ideia formada durante milhares de anos. A descoberta de Hubble de que o universo estava em expansão e a compreensão da insignificância do nosso planeta na sua vastidão foram apenas o ponto de partida. À medida que aumentavam as provas observacionais e teóricas, tornava-se cada vez mais claro que o universo devia ter tido um princípio, até que, em 1970, isso foi finalmente provado por Penrose e por mim com base na teoria da relatividade geral de Einstein. Essa prova mostrou que a relatividade geral é apenas uma teoria incompleta: não pode dizer-nos como surgiu o universo porque prevê que todas as teorias físicas, incluindo ela própria, não se aplicam ao princípio do universo. Assim, a relatividade geral é apenas uma teoria parcial, de modo que o que os teoremas de singularidade mostram realmente é que deve ter havido uma época nos primórdios do universo em que este era tão pequeno que já não podíamos continuar a ignorar os efeitos da outra grande teoria parcial do século XX, a mecânica quântica. Então, no princípio dos anos 70, fomos forçados a voltar as nossas investigações para uma compreensão do universo da teoria do muito grande para a teoria do muito pequeno. Essa teoria da mecânica quântica será descrita a seguir, antes de passarmos aos esforços para combinar as duas teorias parciais numa única teoria quântica da gravidade.

¹ Trata-se das estrelas variáveis cefeides. (N. do R. C.)

² Isotropia e homogeneidade. As propriedades de isotropia e homogeneidade do universo encontram-se encerradas no conteúdo do chamado «princípio cosmológico», talvez o mais importante argumento de toda a cosmologia moderna. (N. do R. C.)

³ Com mais verosimilhança, um balão na superfície do qual se colam papelinhos representando as galáxias a ser inflado. Tal como as heterogeneidades ou irregularidades do universo, os papelinhos não sofrem a inflação. (N. do R. C.)

⁴ Isto é, que, apesar de os modelos generalizados de Friedmann preverem com toda a aproximação pretendida o universo atual, nomeadamente os movimentos laterais das galáxias, outros modelos mais elaborados conduzem-nos também, para trás no tempo, à singularidade inicial. (N. do R. C.)

CAPÍTULO 4

- O PRINCÍPIO DA INCERTEZA -

O êxito das teorias científicas, sobretudo da teoria da gravitação de Newton, levou o cientista francês marquês de Laplace, no início do século XIX, a argumentar que o universo era completamente determinista. Laplace sugeriu que devia haver um conjunto de leis científicas que nos permitissem prever tudo o que viesse a acontecer no universo, bastando para isso sabermos a caracterização completa do seu estado num determinado momento. Por exemplo, se conhecêssemos as posições e velocidades do Sol e dos planetas em determinado momento, podíamos usar as leis de Newton para calcular o estado do sistema solar em qualquer outro momento. O determinismo parece bastante óbvio neste caso, mas Laplace foi mais longe, admitindo que havia leis semelhantes que governavam tudo o mais, incluindo o comportamento humano.

A doutrina do determinismo científico recebeu forte oposição de muitas pessoas, que achavam que ela infringia a liberdade de Deus intervir no mundo, mas manteve-se como hipótese-padrão da ciência até aos primeiros anos deste século. Uma das primeiras indicações de que esta crença teria de ser abandonada surgiu quando cálculos elaborados pelos cientistas britânicos Lord Rayleigh e Sir James Jeans levaram a concluir que um objeto ou corpo quente, como uma estrela, devia radiar quantidades ilimitadas de energia. Segundo as leis em que se acreditava na altura, um corpo quente devia emitir ondas eletromagnéticas (como ondas de rádio, luz visível ou raios X) em quantidades iguais em todas as frequências. Por exemplo, um corpo quente devia radiar a mesma quantidade de energia em ondas com frequências compreendidas entre 1 e 2 milhões de milhões de ondas por segundo, assim como em ondas com frequências compreendidas entre 2 e 3 milhões de milhões de ondas por segundo. Ora, como o número de ondas por segundo não tem limite, isso significaria que a energia total radiada seria infinita.

Para evitar esta conclusão, obviamente ridícula, o cientista alemão Max Planck sugeriu em 1900 que a luz, os raios X e as outras ondas eletromagnéticas não podiam ser emitidas em quantidades arbitrárias, mas apenas em certas quantidades pequenas, a que chamou *quanta*¹. Além disso, cada *quantum* teria uma quantidade de energia que seria tanto maior quanto mais alta fosse a frequência das ondas, de modo que, a uma frequência suficientemente alta, a emissão de um único *quantum* necessitaria de mais energia do que a que se encontrasse disponível. Assim, a radiação a altas frequências seria reduzida e, portanto, a potência de emissão seria finita.

A hipótese dos *quanta* explicava muito bem a emissão de radiação por corpos quentes, mas as suas implicações no determinismo só foram

compreendidas em 1926, quando outro cientista alemão, Werner Heisenberg, formulou o seu famoso princípio da incerteza. Para prever a posição e a velocidade futuras de uma partícula é necessário poder medir com precisão a sua posição e a sua velocidade atuais. A maneira óbvia para conseguir este resultado consiste em fazer incidir luz na partícula. Algumas das ondas luminosas serão dispersas pela partícula, o que indicará a sua posição. Contudo, não conseguiremos determinar a posição da partícula com maior rigor do que a distância entre cristas consecutivas de ondas luminosas², de maneira que é preciso utilizar luz de onda curta para medir com precisão a posição da partícula. Agora, segundo a hipótese do *quantum* de Planck, não pode utilizar-se uma quantidade arbitrariamente pequena de luz; tem de utilizar-se pelo menos um *quantum*. Este *quantum* vai perturbar a partícula e modificar a sua velocidade de um modo que não pode ser previsto. Além disso, quanto maior for a precisão com que se mede a posição, menor será o comprimento de onda necessário e daí maior a energia de um único *quantum*. Portanto, a velocidade da partícula será mais perturbada. Por outras palavras, quanto mais rigorosamente tentarmos medir a posição da partícula, menos precisa será a medida da sua velocidade e vice-versa. Heisenberg mostrou que a incerteza quanto à posição da partícula a multiplicar pela incerteza quanto à sua velocidade e a multiplicar pela massa da partícula nunca pode ser menor do que uma certa quantidade, que é conhecida por constante de Planck. Além disso, este limite não depende do processo de medida da posição, da velocidade ou do tipo de partícula: o princípio da incerteza de Heisenberg é uma propriedade fundamental e inevitável do mundo.

O princípio da incerteza teve implicações profundas na maneira como víamos o mundo que, mais de cinquenta anos depois, ainda não foram devidamente apreciadas por muitos filósofos e continuam a ser objeto de grande controvérsia. O princípio da incerteza marcou o fim do sonho de Laplace de uma teoria científica, um modelo do universo completamente determinista: sem dúvida, é impossível prever acontecimentos futuros com exatidão, pois nem sequer é possível medir com precisão o estado do universo! Podíamos continuar a imaginar que existe um conjunto de leis que determinam completamente os acontecimentos para algum ser sobrenatural, capaz de observar o estado do universo sem o perturbar. Contudo, modelos do universo como esse não são de grande interesse para nós, vulgares mortais. Parece preferível empregar o princípio da parcimónia, conhecido por navalha de Occam, e cortar todos os aspetos da teoria que não podem ser observados. Esta ideia levou Heisenberg, Erwin Schrödinger e Paul Dirac, nos anos 20, a reformularem a mecânica numa nova teoria, chamada mecânica quântica, baseada no princípio da incerteza. Nesta teoria, as partículas deixaram de ter posições e velocidades definidas, que não podiam ser observadas. Em vez disso, tinham um estado quântico resultante da combinação da posição e da velocidade.

Em geral, a mecânica quântica não prevê um único resultado definido para cada observação. Em vez disso, prevê um número de resultados diferentes possíveis e informa-nos sobre a probabilidade de cada um. Ou seja, se uma pessoa executar as mesmas medições num grande número de

sistemas semelhantes, sujeitos a condições iniciais análogas, descobrirá que o resultado das medições será A num determinado número de casos, B num número diferente e assim sucessivamente. Pode prever-se o número aproximado de vezes em que o resultado é A ou B , mas não o resultado específico de uma medição individual. A mecânica quântica introduz, portanto, um elemento inevitável de imprevisibilidade ou acaso na ciência. Einstein protestou fortemente contra esta ideia, apesar do papel importante que desempenhou no seu desenvolvimento. Recebeu o Prémio Nobel pelo seu contributo para a teoria dos *quanta* e, no entanto, nunca aceitou que o universo fosse governado pelo acaso. Os seus sentimentos estão resumidos numa afirmação famosa: «Deus não joga aos dados.» A maior parte dos outros cientistas estava disposta a aceitar a mecânica quântica porque concordava perfeitamente com as experiências. Na realidade, tem sido uma teoria com um êxito notável e está na base de quase toda a ciência e tecnologia modernas. Regula o comportamento de transistores e circuitos integrados, que são componentes essenciais de aparelhos eletrónicos, como televisões e computadores, e é, ao mesmo tempo, a base da química e da biologia modernas. As únicas áreas da física em que a mecânica quântica ainda não foi devidamente incorporada são a gravidade e a estrutura do universo em larga escala.

Embora a luz seja composta por ondas, a hipótese do *quantum* de Planck diz-nos que, de alguma maneira, se comporta como se fosse composta por partículas: só pode ser emitida ou absorvida em pequenas quantidades, ou *quanta*. Do mesmo modo, o princípio da incerteza de Heisenberg implica que as partículas se comportam, em alguns aspetos, como as ondas: não têm uma posição definida, mas estão «espalhadas» com uma certa distribuição de probabilidade. A mecânica quântica baseia-se num tipo inteiramente novo de matemática que já não descreve o mundo real em termos de partículas e ondas; só as observações do mundo podem ser descritas nesses termos. Há, portanto, uma dualidade entre ondas e partículas na mecânica quântica: para alguns fins é útil pensar em partículas como ondas; para outros é melhor pensar em ondas como partículas. Uma consequência importante deste facto consiste em podermos observar aquilo a que chamamos interferência entre dois conjuntos de ondas ou partículas. Ou seja, as cristas de um conjunto de ondas podem coincidir com as cavas de outro conjunto. Os dois conjuntos de ondas anulam-se um ao outro, em vez de se reforçarem para formarem uma onda mais intensa, como poderia esperar-se (fig. 4.1).

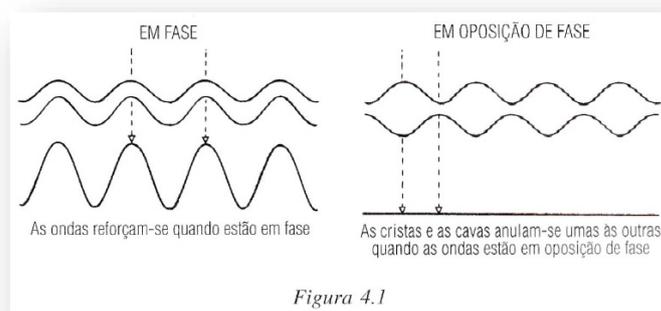
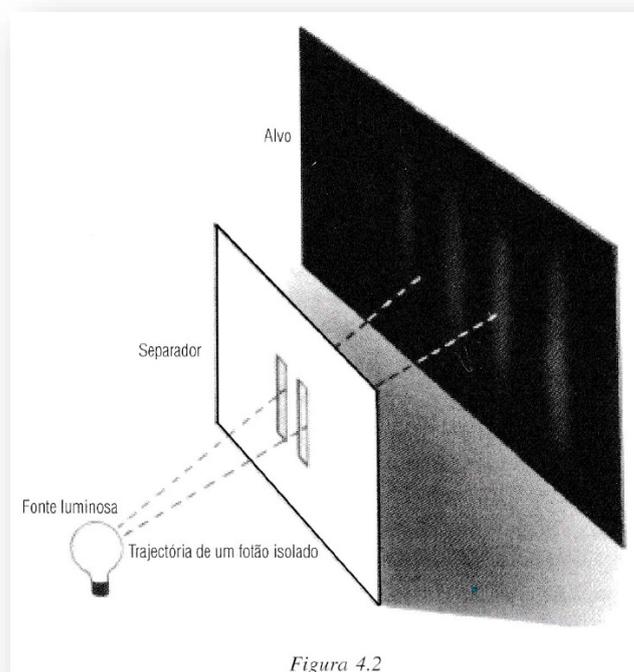


Figura 4.1

Um exemplo familiar de interferência no caso da luz é o das cores que vemos muitas vezes nas bolas de sabão. São causadas por reflexão da luz nos dois lados da fina película de água e sabão que forma a bola. A luz branca consiste em ondas luminosas de comprimentos de onda ou cores diferentes. Para um certo comprimento de onda, as cristas das ondas refletidas de um lado da película de sabão coincidem com as cavas refletidas do outro lado. As cores correspondentes a esses comprimentos de onda estão ausentes da luz refletida, que, portanto, parece ser colorida.

A interferência também pode ocorrer com partículas devido à dualidade introduzida pela mecânica quântica. Um exemplo famoso é o da chamada experiência das duas fendas (fig. 4.2).



Consideremos uma divisória com duas fendas paralelas estreitas. De um dos lados da divisória coloca-se uma fonte de luz de uma cor particular (ou seja, de um comprimento de onda determinado). A maior parte da luz atingirá o separador, mas apenas uma pequena quantidade passará pelas fendas. Suponhamos agora que se coloca um alvo do outro lado da divisória, afastado da luz. Qualquer ponto do alvo receberá ondas das duas fendas. Contudo, no caso geral, a distância que a luz tem de percorrer a partir da fonte até ao alvo, através das duas fendas, será diferente. Isto significa que as ondas estão desfasadas uma da outra quando atingem o alvo: nalguns pontos as ondas anulam-se mutuamente e noutros reforçam-se. O resultado é o padrão característico de franjas claras e escuras.

O que é notável é que se obtém exatamente a mesma espécie de franjas de interferência substituindo a fonte luminosa por uma fonte de partículas, como os eletrões, com uma velocidade definida (o que significa que as ondas correspondentes têm um comprimento definido). Parece ainda mais estranho

porque, quando há só uma fenda, não obtemos franjas, mas sim uma distribuição uniforme de elétrons ao longo do alvo. Poder-se-ia então pensar que a abertura de outra fenda aumentaria apenas o número de elétrons que atingem cada ponto do alvo, mas, devido à interferência, o que realmente acontece é que esse número diminui em alguns pontos. Se os elétrons forem enviados um de cada vez através das fendas, será de esperar que cada um passe por uma ou por outra, e assim se comporte como se a fenda através da qual passou fosse a única, originando uma distribuição uniforme no alvo. Contudo, na realidade, mesmo quando os elétrons são enviados um por um, as franjas continuam a aparecer. Portanto, cada elétron deve estar a passar através de ambas as fendas ao mesmo tempo!

O fenómeno da interferência de partículas tem sido crucial para a nossa compreensão da estrutura dos átomos, as unidades fundamentais da química e da biologia e os blocos de construção de que nós e tudo o que nos rodeia somos formados. No começo deste século pensava-se que os átomos eram bastante parecidos com os planetas em órbita à volta do Sol, com os elétrons (partículas de electricidade negativa) em órbita à volta de um núcleo central com electricidade positiva. Supunha-se que a atração entre a electricidade positiva e a negativa mantinha os elétrons nas suas órbitas, do mesmo modo que a atração gravitacional entre o Sol e os planetas os mantém nas suas órbitas. O problema consistia em que as leis da mecânica e da electricidade, antes da mecânica quântica, previam que os elétrons perderiam energia e se moveriam em espiral para dentro até colidirem com o núcleo. Tal significava o colapso do átomo e, na realidade, de toda a matéria num estado de densidade muito grande. Em 1913, o cientista dinamarquês Niels Bohr encontrou uma solução parcial para o problema. Bohr sugeriu que os elétrons não orbitavam a distâncias arbitrárias do núcleo, mas a determinadas distâncias específicas. Supondo que os elétrons só podem orbitar a qualquer dessas distâncias, está resolvido o problema do colapso do átomo, porque os elétrons já não podem mover-se em espiral senão para preencherem órbitas mais pequenas e de menores energias.

O modelo explicava bastante bem a estrutura do átomo mais simples, o hidrogénio, que tem apenas um elétron em órbita do núcleo. Mas continuava a não ser claro como aplicá-lo a átomos mais complexos. Além disso, a ideia de um conjunto limitado de órbitas permitidas parecia muito arbitrária. A nova teoria da mecânica quântica resolveu esta dificuldade. Revelou que um elétron em órbita do núcleo pode ser considerado uma onda, com um comprimento que depende da sua velocidade. O comprimento de certas órbitas corresponde a um número inteiro (em oposição a um número fracionário) de comprimentos de onda do elétron. Para estas órbitas, a crista de onda está em fase após cada volta, de modo que as ondas se adicionam, correspondendo às órbitas permitidas de Bohr. Contudo, para órbitas cujos comprimentos não são um número inteiro de comprimentos de onda, cada crista é eventualmente anulada por uma cava; essas órbitas não são permitidas.

Uma boa maneira de visualizar a dualidade onda/partícula é através do conceito de soma de histórias, introduzido pelo cientista americano Richard

Feynman. Nesta apresentação do problema, a partícula não tem uma única história ou trajetória no espaço-tempo, como teria numa teoria clássica não quântica. Em vez disso, deverá ir de *A* para *B* seguindo todas as trajetórias possíveis. A cada trajetória estão associados dois números: um representa o tamanho da onda e outro a sua fase (crista ou cava). A probabilidade de ir de *A* para *B* encontra-se somando as ondas para todas as trajetórias. Em geral, ao compararmos trajetórias vizinhas, encontramos grandes diferenças de fase. Isto significa que as ondas associadas a essas trajetórias se anulam umas às outras quase exatamente. Contudo, para alguns conjuntos de trajetórias vizinhas, a fase não varia muito entre elas. As ondas para estas trajetórias não se anulam. São essas as trajetórias que correspondem às órbitas permitidas de Bohr.

Com base nestas ideias, através de uma formulação matemática adequada, foi relativamente fácil calcular as órbitas permitidas aos elétrons em átomos complexos e até em moléculas, estas constituídas por um certo número de átomos unidos por elétrons que partilham mais de um núcleo. Uma vez que a estrutura das moléculas e as suas reações mútuas estão na base de toda a química e de toda a biologia, a mecânica quântica permite-nos, em princípio, prever quase tudo o que vemos à nossa volta, dentro dos limites impostos pelo princípio da incerteza. (Na prática, os cálculos relativos a sistemas com mais do que alguns elétrons são tão complicados que não podemos fazê-los.)

A teoria da relatividade geral de Einstein parece governar a estrutura do universo a grande escala. É uma teoria clássica, ou seja, não faz caso do princípio da incerteza da mecânica quântica, como devia para ser coerente com outras teorias. O motivo pelo qual isto não leva a qualquer discrepância com as observações é que todos os campos gravitacionais que normalmente encontramos são muito fracos. Contudo, os teoremas que mencionámos atrás, relativos a singularidades, indicam que o campo gravitacional deve tornar-se muito intenso em pelo menos duas situações: nas vizinhanças de um buraco negro e no *big bang*. Em campos tão fortes como estes, os efeitos da mecânica quântica devem ser significativos. Assim, em certo sentido, a relatividade geral, ao prever pontos de densidade infinita, demonstrou as suas próprias limitações, tal como a mecânica clássica (ou seja, não quântica) demonstrou as suas limitações ao sugerir que os átomos colapsariam. Não temos ainda uma teoria coerente completa que unifique a relatividade geral e a mecânica quântica, mas conhecemos algumas das propriedades que deve possuir. As consequências que se preveem para os buracos negros e para o *big bang* serão descritas nos capítulos finais. Por agora, porém, voltaremos a nossa atenção para as tentativas recentes de compreensão das outras forças da natureza: a teoria quântica unificada.

¹ Plural de *quantum*. (N. do R. C.)

² Comprimento de onda. (N. do R. C.)

CAPÍTULO 5

- AS PARTÍCULAS ELEMENTARES E AS FORÇAS DA NATUREZA -

Aristóteles acreditava que toda a matéria do universo era constituída por quatro elementos fundamentais: terra, ar, fogo e água. Estes elementos sofriam o efeito de duas forças: a gravidade (tendência da terra e da água para descenderem)¹ e a levitação (tendência do ar e do fogo para subirem). Esta divisão do conteúdo do universo em matéria e forças ainda hoje é utilizada.

Aristóteles acreditava também que a matéria era contínua, ou seja, que podia dividir-se um pedaço de matéria em bocadinhos cada vez mais pequenos, sem limite: nunca se chegava a um grão de matéria que não pudesse ser dividido mais uma vez. Alguns gregos, no entanto, como Demócrito, asseguravam que a matéria era granulosa e que tudo era constituído por grandes quantidades de várias espécies de átomos. (A palavra átomo significa em grego «indivisível».) Durante séculos, a discussão manteve-se sem qualquer prova efetiva a favor de qualquer dos lados, mas em 1803 o químico e físico britânico John Dalton chamou a atenção para o facto de os compostos químicos se combinarem sempre em certas proporções, o que só podia explicar-se pelo agrupamento de átomos em unidades chamadas moléculas. Contudo, a discussão entre as duas escolas só foi resolvida a favor dos atomistas nos primeiros anos do século XX. Uma das provas foi fornecida por Einstein. Num artigo escrito em 1905, algumas semanas antes do famoso trabalho sobre a relatividade restrita, Einstein demonstrou que aquilo a que se chamava o movimento browniano — o movimento irregular e ocasional de pequenas partículas de poeira suspensas num líquido — podia ser explicado como o efeito da colisão das partículas do líquido com os grãos de poeira.

Por essa altura havia já suspeitas de que os átomos não eram, afinal, indivisíveis. Alguns anos antes, um membro do corpo diretivo do Trinity College, de Cambridge, J. J. Thomson, tinha demonstrado a existência de uma partícula de matéria, chamada eletrão, que tinha uma massa de cerca de um milionésimo da do átomo mais leve. Utilizou o que se parecia muito com um moderno aparelho de televisão: um filamento de metal aquecido ao rubro emitia eletrões e, como estes têm uma carga elétrica negativa, podia ser usado um campo elétrico para os acelerar em direção a um alvo revestido de fósforo. Quando atingiam o alvo, produziam clarões de luz. Depressa se compreendeu que esses eletrões deviam provir dos próprios átomos, e em 1911 o físico neozelandês Ernest Rutherford mostrou, finalmente, que os átomos têm realmente uma estrutura interna: são constituídos por um núcleo incrivelmente pequeno e de carga positiva, em torno do qual orbitam os

eletrões. Chegou a esta dedução analisando a maneira como as partículas alfa, de carga positiva, emitidas por átomos radioativos, são defletidas quando colidem com os átomos.

Inicialmente pensava-se que o núcleo do átomo era constituído por eletrões e diferentes quantidades de uma partícula de carga positiva, chamada protão (da palavra grega que significa «primeiro»), porque se julgava tratar-se da unidade fundamental da matéria. Contudo, em 1932, um colega de Rutherford em Cambridge, James Chadwick, descobriu que o núcleo continha outra partícula, chamada neutrão, que tinha praticamente a massa do protão, mas não tinha carga elétrica. Chadwick recebeu o Prémio Nobel pela sua descoberta e foi eleito reitor da Faculdade de Gonville e Caius da Universidade de Cambridge (faculdade de cujo corpo diretivo faço atualmente parte). Mais tarde pediu a demissão desse cargo por desentendimentos com os membros da direção. Tinha havido uma amarga discussão na faculdade desde que um grupo de jovens diretores regressados da guerra se tinha juntado para, por votação, retirar muitos dos velhos membros da direção dos cargos que ocupavam há muito tempo. Isto passou-se antes do meu tempo, pois entrei para a faculdade em 1965, já na reta final de todo o mal-estar, quando desentendimentos semelhantes forçaram outro detentor do Prémio Nobel, Sir Nevill Mott, a pedir a demissão.

Há trinta anos pensava-se que os protões e os neutrões eram partículas «elementares», mas experiências em que se fizeram colidir protões com outros protões ou com eletrões a grandes velocidades revelaram que eram, de facto, constituídos por partículas mais pequenas. A estas partículas atribuiu Murray Gell-Mann, físico do Caltech², que recebeu o Prémio Nobel em 1969 por esse seu trabalho, o nome de *quarks*. A origem do nome é uma citação enigmática do escritor James Joyce: «Three quarks for muster Mark³!» A palavra *quark* devia pronunciar-se como *quart*, mas com um *k* no fim, em vez de *t*, embora seja geralmente utilizada para rimar com *lark*⁴.

Há muitas variedades de quarks: pensa-se que existem seis «sabores», a que chamamos «ascendente» (*up*), «descendente» (*down*), «estranho» (*strange*), «atraente» (*charmed*), «fundo» (*bottom*) e «topo» (*top*). Os três primeiros «sabores» são conhecidos desde 1960, mas «atraente» apenas foi descoberto em 1974, «fundo» em 1977 e «topo» em 1995. Cada sabor surge em três «cores»: vermelho, verde e azul. (Deve acentuar-se que estes termos não passam de rótulos: os quarks são muito mais pequenos do que o comprimento de onda da luz visível e, portanto, não têm qualquer cor no sentido normal da palavra. O que acontece é que os físicos modernos parece terem arranjado maneiras mais imaginativas de batizar novas partículas e fenómenos — já não ficam agarrados ao grego!) Um protão ou um neutrão é constituído por três quarks, um de cada cor. Um protão contém dois quarks *ascendentes* e um *descendente*; um neutrão contém dois *descendentes* e um *ascendente*. Podemos criar partículas constituídas pelos outros quarks (*estranho*, *atraente*, *fundo* e *topo*), mas têm todas uma massa muito maior e decaem muito depressa em protões e neutrões.

Sabemos agora que nem os átomos, nem os prótons, nem os neutrões, são indivisíveis. Portanto, a pergunta que se impõe é esta: quais são as partículas verdadeiramente elementares, os blocos de construção fundamentais a partir dos quais tudo é feito? Uma vez que o comprimento de onda da luz visível é muito maior do que o tamanho de um átomo, não podemos esperar «olhar» para as partes de um átomo no sentido comum. Temos de usar qualquer coisa com um comprimento de onda muito mais pequeno. Como vimos no capítulo anterior, a mecânica quântica diz-nos que todas as partículas são, na realidade, ondas e que, quanto mais elevada é a energia de uma partícula, menor é o comprimento de onda correspondente. Portanto, a melhor resposta que podemos dar à pergunta depende da quantidade de energia de que dispomos, porque é isso que determina a escala a que podemos observá-la. As energias destas partículas são geralmente medidas em unidades chamadas eletrão-volt. (Nas experiências de Thomson com eletrões vimos que ele utilizou um campo elétrico para acelerar os eletrões. A energia que um eletrão ganha num campo elétrico de um volt é um eletrão-volt.) No século XIX, quando as únicas energias de partículas que as pessoas sabiam utilizar eram as energias fracas de uns meros eletrões-volts geradas nas reações químicas, como a reação que se processa numa chama, pensava-se que os átomos eram as unidades mais pequenas. Na experiência de Rutherford, as partículas tinham energias de milhões de eletrões-volts. Mais recentemente aprendemos a utilizar campos eletromagnéticos para darmos às partículas energias de milhões ou mesmo milhares de milhões de eletrões-volts. E, assim, sabemos que algumas partículas que há vinte anos pensávamos serem «elementares» são, na realidade, constituídas por partículas mais pequenas. Será que estas, à medida que temos acesso a energias maiores, serão, por sua vez, reconhecidas como sendo formadas por partículas ainda mais pequenas? É absolutamente possível, mas há algumas razões teóricas para crermos que temos, ou estamos muito perto de ter, um conhecimento dos blocos fundamentais de construção da matéria.

Utilizando a dualidade onda/partícula discutida no capítulo anterior, tudo no universo, incluindo a luz e a gravidade, pode ser descrito em termos de partículas. Estas partículas têm uma propriedade chamada *spin*. Uma maneira de pensar no *spin* consiste em imaginar as partículas como pequenos piões a girar em torno de um eixo. Contudo, isso pode ser enganador, porquanto a mecânica quântica nos diz que as partículas não têm qualquer eixo bem definido. O que o *spin* de uma partícula nos diz na realidade é o aspeto da partícula de diferentes lados. Uma partícula de *spin* 0 é como um ponto: tem o mesmo aspeto vista de qualquer lado (fig. 5.1-I). Por outro lado, uma partícula de *spin* 1 é como uma seta: parece diferente de todos os lados (fig. 5.1-II); só se a fizermos rodar 360 graus é que retoma o mesmo aspeto. Uma partícula de *spin* 2 parece-se com uma seta de duas pontas (fig. 5.1-III).

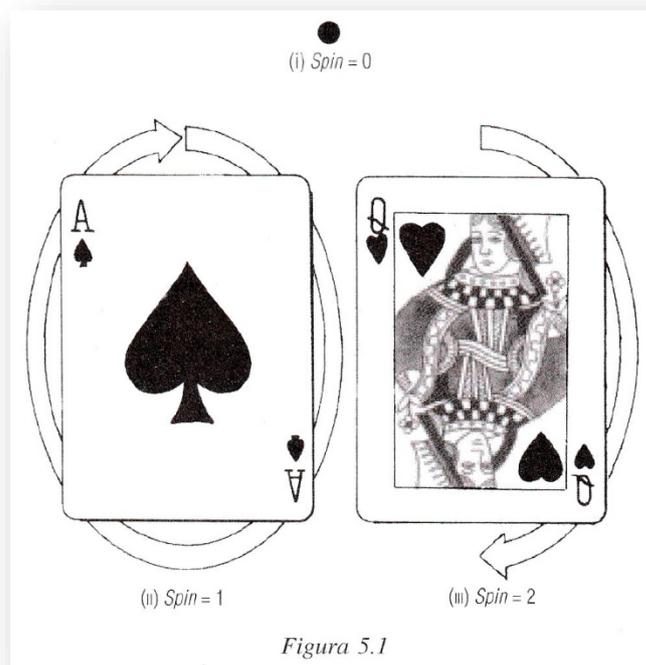


Figura 5.1

Terá o mesmo aspeto se a fizermos dar meia volta (180 graus). Do mesmo modo, partículas de *spin* mais elevado têm o mesmo aspeto se as fizermos girar frações mais pequenas de um giro de 360 graus. Tudo isto parece muito simples, mas o facto notável é existirem partículas que não têm o mesmo aspeto se as fizermos dar apenas uma rotação: temos de obrigá-las a executar duas rotações completas! Diz-se que estas partículas têm $spin\ 1/2$.

Todas as partículas conhecidas no universo podem ser divididas em dois grupos: partículas de $spin\ 1/2$, que constituem a matéria do universo, e partículas de $spin\ 0, 1$ ou 2 , que, como veremos, dão origem a forças entre as partículas de matéria. As partículas de matéria obedecem ao que se chama o princípio da exclusão de Pauli. Este princípio foi descoberto em 1925 por um físico austríaco, Wolfgang Pauli, pelo qual recebeu o Prémio Nobel em 1945. Era o arquétipo do físico teórico: dizia-se dele que a sua simples presença numa cidade faria que todas as experiências aí realizadas resultassem mal! O princípio da exclusão de Pauli diz que duas partículas semelhantes não podem existir no mesmo estado, ou seja, não podem ter ambas a mesma posição e a mesma velocidade, dentro dos limites do princípio da incerteza. O princípio da exclusão é crucial porque explica o motivo pelo qual as partículas de matéria não colapsam num estado de densidade muito elevada sob a influência das forças transmitidas pelas partículas de $spin\ 0, 1$ ou 2 : se as partículas tiverem praticamente as mesmas posições, têm de ter velocidades diferentes, o que significa que não permanecerão na mesma posição durante muito tempo. Se o mundo tivesse sido criado sem o princípio da exclusão, os quarks não formariam prótons e neutrões individualizados e bem definidos. Nem estes, juntamente com os eletrões, formariam átomos individualizados e bem definidos. A matéria colapsaria num «caldo» espesso e grosseiramente uniforme.

A verdadeira compreensão do elétron e de outras partículas de $spin\ 1/2$ só foi atingida em 1928, com a teoria de Paul Dirac, que mais tarde foi eleito professor de Matemática em Cambridge (o mesmo curso que Newton regeu e de que eu agora sou responsável). A teoria de Dirac foi a primeira no gênero consistente tanto com a mecânica quântica como com a teoria da relatividade restrita. Explicou matematicamente por que motivo o elétron tinha $spin\ 1/2$, ou seja, porque não tinha o mesmo aspeto após uma rotação completa, mas sim ao fim de dois giros. Também previu que o elétron devia ter um companheiro, o antieletrão, ou positrão. A descoberta do positrão, em 1932, confirmou a teoria de Dirac e levou a que fosse galardoado com o Prémio Nobel da Física em 1933. Sabemos hoje que toda a partícula tem uma antipartícula, com a qual pode aniquilar-se. (No caso de partículas portadoras de força, as antipartículas coincidem com as próprias partículas.) Pode haver antimundos e antipessoas feitos de antipartículas. Contudo, se encontrar o seu antieu, não lhe aperte a mão! Desapareceriam ambos num grande clarão de luz. A questão de parecer haver mais partículas do que antipartículas à nossa volta é extremamente importante, pelo que voltarei ao assunto ainda neste capítulo.

Na mecânica quântica, as forças ou interações entre partículas de matéria devem ser todas transmitidas por partículas de $spin$ inteiro: 0, 1 ou 2. O que acontece é que uma partícula de matéria, como um elétron ou um quark, emite uma partícula portadora de força. O recuo⁵ provocado por esta emissão altera a velocidade da partícula. A partícula portadora de força colide então com outra partícula de matéria e é absorvida. Esta interação altera também a velocidade da segunda partícula, como se se tivesse manifestado uma força entre as duas partículas.

Uma propriedade importante das partículas portadoras de força consiste em não obedecerem ao princípio da exclusão. Isto significa que não há limite para o número que pode ser permutado e, por isso, podem originar uma força intensa. Contudo, se as partículas portadoras de força tiverem uma grande massa, será difícil produzi-las e permutá-las ao longo de um trajeto grande. Portanto, as forças que transmitem terão apenas um alcance limitado. Por outro lado, se as partículas que transmitem força não tiverem massa própria, as forças serão de longo alcance. As partículas que transportam força entre partículas de matéria chamam-se partículas virtuais porque, ao contrário das partículas «reais», não podem ser detetadas diretamente por um detetor de partículas. Sabemos, contudo, que existem, porque têm um efeito mensurável: originam forças entre partículas de matéria. Partículas de $spin\ 0$, 1 ou 2 também existem em algumas circunstâncias como partículas reais, quando podem ser detetadas diretamente. Surgem-nos então como aquilo a que um físico clássico chamaria ondas, como ondas de luz ou de gravitação. Podem, por vezes, ser emitidas quando as partículas de matéria interagem trocando partículas virtuais que transportam força. Por exemplo, a força de repulsão elétrica entre dois elétrons é devida à troca de fótons virtuais, que não podem ser detetados diretamente, mas, se um elétron passa por outro, podem ser emitidos fótons reais, que detetamos na forma de ondas luminosas.)

As partículas que transportam força podem ser agrupadas em quatro categorias, de acordo com a intensidade da força e as partículas-alvos. Deve acentuar-se que esta divisão em quatro classes é elaborada pelo homem, por ser conveniente para a construção de teorias parciais, mas não pode corresponder a qualquer coisa de mais profundo. O que a maior parte dos físicos espera encontrar é uma teoria unificada que explique as quatro forças como diferentes manifestações de uma única força. Na realidade, muitos diriam que atualmente é esse o objetivo principal da física. Recentemente, foram efetuadas tentativas coroadas de êxito para unificar três das quatro forças — que descreverei neste capítulo. A questão da unificação da outra força, a gravitação, será deixada para mais tarde.

A primeira força é a força da gravitação. Esta força é universal, ou seja, todas as partículas a sentem, conforme a sua massa ou energia. A gravidade é, de longe, a mais fraca das quatro forças; é tão fraca que nem daríamos por ela se não fossem duas propriedades especiais: pode agir a grandes distâncias e é sempre atrativa, o que significa que as forças de gravitação fraquíssimas que atuam entre partículas individuais de dois corpos grandes, como a Terra e o Sol, podem somar-se para produzir uma força significativa. As outras três forças são ou de curto alcance ou umas vezes atrativas e outras repulsivas, tendendo a anular-se. Segundo o modo como na mecânica quântica se encara o campo gravitacional, a força entre duas partículas de matéria é representada como se fosse transportada por uma partícula de *spin 2*, chamada gravitão. Esta não tem massa própria, de modo que a força que transmite é de longo alcance. A força gravitacional entre o Sol e a Terra é atribuída à troca de gravitões entre as partículas que constituem estes dois corpos. Embora as partículas permutadas sejam virtuais, produzem realmente um efeito mensurável: fazem com que a Terra orbite em torno do Sol! Os gravitões reais provocam aquilo a que os físicos chamariam ondas gravitacionais, que são muito fracas e tão difíceis de detetar que ainda não foram observadas.

A força seguinte é a força eletromagnética, que interatua com partículas carregadas de eletricidade, como os eletrões e os quarks, mas não com partículas sem carga, como os gravitões. É muito mais forte do que a força da gravitação: a força eletromagnética entre dois eletrões é de cerca de 1 milhão de milhões de milhões de milhões de milhões de milhões (1 seguido de 42 zeros) de vezes maior do que a força da gravitação. Contudo, há duas espécies de carga elétrica: positiva e negativa. A força entre duas cargas positivas é repulsiva, tal como a força entre duas cargas negativas, mas entre uma carga negativa e uma carga positiva a força é atrativa. Um corpo grande, como a Terra e o Sol, contém quase o mesmo número de cargas positivas e negativas. Assim, as forças atrativas e repulsivas entre partículas individuais quase se anulam e há pouquíssima força eletromagnética disponível. No entanto, à escala dos átomos e das moléculas, as forças eletromagnéticas dominam. A atração eletromagnética entre eletrões de carga negativa e prótons de carga positiva no núcleo obriga os eletrões a orbitarem o núcleo do átomo, tal como a atração gravitacional obriga a Terra a girar à volta do Sol. A atração eletromagnética é vista como sendo causada pela troca de um grande número de partículas virtuais sem massa, de *spin 1*, chamadas fotões. De

novo, os fótons trocados são partículas virtuais. Contudo, quando um elétron transita de uma órbita possível para outra mais próxima do núcleo, é libertada energia e é emitido um fóton real, que pode ser observado como luz visível, se tiver o comprimento de onda certo, ou por um detetor de fótons, como um filme fotográfico. Da mesma maneira, se um fóton real colidir com um átomo, pode deslocar um elétron de uma órbita mais próxima do núcleo para outra mais afastada. Isto gasta a energia do fóton, que é absorvido.

A terceira força chama-se força nuclear fraca e é responsável pela radioatividade, agindo sobre todas as partículas de matéria de *spin* 1/2, mas não sobre partículas de *spin* 0, 1 ou 2, como os fótons ou os grávitons. A força nuclear fraca não era bem compreendida antes de 1967, quando Abdus Salam, no Imperial College de Londres, e Steven Weinberg, em Harvard, propuseram teorias que unificavam esta interação com a força eletromagnética, como Maxwell unificara a eletricidade e o magnetismo cerca de cem anos antes. Sugeriram que, para além do fóton, havia três outras partículas de *spin* 1, conhecidas coletivamente por bósons vetoriais maciços, que transmitiam a força fraca. Foram chamadas W^+ (pronuncia-se W mais), W^- (pronuncia-se W menos) e Z^0 (pronuncia-se Z zero) e cada um tinha uma massa de cerca de 100 GeV (GeV significa GigaEletrão-volt, ou 1000 milhões de eletrões-volts). A teoria de Weinberg e Salam exhibe uma propriedade conhecida por quebra espontânea de simetria. Significa que partículas aparentemente diferentes a baixas energias são manifestações da mesma partícula, mas em estados diferentes. A altas energias, todas estas partículas se comportam de modo semelhante. O efeito é muito parecido com o de uma bolinha de roleta a girar. A altas energias (quando a roleta gira rapidamente), a bolinha tem o mesmo comportamento: não pára de girar. Mas, quando a roleta vai abrandando, a energia da bolinha diminui e acaba por fazê-la cair numa das trinta e sete cavas da roleta. Por outras palavras, a baixas energias há trinta e sete estados diferentes em que a bolinha pode existir. Se, por qualquer razão, só pudessemos observar a bolinha a baixas energias, pensaríamos que havia trinta e sete tipos diferentes de bolinhas!

A teoria de Weinberg e Salam prevê que, a energias muito maiores do que 100 GeV, as três novas partículas e o fóton se comportam todos da mesma maneira. Mas às energias mais baixas que ocorrem na maioria das situações normais esta simetria entre as partículas seria desfeita. W^+ , W^- e Z^0 ficariam com grandes massas, fazendo com que as forças que transportam tivessem um alcance muito curto. Na altura em que Weinberg e Salam propuseram a sua teoria, poucas pessoas acreditaram neles e os aceleradores de partículas não eram suficientemente potentes para alcançarem energias de 100 GeV, necessárias para produzir partículas W^+ , W^- ou Z^0 reais. Contudo, durante os dez anos seguintes, mais ou menos, as outras previsões da teoria a energias mais baixas coincidiram de tal maneira com as experiências que em 1979 Weinberg e Salam receberam o Prémio Nobel da Física, juntamente com Sheldon Glashow, de Harvard, que sugerira teorias unificadas semelhantes para as forças eletromagnética e nuclear fraca. A comissão Nobel foi poupada a um erro embaraçoso devido à descoberta, em 1983, no CERN (Centro Europeu de Pesquisa Nuclear) dos três parceiros maciços do fóton, com massas

e outras propriedades corretamente previstas. Carlo Rubbia, que dirigiu a equipa de várias centenas de físicos que fizeram a descoberta, recebeu o Prémio Nobel em 1984, juntamente com Simon van der Meer, o engenheiro do CERN que tinha desenvolvido o sistema utilizado para armazenar antimatéria. (É muito difícil hoje em dia marcar pontos em física experimental, a não ser que já se esteja no topo!)

A quarta força é a força nuclear forte, que mantém os quarks unidos no próton e no neutrão e os prótons e os neutrões juntos no núcleo de um átomo. Crê-se que esta força é transmitida por uma outra partícula de *spin* 1, chamada gluão, que interatua só consigo própria e com os quarks. A força nuclear forte tem uma propriedade curiosa, chamada confinamento, que mantém as partículas sempre unidas em combinações sem cor. Um quark não pode estar isolado, porque teria cor (vermelha, verde ou azul). Em vez disso, um quark vermelho tem de estar junto de um verde e de um azul por uma «corda» de gluões (vermelho + verde + azul = branco). Este tripleto constitui um próton ou um neutrão. Outra possibilidade é um par formado por um quark e um antiquark (vermelho + antivermelho, ou verde + antiverde, ou azul + antiazul = branco). Estas combinações constituem as partículas conhecidas por mesões, que são instáveis, porque um quark e um antiquark podem aniquilar-se, originando eletrões ou outras partículas. Do mesmo modo, o confinamento evita que se tenha um único gluão, porque os gluões também têm cor. Em vez disso, é preciso ter um conjunto de gluões cujas cores conjuntas produzam o branco. Esse conjunto forma uma partícula instável, chamada *glueball*⁶.

O facto de o confinamento não permitir que se observe um quark ou um gluão isolados podia fazer crer que os quarks e os gluões são partículas um tanto metafísicas. No entanto, há outra propriedade da força nuclear forte, chamada liberdade assintótica, que torna os conceitos de quark e de gluão bem definidos. A energias normais, a força nuclear forte é realmente forte e mantém os quarks unidos. Contudo, experiências com grandes aceleradores de partículas indicam que a energias elevadas a força forte se torna muito mais fraca e os quarks e os gluões se comportam quase como partículas livres. A fig. 5.2 mostra uma fotografia de uma colisão entre um próton e um antipróton a altas energias.

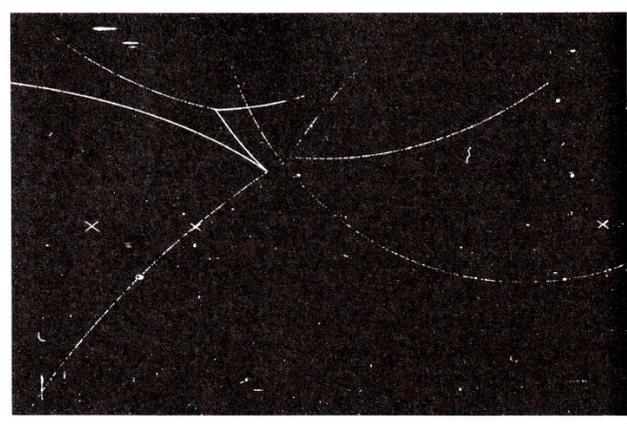


Figura 5.2 – Um próton e um antipróton colidem a alta energia produzindo um conjunto de quarks quase livres.

[Foram produzidos vários quarks quase livres que deram origem aos «jatos» de trajetórias vistos na fotografia] O êxito da unificação das forças eletromagnética e nuclear fraca levou a várias tentativas para combinar estas duas forças com a força nuclear forte naquilo a que se chamou teoria da grande unificação, ou GUT⁷. Este título é um tanto exagerado: as teorias resultantes não são assim tão grandes, nem completamente unificadas, porque não incluem a gravitação. Nem são teorias realmente completas, porque contêm um número de parâmetros cujos valores não podem ser previstos a partir da teoria, mas têm de ser escolhidos para se harmonizarem com as experiências. Apesar disso, podem ser um passo no sentido de uma teoria completa, totalmente unificada. A ideia fundamental das GUT é a seguinte: como já foi mencionado, a força nuclear forte torna-se mais fraca a altas energias. Por outro lado, as forças eletromagnética e nuclear fraca, que não são assintoticamente livres, tornam-se mais fortes a energias elevadas. A uma energia muito elevada, chamada energia da grande unificação, essas três forças teriam todas a mesma intensidade e poderiam, portanto, ser apenas diferentes aspetos de uma única força. As GUT preveem também que as diferentes partículas da matéria de *spin* $1/2$, como os quarks e os elétrons, seriam essencialmente aspetos das mesmas partículas, obtendo-se, assim, outra unificação.

Não se conhece muito bem o valor da energia da grande unificação, mas teria, provavelmente, de ser, pelo menos, de 1000 milhões de milhões de GeV. Os atuais aceleradores de partículas podem fazer colidir partículas a energias de cerca de algumas centenas de GeV e estão planeadas máquinas que elevarão esta energia a alguns milhares de GeV. Mas uma máquina suficientemente potente para acelerar partículas até à energia da grande unificação teria de ser tão grande como o sistema solar — e seria pouco provável haver fundos para ela no atual contexto económico. Portanto, é impossível testar diretamente em laboratório teorias da grande unificação. No entanto, tal como no caso da teoria da unificação eletromagnética e fraca, há consequências a baixas energias que podem ser testadas.

Mais interessante é a previsão de que os prótons, que constituem grande parte da massa da matéria vulgar, podem desintegrar-se espontaneamente em partículas mais leves, como os positrões. Tal é possível porque na energia da grande unificação não há qualquer diferença essencial entre um quark e um elétron. Os três quarks dentro de um próton não têm normalmente energia suficiente para se transformarem em positrões, mas muito ocasionalmente um deles pode adquirir energia suficiente para provocar a transição, porque o princípio da incerteza determina que a energia dos quarks dentro do próton não pode ser exatamente fixada. O próton desintegrar-se-ia. A probabilidade de um quark adquirir energia suficiente é tão pequena que o mais provável é termos de esperar, pelo menos, 1 milhão de milhões de milhões de milhões de milhões de anos (1 seguido de 30 zeros), o que é muito superior ao tempo que passou desde o *big bang*, que aconteceu mais ou menos há 10.000 milhões de

anos (1 seguido de 10 zeros). Portanto, poderíamos pensar que o decaimento espontâneo do próton não poderia ser testado experimentalmente. Contudo, podemos aumentar as nossas hipóteses de detetarmos uma desintegração observando uma grande porção de matéria que contenha um grande número de prótons. (Se, por exemplo, observarmos um número de prótons igual a 1 seguido de 31 zeros durante o período de um ano, podemos esperar observar, segundo a mais simples GUT, mais de um próton a desintegrar-se.)

Têm sido feitas várias experiências deste gênero, mas nenhuma forneceu ainda provas convincentes da desintegração de prótons ou nêutrons. Uma experiência realizada na mina de sal Morton, no Ohio, utilizou 8000 toneladas de água (para evitar que acontecessem outros fenômenos causados por raios cósmicos, que podiam ser confundidos com o decaimento de prótons). Uma vez que não foi observada qualquer desintegração espontânea durante a experiência, podemos calcular que a vida provável do próton deve ser maior do que 10 milhões de milhões de milhões de milhões de milhões de anos (1 seguido de 31 zeros). Este número é superior ao tempo de vida previsto pela teoria da grande unificação mais simples, mas existem teorias mais elaboradas, nas quais os tempos de vida previstos são mais longos. Serão necessárias experiências ainda mais sensíveis com quantidades ainda maiores de matéria para as testar.

Embora seja muito difícil observar o decaimento espontâneo do próton, talvez a nossa própria existência seja uma consequência do processo contrário, a produção de prótons, ou mais simplesmente de quarks, a partir de uma situação inicial em que não havia mais quarks do que antiquarks, que é a maneira mais natural de imaginar o começo do universo. A matéria na Terra é constituída principalmente por prótons e nêutrons, que, por seu turno, são constituídos por quarks. Não há antiprótons, nem antineutrões constituídos a partir de antiquarks, exceto aqueles que são produzidos pelos físicos em grandes aceleradores de partículas. Temos provas, a partir de raios cósmicos, de que o mesmo se passa com a matéria da Galáxia: não estão presentes antiprótons ou antineutrões, para além de um pequeno número de pares partícula/antipartícula obtidos em colisões a altas energias. Se houvesse grandes regiões de antimatéria na Galáxia, esperaríamos ver grandes quantidades de radiação provenientes do contato entre as regiões de matéria e de antimatéria, onde muitas partículas estariam a colidir com as suas antipartículas, aniquilando-se mutuamente e emitindo radiação de energia elevada.

Não temos qualquer prova direta sobre se a matéria nas outras galáxias é constituída por prótons e nêutrons ou antiprótons e antineutrões, mas tem de ser uma coisa ou outra: não pode haver uma mistura numa única galáxia porque nesse caso observaríamos uma grande quantidade de radiação proveniente de aniquilações. Além disso, cremos que todas as galáxias são compostas de quarks e não de antiquarks; não parece plausível que algumas galáxias sejam matéria e outras antimatéria.

Por que haverá mais quarks do que antiquarks? Por que não há um número igual de cada um? É certamente uma sorte para nós os números não serem iguais, pois, se assim fosse, quase todos os quarks e antiquarks se teriam aniquilado mutuamente no princípio do Universo, deixando-o cheio de radiação, mas com muito pouca matéria. Não teria então havido galáxias, estrelas ou planetas onde a vida humana viesse a desenvolver-se. Felizmente, as teorias da grande unificação podem explicar o motivo pelo qual o universo deve conter agora mais quarks do que antiquarks, mesmo que tenha começado com um número igual de ambos. Como vimos, as GUT admitem a mudança de quarks em positrões a altas energias. E também admitem o processo contrário, antiquarks a transformarem-se em elétrons e elétrons e positrões a transformarem-se em antiquarks e quarks. Houve um tempo, logo no princípio do universo, em que havia tanto calor que as energias das partículas seriam suficientemente elevadas para estas transformações ocorrerem. Mas por que havia isso de produzir mais quarks do que antiquarks? A razão consiste em que as leis da física não são exatamente as mesmas para as partículas e para as antipartículas.

Até 1956 acreditava-se que as leis da física obedeciam a três simetrias, denominadas C, P e T. A simetria C significa que as leis são as mesmas para partículas e antipartículas. A simetria P significa que as leis são as mesmas para qualquer situação e a sua imagem num espelho (a imagem num espelho de uma partícula rodando sobre si mesma num sentido é a de uma partícula que gira no outro sentido). A simetria T significa que, se invertermos o sentido do movimento de todas as partículas e antipartículas, o sistema deveria voltar a ser o que era antes; por outras palavras, as leis são as mesmas para diante e para trás no tempo. Em 1956 dois físicos americanos, Tsung-Dao Lee e Chen Ning Yang, sugeriram que a força fraca não obedece à simetria P. Por outras palavras, a força faria que o universo se desenvolvesse de um modo diferente da sua imagem no espelho. No mesmo ano, uma colega, Chien-Shiung Wu, provou que esta teoria estava correta: conseguiu-o alinhando núcleos de átomos radioativos num campo magnético, de modo a ficarem todos a girar sobre si mesmos no mesmo sentido, e mostrou que os elétrons eram produzidos mais num sentido do que no outro. No ano seguinte, Lee e Yang receberam o prémio Nobel. Descobriu-se também que a força fraca não obedecia à simetria C. Ou seja, originaria um universo composto de antipartículas que se comportaria de maneira diferente do nosso universo. Não obstante, parecia que a força fraca obedecia realmente à simetria composta CP. Ou seja, o universo desenvolver-se-ia da mesma maneira que a sua imagem num espelho se, além disso, cada partícula fosse trocada pela sua antipartícula! Contudo, em 1964, mais dois americanos, J. W. Cronin e Val Fitch, descobriram que mesmo a simetria CP não se verificava no decaimento de certas partículas, chamadas mesões *K*. Cronin e Fitch acabaram por receber o Prémio Nobel em 1980. (Muitos prémios têm sido concedidos por se mostrar que o universo não é tão simples como poderia pensar-se!)

Há um teorema matemático que afirma que qualquer teoria que obedeça à mecânica quântica e à relatividade tem de obedecer sempre à simetria composta CPT. Por outras palavras, o universo teria de se comportar da

mesma maneira se substituíssemos as partículas por antipartículas, tomássemos a sua imagem no espelho e ainda invertêssemos o sentido do tempo. Mas Cronin e Fitch demonstraram que, se substituíssemos partículas por antipartículas e considerássemos a imagem no espelho, mas não invertêssemos o sentido do tempo, o universo *não* se comportaria da mesma maneira. As leis da física, portanto, devem ser alteradas quando se inverte o sentido do tempo — não obedecem à simetria T.

Certamente o universo primitivo não obedece à simetria T: à medida que o tempo passa, o universo expande-se: se andasse para trás, o universo ter-se-ia contraído. E, como existem forças que não obedecem à simetria T, segue-se que, enquanto o universo se expande, essas forças podem originar que mais positrões se transformem em quarks do que elétrons em antiquarks. Como o universo se expandiu e arrefeceu, os antiquarks e os quarks aniquilaram-se e, como havia mais quarks do que antiquarks, restou um pequeno excesso de quarks. São eles que constituem a matéria que hoje vemos e da qual nós próprios somos feitos. Assim, a nossa existência real poderia ser considerada uma confirmação das teorias da grande unificação, embora apenas de uma forma qualitativa; as incertezas são tais que é impossível prever o número de quarks que sobreviveriam à aniquilação ou até se o que restaria seriam quarks ou antiquarks. (Se, no entanto, o excesso fosse de antiquarks, teríamos muito simplesmente chamado quarks aos antiquarks e vice-versa).

As teorias da grande unificação não incluem a força da gravidade. Isto não tem muita importância, porque a gravidade é uma força tão fraca que os seus efeitos podem geralmente ser desprezados quando lidamos com partículas elementares ou átomos. Contudo, o facto de ser de longo alcance e sempre atrativa significa que todos os seus efeitos se juntam. Portanto, para um número suficientemente grande de partículas de matéria, as forças gravitacionais podem dominar todas as outras forças. É por isso que a gravidade determina a evolução do universo. Mesmo para objetos do tamanho de estrelas, a força atrativa da gravidade pode vencer todas as outras forças e fazer com que a estrela sofra um colapso. O meu trabalho nos anos 70 incidiu sobre os buracos negros que podem resultar desses colapsos estelares e dos campos gravitacionais que os rodeiam. Foi isso que levou aos primeiros indícios de como a teoria da mecânica quântica e a teoria da relatividade geral podiam influenciar-se uma à outra — um vislumbre de uma teoria quântica da gravidade ainda por encontrar.

¹ O movimento da oitava esfera, a esfera das fixas, comunicava-se à esfera da Lua. O movimento da Lua, por sua vez, causava, por fricção, a mistura dos quatro elementos que compunham o mundo sublunar. Assim, o movimento da oitava esfera era responsável pelo movimento no sentido lato aristotélico: pela génese e pela corrupção. (N. do R. C.)

² Instituto de Tecnologia da Califórnia. (N. do R. C.)

³ «Três litros para o patrão Mark!» (N. do R. C.)

⁴ Cotovia. (N. do R. C.)

⁵ O «recuo», ou o «avanço», consoante a força seja repulsiva ou atrativa. Há uma ideia muito intuitiva sobre estas partículas portadoras de força. Imaginemos dois patinadores no gelo que seguem lado a lado e suponhamos que num dado momento eles decidem arremessar pedras um ao outro: separam-se por efeito do recuo, mas, para um observador para o qual as pedras

fossem invisíveis, tudo se passa como se se repelisse por intermédio de uma força. No caso de os patinadores de decidirem a arremessar bumerangues, em vez de pedras, o efeito seria inverso, aproximar-se-iam, como que sujeitos a uma força atrativa mediada pelos bumerangues (N. do R. C.)

⁶ Bola de grude. (N. do R. C.)

⁷ Grand unification theory. (N. do R. C.)

CAPÍTULO 6

- BURACOS NEGROS -

A expressão *buraco negro* tem uma origem muito recente. Foi forjada em 1969 pelo cientista americano John Wheeler, como descrição gráfica de uma ideia que data, pelo menos, de há duzentos anos, do tempo em que havia duas teorias sobre a luz: uma, que Newton preferia, sustentava que a luz era composta por partículas; a outra defendia que a luz era de natureza ondulatória. Sabemos agora que, na realidade, ambas as teorias estão corretas. Pela dualidade onda/partícula da mecânica quântica, a luz pode ser considerada uma onda ou uma partícula. A teoria ondulatória não explicava como a luz reagia à gravidade. Admitindo que a luz é composta por partículas, compreende-se que seja afetada pela gravidade do mesmo modo que as balas de canhão, os foguetões e os planetas. Ao princípio pensava-se que as partículas de luz se deslocavam com uma velocidade infinita, de maneira que a gravidade não seria capaz de as retardar, mas a descoberta de Roemer de que a luz se propaga com velocidade finita significava que a gravidade podia ter um efeito importante.

Partindo desta suposição, um catedrático de Cambridge, John Michell, escreveu em 1783 um artigo, publicado nos *Philosophical Transactions of the Royal Society of London*, em que chamava a atenção para o facto de uma estrela suficientemente maciça e compacta poder ter um campo gravitacional tão forte que a sua luz não poderia escapar: a luz emitida à superfície seria puxada para trás, pela atração gravitacional da estrela, antes de poder escapar. Michell sugeriu que poderia haver um grande número de estrelas como esta. Embora não pudéssemos vê-las, porque a sua luz não nos alcançaria, não deixaríamos de sentir a sua atração gravitacional. Esses objetos são aquilo a que agora chamamos buracos negros, porque é isso mesmo que são: vazios negros no espaço. Alguns anos mais tarde o cientista francês marquês de Laplace apresentou uma sugestão semelhante, segundo tudo indica, independentemente de Michell. É interessante o facto de Laplace, apesar de ter incluído apenas nas duas primeiras edições do seu livro *Sistema do Mundo*¹, nada referiu nas edições seguintes, talvez por ter decidido que era uma ideia disparatada. (Também a teoria das partículas de luz caiu em desagrado durante o século XIX. Parecia que tudo podia ser explicado com a teoria ondulatória, mas, segundo esta, não era claro se a luz seria ou não afetada pela gravidade.)

De facto, não é adequado tratar a luz como balas de canhão na teoria da gravitação de Newton, porque a velocidade da luz é constante. (Uma bala de canhão disparada verticalmente para cima é desacelerada pela gravidade até acabar por cair. Um fotão, porém, continua para cima a uma velocidade

constante. Como pode a gravidade afetar a luz?) Uma teoria coerente sobre a maneira como a gravidade afeta a luz só surgiu quando Einstein propôs a relatividade geral, em 1915. E, mesmo assim, foi preciso esperar mais tempo até que fossem compreendidas todas as implicações da teoria na evolução das estrelas maciças.

Para compreendermos como pode formar-se um buraco negro precisamos primeiro de compreender o ciclo de vida de uma estrela. Uma estrela forma-se quando uma grande porção de gás (sobretudo hidrogénio) se contrai por causa da atração gravitacional das suas partes. À medida que o gás se contrai, aquece, pois os seus átomos colidem uns com os outros, cada vez com maior frequência e a velocidades progressivamente maiores. A certa altura, o gás estará tão quente que, quando os átomos de hidrogénio² colidem, já não ressaltam, mas juntam-se para formar hélio. O calor libertado nesta reação, que é como a explosão controlada de uma bomba de hidrogénio, faz a estrela brilhar. O calor adicional aumenta a pressão do gás, que acaba por equilibrar a atração gravitacional e deixa de se contrair. É parecido com um balão: há um equilíbrio entre a pressão do ar dentro dele, que tende a dilatar o balão, e a tensão da borracha, que tende a tornar o balão mais pequeno. As estrelas permanecem estáveis durante muito tempo, com o calor das reações nucleares a equilibrar a atração gravitacional. A certa altura, porém, a estrela esgotará o hidrogénio e outros combustíveis nucleares. Paradoxalmente, quanto maior for a porção de combustível com que a estrela começa, mais depressa este se esgotará. Isto sucede porque, quanto mais maciça for a estrela, mais quente precisará de estar para equilibrar a atração gravitacional. E, quanto mais quente estiver, mais depressa gastará o combustível. O nosso Sol tem, provavelmente, combustível suficiente para mais 5000 milhões de anos, mas as estrelas mais maciças podem esgotar o seu combustível em tão pouco tempo como 100 milhões de anos, muito menos do que a idade do universo. Quando uma estrela esgota o combustível, começa a arrefecer e, portanto, a contrair-se. O que pode acontecer-lhe só foi compreendido pela primeira vez no fim dos anos 20.

Em 1928, um estudante indiano finalista, Subrahmanyan Chandrasekhar, foi estudar para Inglaterra, em Cambridge, com o astrónomo britânico Sir Arthur Eddington, especialista em relatividade geral. (Segundo alguns relatos, um jornalista disse a Eddington, no princípio dos anos 20, que tinha ouvido dizer que só havia três pessoas no mundo que compreendiam a relatividade geral. Eddington fez uma pausa e depois respondeu: «Estou a tentar ver se me lembro de quem é a terceira pessoa.») Durante a viagem, Chandrasekhar descobriu qual poderia ser o tamanho máximo de uma estrela para que se sustivesse contra a própria gravidade depois de ter esgotado todo o combustível. A ideia foi esta: quando a estrela se contrai, as partículas de matéria aproximam-se muito umas das outras e, portanto, segundo o princípio de exclusão de Pauli, têm de ter velocidades muito diferentes. Isto leva-as a afastarem-se umas das outras, fazendo que a estrela se expanda. Uma estrela pode então manter-se com raio constante em virtude do equilíbrio entre a atração da gravidade e a repulsão que resulta do princípio de exclusão, tal como anteriormente a gravidade era equilibrada pelo calor.

Chandrasekhar compreendeu, contudo, que há um limite para a repulsão que pode resultar do princípio de exclusão. A teoria da relatividade impõe um limite máximo à velocidade das partículas na estrela: a velocidade da luz. Isto quer dizer que, quando a estrela fica suficientemente densa, a repulsão induzida pelo princípio da exclusão é menor do que a atração gravitacional. Chandrasekhar calculou que uma estrela fria, com cerca de uma vez e meia a massa do Sol, não poderia manter-se contra a própria gravidade. (Esta massa é agora conhecida como limite de Chandrasekhar.) Mais ou menos na mesma altura, o cientista russo Lev Davidovich Landau fez uma descoberta semelhante. Isto teve sérias implicações na investigação sobre o destino final de estrelas maciças. Se a massa de uma estrela for inferior ao limite de Chandrasekhar, então pode eventualmente deixar de contrair-se e manter-se num possível estado final de anã branca, com um raio de alguns milhares de quilómetros e uma densidade de centenas de toneladas por centímetro cúbico. Uma anã branca é sustida pela repulsão induzida pelo princípio de exclusão aplicado aos eletrões da matéria de que é formada. Podemos observar um grande número destas estrelas anãs brancas. Uma das primeiras a serem descobertas foi uma estrela satélite de Sírio, a estrela mais brilhante do céu noturno.

Landau fez notar que havia outro estado final possível para uma estrela com uma massa-limite de cerca de uma ou duas vezes a massa do Sol, mas muito mais pequena do que uma anã branca. Estas estrelas seriam sustidas pela repulsão, induzida pelo princípio de exclusão, entre neutrões e prótons³. Designadas por estrelas de neutrões, teriam um raio de apenas 20 quilómetros, aproximadamente, e uma densidade de centenas de milhões de toneladas por centímetro cúbico. Quando foram previstas pela primeira vez, não era possível observá-las, pelo que só foram detetadas muito mais tarde.

Por outro lado, as estrelas com massas acima do limite de Chandrasekhar enfrentam um grande problema quando se lhes acaba o combustível. Em alguns casos podem explodir ou disparar para o espaço matéria suficiente para reduzirem a massa abaixo do limite e, assim, evitarem o colapso gravitacional catastrófico, mas é difícil acreditar que isto aconteça sempre, seja qual for o tamanho da estrela. Como sabe a estrela que tem de perder massa? E, mesmo que todas as estrelas conseguissem perder massa suficiente para evitarem o colapso, que aconteceria se se acrescentasse massa a uma anã branca ou a uma estrela de neutrões de modo a ultrapassar o limite? Contrair-se-ia até atingir uma densidade infinita? Eddington ficou chocado com a ideia e recusou-se a acreditar nos resultados obtidos por Chandrasekhar. Pensava que era, pura e simplesmente, impossível uma estrela contrair-se até acabar num pontinho. E esta era a opinião da maior parte dos cientistas: o próprio Einstein escreveu um artigo em que afirmava que as estrelas não podiam contrair-se até zero. A hostilidade dos outros cientistas, particularmente de Eddington, seu antigo professor e a maior autoridade na estrutura das estrelas, persuadiu Chandrasekhar a abandonar aquele caminho e a voltar-se para outros problemas de astronomia, como o movimento de aglomerados de estrelas. Contudo, quando recebeu o Prémio Nobel, em 1983, foi, pelo menos em parte, pelo seu trabalho anterior sobre o limite de massa das estrelas frias.

Chandrasekhar tinha mostrado que o princípio de exclusão não podia travar o colapso de uma estrela com massa superior ao limite, mas o problema de compreender o que poderia acontecer a uma dessas estrelas, de acordo com a relatividade geral, foi resolvido por um jovem americano, Robert Oppenheimer, em 1939. No entanto, o resultado que obteve sugeria que não haveria consequências que pudessem ser observadas com os telescópios da época. Depois veio a segunda guerra mundial e o próprio Oppenheimer foi envolvido no projeto da bomba atômica. Depois da guerra, o problema do colapso gravitacional foi praticamente esquecido, porque a maioria dos cientistas preocupava-se com o que acontecia à escala do átomo e do seu núcleo. Nos anos 60, porém, o interesse pelos problemas da astronomia e da cosmologia em macroescala foi reavivado por um aumento considerável do número e da importância das observações astronômicas que surgiram com a tecnologia moderna. O trabalho de Oppenheimer foi então redescoberto e continuado.

A imagem que agora temos do trabalho de Oppenheimer é a seguinte: o campo gravitacional da estrela altera as trajetórias dos raios luminosos no espaço-tempo em relação ao que deviam ser se a estrela não estivesse presente. Os cones de luz, que indicam as trajetórias seguidas no espaço e no tempo por clarões luminosos emitidos nos seus vértices são ligeiramente curvados para dentro, perto da superfície da estrela. Isto pode ser observado na deflexão da luz de estrelas distantes observadas durante um eclipse do Sol. À medida que a estrela se contrai, o campo gravitacional na sua superfície torna-se mais intenso e os cones de luz encurvam-se mais para dentro. Isto torna mais difícil que a luz da estrela se escape e a luz parece mais fraca e mais vermelha a um observador colocado à distância. Quando a estrela se contrai, atingindo um determinado raio crítico, o campo gravitacional na superfície acaba por se tornar tão intenso que os cones de luz se encurvam para dentro de tal modo que a luz já não pode escapar (fig. 6.1).

Segundo a teoria da relatividade, nada pode deslocar-se mais depressa do que a luz. Portanto, se a luz não consegue escapar-se, mais nada o consegue: tudo é arrastado para trás pelo campo gravitacional. Por conseguinte, há um conjunto de acontecimentos, uma região do espaço-tempo, donde não é possível escapar para alcançar um observador distante. Esta região é aquilo a que chamamos buraco negro. A sua fronteira é o horizonte de acontecimentos e coincide com as trajetórias dos raios de luz que não conseguem escapar-se do buraco negro.

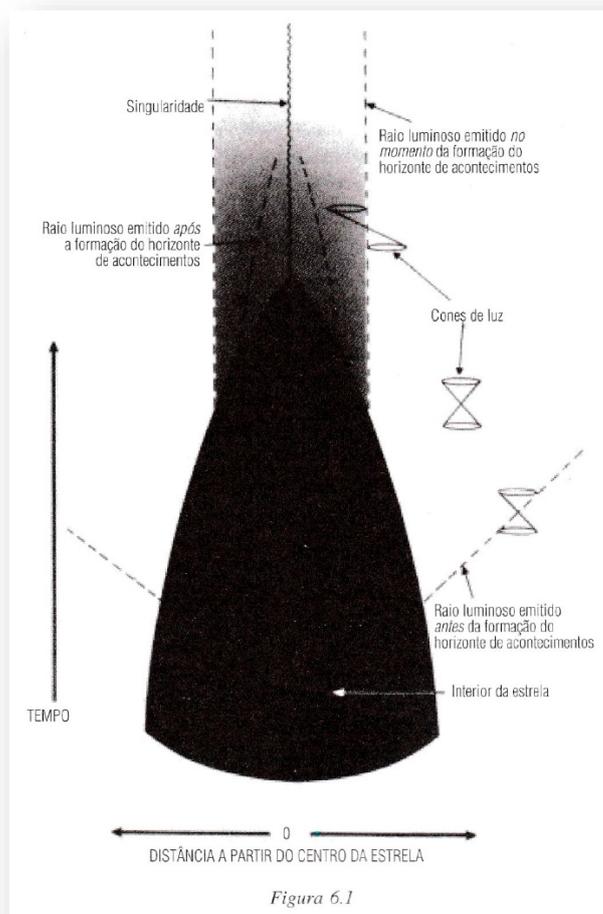


Figura 6.1

Para compreendermos o que veríamos se estivéssemos a observar o colapso de uma estrela e a consequente formação de um buraco negro temos de lembrar que na teoria da relatividade não há tempo absoluto. Cada observador tem a sua medida própria do tempo. O tempo para alguém numa estrela será diferente do tempo para alguém colocado à distância devido ao campo gravitacional da estrela. Suponhamos que um astronauta intrépido, na superfície de uma estrela em colapso, se deixa arrastar para dentro com ela, enviando um sinal de segundo a segundo, de acordo com o seu relógio, para a nave espacial em órbita em redor da estrela. Num dado momento no seu relógio, às 11 horas, por exemplo, a estrela contrai-se aquém do raio crítico em que o campo gravitacional se torna tão forte que nada consegue escapar-se e os sinais já não chegam à nave espacial. À medida que as 11 horas se aproximam, os companheiros que o observam da nave acham que os sinais estão a ser enviados com intervalos cada vez maiores, embora este efeito seja insignificante antes das 10.59.59. Têm de esperar apenas um nadinha mais do que um segundo entre o sinal do astronauta das 10.59.58 e o das 10.59.59, mas teriam de esperar para sempre pelo sinal das 11 horas. As ondas de luz emitidas da superfície da estrela entre as 10.59.59 e as 11 horas, segundo o relógio do astronauta, distribuem-se num período de tempo infinito, no caso da observação feita a partir da nave espacial. O intervalo de tempo entre a chegada de ondas sucessivas à nave é cada vez mais longo, de modo que a luz

da estrela parece cada vez mais vermelha e mais fraca. A certa altura, a estrela está tão apagada que já não pode ser vista da nave: a única coisa que resta é um buraco negro no espaço. A estrela, contudo, continua a exercer a mesma força gravitacional sobre a nave em órbita.

Este cenário não é inteiramente realista: a gravidade é tanto mais fraca quanto mais longe se está da estrela, de modo que a força gravitacional exercida nos pés do astronauta é sempre maior do que a força exercida na sua cabeça. Esta diferença de forças estica o astronauta como se fosse esparguete e destroça-o antes de se atingir o raio crítico e o horizonte de acontecimentos se formar!

Acreditamos que existem no universo objetos muito maiores, como as regiões centrais das galáxias, que podem sofrer colapso gravitacional para produzir buracos negros, mas, nestas circunstâncias, o astronauta não seria desfeito antes da formação do buraco negro⁴. Não sentiria mesmo nada de especial quando alcançasse o raio crítico e podia passar o ponto sem regresso sem dar por ele. Porém, dentro de muito poucas horas, a diferença das forças gravitacionais na cabeça e nos pés do astronauta seria tão grande que ele seria desfeito.

O trabalho que Roger Penrose e eu realizámos entre 1965 e 1970 mostrou que, segundo a relatividade geral, deve haver uma singularidade, uma densidade e uma curvatura infinitas no espaço-tempo interior de um buraco negro. É uma situação muito parecida com o *big bang* do começo dos tempos, mas aqui é o fim do tempo para o corpo em colapso e para o astronauta. Com esta singularidade, as leis da física e a nossa capacidade de prever o futuro são invalidadas. Todavia, qualquer observador exterior ao buraco negro não é afetado por esta falha de prognóstico, porque nem a luz nem qualquer outro sinal proveniente da singularidade podem chegar a ele. Este facto notável levou Penrose a propor a hipótese da censura cósmica, que podia ser parafraseada como «Deus detesta uma singularidade nua». Por outras palavras, as singularidades produzidas por colapso gravitacional ocorrem apenas em regiões como os buracos negros, onde ficam convenientemente escondidas de vistas exteriores por um horizonte de acontecimentos. Isto é censura cósmica fraca: protege os observadores exteriores ao buraco negro das consequências da falha de prognóstico que ocorre na singularidade, mas nada faz pelo infeliz astronauta que cai no buraco.

Existem algumas soluções das equações da relatividade geral em que o nosso astronauta pode ver uma singularidade nua: pode evitar atingi-la se cair através de um «buraco de verme» e sair noutra região do universo. Isto ofereceria grandes possibilidades para viajar no espaço e no tempo, mas, infelizmente, parece que estas soluções são altamente instáveis; a menor perturbação, como a presença de um astronauta, pode causar tais modificações que este só conseguiria ver a singularidade ao atingi-la e chegaria ao fim dos seus dias. Por outras palavras, a singularidade estaria sempre no seu futuro, e nunca no seu passado. A versão forte da hipótese de censura cósmica afirma que, numa solução realista, as singularidades estão

sempre no futuro (como as singularidades decorrentes de um colapso gravitacional) ou inteiramente no passado (como o *big bang*). Acredito sinceramente na censura cósmica, pelo que apostei com Kip Thorne e John Preskill, do Caltech, que ela seria sempre tecnicamente válida. Perdi a aposta porque foram encontrados exemplos de soluções com singularidades visíveis a grandes distâncias. Tive assim de pagar, o que, nos termos da aposta, significava que tinha de vestir a sua nudez. Posso, no entanto, gabar-me de uma vitória moral. As singularidades nuas eram instáveis: a mais pequena perturbação fá-las-ia desaparecer ou ocultar-se atrás de um horizonte de acontecimentos. Deste modo, não poderiam ocorrer em situações reais.

O horizonte de acontecimentos, a fronteira da região do espaço-tempo da qual não é possível escapar, atua como uma membrana de sentido único em redor do buraco negro: os objetos, como os astronautas desprevenidos, podem cair através do horizonte de acontecimentos para dentro do buraco negro, mas nada pode sair dele através do horizonte de acontecimentos (convém lembrar que o horizonte de acontecimentos é o caminho que a luz segue no espaço-tempo ao tentar escapar do buraco negro e que nada se propaga mais depressa do que a luz). Podia perfeitamente dizer-se do horizonte de acontecimentos o mesmo que o poeta Dante disse na porta do Inferno: «Deixai toda a esperança, vós que entraís.» Qualquer coisa ou pessoa que caia através do horizonte de acontecimentos depressa alcançará a região de densidade infinita e o fim do tempo.

A relatividade geral prevê que objetos pesados em movimento causarão a emissão de ondas gravitacionais: são ondas de curvatura do espaço que se propagam à velocidade da luz. Embora semelhantes às ondas luminosas, que são ondas do campo eletromagnético, são muito mais difíceis de detetar. Podem ser, no entanto, observadas em virtude das perturbações ligeiras que produzem na distância entre corpos vizinhos que se movam livremente. Nos Estados Unidos, na Europa e no Japão estão em construção detetores que poderão medir deslocamentos de uma parte em 1000 milhões de milhões de milhões (1 seguido de 21 zeros) ou de menos de um núcleo de um átomo a uma distância de 16 quilómetros.

Tal como a luz, as ondas gravitacionais transportam energia proveniente dos objetos que as emitem. Esperar-se-ia, portanto, que um sistema de corpos maciços atingisse eventualmente um estado estacionário, porque a energia do movimento seria consumida na emissão de ondas gravitacionais. (Um fenómeno semelhante ocorre quando se deixa cair uma rolha de cortiça na água: primeiro anda para baixo e para cima, mas, à medida que a ondulação lhe retira energia, acaba por se restabelecer o equilíbrio.) Por exemplo, o movimento da Terra na sua órbita em redor do Sol produz ondas gravitacionais. O efeito da perda de energia será o de mudar a órbita da Terra, de modo que esta se aproxime gradualmente do Sol, colidindo eventualmente com ele, atingindo-se então um estado estacionário. A taxa de energia perdida, no caso da Terra e do Sol, é muito pequena: cerca do suficiente para fazer funcionar um pequeno aquecedor elétrico. Isto significa que serão precisos 1000 milhões de milhões de milhões de milhões de anos para a Terra

colidir com o Sol, pelo que não há motivo imediato para preocupação! A perturbação da órbita da Terra é demasiado pequena para poder ser observada, mas este mesmo efeito tem sido observado nos últimos anos no sistema chamado PSR 1913+16 (PSR é a sigla de «pulsar», um tipo especial de estrela de neutrões que emite ondas de rádio). Este sistema contém duas estrelas de neutrões que orbitam em torno uma da outra, e a energia que estão a perder pela emissão de ondas gravitacionais faz com que tenham um movimento espiralado na direção uma da outra. J. H. Taylor e R. A. Hulse ganharam o Prémio Nobel em 1993 por esta confirmação da relatividade geral. Serão necessários 10.000 anos para que a colisão ocorra. Antes disso, orbitarão tão rapidamente que emitirão ondas gravitacionais suficientes para que detetores como o LIGO possam captá-las.

Durante o colapso gravitacional de uma estrela para formar um buraco negro os movimentos seriam muito mais rápidos, pelo que a perda de energia seria muito elevada. Por conseguinte, não seria preciso muito tempo para atingir um estado de equilíbrio. Qual seria o aspeto desse estado final? Dependeria de todas as propriedades da estrela original: não só da massa e da velocidade de rotação, mas também das diferentes densidades das várias camadas da estrela e dos complicados movimentos de gases no seu interior? Se os buracos negros fossem tão variados como os objetos que entraram em colapso para lhes darem origem, então poderia ser muito difícil prever as características dos buracos negros em geral.

Em 1967, contudo, o estudo dos buracos negros foi revolucionado por Werner Israel, cientista canadiano (nascido em Berlim, criado na África do Sul e doutorado na Irlanda). Israel mostrou que, de acordo com a relatividade geral, os buracos negros sem rotação deviam ser muito simples: perfeitamente esféricos, o tamanho dependeria apenas da massa, e buracos negros com a mesma massa seriam idênticos. Podiam, de facto, ser descritos por uma solução particular das equações de Einstein conhecida desde 1917 e descoberta por Karl Schwarzschild logo após o advento da relatividade geral. Ao princípio, muitos, incluindo o próprio Israel, argumentaram que, uma vez que os buracos negros tinham de ser perfeitamente esféricos, um buraco negro apenas poderia formar-se a partir do colapso de um objeto perfeitamente esférico. Uma estrela real, que não é *perfeitamente* esférica, só poderia entrar em colapso para formar uma singularidade nua.

Havia, no entanto, uma interpretação diferente do resultado de Israel que foi advogada por Penrose e John Wheeler. Argumentavam que os movimentos rápidos inerentes ao colapso de uma estrela produziam ondas gravitacionais que tornavam a estrela cada vez mais esférica e, uma vez atingido o estado estacionário, ela seria exatamente esférica. Segundo esta opinião, qualquer estrela sem rotação, por mais complicada que fosse a sua forma e a sua estrutura interna, acabaria, depois do colapso gravitacional, num buraco negro perfeitamente esférico, cujo tamanho dependeria apenas da sua massa. Cálculos ulteriores confirmaram esta opinião, que depressa foi adotada.

O resultado de Israel arrumava o caso de buracos negros formados apenas a partir de corpos sem movimento de rotação. Em 1963, o neozelandês Roy Kerr investigou buracos negros em rotação. Estes buracos negros de Kerr rodam a velocidade constante, dependendo o seu tamanho e a sua forma apenas da sua massa e da sua velocidade de rotação. Se esta for nula, o buraco negro é perfeitamente esférico e a solução é idêntica à de Schwarzschild. Se não for nula, o buraco negro é bojudo perto do equador (tal como a Terra e o Sol, devido às suas rotações) e, quanto mais depressa rodar, mais bojudo se torna. Portanto, para generalizar o resultado de Israel de modo a incluir corpos com rotação, conjecturou-se que qualquer corpo em rotação que entrasse em colapso para formar um buraco negro acabaria eventualmente por atingir o estado estacionário descrito pela solução de Kerr.

Em 1970 um colega e meu aluno de investigação em Cambridge, Brandon Carter, deu o primeiro passo para provar esta conjectura. Mostrou que, desde que um buraco negro em rotação estacionário tivesse um eixo de simetria, como o de um pião, o seu tamanho e a sua forma dependeriam apenas da massa e da velocidade de rotação. Depois, em 1971, provei que qualquer buraco negro em rotação estacionário teria realmente esse eixo de simetria. Finalmente, em 1973, David Robinson, do Kings College, em Londres, usou os resultados de Carter e os meus para mostrar que a conjectura estava correta: um buraco negro com essas características tinha mesmo de ser a solução de Kerr. Portanto, depois do colapso gravitacional, um buraco negro devia fixar-se num estado em que podia rodar, mas não pulsar. Além disso, o tamanho e a forma dependeriam apenas da massa e da velocidade de rotação, e não da natureza do corpo que entrou em colapso para lhe dar origem. O resultado tornou-se conhecido pela máxima «um buraco negro não tem cabelo». Este teorema é de grande importância prática, por restringir muito os tipos possíveis de buracos negros. Podemos, pois, elaborar modelos de buracos negros e comparar as previsões dos modelos com as observações. Significa também que uma grande quantidade de informação sobre o corpo que entrou em colapso deve ter-se perdido quando se formou o buraco negro, porque depois apenas pode medir-se a massa do corpo e a sua velocidade de rotação. O significado disto será explicado no próximo capítulo.

Os buracos negros são apenas um exemplo de um pequeno número de casos na história da ciência em que uma teoria foi desenvolvida em pormenor recorrendo a um modelo matemático⁵ antes de haver qualquer evidência a partir de observações de que estava certa. Na realidade, este costumava ser o argumento principal dos que se opunham aos buracos negros: como podia acreditar-se em objetos de cuja existência a única prova eram cálculos baseados na duvidosa teoria da relatividade geral? Em 1963, porém, Maarten Schmidt, um astrónomo do Observatório de Palomar, na Califórnia, mediu o desvio para o vermelho de um objeto ténue parecido com uma estrela na direção da fonte de ondas de rádio designada por 3C273 (ou seja, a fonte n.º 273 do terceiro catálogo de fontes de rádio de Cambridge). Descobriu que era demasiado grande para ser causado por um campo gravitacional: se se tratasse de um desvio gravitacional para o vermelho, o objeto teria de ser tão maciço e de estar tão próximo de nós que perturbaria as órbitas dos planetas

do sistema solar. Isto sugeria que o desvio para o vermelho era então causado pela expansão do universo, o que, por sua vez, significava que o objeto estava a uma distância muito grande. E, para ser visível a uma distância tão grande, o objeto tinha de ser muito brilhante, ou, por outras palavras, tinha de emitir uma enorme quantidade de energia. O único mecanismo em que se conseguiu pensar ser capaz de produzir tão grandes quantidades de energia era o colapso gravitacional não apenas de uma estrela, mas de toda a região central de uma galáxia. Foram descobertos vários outros objetos «quase-estelares» ou quasares, todos com grandes desvios para o vermelho. Mas estão todos demasiado afastados e, portanto, são demasiado difíceis de observar, de modo a fornecerem provas conclusivas sobre a existência de buracos negros.

Em 1967 houve um novo encorajamento para a existência de buracos negros: uma estudante de investigação de Cambridge, Jocelyn Bell-Burnell, descobriu no céu objetos que emitiam pulsações regulares de ondas de rádio. Ao princípio, Bell e o seu orientador, Antony Hewish, pensaram que tinham estabelecido contato com uma civilização estranha na Galáxia! Na realidade, no seminário em que anunciaram a descoberta chamaram às primeiras quatro fontes encontradas LGM 1-4, sendo LGM as iniciais de *little green men*. Contudo, no fim, eles e todos os outros chegaram à conclusão, menos romântica, de que esses objetos, a que foi dado o nome de *pulsares*, eram, na realidade, estrelas de neutrões, que emitiam pulsos de ondas de rádio devido a uma complicada interação dos seus campos magnéticos com a matéria circundante. Foram notícias más para os escritores de aventuras espaciais, mas davam grande esperança ao pequeno número de quantos naquela época acreditavam nos buracos negros: era a primeira prova positiva da existência de estrelas de neutrões. Uma estrela de neutrões tem um raio de cerca de 20 quilómetros, apenas algumas vezes o raio crítico com que uma estrela se torna um buraco negro. Se uma estrela podia entrar em colapso até dimensões tão pequenas, não seria disparate esperar que outras estrelas pudessem entrar em colapso até um tamanho ainda mais pequeno e originar buracos negros.

Como podemos esperar detetar um buraco negro se pela própria definição não emite luz? É um bocado como andar à procura de um gato preto num depósito de carvão. Felizmente, há uma maneira. Como John Michell fez notar no seu trabalho pioneiro de 1783, um buraco negro continua a exercer força gravitacional sobre os objetos próximos. Os astrónomos têm observado muitos sistemas em que duas estrelas orbitam à volta uma da outra atraídas pela gravidade. Também foram observados sistemas em que há uma estrela visível em órbita de uma companheira invisível. Como é evidente, não pode concluir-se imediatamente que a companheira seja um buraco negro: pode ser apenas uma estrela demasiado fraca para ser vista. Contudo, alguns destes sistemas, como o chamado Cygnus X-1 (fig. 6.2), são também fontes intensas de raios X. A melhor explicação para este fenómeno consiste em supor que a matéria é aspirada da superfície da estrela visível. Ao mesmo tempo que cai em direção à companheira invisível desenvolve um movimento espiral (como a água a sair pelo ralo da banheira) e torna-se muito quente, emitindo raios X (fig.6.3). Para este mecanismo funcionar, o objeto invisível tem de ser muito pequeno, como uma anã branca, uma estrela de neutrões ou um buraco negro.

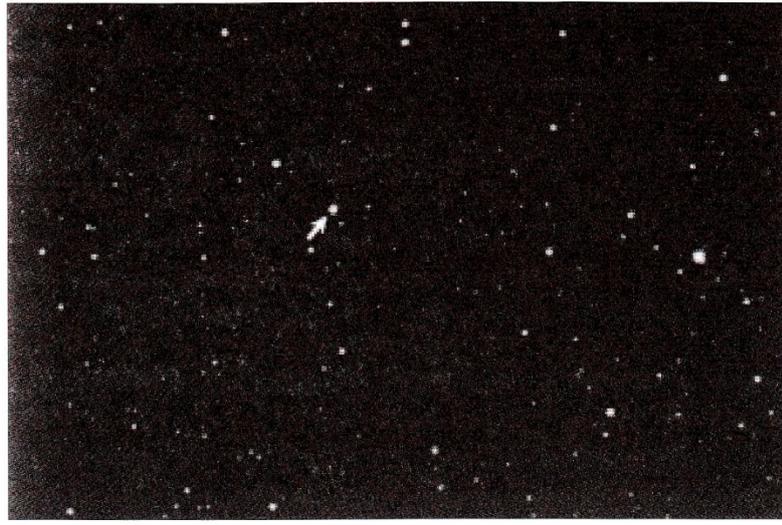


Figura 6.2 – A mais brilhante das duas estrelas perto do centro da fotografia é Cygnus X-1, que, segundo se pensa, consiste num buraco negro e numa estrela normal que orbitam em torno um do outro.

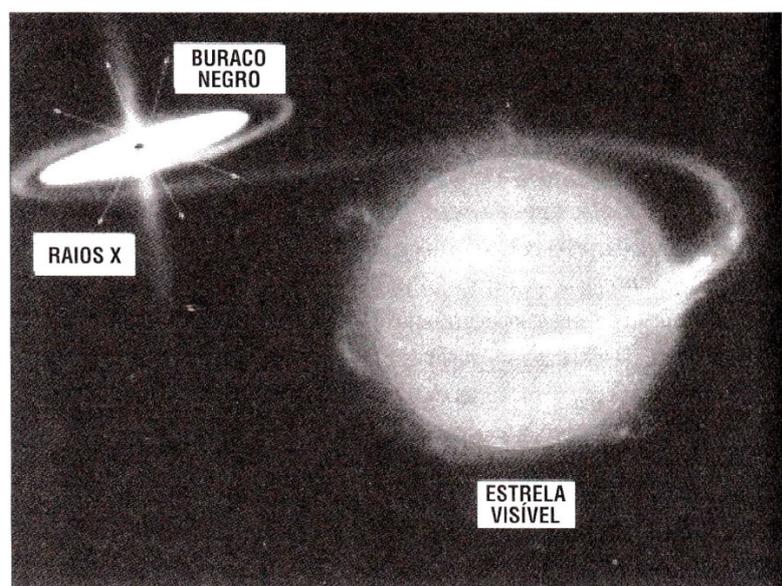


Figura 6.3

A partir da órbita observada da estrela visível podemos determinar a menor massa possível do objeto invisível. No caso de Cygnus X-1, esta é cerca de seis vezes a massa do Sol, o que, segundo o resultado de Chandrasekhar, é demasiado grande para que o objeto invisível seja uma anã branca. É também uma massa demasiadamente grande para ser uma estrela de neutrões. Parece, portanto, que deve ser um buraco negro.

Há outros modelos que explicam Cygnus X-1 e não incluem um buraco negro, mas são todos bastante rebuscados. Um buraco negro parece ser a única explicação realmente natural das observações. Apesar disso, apostei com Kip Thorne, do Caltech, que, na realidade, Cygnus X-1 não contém um buraco negro! Foi uma espécie de seguro que arranjei! Tenho trabalhado muito em buracos negros e todo esse trabalho estaria perdido se acabássemos por verificar que não existem. Mas nesse caso teria a consolação de ganhar a minha aposta, o que me daria uma assinatura da revista *Private Eye* durante quatro anos. De facto, embora a situação relativamente a Cygnus X-1 não se tenha modificado muito desde que fizemos a aposta, em 1975, há agora muitas outras evidências observacionais a favor da existência dos buracos negros, pelo que perdi a aposta. Paguei o estipulado, que era uma assinatura de um ano da revista *Penthouse*, para desgosto da mulher de Kip.

Temos também agora provas relativas a vários outros buracos negros em sistemas como Cygnus X-1 na Galáxia e em duas galáxias vizinhas chamadas nuvens de Magalhães. O número de buracos negros, contudo, é quase de certeza muito mais elevado; na longa história do universo muitas estrelas devem ter gasto todo o seu combustível nuclear e entraram em colapso. O número de buracos negros pode muito bem ser ainda maior do que o número de estrelas visíveis, que totalizam cerca de 100.000 milhões só na nossa galáxia. A atração gravitacional extra de um número tão grande de buracos negros podia explicar por que motivo a Galáxia roda à velocidade a que roda: a massa das estrelas visíveis é insuficiente para o justificar. Também temos algumas evidências de que existe um buraco negro muito maior, com uma massa de cerca de 100.000 vezes a do Sol, no centro da Galáxia. As estrelas da Galáxia que se aproximem demais desse buraco negro serão desfeitas pela diferença entre forças gravitacionais nos seus lados próximo e afastado. Os seus resíduos e o gás que é expedido de outras estrelas cairão em direção ao buraco negro. Como no caso de Cygnus X-1, o gás terá um movimento em espiral para o interior e aquecerá, embora não tanto como naquele caso. Não ficará suficientemente quente para emitir raios X, mas poderá justificar a fonte muito compacta de ondas de rádio e de raios infravermelhos que se observa no centro galáctico.

Pensa-se que buracos negros semelhantes, mas ainda maiores, com uma massa de cerca de 1000 milhões de vezes a massa do Sol, ocorram no centro dos quasares. Por exemplo, observações com o telescópio *Hubble* da galáxia conhecida como M87 mostraram que contém um disco de gás de 130 anos-luz de diâmetro orbitando um objeto central cuja massa é 2000 milhões de vezes a massa do Sol, o que não pode ser senão um buraco negro.

A matéria que caísse num buraco negro superdenso como aquele seria a única fonte capaz de explicar as quantidades enormes de energia que esses objetos emitem. À medida que a matéria entra em espiral no buraco negro, faz com que este rode no mesmo sentido, desenvolvendo um campo magnético muito parecido com o da Terra. Partículas de energia muito elevada seriam geradas perto do buraco negro pela matéria que cai dentro dele. O campo magnético seria tão forte que poderia forçar essas partículas a confluírem em

jetos dirigidos para fora, ao longo do eixo de rotação do buraco negro, ou seja, na direção dos seus polos norte e sul. Jetos como esses são, na realidade, observados em várias galáxias e quasares.

Podemos ainda considerar a possibilidade da existência de buracos negros com massas menores do que a do Sol. Tais buracos negros não podiam ser formados por colapsos gravitacionais porque as suas massas estão abaixo do limite de Chandrasekhar: estrelas com massa assim tão pequena podem aguentar-se contra a força da gravidade, mesmo quando já esgotaram o seu combustível nuclear. Buracos negros de massa reduzida só poderiam formar-se se a matéria fosse comprimida a densidades enormes, forçada por pressões exteriores muito intensas. Tais condições podiam ocorrer numa bomba de hidrogénio muito grande: o físico John Wheeler calculou que, se alguém juntasse toda a água pesada⁶ de todos os oceanos do mundo, podia construir uma bomba de hidrogénio capaz de comprimir a matéria de tal modo que seria criado um buraco negro. (É claro que não ficaria ninguém para observar o fenómeno!) Uma possibilidade mais viável é que esses buracos negros de massa reduzida possam ter sido formados nas altas temperaturas e pressões dos princípios do universo. Os buracos negros só teriam sido formados se o universo, no princípio, não tivesse sido perfeitamente liso e uniforme, porque uma pequena região mais densa do que a média podia ser comprimida desta maneira para formar um buraco negro. Mas sabemos que devia haver algumas irregularidades, porque, de outro modo, a matéria do universo continuaria a estar uniformemente distribuída na época atual, em vez de amontoada em estrelas e galáxias.

A questão de as irregularidades necessárias para justificar estrelas e galáxias terem levado à formação de um número significativo de buracos negros primevos está dependente de pormenores das condições iniciais do universo. Portanto, se pudéssemos determinar quantos buracos negros primevos existem atualmente, aprenderíamos muito sobre o princípio do universo. Buracos negros primevos com massas de mais de 1000 milhões de toneladas (a massa de uma grande montanha) só podiam ser detetados pela sua influência gravitacional sobre a matéria visível ou sobre a expansão do universo. No entanto, como veremos no próximo capítulo, os buracos negros, afinal, não são realmente negros: brilham como um corpo quente e, quanto mais pequenos são, mais brilham. Portanto, paradoxalmente, os buracos negros mais pequenos acabam por ser realmente mais fáceis de detetar do que os maiores!

¹ Houve, pelo menos, dez edições diferentes do *Exposition du système du monde*, publicadas entre 1796 e 1835. Nas primeiras edições Laplace apresentou o seu argumento, sem demonstração, algumas páginas antes do fim do livro v, capítulo 6. (N. do R. C.)

² Em rigor, os protões. Às temperaturas reinantes nos interiores das estrelas, os núcleos dissociam-se dos respetivos eletrões e as estruturas atómicas dissolvem-se em plasma. (N. do R. C.)

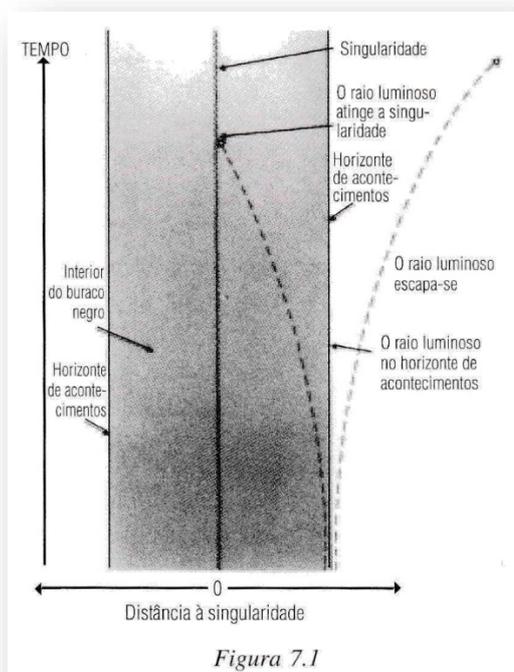
³ Eletrões, protões e neutrões são partículas que respeitam o princípio de exclusão. No interior de uma estrela podemos pensar nestas partículas como se de gases se tratasse. Uma estrela que culmina num estado de equilíbrio entre a gravidade e a pressão do gás de eletrões é uma anã branca. Uma estrela que culmina num estado em que o seu colapso é detido pela pressão das partículas nucleares é uma estrela de neutrões. (N. do R. C.)

- ⁴ Se o leitor fica intrigado com o facto de uma estrela se comprimir até dimensões absurdamente pequenas para formar um buraco negro, tem aqui um exemplo de como pode formar-se um buraco negro em condições bem mais «aceitáveis»: um núcleo galáctico em colapso origina um horizonte de acontecimentos antes de a sua densidade ser superior à densidade da água. (N. do R. C.)
- ⁵ A teoria dos buracos negros desenvolveu-se tanto e está de tal forma enraizada na relatividade geral que levou um certo cientista a concluir que «ou há buracos negros no espaço ou buracos na teoria da relatividade». (N. do R. C.)
- ⁶ Designação da água em cujas moléculas o hidrogénio surge na sua variedade isotópica designada por deutério, mais pesada do que o isótopo comum — o prótio — e muito mais rara. (N. do R. C.)

CAPÍTULO 7

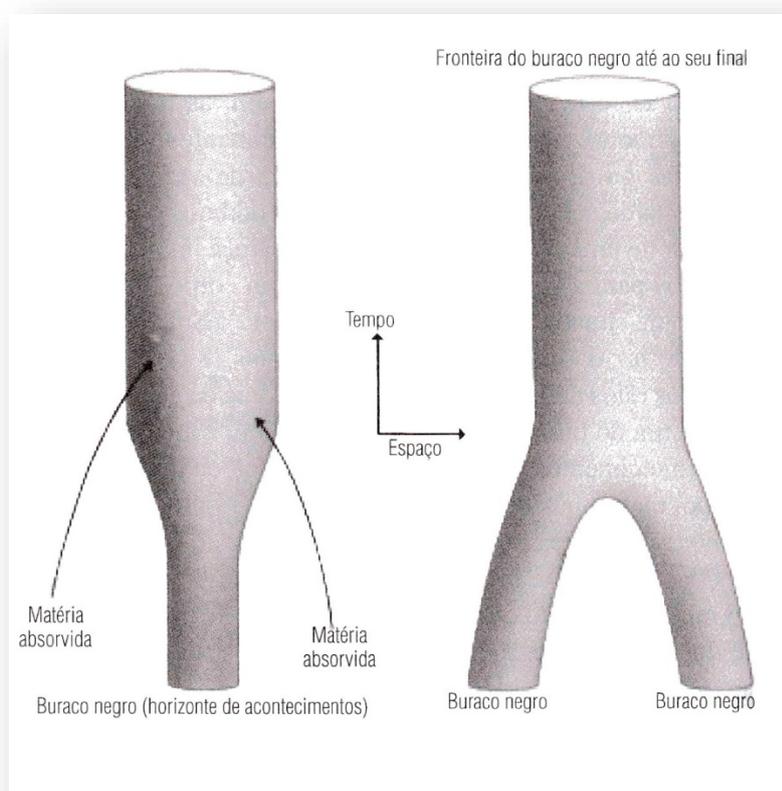
- OS BURACOS NEGROS NÃO SÃO ASSIM TÃO NEGROS -

Antes de 1970 as minhas investigações sobre a relatividade geral tinham-se concentrado sobretudo na questão de ter existido ou não uma singularidade inicial — o *big bang*. Todavia, numa noite de Novembro desse ano, logo a seguir ao nascimento da minha filha Lucy, comecei a pensar em buracos negros quando estava a meter-me na cama. O meu problema físico torna este processo bastante demorado, pelo que tive bastante tempo. Nessa altura não existia qualquer delimitação precisa dos pontos do espaço-tempo que ficavam dentro ou fora do buraco negro¹. Já tinha discutido com Roger Penrose a ideia de definir um buraco negro como o conjunto de acontecimentos dos quais não era possível escapar para o infinito, que é a definição agora geralmente aceite. Significa isto que a fronteira do buraco negro, o horizonte de acontecimentos, é formada pelas trajetórias no espaço-tempo de raios de luz que não escapam ao buraco negro, ficando para sempre na fronteira (fig. 7.1)² É um pouco como fugir da polícia e conseguir ir sempre um passo à frente sem nunca fugir completamente!



De repente compreendi que as trajetórias desses raios de luz nunca podiam aproximar-se umas das outras. Se isso acontecesse, encontrar-se-iam eventualmente, o que seria como encontrar alguém a correr fugindo da polícia no sentido contrário. Seriam ambos apanhados! (Ou, neste caso, cairiam num buraco negro.) Mas, se os raios de luz fossem engolidos pelo buraco negro, não podiam ter estado na sua fronteira. Assim, as trajetórias dos raios luminosos no horizonte de acontecimentos tinham de ser paralelas ou divergentes. Outra maneira de ver é imaginar o horizonte de acontecimentos, a fronteira do buraco negro, como o limiar de uma sombra: a sombra da desgraça iminente. Quando olhamos para a sombra projetada por um corpo iluminado por uma fonte situada a grande distância, como o Sol, vemos que os raios de luz da orla não se aproximam uns os outros.

Se os raios de luz que formam o horizonte de acontecimentos, a fronteira do buraco negro, não convergem, a área do horizonte de acontecimentos pode manter-se a mesma ou aumentar com o tempo, mas nunca diminuir — porque isso significaria que, pelo menos, alguns dos raios de luz do limiar teriam de aproximar-se uns dos outros. De facto, a área aumentaria sempre que matéria ou radiação fossem absorvidas pelo buraco negro (fig. 7.2). Ou, se dois buracos negros colidissem e se unissem para formarem um só, a área do horizonte de acontecimentos do buraco negro final seria maior ou igual à soma das áreas dos horizontes de acontecimentos dos buracos negros iniciais (fig. 7.3).



O facto de a área não se reduzir impõe uma restrição importante ao comportamento dos buracos negros. Fiquei tão excitado com a minha descoberta que quase não dormi nessa noite. No dia seguinte telefonei a Roger Penrose e ele concordou comigo. Acho que, na realidade, ele já tinha a noção desta propriedade da área. No entanto, estava a utilizar uma definição ligeiramente diferente de buraco negro. Não tinha compreendido que a fronteira do buraco negro, segundo as duas definições, seria a mesma, e portanto também as suas áreas, desde que o buraco negro estivesse num estado estacionário.

Este comportamento de um buraco negro era muito remanescente do comportamento da quantidade física, chamada entropia, que mede o grau de desordem de um sistema. É um dado da experiência comum que a desordem tem tendência para aumentar se as coisas forem deixadas entregues a si próprias. (Basta deixarmos de arranjar o que se estraga nas nossas casas para vermos que isso é verdade!) Podemos criar ordem a partir da desordem (por exemplo, podemos pintar a casa), mas isso requer dispêndio de esforço e, portanto, diminui a quantidade de energia ordenada disponível.

Uma asserção precisa desta ideia é conhecida por segunda lei da termodinâmica. Afirma que a entropia de um sistema isolado aumenta sempre e que, quando dois sistemas se unem, a entropia do sistema resultante é maior do que a soma das entropias dos sistemas individuais. Por exemplo, consideremos um sistema de moléculas de gás dentro de uma caixa. Podemos pensar nas moléculas como se fossem pequenas bolas de bilhar colidindo continuamente umas com as outras e ricocheteando nas paredes da caixa. Quanto mais alta for a temperatura do gás, mais depressa as moléculas se movem e, portanto, mais frequentemente e com mais força colidem com as paredes da caixa e maior é a pressão que exercem nestas. Suponhamos que inicialmente as moléculas estão todas confinadas no lado esquerdo da caixa por uma divisória. Se a divisória for retirada, as moléculas terão tendência para se espalhar e ocupar ambas as metades da caixa.

Em qualquer momento ulterior podiam estar, por acaso, todas na metade direita ou de novo na metade esquerda, mas é muitíssimo mais provável haver mais ou menos o mesmo número em ambas as metades. Tal estado é menos ordenado, ou mais desordenado, do que o estado original em que todas as moléculas estavam numa das metades. Diz-se, portanto, que a entropia do gás aumentou. Do mesmo modo, suponhamos que se começa com duas caixas, uma contendo moléculas de oxigénio e a outra contendo moléculas de azoto. Se se unirem as duas caixas, removendo a parede do meio, as moléculas de oxigénio e de azoto começarão a misturar-se. Mais tarde o estado mais provável será o de uma mistura razoavelmente uniforme de moléculas de oxigénio e de azoto pelas duas caixas. Este estado será menos ordenado e terá maior entropia do que o estado inicial das duas caixas separadas.

A segunda lei da termodinâmica tem um estatuto bastante diferente do das outras leis físicas, como a lei da gravitação de Newton, porque num número muito pequeno de casos pode não se aplicar. A probabilidade de todas

as moléculas de gás na nossa primeira caixa poderem ser encontradas mais tarde confinadas a metade da caixa é de um para muitos milhões de milhões, mas pode acontecer. Contudo, se tivermos um buraco negro «à mão», parece haver uma maneira bastante mais fácil de infringir a segunda lei: basta lançar alguma matéria com grande quantidade de entropia, como uma grande caixa de gás, para dentro dele. A entropia total da matéria no exterior do buraco negro decresceria. Podíamos dizer, é claro, que a entropia total, incluindo a do interior do buraco negro, não tinha diminuído, mas, como não há possibilidade de olhar para dentro dele, não podemos saber a quantidade de entropia que existe na matéria que está lá dentro. Seria bom que houvesse alguma característica do buraco negro através da qual observadores fora dele pudessem conhecer a sua entropia e quanto esta aumentaria sempre que matéria com entropia fosse absorvida. Seguindo a descoberta atrás descrita de que a área do horizonte de acontecimentos aumentava sempre que era absorvida matéria, um estudante de investigação em Princeton, Jacob Bekenstein, sugeriu que a área do horizonte de acontecimentos era uma medida da entropia do buraco negro: à medida que matéria com um certo conteúdo entrópico era absorvida pelo buraco negro, a área do horizonte aumentaria, de modo que a soma da entropia da matéria no exterior de um buraco negro e a área do seu horizonte nunca decresceriam.

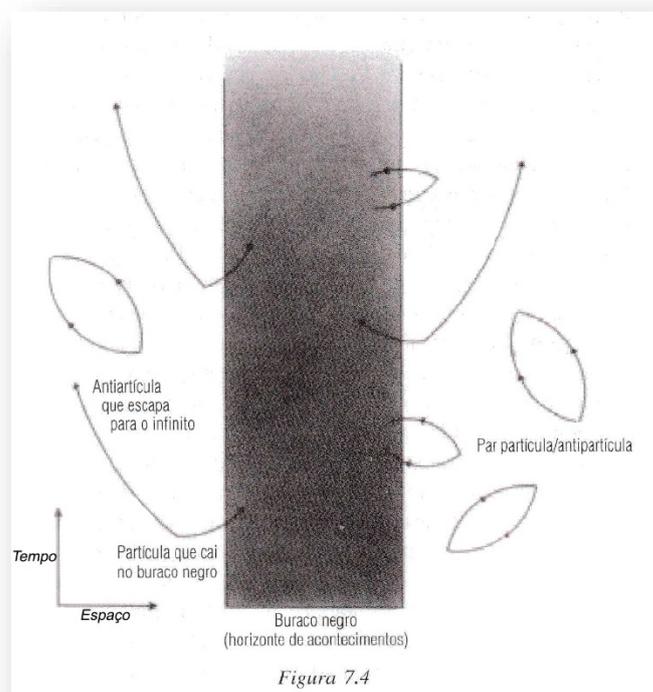
Esta sugestão parecia evitar que a segunda lei da termodinâmica fosse transgredida nestas situações. Havia, no entanto, uma falha fatal. Se um buraco negro tivesse entropia, também devia ter temperatura. Mas um corpo quente emite radiação. É um dado da experiência vulgar que, se aquecermos um atizador numa lareira, ele fica ao rubro e emite radiação, mas corpos a temperaturas mais baixas também emitem radiações; o que acontece é que normalmente não damos por isso, devido ao facto de a quantidade ser bastante pequena. Esta radiação é necessária para evitar a transgressão da segunda lei. Portanto, os buracos negros deviam emitir radiação. Mas, por definição, os buracos negros são objetos que se supõe não emitirem seja o que for. Parecia, pois, que a área do horizonte de um buraco negro não poderia ser considerada uma medida da sua entropia. Em 1972 apresentei um trabalho, em conjunto com Brandon Carter e um colega americano, Jim Bardeen, em que chamávamos a atenção para o facto de, apesar de haver muitas semelhanças entre a entropia e a área do horizonte, existir esta dificuldade aparentemente irresolúvel. Devo admitir que, ao escrever aquele artigo, fui motivado, em parte, pela irritação, pois achava que Bekenstein tinha utilizado mal a minha descoberta do aumento da área do horizonte. No entanto, acabou por se verificar que ele estava fundamentalmente certo, embora de uma maneira que por certo não esperava.

Em Setembro de 1973, durante uma visita a Moscovo, discuti buracos negros com dois especialistas soviéticos importantes, Yakov Zeldovich e Alexander Starobinsky. Convenceram-me de que, segundo o princípio da incerteza da mecânica quântica, buracos negros em rotação deviam criar e emitir partículas. Aceitei os argumentos no campo físico, mas não gostei da forma matemática como calculavam a emissão. Assim, dediquei-me a elaborar um tratamento matemático melhor, que descrevi num seminário informal em

Oxford, no fim de Novembro de 1973. Nessa altura não tinha feito os cálculos para encontrar a taxa de emissão. Estava à espera de descobrir apenas a radiação que Zeldovich e Starobinsky tinham previsto para buracos negros em rotação. Contudo, quando elaborei o cálculo, descobri, para minha surpresa e aborrecimento, que mesmo buracos negros sem rotação deviam aparentemente criar e emitir partículas a um ritmo constante. Primeiro pensei que essa emissão indicasse que um dos métodos que tinha utilizado não fosse válido. Receava que, se Bekenstein encontrasse esse resultado, o usasse como mais um argumento para apoiar as suas teorias da entropia dos buracos negros, das quais continuava a não gostar. Contudo, quanto mais pensava no assunto, mais parecia que realmente estava certo. Mas o que finalmente me convenceu de que a emissão era real foi o facto de o espectro das partículas emitidas ser exactamente o característico de um corpo quente e de o buraco negro estar a emitir partículas exactamente à potência correta de modo a evitar infrações à segunda lei. A partir daí, os cálculos foram repetidos de diversas formas por outras pessoas. Todos confirmam que um buraco negro deve emitir partículas e radiação como se fosse um corpo quente com uma temperatura que depende apenas da sua massa: quanto maior for a massa, mais baixa será a temperatura.

Como é possível que um buraco negro emita partículas quando sabemos que nada consegue escapar de dentro do seu horizonte? A resposta, segundo nos diz a teoria quântica, consiste em que as partículas não vêm de dentro do buraco negro, mas sim do espaço «vazio» contíguo ao horizonte de acontecimentos! Podemos compreender isto da seguinte maneira: aquilo que pensamos ser o espaço «vazio» não pode ser completamente vazio, porque tal significaria que todos os campos, por exemplo o gravitacional e o eletromagnético, tinham de ser exactamente zero. Contudo, o valor de um campo e a sua variação com o tempo assemelham-se à posição e à velocidade de uma partícula: o princípio da incerteza implica que, quanto maior for a precisão com que conhecemos uma destas quantidades, menor será aquela com que podemos conhecer a outra. Portanto, no espaço vazio, o campo não pode ser fixado exactamente em zero, porque então teria um valor preciso (zero) e uma variação precisa (também zero). Tem de existir uma quantidade mínima de incerteza, ou flutuação quântica, no valor do campo. Podemos pensar nessas flutuações como pares de partículas de luz ou de gravitação que surgem juntas em certo momento, se afastam e depois voltam a juntar-se, aniquilando-se uma à outra. Essas partículas são virtuais, como as que transmitem a força gravitacional do Sol: ao contrário das partículas reais, não podem ser observadas directamente com um detetor de partículas. No entanto, os seus efeitos indirectos, tais como variações na energia das órbitas dos electrões nos átomos, podem ser medidos e concordam com as previsões teóricas num grau de precisão notável. O princípio da incerteza também prevê que haverá pares virtuais semelhantes de partículas de matéria, como electrões ou quarks. Neste caso, contudo, um dos membros do par será uma partícula e o outro uma antipartícula (as antipartículas da luz e da gravidade identificam-se com as respectivas partículas).

Como a energia não pode ser criada a partir do nada, uma das companheiras do par partícula/antipartícula terá energia positiva e a outra energia negativa. A que tem energia negativa está condenada a ser uma partícula virtual de curta duração, porque as partículas reais têm sempre energia positiva em situações normais. Deve, pois, procurar a companheira e aniquilarem-se as duas. Contudo, uma partícula real que esteja perto de um corpo maciço tem menos energia do que se estivesse longe, porque é preciso energia para a afastar para longe, contra a atração gravitacional do corpo. Normalmente, a energia da partícula continua a ser positiva, mas o campo gravitacional dentro de um buraco negro é tão forte que mesmo aí uma partícula real pode ter energia negativa. É, portanto, possível, se estiver presente um buraco negro, que a partícula virtual com energia negativa caia dentro dele e se torne uma partícula ou antipartícula real. Neste caso, já não tem de aniquilar-se juntamente com a sua parceira. Esta, abandonada, pode cair também dentro do buraco negro. Ou, se tiver energia positiva, pode também fugir da vizinhança do buraco negro como uma partícula ou antipartícula real (fig. 7.4).



A um observador à distância parecerá ter sido emitida pelo buraco negro. Quanto mais pequeno este for, menor será a distância que a partícula com energia negativa terá de percorrer antes de se tornar uma partícula real e, assim, maior será a potência da emissão e a temperatura aparente do buraco negro.

A energia positiva da radiação que sai seria equilibrada por um fluxo de partículas de energia negativa para dentro do buraco negro. Segundo a equação de Einstein, $E = mc^2$ (em que E é a energia, m a massa e c a velocidade da luz), a energia é proporcional à massa. Um fluxo de energia

negativa dentro do buraco negro reduz, portanto, a sua massa. À medida que o buraco negro perde massa, a área do seu horizonte de acontecimentos torna-se mais pequena, mas esta diminuição na entropia do buraco negro é mais do que compensada pela entropia da radiação emitida, de maneira que a segunda lei nunca é violada.

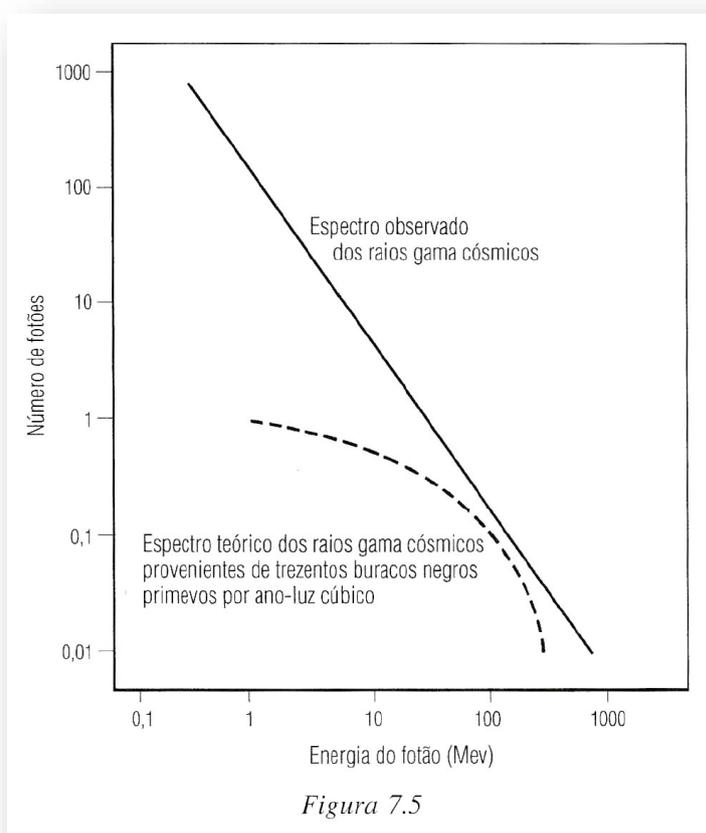
Além disso, quanto menor for a massa do buraco negro, mais alta será a sua temperatura. Portanto, à medida que o buraco negro perde massa, a sua temperatura e a taxa de emissão aumentam, de modo que perde massa mais depressa. Quando a massa do buraco negro eventualmente se torna extremamente pequena, o que acontece não é bem claro, mas a hipótese mais razoável é que desaparecerá completamente numa tremenda explosão final de emissão, equivalente à explosão de milhões de bombas H.

Um buraco negro com uma massa de algumas vezes a massa do Sol teria uma temperatura de apenas dez milionésimos de um grau acima do zero absoluto. Esta temperatura é muito mais baixa do que a da radiação de micro-ondas que enche o universo (cerca de $2,7^{\circ}$ acima do zero absoluto), de modo que estes buracos emitiriam ainda menos do que absorvem. Se o universo estiver destinado a expandir-se para sempre, a temperatura da radiação de micro-ondas diminuirá eventualmente até menos do que a de um destes buracos negros, que começará então a perder massa. Mas, mesmo então, a sua temperatura seria tão baixa que seriam precisos cerca de 1 milhão de milhões de anos (1 seguido de 66 zeros) para se evaporar completamente. Isto é muito mais tempo do que a idade do universo, que é apenas de cerca de 10.000 ou 20.000 milhões de anos (1 ou 2 seguido de 10 zeros). Por outro lado, conforme mencionei no capítulo 6, pode haver buracos negros primevos com uma massa muito mais pequena, formados pelo colapso de irregularidades nos primórdios do universo. Esses buracos negros teriam uma temperatura muito mais alta e estariam a emitir radiação a uma taxa muito maior. Um buraco negro primevo com uma massa inicial de 1000 milhões de toneladas teria um tempo de vida mais ou menos igual à idade do universo. Buracos negros primevos com massas iniciais menores do que este número já se teriam evaporado completamente, mas os que tivessem massas ligeiramente superiores continuariam ainda a emitir radiação na forma de raios X e raios gama. Os raios X e os raios gama são semelhantes a ondas de luz, mas com comprimentos de onda muito mais curtos. Esses buracos não chegam a merecer o epíteto de *negros*: na realidade, são *rubro-brancos* e emitem energia à potência de 10.000 megawatts.

Um buraco negro destes podia fazer funcionar dez grandes centrais geradoras se conseguíssemos domesticar a sua energia. Contudo, isto seria muito difícil: o buraco negro teria a massa de uma montanha comprimida num espaço de dimensão inferior a três milionésimos de milionésimos de centímetro, o tamanho do núcleo de um átomo! Se houvesse um buraco negro destes na superfície da Terra, não haveria maneira de o impedir de cair até ao centro da Terra. Oscilaria através da Terra até eventualmente se instalar no centro. Portanto, o único sítio para colocar um buraco negro desta natureza,

em que poderia utilizar-se a energia por ele emitida, seria em órbita ao redor da Terra e a única maneira de conseguir que orbitasse à volta da Terra seria atraindo-o para aí, arrastando uma enorme massa à sua frente, como uma cenoura à frente de um burro. Não parece uma coisa muito prática, pelo menos no futuro imediato.

Mas, mesmo que não possamos aprisionar um destes buracos negros primevos, quais são as nossas hipóteses de os observarmos? Podíamos procurar os raios gama, que emitem durante a maior parte da sua vida. Embora a radiação da maior parte deles fosse muito fraca, por estarem longe, o total de todos eles podia ser detetável. Conseguimos observar essa radiação de fundo de raios gama: a fig. 7.5 mostra como a intensidade observada difere da intensidade teórica para frequências diferentes (número de ondas por segundo). Porém, essa radiação de fundo podia ter sido, e provavelmente foi, gerada por outros processos diversos.



A linha tracejada na fig. 7.5 mostra como a intensidade deve variar com a frequência dos raios gama emitidos pelos buracos negros primevos, se estes fossem, em média, 300 por ano-luz cúbico. Podemos, portanto, dizer que as observações do ruído de fundo de raios gama não fornecem qualquer prova positiva da existência de buracos negros primevos, mas dizem-nos, isso sim, que, em média, não pode haver mais de 300 por ano-luz cúbico do universo. Este limite significa que os buracos negros primevos podem constituir, no máximo, um milionésimo da matéria do Universo.

Com esta escassez de buracos negros primevos podia parecer inverosímil que existisse um suficientemente perto para podermos observá-lo como fonte individual de raios gama. Mas, uma vez que a gravidade atrairia os buracos negros primevos em direção à matéria, deviam ser muito mais frequentes dentro e em volta das galáxias. Portanto, embora o ruído de fundo de raios gama nos diga que não podem existir mais de 300 buracos negros primevos, em média, por ano-luz cúbico, nada nos diz sobre até que ponto podem ser vulgares na nossa galáxia. Se fossem, digamos, um milhão de vezes mais frequentes, o buraco negro mais próximo de nós estaria, provavelmente, a uma distância de cerca de 1000 milhões de quilômetros, ou seja, mais ou menos tão longe quanto Plutão, o planeta mais afastado. A essa distância seria ainda muito difícil detetar a emissão constante de um buraco negro, mesmo que fosse de 10.000 megawatts. Para poder observar um buraco negro primevo seria preciso detetar vários *quanta* de raios gama vindos da mesma direção num intervalo de tempo razoável, como uma semana. De outra maneira, podiam simplesmente ser parte da radiação de fundo. Mas o princípio quântico de Planck diz-nos que cada *quantum* de raios gama tem energia muito elevada, porque os raios gama têm uma frequência muito alta, de tal modo que não seriam precisos muitos *quanta* para radiarem mesmo 10.000 megawatts. E para os observarmos, vindos da distância de Plutão, seria preciso um detetor de raios gama maior do que qualquer um construído até agora. Além disso, o detetor teria de estar no espaço porque os raios gama não conseguem penetrar na atmosfera.

É evidente que, se um buraco negro tão perto como Plutão chegasse ao fim da sua vida e explodisse, seria fácil detetar o rebentamento final de emissão. Mas, se o buraco negro esteve a emitir durante os últimos 10.000 ou 10.000 milhões de anos, a hipótese de chegar ao fim da sua vida dentro dos próximos anos é realmente bastante remota! Portanto, para haver uma hipótese razoável de ver uma explosão antes de se acabar o dinheiro da bolsa de estudos seria preciso descobrir uma maneira de detetar quaisquer explosões a uma distância de cerca de um ano-luz. De facto, têm sido detetadas explosões de raios gama no espaço através de satélites originalmente construídos para observar as violações do Tratado de Proibição de Testes. Parece que isto ocorre cerca de 16 vezes por mês, com uma distribuição uniforme em todas as direções do céu. Isto indica que vêm do exterior do sistema solar, visto que de outra forma se esperaria que se concentrassem na direção do plano das órbitas dos planetas. A distribuição uniforme também indica que as fontes estão razoavelmente próximas de nós na Galáxia, ou mesmo fora dela, a distâncias cosmológicas, porque, de outra maneira, se concentrariam de novo na direção do plano da Galáxia. Neste caso, a energia necessária para as explosões seria demasiado elevada para ter sido produzida por pequenos buracos negros, mas, se as suas fontes estivessem perto em termos galáticos, seria possível que os buracos negros estivessem a explodir. Gostaria muito de que fosse este o caso, mas tenho de reconhecer que há outras explicações possíveis para as explosões de raios gama, como a colisão de estrelas de neutrões. Novas observações nos próximos anos, em particular por meio de detetores de ondas gravitacionais,

como o LIGO, habilitar-nos-ão a descobrir a origem das explosões de raios gama.

Mesmo que a caça aos buracos negros primevos fracassasse, como parece provável, dar-nos-á informações importantes sobre as fases iniciais do universo. Se este tivesse sido caótico ou irregular, ou se a pressão de matéria tivesse sido baixa, então ter-se-iam produzido muitos mais buracos negros primevos do que o limite já determinado pelas nossas observações da radiação gama de fundo. Só se explica a ausência de buracos negros primevos observáveis se o universo primordial tiver sido muito regular e uniforme, com alta pressão de matéria.

A ideia da radiação proveniente de buracos negros é o primeiro exemplo de uma previsão que depende essencialmente das duas grandes teorias deste século: a relatividade geral e a mecânica quântica. Ao princípio suscitou muita oposição, porque vinha perturbar pontos de vista existentes. «Como pode um buraco negro emitir qualquer coisa?» Quando anunciei os resultados dos meus cálculos numa conferência no Laboratório Rutherford-Appleton, perto de Oxford, fui recebido com incredulidade geral. No fim da minha conferência, o presidente da sessão, John G. Taylor, do Kings College, de Londres, afirmou que era tudo um chorrilho de disparates. Até escreveu umas considerações nesse sentido. No entanto, no fim, a maior parte das pessoas, incluindo John Taylor, acabaram por chegar à conclusão de que os buracos negros deviam radiar, tal como os corpos quentes, se as nossas concepções da relatividade geral e da mecânica quântica estiverem certas. Assim, mesmo sem termos ainda conseguido encontrar um buraco negro primevo, há um consenso razoável de que, se o encontrarmos, terá de estar a emitir uma grande porção de raios X e gama.

A existência de radiação proveniente de buracos negros parece implicar que o colapso gravitacional não é tão final e irreversível como chegámos a pensar. Se um astronauta cair num buraco negro, a massa deste aumentará, mas, eventualmente, a energia equivalente a essa massa extra será devolvida ao universo sob a forma de radiação. Assim, num certo sentido, o astronauta será «reciclado». Seria, contudo, uma pobre espécie de imortalidade, porque qualquer conceito pessoal de tempo para o astronauta chegaria, com certeza, ao fim quando fosse feito em pedaços dentro do buraco negro! Mesmo o tipo de partículas eventualmente emitidas pelo buraco negro seria em geral diferente das que constituíam o astronauta: a única característica do astronauta que sobreviveria seria a sua massa ou a sua energia.

As aproximações que usei para obter a emissão dos buracos negros funcionam bem quando o buraco negro tem uma massa maior do que uma fração de grama. Contudo, falham no fim da vida do buraco negro, quando a sua massa se torna muito pequena. O resultado mais provável parece ser o de o buraco negro acabar por desaparecer, pelo menos da nossa região do universo, levando consigo o astronauta e qualquer singularidade que possa estar lá dentro, se realmente existir alguma. Esta foi a primeira indicação de que a mecânica quântica podia retirar as singularidades previstas pela

relatividade geral. Todavia, os métodos que eu e outros usámos em 1974 não se destinavam a dar resposta a questões como a de haver singularidades em gravidade quântica. A partir de 1975 comecei então a desenvolver uma nova aproximação da gravidade quântica, baseando-me na ideia de Richard Feynman da soma de histórias. As respostas que esta aproximação sugere para a origem e destino último do universo e do seu conteúdo, como os astronautas, serão descritas nos dois capítulos seguintes. Veremos que, embora o princípio da incerteza limite a precisão de todas as nossas previsões, pode, ao mesmo tempo, remover a imprevisibilidade fundamental que ocorre numa singularidade do espaço-tempo.

¹ O problema era mais intrincado com os buracos negros com rotação. Podemos imaginar que, neste caso, o espaço roda à medida que flui para o interior do buraco negro. Há uma região em que o espaço flui com velocidade superior à da luz, embora a componente centrípeta desta velocidade geométrica seja inferior à velocidade da luz. (N. do R. C.)

² Este efeito de dragagem da luz pode ser entendido como se o espaço fluísse para o interior do buraco negro arrastando consigo a luz. No horizonte a luz está a ser arrastada precisamente à velocidade da luz. (N. do R. C.)

CAPÍTULO 8

- A ORIGEM E O DESTINO DO UNIVERSO -

A teoria da relatividade geral de Einstein previa que o espaço-tempo começara com a singularidade do *big bang* e terminaria com a singularidade do *big crunch* (isto se o universo viesse a entrar em colapso) ou numa singularidade no interior de um buraco negro (se uma estrela próxima entrasse em colapso). A matéria que fluísse para o buraco seria destruída na singularidade e apenas o efeito gravitacional da massa continuaria a ser sentido fora dele. Por outro lado, quando os efeitos quânticos foram tidos em conta, parecia que a massa, ou a energia da matéria, acabaria por ser eventualmente devolvida ao resto do universo e que o buraco negro, conjuntamente com a singularidade no seu interior, se evaporaria e, finalmente, desapareceria. Poderia a mecânica quântica ter um efeito igualmente dramático sobre as singularidades do *big bang* e do *big crunch*? Que acontece realmente durante os primeiros ou os últimos estádios do universo, quando os campos gravitacionais são tão intensos que os efeitos quânticos não podem ser ignorados? O universo tem, de facto, um princípio ou um fim? E, se assim é, como são?

Durante os anos 70 dediquei-me ao estudo dos buracos negros, mas em 1981 o meu interesse pelas questões da origem e do destino do universo reacendeu-se quando assisti a uma conferência organizada pelos Jesuítas no Vaticano. A igreja católica tinha cometido um grave erro com Galileu, quando tentou impor a lei numa questão científica, declarando que o Sol girava à volta da Terra. Agora, séculos volvidos, decidira convidar alguns especialistas para a aconselharem sobre cosmologia. No fim da conferência os participantes foram recebidos em audiência pelo papa, que lhes disse que estava certo estudarem a evolução do universo desde o *big bang*¹, mas que não deviam inquirir acerca da natureza do *big bang*, porque tinha sido o momento da criação e, portanto, trabalho de Deus. Nessa altura fiquei satisfeito por ele ignorar qual havia sido a minha contribuição para a conferência: a possibilidade de o espaço-tempo ser finito, mas ilimitado, o que significaria que não tinha tido um princípio e que não havia qualquer momento de criação. Não tinha qualquer desejo de partilhar a sorte de Galileu, com quem me sinto fortemente identificado, em parte devido à coincidência de ter nascido exatamente trezentos anos depois da sua morte!

Para explicar as ideias que eu e outros temos sobre a maneira como a mecânica quântica pode afetar a origem e o destino do universo é necessário compreender primeiro a história do universo geralmente aceite, segundo o que

é conhecido por «modelo quente do *big bang*». Este presume que o universo é descrito por um modelo de Friedmann, mesmo até ao *big bang*. Em modelos deste género descobre-se que, à medida que o universo se expande, a matéria e a radiação arrefecem. (Quando o universo duplica de tamanho, a temperatura reduz-se a metade.)² Uma vez que a temperatura é simplesmente uma medida da energia média, ou velocidade das partículas, este arrefecimento do universo teria um efeito importante sobre a matéria nele existente. A temperaturas muito elevadas, as partículas movem-se tão depressa que anulam o efeito da atração mútua devida a forças nucleares ou eletromagnéticas, mas, quando arrefecem, as partículas atraem-se e começam a juntar-se em amontoados. Além disso, mesmo os tipos de partículas que podem existir no universo dependem da temperatura. A temperaturas suficientemente elevadas, as partículas têm tanta energia que, sempre que colidem, originam muitos pares diferentes partícula/antipartícula e, embora alguns desses pares se aniquilem, a produção é maior do que a aniquilação. Contudo, a temperaturas mais baixas, quando as partículas colidem com menor energia, os pares partícula/antipartícula são produzidos a uma taxa menor e a aniquilação é maior do que a produção.

Pensa-se que, no momento do *big bang*, o tamanho do universo seria zero e, portanto, estaria infinitamente quente. Mas, à medida que se expandiu, a temperatura da radiação baixou. Um segundo após o *big bang* teria baixado para cerca de 10.000 milhões de graus, o que é cerca de mil vezes a temperatura no centro do Sol; temperaturas tão elevadas como estas são alcançadas em explosões de bombas H. Nesse tempo o universo teria contido sobretudo fótons, eletrões e neutrinos (partículas extremamente leves que só são afetadas pela força fraca e pela gravidade) e as suas antipartículas juntamente com alguns prótons e neutrões. À medida que o universo continuou a expandir-se e a temperatura a baixar, a razão de produção de pares eletrão/positrão em colisões deve ter descido abaixo daquela a que eram destruídos pela aniquilação. Portanto, a maior parte dos eletrões e positrões ter-se-iam destruído uns aos outros, produzindo mais fótons e deixando apenas um remanescente de eletrões. No entanto, os neutrinos e os antineutrinos não se teriam aniquilado uns aos outros porque estas partículas interagem muito fracamente entre si e com outras partículas. Por conseguinte, ainda devem andar por aí. Se pudéssemos observá-las, teríamos uma boa experiência da imagem de um estado muito quente do princípio do universo. Infelizmente, as suas energias estariam agora demasiado baixas para podermos observá-las diretamente. Contudo, se os neutrinos tiverem massa, nem que seja uma pequena massa própria, como foi sugerido por algumas experiências recentes, talvez possamos detetá-los indiretamente: podem ser uma forma de «matéria escura», como a que atrás mencionei, com atração gravitacional suficiente para suspender a expansão do universo e obrigá-lo a entrar em colapso.

Cerca de 100 segundos após o *big bang*, a temperatura deve ter baixado para 1000 milhões de graus, que é a temperatura no interior das estrelas mais quentes. A esta temperatura, os prótons e os neutrões já não teriam energia suficiente para escapar à ação da força nuclear forte e teriam começado a



combinar-se uns com os outros para produzirem os núcleos de átomos de deutério (hidrogénio pesado), que contém um protão e um neutrão. Os núcleos de deutério ter-se-iam combinado com mais protões e neutrões para formarem núcleos de hélio, que contém dois protões e dois neutrões, além de pequenas quantidades de dois elementos mais pesados, o lítio e o berílio. Pode imaginar-se que, no modelo quente do *big bang*, cerca de um quarto dos protões e dos neutrões se teriam convertido em núcleos de hélio juntamente com uma pequena quantidade de hidrogénio pesado e outros elementos. Os neutrões restantes decaíam em protões, que são os núcleos dos átomos de hidrogénio vulgares.

Esta imagem de um estado quente do universo primitivo foi sugerida pela primeira vez pelo cientista George Gamow num famoso trabalho apresentado em 1948, juntamente com um aluno seu, Ralph Alpher. Gamow tinha bastante sentido de humor e persuadiu o cientista nuclear Hans Bethe a acrescentar o seu nome ao trabalho para que a lista dos autores fosse «Alpher, Bethe e Gamow», como as três primeiras letras do alfabeto grego, alfa, beta e gama, particularmente apropriadas para um trabalho sobre o começo do universo! Neste trabalho apresentaram a notável previsão de que a radiação (sob a forma de fótons) proveniente dos primeiros tempos do universo devia andar por aí, mas com a temperatura reduzida a apenas uns graus acima do zero absoluto (-273°C). Foi esta radiação que Penzias e Wilson descobriram em 1965. Na altura em que Alpher, Bethe e Gamow escreveram o seu artigo pouco se sabia sobre as reações nucleares dos protões e dos neutrões. Os prognósticos elaborados para as proporções dos vários elementos no universo primevo eram, portanto, bastante imprecisos, mas os cálculos têm sido repetidos já com conhecimentos atualizados e agora concordam perfeitamente com a observação. Além do mais, é muito difícil explicar de outra maneira porque há tanto hélio no universo. Estamos, portanto, razoavelmente confiantes de termos encontrado a imagem certa, pelo menos até cerca de um segundo depois do *big bang*.

Poucas horas depois do *big bang*, a produção de hélio e dos outros elementos deve ter parado. E depois disso, durante o milhão de anos seguinte, mais ou menos, o universo deve ter continuado a expandir-se, sem ter acontecido nada de especial. Eventualmente, quando a temperatura desceu a alguns milhares de graus e os eletrões e os núcleos já não tinham energia suficiente para dominarem a atração eletromagnética entre si, devem ter começado a combinar-se para formarem átomos³. O universo, no seu conjunto, deve ter continuado a expandir-se e a arrefecer, mas, em regiões ligeiramente mais densas do que a média, a expansão deve ter sido mais lenta por causa da atração gravitacional extra. Este facto deve ter suspenso a expansão em algumas regiões e tê-las feito entrar em colapso. Enquanto sofriam colapso, a atração gravitacional da matéria fora destas regiões talvez a tivesse feito entrar em rotação ligeira. À medida que a região em colapso ficava mais pequena, rodava mais depressa, tal como patinadores que rodopiam no gelo com maior velocidade quando encolhem os braços. Eventualmente, quando a região ficou suficientemente pequena, deve ter girado suficientemente depressa para contrabalançar a atração da gravidade e,

desse modo, nasceram galáxias em rotação com a forma de disco. Outras regiões, que não começaram a rodar, originaram objetos de forma oval, chamados galáxias elípticas⁴. Nestas, o colapso deve ter sido sustido porque partes individuais da galáxia estariam em órbita estável em redor do seu centro, mas a galáxia como um todo não teria rotação.

À medida que o tempo foi passando, o hidrogénio e o hélio existentes nas galáxias ter-se-iam dividido em nuvens mais pequenas que entraram em colapso sob o efeito da própria gravidade. Enquanto estas se contraíram e os átomos colidiam uns com os outros no seu interior, a temperatura do gás aumentava, até acabar por se tornar suficientemente elevada para se iniciarem reações de fusão nuclear. Estas convertiam o hidrogénio em mais hélio e o calor emitido aumentava a pressão e evitava, assim, que as nuvens se contraíssem mais. Terão ficado estáveis durante muito tempo, como estrelas iguais ao nosso Sol, transformando hidrogénio em hélio e radiando a energia resultante na forma de calor e de luz. As estrelas mais maciças precisavam de ser mais quentes para poderem equilibrar a sua atração gravitacional mais forte, fazendo com que as reações de fusão nuclear se sucedessem de tal maneira mais depressa que gastavam o seu hidrogénio em apenas 100 milhões de anos. Então contraíam-se ligeiramente e, quando aqueciam mais, começavam a transformar hélio em elementos mais pesados, como o carbono ou o oxigénio. Isto, contudo, não libertava muito mais energia, de maneira que surgia uma crise, como foi descrito no capítulo referente aos buracos negros. O que acontecia depois não é perfeitamente claro, mas parece provável que as regiões centrais da estrela entrassem em colapso até um estado muito denso, como uma estrela de neutrões ou um buraco negro. As regiões exteriores da estrela podem, por vezes, ser destruídas numa tremenda explosão, chamada supernova, que brilha mais do que qualquer outra estrela na sua galáxia. Alguns dos elementos mais pesados, produzidos perto do fim da vida da estrela, seriam ejetados de novo para o gás da galáxia⁵, fornecendo matéria-prima para a nova geração de estrelas. O nosso próprio Sol contém cerca de 2% destes elementos mais pesados, porque é uma estrela de segunda ou terceira geração, formada há cerca de 5000 milhões de anos, a partir de uma nuvem de gás em rotação que continha os detritos de supernovas mais antigas. A maior parte do gás dessa nuvem formou o Sol ou foi «soprado» para o espaço exterior, mas uma pequena quantidade de elementos mais pesados juntou-se para formar os corpos que agora orbitam em torno do Sol, os planetas, entre os quais a Terra.

A Terra, ao princípio, estava muito quente e sem atmosfera. Com o tempo arrefeceu e adquiriu uma atmosfera a partir de gases emitidos pelas rochas. A atmosfera primeva não nos teria permitido sobreviver. Não continha oxigénio, mas sim uma quantidade de outros gases venenosos para nós, como o ácido sulfídrico (o gás que cheira a ovos podres). Há, no entanto, outras formas primitivas de vida que podem desenvolver-se perfeitamente em tais condições. Pensa-se que se desenvolveram nos oceanos, possivelmente como resultado de combinações casuais de átomos em grandes estruturas, chamadas macromoléculas, que seriam capazes de produzir estruturas semelhantes a partir de outros átomos no oceano. Ter-se-iam então reproduzido e

multiplicado. Em alguns casos terá havido erros na reprodução. Na sua maior parte, esses erros terão sido tais que a nova macromolécula não podia reproduzir-se e eventualmente terá acabado por ser destruída. Contudo, alguns dos erros terão produzido novas macromoléculas que se reproduziam ainda melhor. Apresentavam, portanto, vantagens e tinham tendência para substituir as macromoléculas originais. Deste modo se iniciou um processo de evolução que levou ao desenvolvimento de organismos autorreprodutores cada vez mais complicados. As primitivas formas de vida consumiam diversos materiais, incluindo o ácido sulfídrico, e libertavam oxigénio. Isto modificou gradualmente a atmosfera até à composição que hoje tem e permitiu o desenvolvimento de formas de vida mais complexas, como os peixes, os répteis, os mamíferos e, por fim, a raça humana.

Esta imagem de um universo que começou muito quente e arrefeceu enquanto se expandia está de acordo com todas as provas observacionais que hoje possuímos. No entanto, deixa um importante número de perguntas sem resposta:

1. Por que motivo era o universo tão quente no princípio?
2. Por que razão é o universo tão uniforme em macroescala? Por que parece o mesmo em todos os pontos do espaço e em todas as direções? Em particular, por que é a temperatura da radiação de fundo de micro-ondas tão semelhante quando olhamos em direções diferentes? Parecem perguntas de exame a um grupo de estudantes. Se todos dão a mesma resposta, podemos ter quase a certeza de que copiaram. No entanto, no modelo que acabámos de descrever não terá havido tempo para a luz ir de uma região distante até outra, depois do *big bang*, embora as regiões estivessem perto umas das outras no princípio do universo. Segundo a teoria da relatividade, se a luz não consegue ir de uma região a outra, então nenhuma outra informação o consegue. Assim, não havia hipótese de regiões diferentes no começo do universo terem a mesma temperatura, a não ser que, por qualquer razão inexplicável, tivessem tido a mesma temperatura inicial.
3. Por que se iniciou a expansão do universo com uma velocidade tão próxima da velocidade crítica, que separa a classe dos universos que entram em colapso da classe dos que se expandem indefinidamente, de tal maneira que mesmo hoje, decorridos 10 milhões de anos, continua a expandir-se aproximadamente à mesma velocidade? Se a velocidade de expansão, um segundo após o *big bang*, tivesse sido menor, nem que fosse uma parte em 100.000 milhões de milhões, o universo teria voltado a entrar em colapso antes de ter alcançado o seu tamanho atual.
4. Apesar de o universo ser tão uniforme e homogéneo em macroescala, contém irregularidades locais, como as estrelas e as galáxias. Pensa-se que estas se desenvolveram a partir de pequenas

diferenças na densidade do universo inicial de região para região. Qual a origem destas flutuações de densidade?

A teoria da relatividade geral, por si própria, não pode explicar estas características ou responder a estas perguntas por causa do seu prognóstico de que o universo começou com densidade infinita na singularidade do *big bang*. Na singularidade, a relatividade geral e todas as outras leis físicas perdem a validade: não é possível prever o que vai sair da singularidade. Como já foi explicado, isto significa que pode perfeitamente excluir-se a singularidade da teoria do *big bang*, bem como quaisquer acontecimentos anteriores, porque não podem ter efeito no que observamos. O espaço-tempo teria um limite: um começo no *big bang*.

A física parece ter descoberto um conjunto de leis que, dentro dos limites impostos pelo princípio da incerteza, nos dizem como o universo vai evoluir se soubermos em que estado está em determinado momento. Estas leis podem ter sido decretadas por Deus, mas, depois, parece que ele deixou o universo evoluir de acordo com elas, sem mais intervir. Mas como escolheu ele o estado ou configuração inicial do universo? Quais foram as «condições de fronteira» no começo do tempo?

Uma resposta possível é dizer que Deus escolheu a configuração inicial do universo por razões que nunca compreenderemos. Isso estaria certamente dentro dos poderes de um ser onipotente, mas, se o criou de uma maneira tão incompreensível, por que o deixou depois evoluir segundo leis inteligíveis? Toda a história da ciência tem sido a compreensão gradual de que os acontecimentos não ocorrem de maneira arbitrária, mas refletem uma certa ordem subjacente, que pode ou não ser de inspiração divina. Nada mais natural do que supor que essa ordem se aplica não apenas às leis, mas também às condições que na fronteira do espaço-tempo especificam o estado inicial do universo. Pode haver um grande número de universos que evoluíram a partir de condições iniciais diferentes, mas que obedecem às mesmas leis. Devia haver algum princípio que determinasse um estado inicial, e daí um modelo para representar o nosso universo.

Uma possibilidade consiste em admitir condições de fronteira caóticas. Estas presumem implicitamente que o universo é espacialmente infinito ou que há infinitamente muitos universos. Sob condições de fronteira caóticas, a probabilidade de descobrir uma região particular do espaço numa dada configuração logo após o *big bang* é, em certo sentido, igual à probabilidade de a descobrir em qualquer outra configuração: o estado inicial do universo é escolhido puramente ao acaso. Isto significa que o universo, no princípio, era, provavelmente, muito irregular e caótico, por haver muito mais configurações caóticas e irregulares para o universo do que regulares e ordenadas. (Se cada configuração é igualmente provável, é admissível que o universo tenha tido início num estado desordenado e caótico simplesmente por haver mais soluções deste tipo.) É difícil ver como tais condições iniciais caóticas podem ter dado origem a um universo tão regular e sem acidentes, em macroescala, como é o nosso atualmente. Também se esperaria que as flutuações de

densidade em tal modelo tivessem levado à formação de mais buracos negros primários do que o limite máximo imposto por observações da radiação gama de fundo.

Se o universo é espacialmente infinito (ou se há infinitamente muitos universos), então há, provavelmente, muitas regiões (muitos universos) regulares e uniformes. É um pouco como o bem conhecido bando de macacos a escrever à máquina; a maior parte do que escrevem não presta, mas ocasionalmente, só por puro acaso, poderão escrever um soneto de Shakespeare. No caso do universo, será que estamos a viver numa região que, por mero acaso, é regular e uniforme? À primeira vista, isto pode parecer muito pouco provável, porque regiões assim regulares seriam largamente ultrapassadas em número por regiões irregulares e caóticas. Vamos, porém, supor que só nas regiões regulares se formaram galáxias e estrelas e as condições eram adequadas ao desenvolvimento de complicados organismos autorreplicativos, como nós, que fossem capazes de fazer a pergunta: por que motivo é o universo tão regular? Este é um exemplo da aplicação do chamado princípio antrópico, que pode ser parafraseado do seguinte modo: «Vemos o universo da maneira que é porque existimos.»

Há duas versões do princípio antrópico, a fraca e a forte. O princípio antrópico fraco afirma que, num universo muito grande ou infinito no espaço e/ou no tempo, as condições necessárias para o desenvolvimento de vida inteligente só se verificam em certas regiões limitadas no espaço e no tempo. Os seres inteligentes que habitam essas regiões não devem, portanto, admirar-se por observarem que, nas suas imediações, o universo satisfaz as condições necessárias à sua existência. É um pouco como uma pessoa muito rica que vive numa área de pessoas ricas e nunca vê a pobreza.

Um exemplo da utilização do princípio antrópico fraco consiste em explicar por que motivo o *big bang* ocorreu há cerca de 10.000 milhões de anos: porque é esse o tempo necessário para a evolução de seres inteligentes. Como já expliquei, começou por se formar uma primeira geração de estrelas. Essas estrelas converteram algum do hidrogénio e do hélio originais em elementos como o carbono e o oxigénio, dos quais somos formados. As estrelas então explodiram como supernovas e os seus detritos formaram outras estrelas e os planetas, entre eles os do nosso sistema solar, que tem cerca de 5000 milhões de anos de idade. Os primeiros 1000 ou 2000 milhões de anos de existência da Terra foram demasiadamente quentes para que qualquer coisa complicada pudesse desenvolver-se. Os restantes 3000 milhões de anos, ou coisa parecida, foram ocupados pelo processo lento da evolução biológica, começando nos organismos mais simples e originando seres capazes de medir o tempo para trás, até ao *big bang*.

Poucas pessoas argumentariam contra a validade ou utilidade do princípio antrópico fraco. Algumas, no entanto, vão muito mais além e propõem uma versão forte do princípio. De acordo com esta teoria, ou existem muitos universos diferentes ou muitas regiões diferentes num único universo, cada um com a sua configuração inicial e, talvez, com leis físicas próprias. Na maior

parte destes universos, as condições não seriam adequadas ao desenvolvimento de organismos complexos; só em poucos universos semelhantes ao nosso se desenvolveriam seres inteligentes que fariam a pergunta: «Por que é o universo como o vemos?» A resposta então é simples: se fosse diferente, não estaríamos aqui!

As leis da ciência, como as conhecemos atualmente, contêm muitos números fundamentais, como a magnitude da carga elétrica do elétron e a proporção entre as massas do próton e do elétron. Pelo menos de momento não podemos prever os valores destes números a partir da teoria; temos de medi-los por observação. Talvez um dia descubramos uma teoria unificada completa que preveja tudo isso, mas também é possível que alguns, ou todos eles, variem de universo para universo, ou dentro do mesmo universo. O que é notável é que os valores destes números parecem ter sido muito bem ajustados para tornarem possível o desenvolvimento da vida. Por exemplo, se a carga elétrica do elétron fosse apenas ligeiramente diferente, as estrelas ou seriam incapazes de queimar hidrogénio e hélio, ou então não teriam explodido. É claro que pode haver outras formas de vida inteligente, nem sequer sonhadas pelos escritores de ficção científica, que não precisem da luz de uma estrela, como o Sol, ou dos elementos químicos pesados, sintetizados nas estrelas e devolvidos ao espaço quando as estrelas explodem. Todavia, parece claro que há espectros relativamente estreitos de valores para os números que permitem o desenvolvimento da vida inteligente. Valores arbitrários originam universos que, embora pudessem ser muito bonitos, não conteriam ninguém que se maravilhasse com essa beleza. Podemos tomar isto como prova de um propósito divino na criação e na escolha das leis da natureza ou como suporte do princípio antrópico forte.

Várias objeções podem ser levantadas contra o princípio antrópico forte como explicação do estado observável do universo. Primeiro, em que sentido pode dizer-se que existem todos esses universos? Se estão realmente separados uns dos outros, o que acontece noutra universo não pode ter consequências observáveis no nosso. Devemos, portanto, utilizar o princípio da parcimónia e eliminá-los da teoria. Se, por outro lado, não são mais do que regiões diferentes de um único universo, as leis físicas seriam as mesmas em todas as regiões, porque, se assim não fosse, não seria possível andar continuamente de uma região para outra. Neste caso, a única diferença entre as regiões seria a sua configuração inicial e, portanto, o princípio antrópico forte reduzir-se-ia ao fraco.

Uma segunda objeção ao princípio antrópico forte é o facto de correr contra a maré da história da ciência. Desenvolvemo-lo a partir das cosmologias geocêntricas de Ptolemeu e dos seus antepassados, através da cosmologia heliocêntrica de Copérnico e de Galileu, até ao quadro moderno, em que a Terra é um planeta de tamanho médio orbitando em redor de uma estrela média nos subúrbios de uma galáxia espiralada comum, que é apenas uma entre cerca de 1 milhão de milhões de galáxias no universo observável. Contudo, o princípio antrópico forte afirmaria que toda esta vasta construção existe simplesmente por nossa causa. O que é muito difícil de acreditar. O

sistema solar é certamente um pré-requisito para a nossa existência, e podemos abranger com isto toda a galáxia para justificar uma geração anterior de estrelas que criaram os elementos mais pesados. Mas não parece haver qualquer necessidade para todas as outras galáxias nem para o universo em macroescala de ser tão uniforme e semelhante em todas as direções.

Podíamos sentir-nos mais felizes quanto ao princípio antrópico, pelo menos na sua versão fraca, se pudéssemos mostrar que um vasto número de diferentes configurações iniciais para o universo podiam ter evoluído para produzir um universo como o que observamos. Se fosse esse o caso, um universo que se desenvolveu ao acaso a partir de condições iniciais arbitrárias devia conter um número de regiões regulares, uniformes e adequadas à evolução da vida inteligente. Por outro lado, se o estado inicial do universo teve de ser escolhido com todo o cuidado para conduzir a qualquer coisa como o que vemos à nossa volta, seria improvável que o universo contivesse qualquer região onde tivesse aparecido vida. No modelo quente⁶ do *big bang*, já descrito, não houve tempo suficiente no estado inicial do universo para o calor ir de uma região para outra. Isto significa que o estado inicial do universo deveria ter tido exatamente a mesma temperatura por toda a parte para explicar o facto de a radiação de micro-ondas de fundo ter a mesma temperatura para onde quer que olhemos. A velocidade inicial de expansão também deveria ter sido escolhida com grande precisão para que se mantivesse tão próxima da velocidade crítica necessária para evitar o colapso. Isto significa que, se o modelo quente do *big bang* está certo até ao começo do tempo, o estado inicial do universo deve ter sido realmente escolhido com grande cuidado. Seria muito difícil explicar o motivo pelo qual o universo deve ter começado exatamente assim, exceto como ato de um deus que tencionava criar seres como nós.

Numa tentativa para descobrir um modelo do universo em que muitas configurações iniciais diferentes podiam ter evoluído até qualquer coisa como o universo atual, um cientista do Instituto de Tecnologia do Massachusetts, Alan Guth, sugeriu que o universo, no princípio, podia ter passado por um período de expansão muito rápida. Esta expansão diz-se «inflacionária», significando que o universo, a certa altura, se expandiu com velocidade crescente, e não com velocidade decrescente, como acontece hoje. Segundo Guth, o raio do universo aumentou 1 milhão de milhões de milhões de milhões de milhões de vezes (1 seguido de 30 zeros) em apenas uma minúscula fração de segundo.

Guth sugeriu que o universo começou num estado muito quente, mas bastante caótico. Essas temperaturas elevadas teriam significado que as partículas do universo se moviam muito depressa e tinham energias elevadas. Como já discutimos, esperar-se-ia que, a temperaturas tão elevadas, as forças nucleares fraca e forte e a força eletromagnética estivessem unificadas numa força única. O universo em expansão arrefeceria e a energia das partículas baixaria. Eventualmente, chegar-se-ia ao que se chama uma transição de fase e a simetria entre as forças seria quebrada: as forças forte, fraca e eletromagnética tornar-se-iam distintas. Um exemplo vulgar de uma transição de fase é o congelamento da água, quando arrefece. A água líquida é

contra a força da gravidade, que os atrai um para o outro. Portanto, em certo sentido, o campo gravitacional tem energia negativa. No caso de um universo que seja aproximadamente uniforme no espaço pode mostrar-se que esta energia gravitacional negativa anula exatamente a energia positiva representada pela matéria. Portanto, a energia total do universo é zero.

Ora, duas vezes zero é zero. Assim, o universo pode duplicar a quantidade de energia positiva da matéria e também a energia gravitacional negativa sem violar a conservação da energia. Isto não acontece na expansão normal do universo, em que a densidade de energia da matéria diminui à medida que o universo se torna maior. Acontece sim na expansão inflacionária, porque a densidade de energia do estado sobrearrefecido permanece constante enquanto o universo se expande: quando o universo duplica o tamanho, a energia positiva da matéria e a energia gravitacional negativa também duplicam, de modo que a energia total continua a ser zero. Durante a fase inflacionária, o universo aumenta muito o seu tamanho, o que faz com que a quantidade total de energia disponível para produzir partículas se torne muito grande. Como Guth afirmou: «Diz-se que não há nada como um almoço de graça. Mas o universo é o derradeiro almoço de graça.»

Atualmente, o universo não está a expandir-se de uma maneira inflacionária. Portanto, tem de haver um mecanismo qualquer que elimine o enorme efeito da constante cosmológica e, assim, modifique a taxa de expansão acelerada para uma taxa desacelerada pela gravidade, como a que temos agora. Na expansão inflacionária pode esperar-se que, eventualmente, a simetria entre as forças seja quebrada, tal como a água sobrearrefecida acaba sempre por gelar. A energia extra do estado de simetria não quebrada seria então libertada e voltaria a aquecer o universo até uma temperatura imediatamente abaixo da temperatura crítica para a simetria entre as forças. O universo continuaria a expandir-se e a arrefecer, tal como o universo-padrão do *big bang*, mas haveria agora uma explicação para o facto de o universo estar a expandir-se exatamente à velocidade crítica e de diferentes regiões terem a mesma temperatura.

Na proposta original de Guth supunha-se que a transição de fase ocorria repentinamente, um pouco como o aparecimento de cristais de gelo em água muito fria. A ideia consistia em que «bolhas» da nova fase de simetria quebrada se teriam formado na fase antiga, como bolhas de vapor em água a ferver. As bolhas expandir-se-iam e juntar-se-iam umas às outras até todo o universo se encontrar na nova fase. O contratempo era, como eu e outros sublinhámos, que o universo estava a expandir-se tão depressa que, mesmo que as bolhas crescessem à velocidade da luz, se afastariam umas das outras e, assim, nunca poderiam juntar-se. O universo ficaria num estado não uniforme, onde algumas regiões conservariam ainda simetria entre as diferentes forças. Tal modelo de universo não corresponderia ao que vemos.

Em outubro de 1981 fui a Moscovo para uma conferência sobre a gravidade quântica. Depois da conferência dirigi um seminário sobre o modelo inflacionário e os seus problemas no Instituto Astronómico Sternberg. Antes

disso tinha arranjado alguém para me substituir nas minhas intervenções, em virtude de a maior parte das pessoas não conseguirem perceber a minha voz. Mas não havia tempo para preparar o seminário, pelo que tive de ser eu a ministrá-lo, com um dos meus alunos já graduados a repetir as minhas palavras. Entre o público estava um jovem russo, Andrei Linde, do Instituto Lebedev de Moscovo, que disse que a dificuldade de as bolhas não se juntarem podia ser evitada se as bolhas fossem tão grandes que a nossa região do universo estivesse toda contida numa só. Para isto resultar, a quebra de simetria devia ter ocorrido muito lentamente dentro da bolha, o que é absolutamente possível segundo as teorias da grande unificação. A ideia de Linde de uma quebra lenta da simetria era muito boa, mas mais tarde cheguei à conclusão de que as bolhas precisavam de ter sido maiores do que o universo! Mostrei que, afinal, a simetria teria sido quebrada por toda a parte ao mesmo tempo, em vez de apenas dentro das bolhas. Isto conduzia a um universo uniforme, tal como o observamos. Fiquei muito excitado com a ideia e discuti-a com um dos meus alunos, Ian Moss. Como amigo de Linde, fiquei bastante embaraçado quando mais tarde me enviaram o seu trabalho através de uma revista científica, perguntando-me se tinha qualidade para ser publicado. Respondi que havia essa falha acerca de as bolhas serem maiores do que o universo, mas que a ideia fundamental de uma quebra lenta de simetria era muito boa. Recomendei, no entanto, que o artigo fosse publicado tal como estava, porque Linde levaria alguns meses a corrigi-lo, visto que tudo o que ele mandasse para o Ocidente teria de passar pela censura soviética, que não era muito hábil nem muito rápida com os artigos científicos. Entretanto escrevi um curto artigo, juntamente com Ian Moss, que foi publicado na mesma revista em que chamávamos a atenção para o problema da bolha e mostrávamos como podia ser resolvido.

Logo no dia seguinte ao do meu regresso de Moscovo parti para Filadélfia, onde ia receber uma medalha no Instituto Franklin. A minha secretária, Judy Fella, tinha usado o seu considerável encanto para persuadir a British Airways a oferecer-nos, a mim e a ela, duas passagens de graça num *Concorde*, a título publicitário. Contudo, atrasei-me no caminho para o aeroporto por causa da chuva e perdi o avião. Mas acabei por chegar a Filadélfia e receber a minha medalha. Pediram-me depois que dirigisse um seminário sobre o universo inflacionário na Universidade de Drexel, em Filadélfia. Passei a maior parte do seminário a falar sobre os problemas do modelo inflacionário, tal como em Moscovo.

Uma ideia muito semelhante à de Linde foi sugerida independentemente, alguns meses depois, por Paul Steinhardt e Andreas Albrecht, da Universidade da Pensilvânia. Atualmente, são considerados, juntamente com Linde, autores do chamado «novo modelo inflacionário», baseado na ideia da quebra lenta de simetria. (O antigo modelo inflacionário era a sugestão original de Guth de uma súbita quebra de simetria com a formação de bolhas.)

O novo modelo inflacionário foi uma boa tentativa para explicar o motivo pelo qual o universo é como é. No entanto, eu e outros demonstrámos que, pelo menos na sua forma original, previa muito maiores variações na

temperatura da radiação de fundo de micro-ondas do que as que são observadas. Trabalhos ulteriores também lançaram dúvidas sobre se podia haver uma transição de fase no começo do universo do tipo necessário. Na minha opinião pessoal, o novo modelo inflacionário está morto como teoria científica, embora muita gente pareça não ter ouvido falar da sua morte e continue a escrever trabalhos como se ele fosse viável. Um modelo melhor, chamado modelo inflacionário caótico, foi proposto por Linde em 1983. Neste não existe qualquer transição de fase ou sobrearrefecimento. Em vez disso há um campo de *spin* 0, que, devido a flutuações quânticas, teria tido valores elevados em algumas regiões do universo inicial. A energia do campo nessas regiões comportar-se-ia como uma constante cosmológica. Teria um efeito gravitacional repulsivo e levaria, portanto, essas regiões a expandirem-se de uma maneira inflacionária. À medida que se expandiam, a energia do campo diminuiria lentamente até que a expansão inflacionária se transformava numa expansão como a do modelo quente do *big bang*. Uma dessas regiões viria a ser aquilo que agora designamos por universo observável. Este modelo tem as vantagens dos primeiros modelos inflacionários, mas não depende de uma fase de transição duvidosa e pode, além disso, fornecer dimensões razoáveis para as flutuações da temperatura do fundo de micro-ondas concordantes com a observação.

Este trabalho sobre modelos inflacionários mostrou que o estado atual do universo podia ter surgido a partir de um número bastante elevado de configurações iniciais diferentes. Isto é importante porque mostra que o estado inicial da parte do universo que habitamos não teve de ser escolhido com grande cuidado. Portanto, podemos, se quisermos, utilizar o princípio antrópico fraco para explicarmos o motivo por que o universo tem o aspeto que hoje tem. O que não pode, no entanto, é cada configuração inicial ter conduzido a um universo como o que observamos. Podemos mostrá-lo considerando um estado muito diferente para o universo na atualidade, digamos um estado muito irregular. Podemos utilizar as leis da física para imaginarmos o universo no passado a fim de determinarmos a sua configuração inicial. Segundo os teoremas de singularidade da relatividade geral, continuaria a ter havido a singularidade do *big bang*. Se imaginarmos a evolução desse universo segundo as leis da física, acabamos com o estado irregular e heterogéneo com que começámos. Portanto, deve ter havido configurações iniciais que não dariam origem a um universo como o que vemos hoje. Nem mesmo o modelo inflacionário nos diz por que motivo a configuração inicial não era de molde a produzir qualquer coisa muito diferente da que observamos. Teremos de nos voltar para o princípio antrópico para obtermos uma explicação? Terá sido apenas uma questão de sorte? Pareceria uma causa perdida, uma negação de todas as nossas esperanças de compreendermos a ordem subjacente do universo.

Para ser possível reconstituir o estado inicial do universo são precisas leis que prevaleçam até ao começo do tempo. Se a teoria clássica da relatividade geral estiver correta, os teoremas de singularidade que Roger Penrose e eu provámos mostram que o começo do tempo devia ter sido um ponto de densidade infinita e curvatura infinita do espaço-tempo. Todas as leis

conhecidas perderiam a validade num tal ponto. Podemos supor que existem novas leis que se mantêm válidas nas singularidades, mas é muito difícil formulá-las em pontos de comportamento tão mau, e não temos qualquer indicador, a partir das observações, que nos diga como devem ser essas leis. Todavia, o que os teoremas de singularidade indicam realmente é que o campo gravitacional se torna tão forte que os efeitos da gravitação quântica se tornam importantes: a teoria clássica deixa de ser uma boa descrição do universo. Assim, é preciso utilizar uma teoria quântica da gravidade para discutir o estado primitivo do universo. Como veremos, é possível, na teoria quântica, que as leis vulgares da ciência se mantenham válidas por toda a parte, inclusive no princípio do tempo: não é necessário postular novas leis para as singularidades, porque as singularidades na teoria quântica não são inevitáveis.

Ainda não temos uma teoria completa e coerente que combine a mecânica quântica com a gravidade. No entanto, temos praticamente a certeza de algumas características necessárias a uma teoria unificada como essa. Uma é a que deve incorporar a proposta de Feynman de formular a teoria quântica em termos de uma soma de histórias. Neste caso, uma partícula não tem apenas uma história, como teria numa teoria clássica. Em vez disso, supõe-se que segue todas as trajetórias possíveis no espaço-tempo e que a cada uma das histórias está associado um par de números, o comprimento de onda e a sua fase. A probabilidade de a partícula passar num certo ponto é encontrada somando as ondas associadas às histórias que contêm esse ponto. Quando tentamos na prática efetuar estas somas, aparecem, contudo, graves problemas técnicos. A única maneira de os evitar é a seguinte receita estranha: somar também as ondas relativas às histórias que não ocorrem no tempo «real» em que nos encontramos, mas que ocorrem no chamado tempo imaginário. Tempo imaginário é qualquer coisa que pode cheirar a ficção científica, mas é, na realidade, um conceito matemático bem definido. Se tomarmos um número ordinário ou «real» e o multiplicarmos por si próprio, o resultado é um número positivo (por exemplo, 2 vezes 2 são 4, mas -2 vezes -2 também são 4). Existem, contudo, números especiais, chamados imaginários, que dão números negativos quando multiplicados por si próprios. (O chamado i quando multiplicado por si próprio, dá -1: $2i$ multiplicado por si próprio dá -4, etc.)

Podemos representar números reais e imaginários como se segue: os números reais podem ser representados por uma linha que vai da esquerda para a direita com o zero no meio, os números negativos, como -1, -2, etc., à esquerda e os números positivos, 1, 2, etc., à direita. Então os números imaginários são representados por uma linha que vai de baixo para cima com o zero no meio, os números como i , $2i$, etc., em cima e os números $-i$, $-2i$, etc., em baixo. Assim, os números imaginários são, num certo sentido, números que fazem ângulos retos com os números reais ordinários.

Para evitar as dificuldades técnicas com a soma de histórias de Feynman é preciso utilizar o tempo imaginário, isto é, para efeitos de cálculo, deve medir-se o tempo utilizando números imaginários, em vez de reais. Isto tem um

efeito interessante sobre o espaço-tempo: a distinção entre espaço e tempo desaparece completamente. Um espaço-tempo no qual os acontecimentos têm valores imaginários da coordenada tempo diz-se euclidiano, do nome do grego Euclides, que fundou o estudo da geometria de superfícies bidimensionais. Aquilo a que agora chamamos espaço-tempo euclidiano é muito semelhante, exceto o facto de ter quatro dimensões, em vez de duas. No espaço-tempo euclidiano não existe diferença entre direção do tempo e direção do espaço. Por outro lado, no espaço real, em que os acontecimentos são identificados por valores ordinários e reais da coordenada tempo, é fácil distinguir a diferença: a direção do tempo em todos os pontos fica dentro do cone de luz e as direções do espaço ficam fora. Em qualquer caso, no que diz respeito à mecânica quântica, podemos considerar a nossa utilização do tempo imaginário e do espaço-tempo euclidiano uma mera artimanha ou truque matemático para calcular respostas acerca do espaço-tempo real.

A segunda característica de uma teoria física acabada é a ideia de Einstein de que o campo gravitacional é representado por um espaço-tempo curvo: as partículas tentam seguir uma coisa parecida com uma trajetória retilínea num espaço curvo, mas, como o espaço-tempo não é plano, as suas trajetórias parecem encurvadas como que por ação do campo gravitacional. Quando aplicamos a soma das histórias de Feynman à teoria da gravidade de Einstein, o que é análogo à história de uma partícula é agora um espaço-tempo completamente curvo, que representa a história global do universo. Para fugir às dificuldades técnicas ao efetuar a soma de histórias, estes espaços-tempos curvos são considerados euclidianos. Ou seja, o tempo é imaginário e não mais se distingue das direções do espaço. Para calcular a probabilidade de encontrar um espaço-tempo com determinada propriedade, tal como ter o mesmo aspeto em qualquer ponto e em qualquer direção, somam-se as ondas associadas a todas as histórias que tenham essa propriedade.

Na teoria clássica da relatividade geral existem muitos espaços-tempos possíveis diferentes, cada um correspondendo a um diferente estado inicial do universo. Se conhecêssemos o estado inicial do universo, conheceríamos toda a sua história. De modo semelhante, na teoria quântica da gravidade há muitos estados quânticos possíveis para o universo. Mais uma vez, se soubéssemos como se comporta a soma de histórias nos espaços-tempos euclidianos primitivos, conheceríamos o estado quântico do universo.

Na teoria clássica da gravidade, que se baseia num espaço-tempo real, só há dois comportamentos possíveis para o universo: ou existe há um tempo infinito, ou então teve um princípio numa singularidade há um tempo finito no passado. Na teoria quântica da gravidade, por seu lado, surge uma terceira possibilidade. Como se utilizam espaços-tempos euclidianos, nos quais as direções do tempo estão em pé de igualdade com as direções do espaço, é possível o espaço-tempo ser finito em extensão e, contudo, não ter quaisquer singularidades a formarem uma fronteira ou um limite. O espaço-tempo seria como a superfície da Terra, mas com mais duas dimensões. A superfície da Terra é finita na sua extensão, mas não tem uma fronteira ou limite. Se navegarmos em direção ao pôr do Sol, não cairemos de nenhuma fronteira

nem se nos deparará uma singularidade. (Eu sei-o porque já dei a volta ao mundo!).

Se o espaço-tempo euclidiano se estende para trás até um tempo infinito imaginário, depara-se-nos o mesmo problema da teoria clássica de saber qual era o estado inicial do universo: Deus pode saber como o universo principiou, mas nós não somos capazes de encontrar uma razão especial para pensarmos que começou de uma maneira ou de outra. Porém, a teoria quântica da gravidade abre uma nova perspectiva: pode não existir fronteira no espaço-tempo e, neste caso, não há necessidade de especificar o comportamento na fronteira. Não há quaisquer singularidades a que as leis da ciência não se apliquem, nem limites no espaço-tempo cujas condições de fronteira seja preciso pedir a Deus ou que alguma nova lei estabeleça. Podíamos então dizer: «A condição de fronteira do universo consiste em não ter fronteira.» O universo seria completamente independente e nunca afetado por qualquer coisa exterior a ele. Não teria sido criado nem seria destruído. SERIA apenas.

Foi na conferência no Vaticano, que já mencionei, que apresentei pela primeira vez a sugestão de que talvez o espaço e o tempo juntos formassem uma superfície finita em extensão, mas sem qualquer fronteira ou limite. O meu trabalho era bastante matemático, pelo que as suas implicações quanto ao papel de Deus na criação não foram percebidas na altura (felizmente para mim). Nessa altura ainda não sabia como utilizar a ideia da «ausência de fronteiras» para elaborar prognósticos sobre o universo. No entanto, passei o Verão seguinte na Universidade da Califórnia, em Santa Bárbara. Aí, um colega e amigo, Jim Hartle, trabalhou comigo, tentando descobrir que condições o universo devia satisfazer para que o espaço-tempo não tivesse limites. Quando regresssei a Cambridge, continuei o trabalho com dois dos meus alunos de investigação, Julian Luttrell e Jonathan Halliwell.

Gostaria de frisar que esta ideia de que o tempo e o espaço devem ser finitos, mas ilimitados, não passa de uma proposta: não pode ser deduzida de qualquer outro princípio. Tal como qualquer outra teoria científica, pode ser inicialmente proposta por razões estéticas ou metafísicas, mas a prova real reside em saber se faz previsões que concordam com a observação. Isto, porém, no caso da gravidade quântica, é difícil de determinar por duas razões: primeiro, como será explicado no capítulo 11, ainda não temos a certeza exata de qual é a teoria que combina com êxito a relatividade geral e a mecânica quântica, embora saibamos bastante sobre a forma que tal teoria deve ter; segundo, qualquer modelo que descreva o universo no seu todo, em pormenor, seria do ponto de vista matemático demasiadamente complexo para sermos capazes de calcular previsões exatas. Portanto, temos de fazer suposições e aproximações simplificativas e, mesmo assim, o problema de extrair previsões mantém-se.

Cada história na soma das histórias descreverá não apenas o espaço-tempo, mas também tudo o que nele existe, incluindo quaisquer organismos, como os seres humanos, que são capazes de observar a história do universo. Isto pode fornecer outra justificação para o princípio antrópico, porque, se

todas as histórias são possíveis, podemos, desde que existamos numa delas, utilizar o princípio antrópico para explicarmos o motivo pelo qual o universo parece ser como é. Que significado exato pode ser dado às outras histórias em que não existimos, não o sabemos. Esta ideia de uma teoria quântica da gravidade seria muito mais satisfatória, no entanto, se pudéssemos mostrar que, usando a soma de histórias, o nosso universo não é apenas uma das histórias possíveis, mas sim uma das mais prováveis. Para o conseguirmos temos de operar a soma de histórias para todos os possíveis espaços-tempos euclidianos ilimitados.

Sob a proposta de ausência de fronteiras, ficamos a saber que a hipótese de o universo estar a seguir a maior parte das histórias possíveis é insignificante, mas existe um grupo especial de histórias que são muito mais prováveis do que as outras. Estas histórias podem ser imaginadas como a superfície da Terra, em que a distância ao Polo Norte representa o tempo imaginário e o perímetro de um círculo, a essa distância constante do Polo Norte, representa o tamanho espacial do universo. O universo começa no Polo Norte como um simples ponto. À medida que nos dirigimos para sul, os círculos de latitude a uma distância constante do Polo Norte tornam-se maiores, correspondendo à expansão do Universo no tempo imaginário (fig. 8.1). O universo chegaria ao seu tamanho máximo no equador e contrair-se-ia no decurso do tempo imaginário num ponto único no Polo Sul. Embora o universo tivesse o tamanho zero nos Polos Norte e Sul, estes pontos não seriam singularidades, tal como os Polos Norte e Sul na Terra não são singularidades. As leis da ciência manter-se-iam válidas, tal como acontece nos Polos Norte e Sul da Terra.

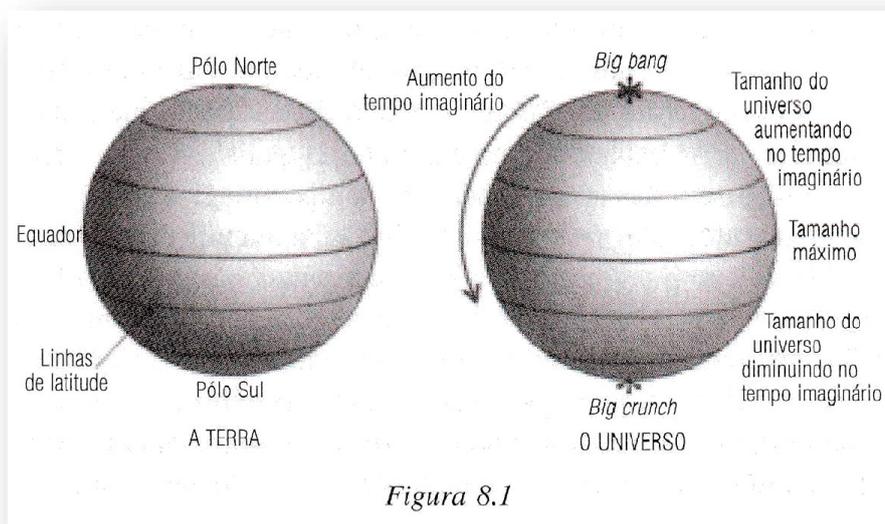


Figura 8.1

A história do universo no tempo real, contudo, teria um aspeto muito diferente. Há cerca de 10 ou 20 milhões de anos teria um tamanho mínimo, que era igual ao raio máximo da história no tempo imaginário. Mais tarde, em tempos reais, o universo expandir-se-ia como o modelo inflacionário caótico proposto por Linde (mas agora já não teríamos de admitir que o universo

tenha sido criado no estado certo). O Universo expandir-se-ia até um tamanho muito grande e eventualmente entraria de novo em colapso para aquilo que parece ser uma singularidade no tempo real. Assim, em certo sentido, continuamos a estar condenados, mesmo que nos mantenhemos afastados de buracos negros. Só se conseguíssemos conceber o universo em termos de tempo imaginário não haveria singularidades.

Se o universo está realmente em tal estado quântico, não haveria singularidades na sua história no tempo imaginário. Podia parecer, pois, que o meu trabalho mais recente tinha desfeito os resultados do meu trabalho anterior sobre as singularidades. Mas, como acabei de indicar, a verdadeira importância dos teoremas sobre singularidades consiste em mostrar que não podem ignorar-se os efeitos quânticos quando o campo gravitacional se torna muito intenso. Isto, por sua vez, levou à ideia de que o universo podia ser finito no tempo imaginário, mas sem limites ou singularidades. Contudo, quando se volta ao tempo real em que vivemos, continua a parecer que existem singularidades. O pobre astronauta que cai num buraco negro continua a ter um fim infeliz; só não encontrava singularidades se vivesse em tempo imaginário.

Isto podia sugerir que o chamado tempo imaginário é realmente o tempo verdadeiro e que aquilo a que chamamos tempo real é produto da nossa imaginação. No tempo real, o universo tem um princípio e um fim em singularidades que formam uma fronteira para o espaço-tempo e perante as quais as leis da ciência deixam de ser válidas. Mas no tempo imaginário não há singularidades nem fronteiras. Assim, talvez aquilo a que chamamos tempo imaginário seja realmente mais básico e aquilo a que chamamos tempo real seja apenas uma ideia que inventámos para nos ajudar a descrever como pensamos que o universo é. Mas, de acordo com a aproximação que descrevi no capítulo 1, uma teoria científica não passa de um modelo matemático que elaboramos para descrever as nossas observações: só existe nas nossas mentes. Portanto, não tem significado perguntar, o que é real, o tempo «real» ou o «imaginário»? Trata-se apenas de saber qual é a descrição mais útil.

Podemos também utilizar a soma das histórias, juntamente com a proposta de ausência de fronteiras, para descobrirmos quais as propriedades do universo mais prováveis de ocorrerem em conjunto. Por exemplo, podemos calcular a probabilidade de o universo estar a expandir-se quase à mesma velocidade em diferentes direções numa altura em que a densidade do universo tem o seu valor atual. Nos modelos simplificados que têm sido examinados até agora, esta probabilidade parece ser grande, ou seja, a condição de ausência de fronteiras leva à conclusão de que é muito provável que a atual velocidade de expansão do universo seja quase a mesma em todas as direções. Isto é coerente com as observações da radiação do fundo de micro-ondas, que mostra que ela tem quase exatamente a mesma intensidade em todas as direções. Se o universo estivesse a expandir-se mais depressa numa direção do que noutras, a intensidade da radiação nessas direções seria reduzida por um desvio adicional para o vermelho.

Estão a ser trabalhadas mais previsões da condição de ausência de fronteiras. Um problema particularmente interessante é a amplitude de pequenos desvios da densidade uniforme no universo primordial que causaram a formação, primeiro, de galáxias, depois, de estrelas e, finalmente, de seres humanos. O princípio da incerteza implica que o universo, no princípio, não pode ter sido completamente uniforme, porque deve ter havido algumas incertezas ou flutuações nas posições e velocidades das partículas. Usando a condição de ausência de fronteiras, descobrimos que o universo deve, de facto, ter começado com o mínimo possível de heterogeneidades permitido pelo princípio da incerteza. O universo teria então passado por um período de expansão rápida, como nos modelos inflacionários. Durante esse período, as heterogeneidades iniciais teriam sido ampliadas até serem suficientemente grandes para explicarem a origem das estruturas que observamos à nossa volta. Em 1992 o satélite *Cosmic Background Explorer* (COBE) detetou pela primeira vez variações muito ligeiras na intensidade do fundo de micro-ondas em função da direção. O modo como estas não uniformidades dependem da direção parece concordar com as previsões do modelo inflacionário e da proposta de ausência de fronteiras. Assim, esta proposta é uma boa teoria científica no sentido de Karl Popper: podia ser falsificada por estas observações, mas as suas previsões foram confirmadas. Num universo em expansão, em que a densidade da matéria variou ligeiramente de local para local, a gravidade deve ter obrigado as regiões mais densas a afrouxarem a sua expansão e a começarem a contrair-se. Isto terá levado à formação de galáxias, estrelas e, eventualmente, até de criaturas insignificantes, como nós. Assim, todas as estruturas complicadas que vemos no universo podiam ser explicadas pela condição de não haver fronteiras conjuntamente com o princípio da incerteza da mecânica quântica.

A ideia de que o espaço e o tempo podem formar uma superfície fechada finita e ilimitada tem também profundas implicações no papel de Deus na criação. Com o êxito das teorias científicas na descrição dos fenómenos, a maioria das pessoas acabou por acreditar que Deus permite que o universo evolua segundo um conjunto de leis e não intervém nele para quebrar essas leis. Contudo, as leis não nos dizem qual era o aspeto do universo primitivo. O ato de dar corda ao relógio e escolher como pô-lo a trabalhar continuaria a ser com Deus. Desde que o universo tenha tido um princípio, podemos supor que teve um criador. Mas, se o universo for na realidade completamente independente, sem qualquer fronteira ou limite, não terá princípio nem fim: existirá apenas. Qual seria então o papel do criador?

¹ Georges Lemaître foi o primeiro a advogar a tese do estado inicial de alta densidade, a que chamou átomo primevo. Lemaître nasceu em 1894 e foi ordenado sacerdote em 1927, no mesmo ano em que obteve o grau de doutor no MIT e publicou o seu trabalho sobre a origem do universo. Parecia a Lemaître, bem como à igreja católica, que este modelo se adaptava particularmente bem ao momento da criação. (N. do R. C.)

² De facto, com a expansão, a radiação arrefece mais devagar do que a matéria. Houve uma Era em que a matéria e a radiação se achavam em equilíbrio térmico. A partir do fim desse tempo, a radiação e a matéria passaram a evoluir separadamente, não mais interagindo uma com a outra. (N. do R. C.)

³ Esta época é conhecida por época da recombinação. (N. do R. C.)

⁴ Em rigor, deveria chamar-se-lhes elipsoidais, pois não são aplanadas. (N. do R. C.)

- ⁵ Este fenómeno cíclico, que possibilita a reciclagem da matéria no espaço interestelar, é designado por *astracção*. (N. do R. C.)
- ⁶ Também designado por modelo-padrão. (N. do R. C.)

CAPÍTULO 9

- A SETA DO TEMPO -

Em capítulos anteriores vimos como as nossas ideias sobre a natureza do tempo se modificaram com o passar dos anos. Até ao começo deste século, as pessoas acreditavam num tempo absoluto. Ou seja, cada acontecimento podia ser rotulado por um número chamado «tempo» de uma maneira única e todos os relógios mediriam o mesmo intervalo de tempo entre dois acontecimentos. Contudo, a descoberta de que a velocidade da luz é a mesma para todos os observadores conduziu à teoria da relatividade, segundo a qual temos de abandonar a ideia de um tempo absoluto único. Em vez disso, cada observador tem a sua própria medida de tempo, registada pelo seu relógio: relógios diferentes não coincidem necessariamente uns com os outros. Deste modo, o tempo tornou-se um conceito mais pessoal, relativo ao observador que o mede.

Quando se tentou unificar a gravidade com a mecânica quântica, foi preciso introduzir a noção de tempo «imaginário». O tempo imaginário não se distingue das direções no espaço. Se pode ir-se para norte, pode voltar-se para trás e ir para sul; do mesmo modo, se se pode avançar no tempo imaginário, também deve poder-se voltar para trás. Isto quer dizer que não pode haver uma diferença importante entre os sentidos para diante e para trás no tempo imaginário. Por outro lado, quando se olha para o tempo «real», há uma diferença enorme entre os sentidos para diante e para trás, como todos sabemos. Donde vem esta diferença entre o passado e o futuro? Por que nos lembramos do passado, mas não do futuro?

As leis da ciência não estabelecem uma diferença entre o passado e o futuro. Mais precisamente, como já foi explicado, as leis da física são invariantes sob a combinação de operações (ou simetrias) conhecidas por C, P e T. (C significa troca de partículas por antipartículas; P significa tomar a imagem no espelho, de modo que a esquerda e a direita fiquem trocadas; T significa trocar o sentido do movimento de todas as partículas, isto é, inverter o sentido do tempo.) As leis da física que governam o comportamento da matéria em todas as condições normais não se modificam sob a composição das duas operações C e P. Por outras palavras, a vida seria igual para habitantes de outro planeta que fossem nossas imagens no espelho constituídas por antimatéria.

Se as leis da física não se alteram com a composição das operações C e P, nem pela combinação de C, P e T, também não deviam alterar-se sob a operação T isolada. Contudo, há uma diferença grande entre os sentidos para a frente e para trás no tempo real na vida de todos os dias. Imaginemos uma chávena de chá que cai de uma mesa e se quebra no chão em mil e um

bocadinhos. Se filmarmos a cena, podemos depois, facilmente, dizer quando é que o filme está a correr para diante ou para trás. Se o fizermos correr para trás, então veremos os bocadinhos juntarem-se subitamente para formarem uma chávena inteira no chão, que ressalta para cima da mesa. Podemos dizer que o filme está a correr ao contrário porque esta espécie de comportamento nunca se observa na vida quotidiana. Se assim fosse, os fabricantes de loiça iam à falência.

A explicação geralmente dada para que não vejamos fragmentos de loiça reunirem-se em chávenas que saltam inteiras do chão para cima de uma mesa provém da segunda lei da termodinâmica. Esta diz que, em qualquer sistema fechado, a desordem ou entropia aumenta sempre com o tempo. Por outras palavras, é uma forma da lei de Murphy: as coisas têm sempre tendência para correr mal! Uma chávena intacta em cima de uma mesa é um estado ordenado, mas uma chávena partida no chão é um estado desordenado. Podemos passar prontamente da chávena em cima da mesa no passado para a chávena partida no chão, mas não o contrário.

O aumento de desordem ou entropia com o tempo é um exemplo daquilo a que se chama uma seta do tempo, qualquer coisa que distingue o passado do futuro, dando um sentido ao tempo. Há, pelo menos, três setas diferentes para o tempo. Primeiro há a seta termodinâmica, o sentido do tempo em que a desordem ou entropia aumenta. Depois há a seta psicológica, ou seja, a direção em que sentimos que o tempo passa, em que nos lembramos do passado, mas não do futuro. Finalmente, há a seta cosmológica, que é o sentido do tempo em que o universo está a expandir-se, e não a contrair-se.

Neste capítulo argumentei que a ausência de fronteiras para o universo conjuntamente com o princípio antrópico fraco podem explicar por que motivo as três setas apontam no mesmo sentido e, além disso, por que deve mesmo existir uma seta do tempo bem definida. Argumentarei que a seta psicológica é determinada pela seta termodinâmica e que estas duas setas apontam necessariamente no mesmo sentido. Se admitirmos que não há limites no universo, veremos que deve haver uma seta termodinâmica bem definida e uma seta cosmológica igualmente bem definida, mas que nem sempre apontarão no mesmo sentido durante a história do universo. Contudo, argumentarei que só quando apontam no mesmo sentido são as condições propícias ao desenvolvimento de seres inteligentes que podem perguntar: por que aumenta a desordem no mesmo sentido do tempo em que o universo se expande?

Discutirei primeiro a seta termodinâmica do tempo. A segunda lei da termodinâmica resulta de haver sempre mais estados desordenados do que ordenados. Consideremos, por exemplo, as peças de um quebra-cabeças dentro de uma caixa: há um, e apenas um, arranjo em que as peças formam uma imagem completa. Por outro lado, há um grande número de arranjos em que as peças estão desordenadas e não formam uma imagem coerente.

Suponhamos que um sistema começa a partir de um pequeno número de estados ordenados. À medida que o tempo passa, o sistema evoluirá de acordo com as leis da física e o seu estado mudará. Mais tarde é mais provável que o sistema esteja num estado desordenado, porque existem mais estados desordenados. Portanto, a desordem tenderá a aumentar com o tempo se o sistema obedecer a uma condição inicial de ordem elevada.

Suponhamos que as peças do quebra-cabeças partem de um arranjo ordenado na caixa, formando uma imagem: se sacudirmos a caixa, as peças originarão outro arranjo, que será provavelmente desordenado, sem formar uma imagem, simplesmente por haver muito mais arranjos desordenados. Alguns grupos de peças podem formar partes da imagem, mas, quanto mais se sacode a caixa, mais provável é que esses grupos se desfaçam e as peças fiquem num estado de completa desordem em que não formem qualquer espécie de imagem. Portanto, a desordem das peças, provavelmente, aumentará com o tempo se as peças obedecerem à condição inicial de ordem elevada.

Suponhamos, no entanto, que Deus decidiu que o universo devia acabar num estado muito ordenado, sem importar o seu estado inicial. Nos primeiros tempos, o universo estaria, possivelmente, num estado desordenado, o que significaria que a desordem diminuiria com o tempo. Veríamos chávenas partidas a juntarem-se de novo a partir dos pedaços e a saltarem inteiras para cima das mesas. Todavia, quaisquer seres humanos que estivessem a observar as chávenas estariam a viver num universo em que a desordem diminuía com o tempo. Argumentarei que esses seres teriam uma seta psicológica do tempo voltada ao contrário. Ou seja, lembrar-se-iam de acontecimentos do futuro e não se lembrariam de acontecimentos do seu passado. Quando a chávena se partisse, lembrar-se-iam de ela ter estado em cima da mesa, mas, quando estivesse realmente lá, não se lembrariam de ter estado no chão.

É bastante difícil falar da memória humana, porque não conhecemos o funcionamento pormenorizado do cérebro. Sabemos, porém, tudo sobre o funcionamento das memórias dos computadores. Assim, discutirei a seta psicológica do tempo dos computadores. Penso que é razoável admitir que a seta para os computadores é a mesma que para os homens. Se não fosse, podíamos ganhar uma fortuna na bolsa com um computador que se lembrasse dos preços de amanhã!

A memória de um computador é fundamentalmente um dispositivo que contém elementos que podem existir num de dois estados. Um exemplo simples é o ábaco. Na sua forma mais simples consiste em alguns arames em que deslizam contas que podem ser colocadas numa de duas posições. Antes de a informação ser registada, a memória de um computador está num estado desordenado, com probabilidades iguais para os estados possíveis. (As contas do ábaco são colocadas ao acaso nos seus arames.) Depois de a memória interagir com o sistema a ser lembrado ficará num estado ou no outro de acordo, com o estado do sistema. (Cada conta do ábaco estará ou à esquerda ou à direita do arame.) Portanto, a memória passou de um estado

desordenado para um estado ordenado. Contudo, para se ter a certeza de que a memória está no estado devido é necessário utilizar uma determinada quantidade de energia (para mover a conta ou para pôr a funcionar o computador, por exemplo). Esta energia é dissipada como calor e aumenta a quantidade de desordem no universo. Pode mostrar-se que este aumento da desordem é sempre maior do que o aumento da ordem da própria memória. Assim, o calor expelido pela ventoinha de arrefecimento do computador significa que, quando um computador regista uma informação na sua memória, a quantidade total de desordem no universo continua a aumentar. O sentido do tempo em que um computador se lembra do passado é o mesmo em que a desordem aumenta.

A nossa noção subjetiva do sentido do tempo, a seta psicológica do tempo é, portanto, determinada dentro do nosso cérebro pela seta termodinâmica. Tal como um computador, devemos lembrar-nos das coisas pela ordem em que a entropia aumenta. Isto torna a segunda lei da termodinâmica quase trivial. A desordem aumenta com o tempo porque medimos o tempo no sentido em que a desordem aumenta. Não pode haver uma aposta mais segura!

Mas por que há-de a seta termodinâmica do tempo existir? Ou, por outras palavras, por que há-de o universo estar num estado de grande ordem de um lado do tempo, daquele lado a que chamamos passado? Por que não está num estado de desordem total em todos os momentos? Afinal, isto poderia ser mais provável. E por que motivo é o sentido do tempo em que a desordem aumenta o mesmo em que o universo se expande? Na teoria clássica da relatividade geral não podemos prever como teria o universo começado, porque todas as leis conhecidas da física deixam de ser válidas na singularidade do *big bang*. O universo podia ter surgido num estado muito regular e ordenado. Isto teria levado às setas termodinâmica e psicológica bem definidas, como observamos. Mas podia igualmente ter surgido num estado muito irregular e desordenado. Neste caso, o universo estaria já num estado de desordem completa, pelo que esta não poderia aumentar com o tempo. Ou se manteria constante, e nesse caso não haveria qualquer seta termodinâmica do tempo bem definida, ou diminuiria, e nesse caso a seta termodinâmica do tempo apontaria no sentido oposto ao da seta cosmológica. Nenhuma dessas possibilidades concorda com o que observamos. Porém, como vimos, a relatividade geral clássica antevê as suas próprias limitações. Quando a curvatura do espaço-tempo se torna grande, os efeitos quânticos da gravitação devem ser tomados em consideração e a teoria clássica deixa de ser uma boa descrição do universo. É preciso utilizar uma teoria quântica da gravidade para se compreender como surgiu o universo.

Numa teoria quântica da gravidade, como vimos no capítulo anterior, para especificar o estado do universo continuaria a ser preciso dizer como é que as possíveis histórias do universo se comportaram no passado nos limites do espaço-tempo. Pode evitar-se esta dificuldade de ter de descrever o que não se sabe e não pode saber-se se as histórias satisfizerem a condição de ausência de fronteiras: são finitas em extensão, mas não têm fronteiras, limites ou singularidades. Nesse caso, o começo do tempo teria sido um ponto

regular do espaço-tempo e o universo teria começado a sua expansão num estado muito regular e ordenado. Poderia não ter sido completamente uniforme, porque isto violaria o princípio da incerteza da teoria quântica, mas ter havido pequenas flutuações na densidade e nas velocidades das partículas. A ausência de fronteiras implica, todavia, que essas flutuações são as mais pequenas possíveis, compatíveis com o princípio da incerteza.

O universo teria começado com um período de expansão exponencial ou inflacionária, em que o seu tamanho teria tido um grande aumento. Durante essa expansão, as flutuações de densidade teriam sido pequenas no princípio, mas mais tarde começariam a aumentar. As regiões em que a densidade fosse ligeiramente superior à média teriam tido a sua expansão retardada pela atração gravitacional da massa extra. Eventualmente, essas regiões teriam deixado de se expandir e entrariam em colapso para formar galáxias, estrelas e seres como nós. O universo teria partido de um estado regular e ordenado e ter-se-ia tornado irregular e desordenado à medida que o tempo ia passando. Isto explicaria a existência da seta termodinâmica do tempo.

Mas que aconteceria se e quando o universo parasse de se expandir e começasse a contrair-se? A seta termodinâmica inverter-se-ia e a desordem começaria a diminuir com o tempo? Isto levaria a toda a espécie de possibilidades do género da ficção científica para pessoas que sobrevivessem à passagem da fase de expansão para a de contração. Veríamos os bocadinhos das chávenas partidas juntarem-se no chão e saltarem de novo para cima das mesas? Seríamos capazes de nos lembrar dos preços do dia seguinte e de ganhar uma fortuna na bolsa? Pode parecer um pouco académico preocuparmo-nos com o que acontecerá quando o universo entrar em colapso, já que não começará a contrair-se antes de decorridos mais 10.000 milhões de anos. Mas há uma maneira mais rápida de descobrir o que vai acontecer: saltar para dentro de um buraco negro. O colapso de uma estrela para formar um buraco negro é bastante parecido com os últimos estados do colapso de todo o universo. Portanto, se a desordem diminuísse na fase de contração do universo, também poderia esperar-se que diminuísse dentro de um buraco negro. Assim, talvez um astronauta que caísse num buraco negro fosse capaz de ganhar dinheiro à roleta por recordar onde a bola ia parar antes de fazer a jogada. (Infelizmente, não teria muito tempo para jogar antes de ser transformado em esparquete. Nem seria capaz de nos falar da inversão de sentido da seta termodinâmica, nem de receber os ganhos, porque ficaria aprisionado por trás do horizonte de acontecimentos do buraco negro.)

Inicialmente acreditava que a desordem diminuiria quando o universo voltasse a entrar em colapso, porque pensava que o universo tinha de voltar a um estado regular e ordenado quando se tornasse de novo pequeno. Isto significaria que a fase de contração seria como o inverso do tempo da fase de expansão. Na fase de contração as pessoas viveriam uma vida às avessas: morreriam antes de terem nascido e ficariam mais novas à medida que o universo se contraísse.

Esta ideia é atraente porque significaria uma boa simetria entre as fases de expansão e de contração. Todavia, não podemos adotá-la independentemente de outras ideias sobre o universo. A questão agora é saber se esta ideia é consistente ou não com a condição de ausência de fronteiras. Como já disse, pensei no começo que a condição de ausência de fronteiras implicava realmente que a desordem diminuiria na fase de contração. Fui induzido em erro em parte pela analogia com a superfície da Terra. Se pensássemos no princípio do universo como correspondente ao Polo Norte, o fim do universo seria semelhante ao princípio, tal como o Polo Sul é semelhante ao Polo Norte. Porém, os Polos Norte e Sul correspondem ao princípio e ao fim do universo no tempo imaginário. O princípio e o fim no tempo real podem ser muito diferentes um do outro. Também fui enganado pelo trabalho que tinha elaborado num modelo simples do universo em que a fase de colapso seria simétrica da fase de expansão. No entanto, um colega meu, Don Page, da Universidade do estado da Pensilvânia, fez-me notar que a condição de ausência de fronteiras não obrigava necessariamente a que a fase de contração fosse simétrica da fase de expansão. Além disso, um dos meus alunos, Raymond Laflamme, concluiu que, num modelo ligeiramente mais complicado, o colapso do universo era muito diferente da expansão. Compreendi que tinha cometido um erro: a condição de não haver fronteiras implicava que a desordem continuaria de facto a aumentar durante a contração. As setas termodinâmica e psicológica do tempo não se inverteriam nem quando o universo começasse a contrair-se nem dentro dos buracos negros.

Que devemos fazer ao descobirmos que cometemos um erro deste género? Algumas pessoas nunca admitem que se enganaram e continuam a descobrir argumentos novos e muitas vezes inconscientes para fazerem valer as suas opiniões: como Eddington fez ao opor-se à teoria dos buracos negros. Outros afirmam que nunca apoiaram realmente a teoria incorreta ou que, se o fizeram, foi apenas para mostrarem que era inconsistente. Parece-me muito melhor e menos confuso admitir por escrito que nos enganámos. Um bom exemplo disto foi-nos deixado por Einstein, que chamou à constante cosmológica, que apresentou quando tentava elaborar um modelo estático do universo, o maior erro da sua vida.

Voltando à seta do tempo, persiste a pergunta: por que observamos as setas termodinâmica e cosmológica a apontarem no mesmo sentido? Ou, por outras palavras, por que aumenta a desordem à medida que o universo se expande? Se acreditarmos que o universo se expande para depois voltar a contrair-se, como a condição de ausência de fronteiras parece implicar, isto suscita a questão do motivo pelo qual devemos estar na fase de expansão, e não na de contração.

Podemos encontrar uma resposta baseada no princípio antrópico fraco. As condições na fase de contração não seriam adequadas à existência de seres inteligentes que pudessem perguntar: por que está a desordem sempre a aumentar à medida que o universo se expande? A inflação no princípio do universo, implicada pela ausência de fronteiras, significa que o universo deve

estar a expandir-se a uma velocidade muito próxima da velocidade crítica, à qual conseguirá evitar o colapso e, portanto, não entrará em colapso durante muito tempo. Por essa altura, as estrelas ter-se-ão apagado e os seus prótons e neutrões ter-se-ão transformado, provavelmente, em partículas de luz e radiação. O universo estaria num estado de quase completa desordem. Não haveria uma seta termodinâmica forte do tempo. A desordem não poderia aumentar muito porque o universo já estaria num estado de desordem quase total. Contudo, uma seta termodinâmica forte é necessária à vida inteligente. Para sobreviverem, os seres humanos têm de consumir alimentos, que constituem uma forma ordenada de energia, e convertê-los em calor, que é uma forma desordenada de energia. Desta maneira, a vida inteligente não poderia existir na fase de contração do universo. Assim se explica que observemos que as setas termodinâmica e cosmológica do tempo apontam no mesmo sentido. Não é a expansão do Universo que provoca o aumento da desordem, mas sim a ausência de fronteiras, que também propicia a vida inteligente na fase de expansão.

Resumindo, as leis da física não distinguem tempo para trás de tempo para diante. Contudo, há pelo menos três setas do tempo que distinguem realmente o passado do futuro: a seta termodinâmica, o sentido do tempo em que a desordem aumenta, a seta psicológica, o sentido do tempo em que nos lembramos do passado, e não do futuro, e a seta cosmológica, o sentido do tempo em que o universo se expande, em vez de se contrair. Já demonstrei que a seta psicológica é essencialmente a mesma que a seta termodinâmica, de modo que ambas apontam sempre no mesmo sentido. A hipótese de ausência de fronteiras no universo deixa prever a existência de uma seta termodinâmica bem definida, porque o universo tem de principiar num estado regular e ordenado. E o motivo pelo qual observamos que a seta termodinâmica coincide com a cosmológica consiste em que os seres inteligentes só podem existir na fase de expansão. A fase de contração não seria apropriada porque não se caracteriza por uma seta termodinâmica forte do tempo.

O progresso do homem na compreensão do universo estabeleceu um pequeno cantinho de ordem num mundo cada vez mais desordenado. Se o leitor se lembrar de todas as palavras deste livro, a sua memória terá registado cerca de 2 milhões de informações: a ordem no seu cérebro terá aumentado em cerca de 2 milhões de unidades. No entanto, enquanto o leitor leu este livro, terá convertido, pelo menos, 1000 calorias de energia ordenada, sob a forma de alimentos, e em energia desordenada, sob forma de calor, que se perde para o ar circundante por convecção e transpiração. Isto aumentará a desordem do universo em cerca de 20 milhões de milhões de milhões de milhões de unidades, ou cerca de 10 milhões de milhões de milhões de vezes o aumento da ordem do cérebro do leitor, se se lembrar de tudo o que está no livro. No próximo capítulo tentarei aumentar um pouco mais a ordem do nosso recanto, explicando como há gente a tentar conjugar as teorias que descrevi para formar uma teoria unificada completa que explique o universo e todo o seu conteúdo.

CAPÍTULO 10

- BURACOS DE VERME E VIAGENS NO TEMPO -

No último capítulo discuti por que razão vemos o tempo avançar, por que aumenta a desordem e por que nos lembramos do passado, e não do futuro. O tempo foi tratado como se nos encontrássemos numa linha de comboio retilínea na qual nos pudéssemos deslocar num sentido ou no outro.

Mas que aconteceria se a linha tivesse voltas e reviravoltas e ramais, de modo que o comboio, avançando sempre, voltasse a uma estação pela qual já tivesse passado? Por outras palavras, será possível viajar para o futuro ou para o passado?

H. G. Wells, em *A Máquina do Tempo*, explorou estas possibilidades, tal como muitos outros autores de ficção científica. Porém, muitas das ideias de ficção científica, como os submarinos e a viagem à Lua, tornaram-se factos científicos. Então quais são as perspetivas para a viagem no tempo?

O primeiro indício de que as leis da física permitiriam viajar no tempo data de 1949, quando Kurt Gödel descobriu um novo espaço-tempo admissível pela teoria da relatividade geral. Gödel era um matemático famoso por ter provado que é impossível demonstrar todas as asserções válidas, por exemplo, todas as asserções válidas de uma teoria tão limitada como a aritmética. Tal como o princípio da incerteza, o teorema da incompletude de Gödel pode ser uma limitação fundamental à nossa capacidade de compreensão e previsão do universo, mas pelo menos até agora não constituiu obstáculo à demanda de uma teoria unificada completa.

Gödel tomou conhecimento da teoria da relatividade geral quando conviveu com Einstein, nos últimos anos da vida deste, no Instituto de Estudos Avançados de Princeton. O seu espaço-tempo tinha a curiosa propriedade de que todo o universo se encontrava em rotação. Podemos perguntar: «Em rotação relativamente a quê?» A resposta é a de que a matéria distante se encontraria em rotação relativamente a direções determinadas por pequenos piões e giroscópios.

O espaço-tempo de Gödel tinha como efeito colateral a possibilidade de alguém partir num foguetão e regressar à Terra antes da partida. Esta propriedade preocupou de facto Einstein, que pensara que a relatividade geral não permitiria viajar no tempo. Porém, dados os antecedentes das mal fundadas oposições de Einstein ao colapso gravitacional e ao princípio da incerteza, a sua preocupação talvez fosse um sinal de encorajamento. A

solução que Gödel encontrou não corresponde ao universo em que vivemos, porque podemos mostrar que o universo não se encontra em rotação. Gödel também assumira um valor não nulo para a constante cosmológica, que Einstein introduzira quando pensara que o universo era imutável. Depois da descoberta de Hubble da expansão do universo, já não era mais necessária uma constante cosmológica, para a qual se toma atualmente o valor zero. Porém, foram encontrados outros espaços-tempos mais razoáveis, também permitidos pela relatividade geral, que admitem a viagem ao passado. Um é característico do interior dos buracos negros em rotação; outro é um espaço-tempo que contém duas cordas cósmicas que se movem uma no passado da outra a grande velocidade. Tal como o nome sugere, as cordas cósmicas são objetos que têm comprimento, mas secção desprezável. De facto, são mais como elásticos, pois encontram-se sob enorme tensão, da ordem de 1 milhão de milhões de milhões de toneladas. Uma corda cósmica ligada à Terra acelerá-la-ia de 0 a 100 quilómetros por hora em 1/30 de segundo. As cordas cósmicas soam a ficção científica, mas há razões para crer que podiam ter-se originado no universo primitivo como resultado da quebra de simetria analisada no capítulo 5. Porque se encontram sob enorme tensão e podem ter sido originadas em qualquer configuração, podem acelerar a muito grandes velocidades quando são relaxadas.

Na solução de Gödel e na solução da corda cósmica o espaço-tempo começa tão distorcido que a viagem ao passado é sempre possível. Deus podia ter criado um universo deformado, mas não temos razões para acreditar que tenha sido assim. As observações do fundo cósmico de micro-ondas e a abundância de elementos leves indicam que o universo primitivo não era dotado da curvatura adequada à viagem no tempo. À mesma conclusão pode chegar-se teoricamente se se admitir que o universo não tem fronteiras. Assim, a questão é a seguinte: se o universo teve início sem a curvatura requerida para a viagem no tempo, poderemos deformar o espaço-tempo de forma a permiti-la?

Um problema estreitamente relacionado com este, que também preocupa os autores de ficção científica, é a viagem interestelar ou intergaláctica rápida. De acordo com a relatividade, nada pode viajar mais depressa do que a luz. Assim, se enviássemos uma nave espacial à estrela mais próxima de nós, Alfa do Centauro, que está a cerca de quatro anos-luz, passar-se-iam cerca de 8 anos até que os viajantes regressassem para nos contarem o que lá encontraram. Se a expedição fosse ao centro da Galáxia, decorreriam, pelo menos, 100.000 anos até que regressassem. A teoria da relatividade é a consolação de cada um. Estas viagens dão origem ao paradoxo dos gémeos, já mencionado no capítulo 2.

Porque não há um padrão único de tempo, e cada observador mede o seu tempo próprio através do seu relógio, é possível que aos viajantes do espaço a viagem pareça ter sido muito mais curta do que aos que ficaram na Terra. No entanto, a alegria deles não será tão grande ao voltarem após uma viagem no espaço, uns tantos anos mais velhos, para descobrirem que os que ficaram na Terra morreram há milhares de anos. Assim, de modo a criarem interesse

pelas suas histórias, os autores de ficção científica imaginaram que um dia descobriremos como viajar mais rapidamente do que a luz. O que a maior parte destes autores ainda não entendeu é que, se for possível viajar mais rapidamente do que a luz, então a teoria da relatividade implicará que podemos viajar para trás no tempo, tal como os versos seguintes parecem anunciar:

Havia uma jovem corajosa

Que viajava mais depressa do que a luz.

Um dia ela partiu

Pelo caminho destinado

E só regressou na noite anterior.

A questão consiste em que a teoria da relatividade diz que não há uma medida única do tempo sobre a qual todos os observadores estejam de acordo. Pelo contrário, cada observador tem a sua própria medida do tempo. Se for possível um foguetão viajar mais depressa do que a luz para, partindo do acontecimento A, por exemplo o final da corrida dos 100 metros nos Jogos Olímpicos de 2012, chegar ao acontecimento B, por exemplo a abertura da centésima milésima quarta reunião do Congresso de Alfa do Centauro, todos os observadores concordarão que o evento A aconteceu antes do evento B, de acordo com os seus relógios. Suponhamos, porém, que a nave espacial tinha de Viajar mais depressa do que a luz para levar as novidades da competição ao congresso.

Então os observadores, que se movem a velocidades diferentes, podem discordar sobre qual dos eventos ocorre primeiro. De acordo com o relógio de um observador que se encontre em repouso relativamente à Terra, pode muito bem acontecer que o congresso abra depois da competição olímpica. Este observador pensará que a nave espacial só pode viajar de A, a B, a uma velocidade superior à da luz. Porém, a um observador que se encontre em Alfa do Centauro, afastando-se da Terra quase à velocidade da luz, parecerá que o acontecimento B, a abertura do Congresso, aconteceu antes do evento A, a competição dos 100 metros. A teoria da relatividade diz que as leis da física são as mesmas relativamente a observadores que se movem com velocidades diferentes.

Todos estes factos foram testados pela experiência e tudo leva a crer que permanecerão válidos mesmo que venhamos a substituir a relatividade por uma teoria mais avançada. Assim, o observador em movimento dirá que, se a viagem a uma velocidade superior à da luz for possível, será igualmente possível viajar do evento B, a abertura do Congresso, para o evento A, a competição dos 100 metros. Se viajarmos um pouco mais depressa, poderemos mesmo regressar antes da competição e fazer uma aposta com a certeza de que ganharemos.

Há um problema com a passagem da barreira da velocidade da luz. Segundo a teoria da relatividade, a potência necessária para acelerar uma nave espacial torna-se cada vez maior à medida que a velocidade desta se aproxima da velocidade da luz. Há evidência experimental deste facto, não com naves espaciais, mas com partículas elementares em aceleradores, no Fermilab ou no CERN (Centro Europeu de Pesquisa Nuclear). Podemos acelerar partículas à velocidade 99,99% da velocidade da luz, mas, por maior que seja a energia disponibilizada no processo, não conseguimos fazê-las passar a barreira da velocidade da luz. O mesmo acontece com as naves espaciais: seja qual for a potência do foguetão, elas não podem adquirir velocidades superiores à velocidade da luz.

Este facto parece arrumar, quer as viagens rápidas no espaço, quer as viagens para trás no tempo. Porém, há uma outra possibilidade. Talvez seja possível deformar o espaço-tempo de forma a atalhar entre A e B. Uma maneira de consegui-lo seria através da criação de um buraco de verme entre A e B. Como o nome sugere, um buraco de verme é um tubo fino de espaço-tempo que pode ligar duas regiões distantes quase planas.

Não há necessariamente uma relação entre a distância percorrida através de um buraco de verme e a separação das suas extremidades no espaço comum. Assim, pode imaginar-se que alguém pode criar ou descobrir um buraco de verme que vá das vizinhanças do sistema solar até Alfa do Centauro. A distância através do buraco de verme pode ser apenas de alguns milhões de quilómetros, apesar de a Terra e Alfa do Centauro distarem cerca de 30 milhões de milhões de quilómetros no espaço comum. Este facto permitiria que as notícias da corrida dos 100 metros chegassem à abertura do Congresso. Então um observador que se movesse em direção à Terra também seria capaz de encontrar outro buraco de verme que lhe permitisse vir da abertura do Congresso em Alfa do Centauro para trás, para a Terra, antes de a corrida começar. Deste modo, um buraco de verme, tal como qualquer outra forma de viajar mais depressa do que a luz, permitiria voltar ao passado.

A ideia de buracos de verme entre diferentes regiões do espaço-tempo não foi uma invenção dos autores de ficção científica; teve origem numa fonte muito respeitável. Em 1935, Einstein e Natham Rosen escreveram um artigo em que mostraram que a relatividade geral permitia aquilo a que chamaram «pontes», agora designadas por buracos de verme. As pontes de Einstein-Rosen não duravam o tempo suficiente para que uma nave espacial passasse através delas: navegaria antes em direção a uma singularidade quando o buraco de verme se cindisse. Contudo, foi sugerido que seria possível que uma civilização avançada mantivesse aberto um buraco de verme. Pode mostrar-se que para o conseguir, ou para deformar o espaço-tempo de qualquer outra maneira de modo a tornar possível viajar no tempo, é necessária uma região do espaço-tempo com curvatura negativa, como a da superfície de uma sela. A matéria comum, que tem densidade de energia positiva, confere ao espaço-tempo uma curvatura positiva, como a da superfície de uma esfera. Deste modo, o que é necessário para deformar o espaço-tempo de maneira a permitir viajar para o passado é matéria com densidade de energia negativa.

A energia é um pouco como o dinheiro: se o saldo é positivo, pode ser distribuída de várias maneiras, mas, de acordo com as leis clássicas em que se acreditava no princípio do século, não era possível dar mais do que se tinha. Assim, estas leis clássicas excluíam a possibilidade de qualquer viagem no tempo. Contudo, como se descreveu nos capítulos anteriores, as leis clássicas foram substituídas pelas leis quânticas, centradas no princípio da incerteza. As leis quânticas são mais liberais e permitem levantar mais do que se tem numa de duas contas, desde que o balanço total seja positivo. Por outras palavras, a teoria quântica permite que a densidade da energia seja negativa nalgumas regiões, desde que seja contrabalançada por uma densidade de energia positiva noutras regiões, de modo que a energia total seja positiva. Um exemplo de como a teoria quântica pode permitir densidades de energia negativas é o efeito de Casimir. Como vimos no capítulo 7, mesmo aquilo que julgamos ser o espaço «vazio» é preenchido com pares de partículas e antipartículas virtuais que aparecem juntas, se afastam e voltam a juntar-se e se aniquilam umas às outras. Consideremos duas lâminas de metal paralelas a curta distância uma da outra. As lâminas atuarão como espelhos para os fótons virtuais ou partículas de luz.

De facto, formarão entre elas uma cavidade, um tanto parecida com um tubo de órgão que apenas ressoa com certas notas. Significa isto que os fótons virtuais só poderão ocorrer no espaço entre as lâminas se a separação das lâminas for um múltiplo dos seus comprimentos de onda (a distância entre cristas sucessivas de uma onda). Se a largura da cavidade for um múltiplo mais uma fração do comprimento de onda, então, após algumas reflexões para trás e para diante entre as lâminas, a crista de uma onda coincidirá com as depressões de outra e as ondas anular-se-ão.

Como os fótons virtuais entre as lâminas apenas podem ter comprimentos de onda ressonantes, haverá ligeiramente menos fótons entre as lâminas do que no espaço exterior às lâminas, onde os fótons virtuais podem ter qualquer comprimento de onda. Assim, haverá ligeiramente menos fótons virtuais a chocar nas superfícies interiores das lâminas do que nas superfícies exteriores. Poder-se-ia esperar então uma força nas lâminas que as comprimiria uma contra a outra. Esta força foi realmente detetada e tem o valor previsto. Temos evidência experimental da existência de partículas virtuais e dos seus efeitos reais.

O facto de haver menos fótons virtuais entre as lâminas significa que a sua densidade de energia será menor do que em qualquer outro lugar. Mas a densidade de energia total no espaço «vazio» longe das lâminas tem de ser nula, porque, de outro modo, a densidade de energia deformaria o espaço, que não seria quase plano. Deste modo, se a densidade de energia no espaço entre as lâminas é menor do que a densidade de energia no espaço longe das lâminas, ela tem de ser negativa.

Temos, assim, evidência experimental de que o espaço-tempo pode ser deformado (da inclinação da luz durante os eclipses) e de que pode ser curvo de forma adequada a permitir viajar no tempo (do efeito de Casimir).

Poderemos esperar então que, à medida que a ciência e a tecnologia progredirem, venhamos a construir uma máquina do tempo. Mas como se compreende que ninguém tenha voltado do futuro para nos contar como fazê-lo? Dado que o nosso estado de desenvolvimento é primitivo, deve haver boas razões para ninguém ser imprudente ao ponto de nos revelar o segredo da viagem no tempo e, a menos que a natureza humana venha a mudar radicalmente, é difícil acreditar que algum viajante do futuro nos desvende o segredo. É claro que algumas pessoas dirão que a visão dos OVNIS é a prova de que somos visitados por alienígenas ou por pessoas vindas do futuro. (Se os alienígenas chegassem aqui em tempo razoável, precisariam de viajar mais rapidamente do que a luz, pelo que as duas possibilidades são equivalentes.)

No entanto, penso que qualquer visita de alienígenas ou de outras pessoas do futuro seria mais óbvia e, provavelmente, muito mais desagradável. Se eles quiserem mostrar-se, então porquê fazê-lo apenas àqueles que não são testemunhas credíveis? Se procuram acautelar-nos de algum perigo, não são muito convincentes.

Uma explicação possível para a ausência de visitantes do futuro consiste em assumir que o passado é fixo porque o observamos e vemos que não tem a espécie de deformação adequada à viagem ao passado. Por outro lado, o futuro é desconhecido e aberto, pelo que pode ter a curvatura adequada, o que significaria que qualquer viagem no tempo estaria confinada ao futuro. Não haveria qualquer hipótese de o capitão Kirk e a nave *Enterprise* voltarem ao presente.

Isto explica por que não fomos abordados por turistas vindos do futuro, mas não evita os problemas com que se confrontaria alguém que, voltando ao passado, viesse a modificar a sua história. Suponhamos, por exemplo, que o leitor voltava atrás e assassinava o seu tetravô quando este era uma criança. Há muitas versões deste paradoxo, mas, na essência, são equivalentes: haveria contradições se alguém pudesse alterar o passado.

Parece haver duas soluções possíveis para os paradoxos da viagem no tempo. Uma delas será designada por abordagem das histórias coerentes. Segundo esta, mesmo que o espaço-tempo fosse deformado, de modo a tornar possível a viagem ao passado, o que viesse a acontecer ao espaço-tempo seria coerente com as leis da física. De acordo com este ponto de vista, só seria possível voltar atrás no tempo se a história evidenciasse que já se tinha estado lá e que não se matara o tetravô ou cometera qualquer ato que viesse a entrar em conflito com o presente. Ao voltar ao passado, não seria possível alterar a história, o que significa que não se seria livre de fazer o que se quisesse. É claro que podemos argumentar que o livre arbítrio é uma ilusão. Se, na realidade, existe uma teoria unificada que tudo rege, é provável que ela determine também os nossos atos. Mas, a ser assim, fá-lo de uma forma que é impossível de calcular para um organismo tão complicado como é o ser humano. A razão por que se diz que o ser humano tem livre arbítrio reside no facto de não ser possível prever o que fará. No entanto, se um ser humano partir numa nave espacial e voltar para trás num momento anterior à partida,

seremos capazes de prever o que fará porque fará parte de uma história recordada. Assim, nesta situação, o viajante do tempo não terá livre arbítrio.

A outra solução para os paradoxos da viagem no tempo será designada por abordagem das histórias alternativas. A ideia consiste em que, quando os viajantes do tempo voltam ao passado, entram em histórias alternativas diferentes das histórias recordadas. Assim, podem agir livremente, sem o constrangimento da coerência com as suas histórias anteriores. Steven Spielberg jogou com esta ideia em *Regresso ao Futuro*: Marty McFly pôde voltar e alterar a relação dos pais para obter uma história mais satisfatória.

A hipótese das histórias alternativas assemelha-se à força como Richard Feynman exprimiu a teoria quântica como uma soma de histórias, que descrevemos nos capítulos 4 e 8. Segundo esta, o universo não tem uma só história: tem várias histórias possíveis, cada uma delas com a sua probabilidade. No entanto, parece haver uma diferença importante entre a hipótese de Feynman e a hipótese das histórias alternativas. Na soma de Feynman cada história é um espaço-tempo mais tudo o que nele existe. O espaço-tempo pode ser tão deformado que é possível viajar ao passado num foguetão. Mas o foguetão permaneceria no mesmo espaço-tempo e a história seria a mesma, teria de ser coerente. Assim, a hipótese da soma de histórias de Feynman parece fundamentar mais a abordagem das histórias coerentes do que a abordagem das histórias alternativas.

A soma de histórias de Feynman permite voltar ao passado à escala microscópica. No capítulo 9 vimos que as leis da física são imutáveis através de combinações das operações C , P e T . Significa isto que uma antipartícula que gira no sentido contrário ao dos ponteiros do relógio e se move de A para B pode ser considerada uma partícula que gira no sentido dos ponteiros do relógio e se move para trás no tempo de A para B . De modo análogo, uma partícula que se move para diante no tempo é equivalente a uma antipartícula que se move para trás no tempo. Como se analisou neste capítulo e no capítulo 7, o espaço «vazio» está cheio de pares de partículas e antipartículas virtuais que aparecem juntas, se afastam e voltam a juntar-se e se aniquilam umas às outras.

Assim, podemos considerar o par de partículas uma só partícula que se move num arco fechado no espaço-tempo. Quando o par se move para a frente no tempo (do evento em que é criado para o evento em que é aniquilado), designa-se por partícula. Mas, quando o par se move para trás no tempo (do evento em que é aniquilado para o evento em que é criado), diz-se que é uma antipartícula que viaja para a frente no tempo.

A explicação (dada no capítulo 7) de como os buracos negros podem emitir partículas e radiação foi a de que um membro de um par virtual partícula/antipartícula (por exemplo, uma antipartícula) pode cair no buraco negro, deixando o outro membro sem o par aniquilante. A partícula abandonada pode cair no buraco, mas também pode escapar da vizinhança do

buraco negro. Se assim acontecer, a um observador à distância parecerá que a partícula foi emitida pelo buraco negro.

No entanto, podemos ter uma representação intuitiva, diferente, mas equivalente, do mecanismo de emissão de um buraco negro. Podemos considerar o membro do par virtual que cai no buraco negro (por exemplo, a antipartícula) uma partícula que viaja para trás no tempo para fora do buraco negro. Quando chega ao ponto de reencontro do par virtual partícula/antipartícula, é dispersada pelo campo gravitacional numa partícula que viaja para a frente no tempo e escapa ao buraco negro. Se, em vez da antipartícula, fosse a partícula a cair no buraco, podíamos considerá-la uma antipartícula que viaja para trás no tempo e sai do buraco negro. Assim, a radiação pelos buracos negros mostra que a teoria quântica permite viajar para trás no tempo numa escala microscópica e, deste modo, essa viagem no tempo pode produzir efeitos observáveis.

Podemos então perguntar: permitirá a teoria quântica a viagem no tempo a uma escala macroscópica, de modo que uma pessoa possa realizá-la? À primeira vista, parece que sim. A hipótese da soma de histórias de Feynman parece estar acima de todas as histórias. Deste modo, deve incluir as histórias em que o espaço-tempo é tão deformado que é possível voltar ao passado. Então por que não estamos preocupados com a história? Suponhamos, por exemplo, que alguém retrocedeu no tempo e deu aos nazis o segredo da bomba atômica.

Poderemos evitar estes problemas se for válida a conjectura de proteção da cronologia. Segundo esta, as leis da física conspiram para impedir que os corpos *macroscópicos* transportem informação para o passado. Tal como a hipótese da censura cósmica, não foi provada, mas há razões para acreditar que seja verdadeira.

A razão para acreditar que a proteção da cronologia é um princípio ativo consiste em que, quando o espaço-tempo é deformado o suficiente para permitir viajar ao passado, as partículas virtuais que se movem em trajetórias fechadas no espaço-tempo podem transformar-se em partículas reais que viajam para a frente no tempo a uma velocidade igual ou inferior à velocidade da luz. Como estas partículas podem andar à volta um grande número de vezes, passam muitas vezes por cada ponto da sua trajetória. Deste modo, a sua energia é continuamente incrementada e a densidade da energia torna-se muito grande, o que poderá dar ao espaço-tempo uma curvatura positiva que não permite voltar ao passado. Ainda não é claro se estas partículas causam curvatura positiva ou negativa ou se a curvatura produzida por algumas espécies de partículas virtuais pode anular a produzida por outras espécies. Assim, a possibilidade de viajar no tempo continua em aberto. Não me arrisco, no entanto, a apostar nela. O meu opositor pode ter a desonesta vantagem de conhecer o futuro.

CAPÍTULO 11

- A UNIFICAÇÃO DA FÍSICA -

Como foi explicado no capítulo 1, seria muito difícil construir de uma só vez uma teoria unificada do universo. Por isso, temos progredido por elaboração de teorias parciais que explicam certos fenômenos, mas não explicam outros ou, quando muito, facultam uma explicação numérica. A química, por exemplo, permite-nos calcular as interações entre os átomos, sem que para isso se torne necessário conhecer a estrutura interna do núcleo de um átomo. No fim, contudo, tem-se a esperança de descobrir uma teoria unificada, coerente e completa, que inclua todas essas teorias parciais como aproximações e que não precise de ser ajustada, através de uma escolha judiciosa de valores de certos parâmetros arbitrários, para condizer com os factos. A procura de uma tal teoria é conhecida por «unificação da física». Einstein passou a maior parte dos seus últimos anos numa busca sem êxito de uma teoria unificada, mas a altura ainda não tinha chegado: havia teorias parciais para a gravidade e para o eletromagnetismo, mas sabia-se muito pouco sobre as forças nucleares. Além disso, Einstein recusava-se a acreditar na realidade da mecânica quântica, apesar do papel importante que ele próprio tivera no seu desenvolvimento. Além disso, parece que o princípio da incerteza é uma característica fundamental do mundo em que vivemos. Uma teoria unificada bem-sucedida deve, pois, incorporar este princípio.

Como descreverei, as perspectivas de descobrir tal teoria parecem ser muito melhores agora, porque sabemos muito mais sobre o universo. Mas devemos ter cuidado com excessos de confiança; já tivemos outros alvoreceres falsos! No começo deste século, por exemplo, pensava-se que tudo podia ser explicado através das propriedades da matéria contínua, como a elasticidade e a condução do calor. A descoberta da estrutura atômica e do princípio da incerteza pôs fim a tudo isso. Ainda mais, em 1928 Max Born, o físico detentor do Prémio Nobel, afirmou a um grupo de visitantes da Universidade de Göttingen: «A física, tal como a conhecemos, estará acabada em seis meses.» A sua confiança tinha como base a recente descoberta de Dirac da equação que comandava o electrão. Pensava-se que uma equação semelhante governaria o protão, que era a única outra partícula conhecida na altura, e que isto seria o fim da física teórica. Contudo, a descoberta do neutrão e das forças nucleares também deu cabo dessa ideia. Tendo dito isto, continuo a acreditar que há motivos para um otimismo cauteloso quanto a podermos estar agora perto do fim da busca das leis definitivas da natureza.

Em capítulos anteriores descrevi a relatividade geral como teoria parcial da gravidade e as teorias parciais que governam as forças fraca, forte e eletromagnética. As últimas três podem ser combinadas nas chamadas teorias

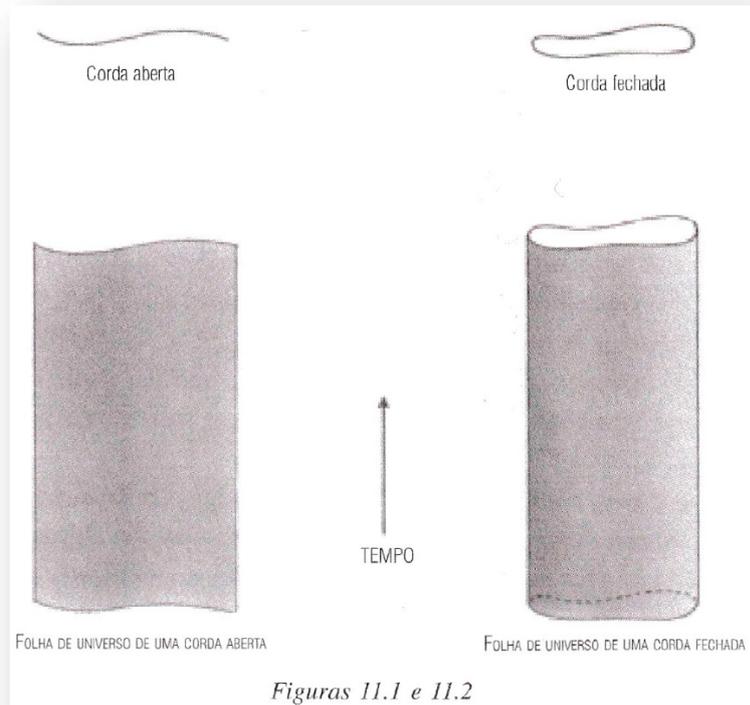
da grande unificação, ou GUT, que não são muito satisfatórias, porque não incluem a gravidade e contêm um conjunto de quantidades, como as massas relativas das diferentes partículas, que não podem ser previstas pela teoria, mas antes têm de ser escolhidas a fim de condizerem com as observações. A dificuldade principal para encontrar uma teoria que unifique a gravidade com as outras forças é o facto de a relatividade geral ser uma teoria «clássica», ou seja, não inclui o princípio da incerteza da mecânica quântica. Por outro lado, as outras teorias parciais dependem da mecânica quântica de uma forma essencial. O primeiro passo a dar consiste em combinar a relatividade geral com o princípio da incerteza. Como vimos, isto pode trazer algumas consequências notáveis, tais como buracos negros não tão negros e universos sem singularidades, que se contêm a si mesmos e não têm fronteiras. O problema reside em que, como expliquei no capítulo 7, o princípio da incerteza significa que mesmo o espaço «vazio» está cheio de pares de partículas virtuais e respetivas antipartículas. Estes pares teriam uma quantidade infinita de energia e, portanto, segundo a famosa equação de Einstein, $E = mc^2$, teriam uma quantidade de massa infinita. A sua atração gravitacional encurvaria, assim, o universo até um tamanho infinitamente pequeno.

De um modo muito semelhante, nas outras teorias parciais surgem infinidades aparentemente absurdas, mas em todos esses casos as infinidades podem ser removidas por um processo designado por renormalização, que consiste na remoção de infinidades por introdução de outras infinidades. Embora esta técnica seja matematicamente bastante duvidosa, parece funcionar na prática e já foi usada nestas teorias para elaborar previsões que concordam com as observações com extraordinária precisão. A renormalização, contudo, tem um inconveniente sério no que diz respeito a encontrar uma teoria universal, porque os valores das massas e das intensidades das forças não podem ser previstos pela teoria, mas devem ser escolhidos a fim de condizerem com as observações.

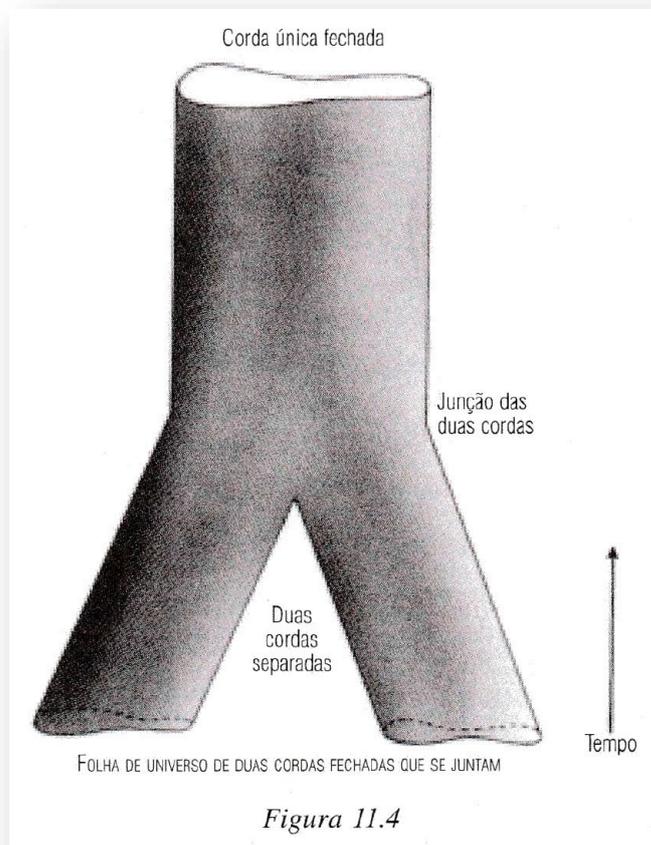
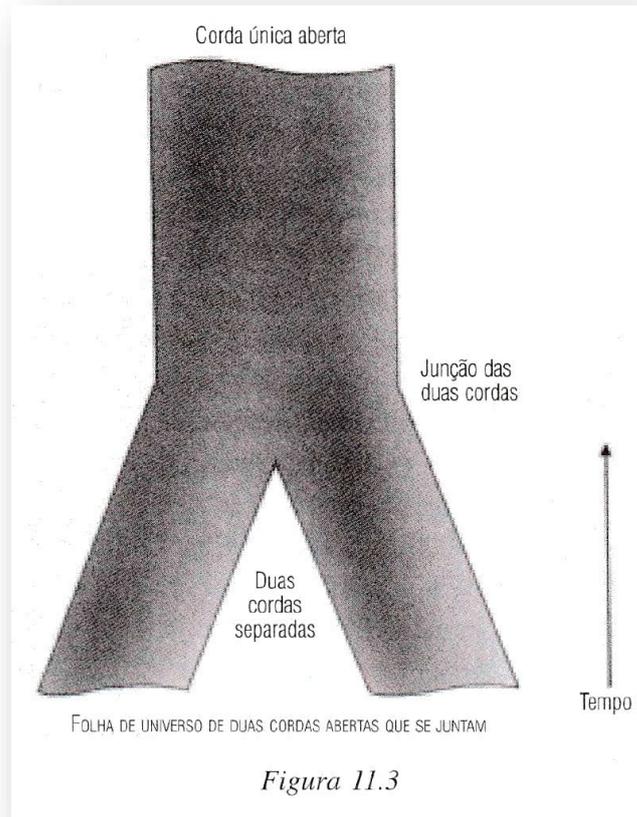
Ao tentarmos incorporar o princípio da incerteza na relatividade geral, só temos duas quantidades que podem ser ajustadas: a força da gravidade e o valor da constante cosmológica. Mas não basta ajustá-las para remover todas as infinidades. Temos, portanto, uma teoria que parece prever que determinadas quantidades, como a curvatura do espaço-tempo, são realmente infinitas, mas essas quantidades podem ser observadas e medidas como se fossem perfeitamente finitas! Esta dificuldade em combinar a relatividade geral com o princípio da incerteza já se adivinhava há algum tempo, mas foi, finalmente, confirmada por cálculos pormenorizados em 1972. Quatro anos depois foi sugerida uma solução possível, chamada «supergravidade». A ideia consistia em combinar a partícula de *spin* 2, chamada gravitão, que transmite a força gravitacional, com algumas outras partículas novas de *spin* $3/2$, 1, $1/2$ e 0. Num certo sentido, todas essas partículas podiam então ser encaradas como aspetos diferentes da mesma «superpartícula», que unificava, assim, as partículas de matéria, de *spin* $1/2$ e $3/2$, e as partículas portadoras de força, de *spin* 0, 1 e 2. Os pares virtuais partícula/antipartícula de *spin* $1/2$ e $3/2$, teriam energia negativa e, deste modo, tenderiam a anular a energia positiva dos

pares virtuais de *spin* 2, 1 ou 0. Isto permitiria a remoção de muitas das possíveis infinidades, mas suspeitou-se de que podiam permanecer algumas. Todavia, os cálculos necessários para verificar se algumas das infinidades tinham ficado ou não por remover eram tão longos e difíceis que ninguém estava preparado para o empreendimento. Mesmo com auxílio de um computador, concluiu-se que seriam necessários, pelo menos, quatro anos e as probabilidades de cometer pelo menos um erro, ou mesmo mais, eram muito grandes. Nestas circunstâncias, só se saberia se a resposta estava certa se alguém mais repetisse o cálculo e chegasse ao mesmo resultado, o que parecia pouco provável!

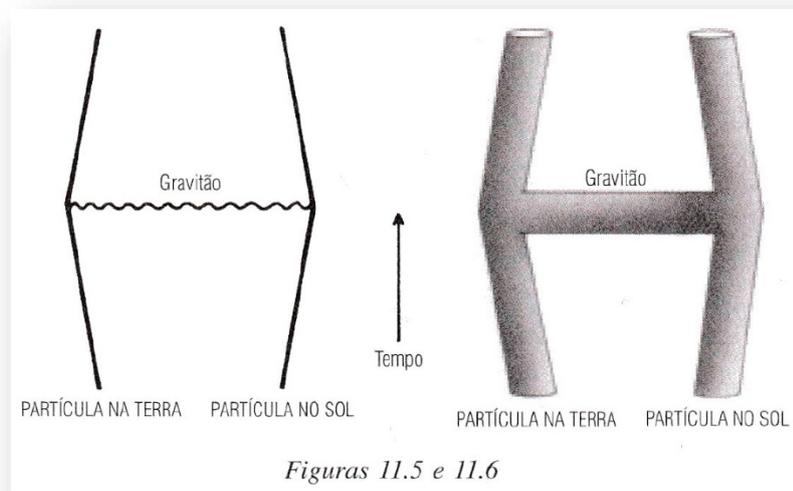
Apesar destes problemas, e do facto de as partículas nas teorias da supergravidade não parecerem condizer com as partículas observadas, muitos cientistas acreditaram que a supergravidade era, provavelmente, a resposta correta para o problema da unificação da física. Parecia o melhor caminho para a unificação da gravidade com as outras forças. Porém, em 1984 houve uma mudança notável de opinião a favor das chamadas teorias das cordas. Nestas teorias os objetos fundamentais não são partículas que ocupam um simples ponto no espaço, mas entidades que só têm comprimento e nenhuma outra dimensão, como um segmento de fio infinitamente fino. Estas cordas podem ter extremidades (as chamadas cordas abertas) ou estar unidas sobre si próprias em anéis fechados (cordas fechadas) (figs. 11.1 e 11.2). Uma partícula ocupa um ponto do espaço em cada instante. Deste modo, a sua história pode ser representada por uma linha no espaço-tempo (*a linha de universo*). Uma corda, por outro lado, ocupa uma linha no espaço em cada momento. Assim, a sua história no espaço-tempo é uma superfície bidimensional chamada *folha de universo*. (Qualquer ponto de uma folha de universo pode ser descrito por dois números: um que especifica o tempo e outro que especifica a posição do ponto na corda.) A folha de universo de uma corda aberta é uma fita; os seus limites representam os trajetos no espaço-tempo dos extremos da corda (fig. 11.1). A folha de universo de uma corda fechada é um cilindro ou tubo (fig. 11.2); um corte do tubo é uma curva fechada que representa a posição da corda num momento particular.



Duas cordas podem ser unidas para formar outra corda; as cordas abertas juntam-se simplesmente nas extremidades (fig. 11.3), ao passo que a junção de cordas fechadas se assemelha à união das duas pernas de um par de calças (fig. 11.4). De modo semelhante, uma corda pode ser dividida em duas. Nas teorias das cordas, as partículas são representadas por ondas que progridem na corda como as ondas no fio vibrante de um papagaio de papel. A emissão ou absorção de uma partícula por outra corresponde à divisão ou junção de cordas. Por exemplo, nas teorias das partículas, a atração gravitacional entre o Sol e a Terra resulta da emissão de gravitões pelas partículas do Sol e da respetiva absorção de gravitões pelas partículas da Terra, e vice-versa (fig. 11.5). Nas teorias das cordas, este processo corresponde a um tubo ou cano em forma de H (fig. 11.6) (as teorias das cordas têm analogias com a canalização).



As duas hastes do H correspondem às partículas do Sol e da Terra e a barra horizontal corresponde ao gravitão que se desloca entre elas.

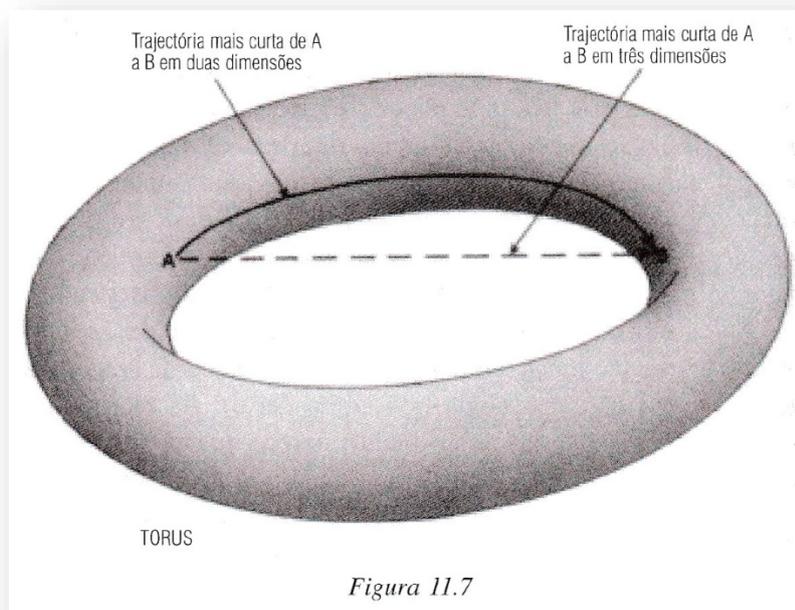


A teoria das cordas tem uma história curiosa. Foi inventada no fim dos anos 60, numa tentativa para descobrir uma teoria que descrevesse a força forte. A ideia era a de que as partículas, como o protão e o neutrão, podiam ser encaradas como ondas numa corda. A força forte entre partículas corresponderia a cordas que ligavam cordas a cordas, como numa teia de aranha. Para esta teoria dar o valor observado da força forte entre as partículas, as cordas tinham de ser como elásticos que suportassem tensões de cerca de 10 toneladas.

Em 1974 Jöel Scherk, de Paris, e John Schwarz, do Instituto de Tecnologia da Califórnia, publicaram um artigo em que mostraram que a teoria das cordas podia descrever a força gravitacional, mas só se a tensão na corda fosse muito mais elevada, da ordem de 1000 milhões de milhões de milhões de milhões de milhões de toneladas (1 seguido de 39 zeros). As previsões da teoria das cordas seriam exatamente as mesmas que as da relatividade geral a distâncias normais, mas seriam diferentes a distâncias muito pequenas, menores do que um milésimo de milionésimo de milionésimo de milionésimo de milionésimo de centímetro (1 centímetro dividido por 1 seguido de 33 zeros). Contudo, o seu trabalho não mereceu grande atenção, porque mais ou menos nessa altura muita gente abandonou a teoria das cordas da força forte a favor da teoria centrada nos quarks e nos gluões, que parecia concordar muito melhor com as observações. Scherk morreu em circunstâncias trágicas (sofria de diabetes e entrou em coma quando não tinha perto dele ninguém que lhe desse uma injeção de insulina). Deste modo, Schwarz ficou só, praticamente como único defensor da teoria das cordas, que advogava agora um valor muito mais elevado para a tensão. Em 1984 o interesse pela teoria reacendeu-se subitamente, por duas razões, parece. Uma deveu-se ao facto de as pessoas não estarem realmente a progredir na demonstração de que a supergravidade era finita ou que podia explicar as variedades de partículas que observamos. A outra foi a publicação de um

artigo de John Schwarz e Mike Green, do Queen Mary College, de Londres, que mostrava que a teoria das cordas podia explicar a existência de partículas de tendência sinistrogira, como algumas das partículas que observamos. Quaisquer que tenham sido as razões, um número apreciável de pessoas começou a trabalhar na teoria das cordas e foi desenvolvida uma nova versão, chamada corda heterótica, que parecia poder explicar os tipos de partículas que observamos.

As teorias das cordas conduzem também a infinitudes, mas pensa-se que são todas removíveis na versão da corda heterótica (embora isto não seja ainda tido como certo). As teorias das cordas, porém, têm um senão maior: parece que são coerentes apenas quando o espaço-tempo tem dez ou vinte e seis dimensões, em vez das quatro habituais! É certo que as dimensões extra do espaço-tempo são um lugar-comum na ficção científica; na verdade, fornecem uma maneira ideal de ultrapassar a restrição normal da relatividade geral segundo a qual não pode viajar-se mais depressa do que a luz ou retroceder no tempo (ver capítulo 10). A ideia da ficção científica consiste em, através das dimensões extra, ser possível encurtar caminho. Podemos representar isto do seguinte modo: imaginemos que o espaço em que vivemos tem apenas duas dimensões e que é curvo como a superfície de um toro (fig. 11.7).

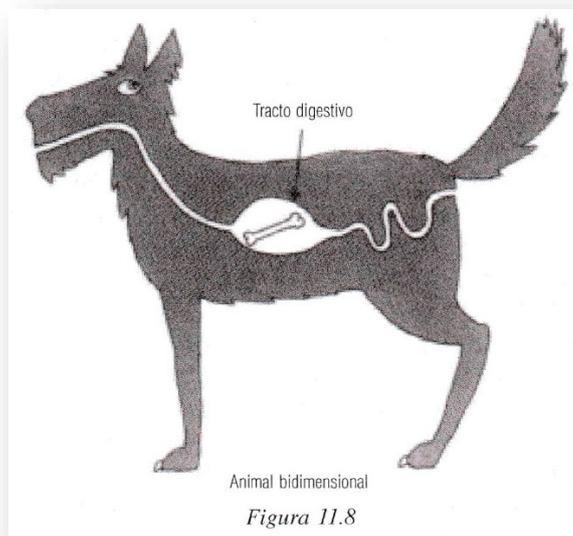


Se estivermos num ponto da superfície do toro e quisermos atingir o ponto diametralmente oposto, teremos de ir ao longo da superfície do toro. Porém, se fôssemos capazes de viajar na terceira dimensão, poderíamos atravessar a direito.

Por que não notamos todas essas dimensões extras se elas existem realmente? Por que vemos somente três dimensões do espaço e uma do tempo? A sugestão é que as outras dimensões estão encurvadas num espaço

muito pequeno, qualquer coisa como 1 milionésimo de milionésimo de milionésimo de milionésimo de centímetro. É tão pequeno que realmente não notamos: vemos somente uma dimensão temporal e três dimensões espaciais, em que o espaço-tempo é razoavelmente plano. É como a superfície de uma palha: se a observarmos de muito perto, verificaremos que é bidimensional (a posição de um ponto na palha é descrita por dois números, o comprimento ao longo da palha e a distância à roda do diâmetro). Mas, se a olharmos a certa distância, não veremos a espessura da palha, que parece ter apenas uma dimensão (a posição de um ponto é especificada apenas pela distância ao longo da palha). É isso que sucede com o espaço-tempo: numa escala muito pequena é decadimensional e fortemente encurvado, mas em escalas maiores não vemos a curvatura nem as dimensões extra. Se esta imagem for a correta, trará notícias desagradáveis aos pretensos viajantes do espaço: as dimensões extra são demasiadamente pequenas para conterem uma nave espacial. Todavia, surge ainda um problema maior: por que estão algumas dimensões, mas não todas, encaracoladas numa só? Presumivelmente, no início do universo todas as dimensões eram muito encurvadas. Por que razão uma dimensão temporal e três espaciais aplanaram, ao passo que as outras dimensões permaneceram fortemente encurvadas?

Uma resposta possível é o princípio antrópico. Duas dimensões espaciais não parecem suficientes para permitirem o desenvolvimento de seres complicados como nós. Por exemplo, animais bidimensionais que vivessem numa terra com uma dimensão teriam de trepar uns por cima dos outros para se ultrapassarem. Se uma criatura bidimensional comesse qualquer coisa que não conseguisse digerir completamente, teria de vomitar os restos da mesma maneira que os engolira, porque, se houvesse uma passagem através do seu corpo, esta dividiria a criatura em duas¹; o ser bidimensional desfar-se-ia (fig. 11.8). Da mesma maneira, é difícil ver como poderia haver circulação sanguínea numa criatura bidimensional. Haveria também problemas em mais de três dimensões espaciais².



A força gravitacional entre dois corpos diminuiria mais rapidamente com a distância do que diminui em três dimensões. (Em três dimensões a força gravitacional diminui para $1/4$ se se duplicar a distância. Em quatro dimensões diminuiria para $1/8$, em cinco para $1/16$, etc.) Isto significa que as órbitas dos planetas, como a da Terra em torno do Sol, seriam instáveis: a mais pequena perturbação numa órbita circular (como a que seria causada pela atração gravitacional dos outros planetas) resultaria num movimento espiralado da Terra, afastando-se ou dirigindo-se para o Sol. Ou gelaríamos ou arderíamos. Na realidade, o mesmo comportamento da gravidade com a distância a mais de três dimensões não permitiria ao Sol existir num estado estável, no qual a pressão equilibra a gravidade. Desfar-se-ia ou entraria em colapso, originando um buraco negro. Em qualquer dos casos, não seria de grande utilidade como fonte de calor e de luz para a vida da Terra. Numa escala mais pequena, as forças elétricas que forçam os eletrões a orbitar à volta do núcleo de um átomo comportar-se-iam como forças gravitacionais. Assim, os eletrões ou se escapavam completamente do átomo ou começavam a mover-se em espiral em direção ao núcleo. Em ambos os casos os átomos não seriam como os conhecemos.

Parece claro que a vida, pelo menos como a conhecemos, só pode existir em regiões do espaço-tempo em que três dimensões espaciais e uma temporal não estão encurvadas. Isto significa que podemos apelar para o princípio antrópico fraco, desde que possamos demonstrar que a teoria das cordas permite a existência de regiões deste tipo no universo — e parece que a teoria das cordas o permite de facto. Pode perfeitamente haver outras regiões do universo, ou outros universos (seja o que for que isso signifique), em que todas as dimensões estão enroscadinhas ou em que mais de quatro dimensões são quase planas, mas onde não há seres inteligentes para observarem o número diferente de dimensões.

Outro problema é que há, pelo menos, quatro teorias das cordas diferentes (a teoria das cordas abertas e três teorias das cordas fechadas) e milhões de maneiras em que as dimensões adicionais se encurvam. Por que seria escolhida uma única teoria das cordas e uma espécie de encurvamento? Durante algum tempo pareceu não haver resposta e o progresso sumiu-se. Então, a partir de 1994, as pessoas começaram a descobrir as chamadas dualidades: diferentes teorias das cordas e diferentes maneiras de encurvar as dimensões extra poderiam levar aos mesmos resultados em quatro dimensões. Além disso, tal como as partículas, que ocupam um simples ponto no espaço, e as cordas, que são linhas, descobriu-se que havia outros objetos, chamados membranas, que ocupam superfícies bidimensionais ou volumes de dimensão mais elevada. (Uma partícula pode ser considerada uma membrana de dimensão 0 e uma corda uma membrana de dimensão 1, mas também há membranas de dimensão p , de $p = 2$ a $p = 9$.) O que isto parece indicar é que há uma espécie de democracia entre supergravidade, cordas e membranas: parecem ajustar-se umas às outras, mas nenhuma pode dizer-se mais fundamental do que as outras. Parecem diferentes aproximações de uma teoria fundamental válidas em situações diversas.

Tem-se procurado esta teoria fundamental, mas sem êxito até agora. Acredito que não há uma formulação única da teoria fundamental, do mesmo modo que Gödel mostrou que não podia formular-se a aritmética em termos de um conjunto finito de axiomas. Em vez disso, deve ser como mapas — não pode usar-se um único mapa para descrever a superfície da Terra ou uma âncora: para cobrir todos os pontos são necessários, pelo menos, dois mapas no caso da Terra e quatro no caso de uma âncora. Cada mapa é apenas válido numa região limitada, mas diferentes mapas podem sobrepor-se em determinadas regiões. A coleção dos mapas fornece uma descrição completa da superfície. De modo análogo, em física pode ser necessário usar diferentes formulações em situações diversas, mas duas formulações diferentes concordarão em todas as situações em que podem ser ambas aplicadas. Um conjunto de formulações diferentes pode ser considerado uma teoria unificada, embora não possa ser expressa em termos de um único conjunto de postulados.

Mas pode realmente haver uma teoria unificada? Ou será que andamos a correr atrás de uma miragem? Há três possibilidades:

- 1) Há realmente uma teoria unificada completa (ou um conjunto de formulações justapostas), que descobriremos se para tal formos suficientemente espertos;
- 2) Não há qualquer teoria fundamental do universo, mas apenas uma sucessão infinita de teorias que descrevem o universo com uma precisão cada vez maior;
- 3) Não há qualquer teoria do universo: os acontecimentos só podem ser previstos em certa medida, além da qual ocorrem aleatória e arbitrariamente.

Algumas pessoas serão a favor da terceira possibilidade, argumentando que, se houvesse um conjunto completo de leis, isso infringiria a liberdade de Deus mudar de opinião e intervir no mundo. É um pouco como o velho paradoxo: Deus poderá tornar uma pedra tão pesada que não seja capaz de levantá-la? Mas a ideia de que Deus pode querer mudar de opinião é um exemplo da falácia, apontada por Santo Agostinho, de imaginarmos Deus como um ser que existe no tempo: o tempo é apenas uma propriedade do universo que Deus criou. Presume-se que sabia o que tencionava fazer quando o criou!

Com o advento da mecânica quântica acabámos por reconhecer que os acontecimentos não podem ser previstos com precisão perfeita e que haverá sempre um grau de incerteza. Se quisermos, podemos atribuir essa incerteza à intervenção de Deus, mas seria uma intervenção muito estranha: não existe qualquer prova de que seja dirigida para algum objetivo. Na realidade, se o fosse, seria, por definição, não aleatória. Nos tempos modernos libertámo-nos definitivamente da terceira hipótese, redefinindo os objetivos da ciência: o que se pretende é formular um conjunto de leis que permitam prever os acontecimentos até ao limite imposto pelo princípio da incerteza.

A segunda hipótese, a de haver uma sequência infinita de teorias cada vez mais aperfeiçoadas, está de acordo com toda a nossa experiência até agora. Em muitas ocasiões aumentámos a sensibilidade das nossas medições ou elaborámos um novo tipo de observações apenas para descobrirmos fenómenos novos que não tinham sido previstos pela teoria existente e para a explicação dos quais tivemos de desenvolver uma teoria mais avançada. Não seria, portanto, muito surpreendente se a atual geração de teorias da grande unificação estivesse errada ao afirmar que nada de essencialmente novo acontecerá entre a energia da unificação eletrofraca de cerca de 100 GeV e a energia da grande unificação de cerca de 1000 milhões de milhões de GeV. Poderíamos realmente esperar encontrar várias camadas novas de estrutura mais fundamental do que quarks e eletrões, que agora consideramos partículas «elementares».

No entanto, parece que a gravidade pode fornecer um limite para esta sequência de «caixas dentro de caixas». Se tivéssemos uma partícula com uma energia acima daquilo a que se chama a energia de Planck, ou seja, 10 milhões de milhões de milhões de GeV (1 seguido de 19 zeros), a sua massa estaria tão concentrada que se separaria do resto do universo e formaria um pequeno buraco negro. Portanto, parece realmente que a sucessão de teorias cada vez mais aperfeiçoadas deve ter um limite, uma vez que caminhamos para energias cada vez mais elevadas, pelo que deve existir alguma teoria definitiva do universo. É evidente que a energia de Planck está muito longe das energias de 100 GeV, que são o máximo que atualmente podemos atingir em laboratório. Não seremos capazes de ultrapassar esse obstáculo com aceleradores de partículas no futuro próximo. Contudo, o universo primitivo é esse laboratório onde tais energias devem ter ocorrido. Penso que existe uma boa probabilidade de o estudo do universo primitivo e os requisitos da coerência matemática levarem a uma teoria unificada completa ainda durante a vida de algumas das pessoas que andam por aí hoje, sempre presumindo que não nos destruiremos antes.

Que significado teria descobrirmos realmente a teoria definitiva do universo? Como explicámos no capítulo 1, nunca podemos ter a certeza absoluta de havermos descoberto realmente a teoria correta, uma vez que as teorias não podem ser provadas. Mas, se a teoria fosse matematicamente coerente e conduzisse sempre a previsões concordantes com a observação, podíamos confiar razoavelmente que seria a teoria certa. Chegaria, assim, o fim de um longo e glorioso capítulo na história da luta intelectual da humanidade para compreender o universo. Mas também revolucionaria a compreensão que as pessoas vulgares têm das leis que governam o universo. No tempo de Newton era possível a uma pessoa instruída abranger todo o conhecimento humano, pelo menos na generalidade. Mas, a partir de então, o desenvolvimento da ciência tornou isso impossível. Como as teorias estão sempre a ser modificadas para darem conta de novas observações, nunca chegam a ser devidamente digeridas ou simplificadas de modo que as pessoas vulgares possam compreendê-las. É preciso ser-se especialista, mas, mesmo assim, só pode esperar-se uma compreensão completa de uma pequena proporção das teorias científicas. Além disso, o progresso é tão rápido que

aquilo que se aprende na escola ou na universidade está sempre um pouco fora da moda. Apenas algumas pessoas conseguem manter-se a par da fronteira do conhecimento, que avança rapidamente, e têm de dedicar todo o seu tempo a isso e a especializar-se numa pequena área. O resto da população tem uma ideia precária dos avanços conseguidos ou da excitação que geram. Há setenta anos, a acreditar em Eddington, só duas pessoas compreendiam a teoria da relatividade geral. Hoje em dia dezenas de milhares de licenciados compreendem-na e muitos milhões de pessoas estão, pelo menos, familiarizadas com a ideia. Se fosse descoberta uma teoria unificada completa, em muito pouco tempo seria da mesma maneira digerida e simplificada e ensinada nas escolas, pelo menos na generalidade. Seríamos então capazes de ter alguma compreensão das leis que regem o universo e são responsáveis pela nossa existência.

Mesmo que cheguemos a descobrir uma teoria unificada completa, tal não significa que venhamos a ser capazes de prever acontecimentos em geral, por duas razões. A primeira é a limitação que o princípio da incerteza da mecânica quântica traz aos nossos poderes de previsão. E não há nada que possamos fazer para nos livrarmos dela. Na prática, no entanto, esta primeira limitação é menos restritiva do que a segunda, que advém do facto de não podermos resolver exatamente as equações da teoria, exceto em situações muito simples. (Nem sequer conseguimos resolver exatamente o movimento de três corpos na teoria da gravitação de Newton e a dificuldade aumenta com o número de corpos e a complexidade da teoria.) Já conhecemos todas as leis que governam o comportamento da matéria sob todas as condições, mesmo as mais extremas. Em particular, conhecemos as leis que estão na base de toda a química e de toda a biologia. E, no entanto, ainda não fomos capazes de reduzir esses temas à condição de problemas resolvidos; por enquanto tivemos pouco êxito com a previsão do comportamento humano a partir de equações matemáticas! Portanto, mesmo que encontremos realmente um conjunto de leis básicas, continuará a existir nos anos seguintes a tarefa intelectualmente desafiadora de desenvolver métodos de aproximação melhores para que possamos elaborar com êxito previsões das consequências prováveis em situações complicadas e realistas. Uma teoria unificada coerente e completa é apenas o primeiro passo: o nosso objetivo é a *compreensão* completa dos acontecimentos à nossa volta e da nossa própria existência.

¹ Há muitos factos notáveis a respeito desta argumentação! Aconselho vivamente o leitor interessado a ler (e verificará que o faz num ápice!) o ensaio de Rudy Rucker, *A Quarta Dimensão*, já publicado pela Gradiva. (N. do R. C.)

² Uma discussão deliciosa acerca da dimensionalidade do espaço encontra-se em *A Criação*, de P. W. Atkins, já publicado em língua portuguesa, numa edição da Editorial Presença. (N. da R. C.)

CAPÍTULO 12

- CONCLUSÃO -

Encontramo-nos num mundo desconcertante. Queremos que o que nos rodeia faça sentido e perguntamos: Qual é a natureza do universo? Qual é o nosso lugar no universo e donde provieram todas as pessoas? Por que é o universo como é?

Para tentarmos responder a estas perguntas adotámos uma «imagem do mundo». Nela, a teoria das cordas é como a torre infinita de tartarugas que sustém a Terra plana. Ambas são teorias do universo, embora a primeira seja muito mais matemática e precisa do que a segunda. A ambas faltam provas observacionais: jamais alguém viu uma tartaruga gigante com a Terra às costas, mas também nunca ninguém viu uma corda. Contudo, a teoria da tartaruga não é uma boa teoria científica porque prevê que as pessoas caiam na fronteira do mundo. Tal não concorda com a experiência, a não ser que venha a ser a explicação para as que se supõe terem desaparecido no triângulo das Bermudas!

As primeiras tentativas teóricas para descrever e explicar o universo envolviam a ideia de que os acontecimentos e os fenómenos naturais eram controlados por espíritos com emoções humanas, que agiam de uma maneira muito humana e imprevisível. Estes espíritos habitavam objetos naturais, como rios e montanhas, incluindo corpos celestes, como o Sol e a Lua. Tinham de ser apaziguados e os seus favores conquistados para se garantir a fertilidade do solo e o ciclo das estações do ano. Gradualmente, contudo, deve ter-se começado a reparar em certas regularidades: o Sol erguia-se sempre a leste e punha-se a oeste, quer se tivesse ou não oferecido um sacrifício ao deus Sol. Além disso, o Sol, a Lua e os planetas seguiam trajetórias precisas no céu, que podiam ser antecipadas com considerável precisão. O Sol e a Lua podiam ser deuses, mas eram deuses que obedeciam a leis estritas, aparentemente sem exceções, se ignorarmos histórias como a do Sol ter parado para Josué.

Inicialmente, estas regularidades e leis surgiram apenas na astronomia e em alguns outros contextos. Contudo, à medida que a civilização se desenvolveu, particularmente nos últimos 300 anos, foram descobertas cada vez mais regularidades e leis. O êxito dessas leis levou Laplace, no princípio do século XIX, a postular o determinismo científico, ou seja, sugeriu que devia haver um conjunto de leis que determinavam com precisão a evolução do universo, dada a sua configuração em determinado momento.

O determinismo de Laplace estava duplamente incompleto. Não dizia como deviam ser escolhidas as leis e não especificava qual teria sido a configuração inicial do universo. Tudo isso era deixado a Deus. Deus teria

escolhido o princípio do universo e a que leis o mesmo obedeceria, mas para não mais intervir¹. Com efeito, Deus foi confinado a áreas que a ciência do século XIX não compreendia.

Sabemos agora que as esperanças do determinismo de Laplace não podem realizar-se, pelo menos nos termos que ele tinha em mente. O princípio da incerteza da mecânica quântica implica que certos pares de quantidades, como a posição e a velocidade de uma partícula, não podem ser ambos previstos com toda a precisão.

A mecânica quântica trata esta situação por intermédio de uma classe de teorias quânticas em que as partículas não têm posições nem velocidades bem definidas, mas são representadas por ondas. Estas teorias quânticas são determinadas no sentido em que proporcionam leis para a evolução da onda com o tempo. Assim, se conhecermos a onda num dado momento, podemos calculá-la em qualquer outro. O elemento imprevisível e de acaso só entra quando tentamos interpretar a onda em termos de posição e de velocidade de uma partícula. Mas talvez seja esse o nosso erro: talvez não existam tais conceitos de posição e de velocidades de uma partícula, mas apenas ondas. Nós é que procuramos encaixar as ondas nas nossas ideias preconcebidas de posição e de velocidade. A má combinação resultante é a causa da aparente imprevisibilidade.

Com efeito, redefinimos a tarefa da ciência como sendo a descoberta de leis que nos permitam prever acontecimentos até aos limites impostos pelo princípio da incerteza. Mantém-se, no entanto, a pergunta: como e por que foram escolhidas as leis e o estado inicial do universo?

Neste livro dei lugar de destaque às leis que regem a gravidade, porque é a gravidade que dá forma à estrutura do universo em macroescala, embora seja a mais fraca das quatro forças. As leis da gravidade eram incompatíveis com a opinião, mantida até há pouco, de que o universo era imutável no tempo: o facto de a gravidade ser sempre atrativa implica que o universo deve estar a expandir-se ou a contrair-se. Segundo a teoria da relatividade geral, deve ter havido um estado de densidade infinita no passado, o *big bang*, que teria sido um começo efetivo do tempo. Do mesmo modo, se todo o universo voltasse a entrar em colapso, devia haver outro estado de densidade infinita no futuro, o *big crunch*, que seria um fim do tempo. Mesmo que todo o universo não entrasse de novo em colapso, haveria singularidades em todas as regiões que entrassem em colapso para formar buracos negros. Estas singularidades seriam um fim do tempo para quem caísse no buraco negro. No *big bang* e nas outras singularidades todas as leis perdiam a validade, de maneira que Deus teria tido liberdade completa para decidir o que acontece e como o universo começou.

Quando combinamos a mecânica quântica com a relatividade geral, parece haver uma nova possibilidade: que o espaço e o tempo juntos podem formar um espaço quadridimensional finito sem singularidades nem fronteiras, como a superfície da Terra, mas com mais dimensões. Parece que esta ideia

pode explicar muitas das características observadas no universo, como a sua uniformidade em macroescala e também os desvios da homogeneidade, em menor escala, como as galáxias, as estrelas e até os seres humanos. Podia até justificar a seta do tempo que observamos. Mas, se o universo é completamente independente, sem singularidades nem fronteiras, e completamente descrito por uma teoria unificada, isso tem implicações profundas quanto ao papel de Deus como Criador.

Einstein perguntou uma vez: «Que capacidade de escolha tinha Deus ao construir o universo?» Se a hipótese de não haver fronteiras está correta, não teve realmente liberdade de escolha quanto às condições iniciais. Teria, é claro, tido a liberdade de escolher as leis a que o universo obedeceria. Pode, no entanto, não ter havido escolha; pode haver só uma teoria, ou um pequeno número de teorias unificadas completas, tal como a teoria heterótica das cordas, que seja autoconsistente e permita a existência de estruturas tão complicadas como a dos seres humanos, que podem investigar as leis do universo e inquirir sobre a natureza de Deus.

Mesmo que haja só uma teoria unificada possível, não passa de um conjunto de normas e equações. O que é que dá vida às equações e forma ao universo por elas descrito? O propósito da ciência, quando se constrói um modelo matemático, não é responder à pergunta se existe um universo descrito pelo modelo. Por que se dá o universo ao trabalho de existir? A teoria unificada é tão imperativa que dá origem à sua própria existência? Ou precisa de um criador e, nesse caso, terá ele outro efeito sobre o universo? E quem o criou a ele?

Até agora a maior parte dos cientistas têm estado demasiado ocupados com o desenvolvimento de novas teorias que descrevem *o que* é o universo para perguntarem *porquê*. Por outro lado, as pessoas que deviam perguntar *porquê*, os filósofos, não foram capazes de acompanhar o avanço das teorias científicas. No século XVIII, os filósofos consideravam todo o conhecimento humano, incluindo a ciência, como campo seu e discutiam questões como: terá o universo tido um princípio? No entanto, nos séculos XIX e XX a ciência tornou-se demasiado técnica e matemática para os filósofos ou para qualquer outra pessoa, com exceção de alguns especialistas. Os filósofos reduziram o objetivo das suas pesquisas de tal modo que Wittgenstein, o filósofo mais famoso deste século, afirmou: «A única tarefa que resta à filosofia é a análise da linguagem.» Que queda para a grande tradição da filosofia desde Aristóteles até Kant!

Todavia, se descobirmos uma teoria completa, deve acabar por ser compreensível, na generalidade, para toda a gente, e não apenas para alguns cientistas. Então poderemos todos, filósofos, cientistas e pessoas comuns, tomar parte na discussão do porquê da nossa existência e da do universo. Se descobirmos a resposta, será o triunfo máximo da razão humana, porque nessa altura conheceremos o pensamento de Deus.

¹ A ideia de um deus relojoeiro: deu corda ao mundo por ele recém-criado e afastou-se para não mais intervir! No princípio do século XIX, o universo newtoniano está um pouco às avessas: o espaço substancial cartesiano, substituído depois pelo espaço newtoniano, através do qual Deus sentia o mundo e atuava nele continuamente, foi cedendo lugar ao espaço dos atomistas, vazio de substância e da presença de Deus. (N. do R. C.)

ALBERT EINSTEIN

A ligação de Einstein com a política da bomba nuclear é bem conhecida: ele assinou a famosa carta dirigida ao presidente Franklin Roosevelt que persuadiu os Estados Unidos a levarem a ideia a sério e participou nos esforços do pós-guerra para evitar a guerra nuclear. Mas estas não foram apenas ações isoladas de um cientista arrastado para o mundo da política. A vida de Einstein foi, na realidade, utilizando as suas próprias palavras, «dividida entre a política e as equações».

As primeiras atividades políticas de Einstein surgiram durante a Primeira Guerra Mundial, quando era professor em Berlim. Perturbado com o que considerava um desperdício de vidas humanas, envolveu-se em demonstrações contra a guerra. O seu apoio à desobediência civil e o encorajamento público contra o recrutamento pouco contribuíram para o tornar querido dos colegas. Depois, a seguir à guerra, canalizou os seus esforços no sentido da reconciliação e do melhoramento das relações internacionais. Também isto não o tornou popular e depressa a sua política lhe dificultou as visitas aos Estados Unidos, mesmo para fazer conferências.

A segunda grande causa de Einstein foi o sionismo. Embora fosse de ascendência judaica, Einstein rejeitava a ideia bíblica de Deus. Contudo, uma grande tomada de consciência do antissemitismo, quer antes, quer durante a Primeira Guerra Mundial, levou-o a identificar-se gradualmente com a comunidade judaica e, mais tarde, a tornar-se adepto do sionismo. Mais uma vez a impopularidade não o impediu de dizer o que pensava. As suas teorias começaram a ser atacadas e chegou a formar-se uma organização anti-Einstein. Um certo homem foi condenado por incitar outros a assassinar Einstein (e multado em apenas 6 dólares). Mas Einstein comentou fleumaticamente, quando foi publicado um livro intitulado *Cem Autores Contra Einstein*: «Se eu não tivesse razão, um teria sido suficiente.»

Em 1933 Hitler subiu ao poder. Einstein estava na América e declarou que não voltaria à Alemanha. Então, enquanto as milícias nazis faziam uma busca à sua casa e lhe confiscavam o dinheiro que tinha no banco, um jornal de Berlim publicava o cabeçalho «Boas notícias de Einstein – não regressa». Perante a ameaça nazi, Einstein renunciou ao pacifismo e, receando eventualmente que os cientistas alemães construíssem uma bomba nuclear, propôs que os Estados Unidos fizessem a sua. Mas, justamente antes de a primeira bomba atômica ser detonada, chamou publicamente a atenção para os perigos de uma guerra nuclear e propôs o controle internacional das armas nucleares.

Durante toda a vida, os esforços de Einstein em prol da paz tiveram, provavelmente, poucos êxitos duráveis e fizeram-no certamente ganhar poucos amigos. O seu apoio verbal à causa sionista, contudo, foi devidamente reconhecido em 1952, quando lhe ofereceram a presidência de Israel. Recusou, afirmando julgar-se demasiado ingênuo em matéria política. Mas

talvez a verdadeira razão tenha sido diferente, parafraseando-se de novo: «As equações são mais importantes para mim, porque a política é do presente e uma equação é qualquer coisa para a eternidade.»

GALILEU GALILEI

Galileu, talvez mais do que qualquer outra pessoa, foi responsável pelo nascimento da ciência moderna. O célebre conflito que travou com a Igreja católica foi o centro da sua filosofia, porque Galileu foi um dos primeiros a defender que o homem podia ter esperança de compreender o funcionamento do mundo e, além disso, poderia fazê-lo observando o mundo real.

Galileu acreditava na teoria de Copérnico (de que os planetas orbitavam em torno do Sol) desde o princípio, mas foi só quando descobriu as provas necessárias para apoiar esta ideia que começou a defendê-la publicamente. Escreveu sobre a teoria de Copérnico em italiano (não no usual latim acadêmico) e depressa as suas opiniões foram largamente apoiadas fora das universidades. Isto aborreceu os professores aristotélicos, que se uniram contra ele, tentando persuadir a Igreja católica a banir as ideias de Copérnico.

Galileu, preocupado, foi a Roma para falar com as autoridades eclesiásticas. Argumentou que a Bíblia não queria dizer-nos nada sobre as teorias científicas e que era costume admitir-se que, onde a Bíblia entrasse em conflito com o senso comum, estava a ser alegórica. Mas a Igreja tinha medo de um escândalo que pudesse prejudicá-la na sua luta contra o protestantismo e tomou medidas repressivas. Em 1616 declarou o copernicanismo «falso e errado» e deu ordens a Galileu para que nunca mais «defendesse ou apoiasse» a doutrina. Galileu concordou.

Em 1623 um amigo de longa data de Galileu tornou-se papa. Imediatamente Galileu tentou que o decreto de 1616 fosse revogado. Não o conseguiu, mas acabou por obter licença para escrever um livro que abordasse as teorias aristotélica e copernicana, com duas condições: não tomaria qualquer partido e chegaria à conclusão de que o homem não podia, em qualquer dos casos, determinar como funcionava o mundo, porque Deus podia causar os mesmos efeitos de maneiras não imagináveis pelo homem, que não podia restringir a sua onipotência.

O livro *Diálogo sobre os Dois Principais Sistemas do Mundo* ficou pronto e foi publicado em 1632, com todo o apoio dos censores, e foi imediatamente recebido por toda a Europa como uma obra-prima literária e filosófica. Logo a seguir, o papa, compreendendo que as pessoas viam o livro como um argumento convincente a favor do copernicanismo, arrependeu-se de o ter deixado publicar. Argumentou que, embora o livro tivesse a bênção oficial dos censores, Galileu tinha desobedecido ao decreto de 1616. Obrigou-o a apresentar-se perante a Inquisição, que o condenou a prisão domiciliária perpétua e lhe ordenou que renunciasse publicamente ao copernicanismo. Pela segunda vez Galileu concordou. E permaneceu católico fiel, mas a sua crença na independência da ciência não foi por isso esmagada. Quatro anos antes de morrer, em 1642, sempre sob prisão domiciliária, o manuscrito do seu segundo livro foi passado clandestinamente para um editor na Holanda. Era a

obra conhecida por *Duas Novas Ciências* que, ainda mais do que o seu apoio ao copernicanismo, haveria de tornar-se o germe da física moderna.

ISAAC NEWTON

Isaac Newton não era um homem agradável. As suas relações com os outros académicos ficaram célebres, pois, sobretudo nos últimos anos da sua vida, envolveu-se em acasas discussões. A seguir à publicação dos *Principia Mathematica*, certamente o livro de física mais influente que alguma vez foi escrito, Newton subiu rapidamente à proeminência pública. Foi nomeado presidente da Royal Society e foi o primeiro cientista a ser designado cavaleiro.

Newton depressa começou às turras com o astrónomo real, John Flamsteed, que lhe tinha fornecido muitas informações para os seus *Principia*, mas que começou então a negar-lhe as informações que ele queria. Newton não aceitava um não como resposta e tratou de se fazer nomear para a direção do Observatório Real, tentando logo a seguir obrigar à publicação das informações. Eventualmente, conseguiu que o trabalho de Flamsteed fosse confiscado e preparado para publicação pelo inimigo mortal deste, Edmond Halley. Mas Flamsteed levou o caso aos tribunais e, num instante, obteve destes uma ordem que proibia a distribuição do trabalho roubado. Newton ficou furioso e procurou vingar-se, retirando sistematicamente quaisquer referências a Flamsteed das edições seguintes dos *Principia*.

Discussão mais séria foi a que teve com o filósofo alemão Gottfried Leibniz. Tanto este como Newton tinham desenvolvido independentemente o ramo da matemática chamado cálculo, que está na base da maior parte da física moderna. Embora agora saibamos que Newton descobriu o cálculo antes de Leibniz, publicou o seu trabalho muito mais tarde. Começou assim uma enorme discussão sobre quem tinha sido o primeiro, com cientistas a defenderem vigorosamente os dois oponentes. É notável, porém, que a maioria dos artigos que apareceram a defender Newton tivessem sido escritos por ele mesmo e apenas publicados em nome dos amigos! No calor da discussão, Leibniz cometeu o erro de apelar à Royal Society para decidir a disputa. Newton, como presidente, nomeou uma comissão «imparcial» para investigar, formada, por coincidência, apenas por amigos seus! Mas isso não foi tudo: Newton escreveu depois o relatório da comissão e fez com que a Royal Society o publicasse, acusando oficialmente Leibniz de plágio. Como ainda não estava satisfeito, escreveu uma crítica anónima do relatório na publicação privada da Royal Society. Após a morte de Leibniz, diz-se que Newton declarou que tinha ficado radiante «por ter desfeito o coração de Leibniz». Durante estas duas discussões, Newton já tinha deixado Cambridge e a vida académica. Tivera um papel ativo na política anticatólica em Cambridge e mais tarde no Parlamento, tendo acabado por ser recompensado com o lucrativo lugar de administrador da Real Casa da Moeda. Aí usou os seus talentos tortuosos e vitriólicos de maneira mais aceite socialmente, orquestrando com êxito uma campanha contra a moeda falsa, chegando a mandar vários homens para a forca.

GLOSSÁRIO

Aceleração: Razão a que a velocidade de um objeto varia.

Acelerador de partículas: Máquina que, por meio de eletromagnetos, pode acelerar partículas carregadas em movimento, comunicando-lhes maior energia.

Acontecimento: Ponto do espaço-tempo especificado pelas suas coordenadas de lugar e de tempo.

Anã branca: É uma estrela fria, estável, mantida pela repulsão decorrente da aplicação do princípio de exclusão aos elétrons.

Antipartículas: A cada tipo de partícula de matéria corresponde um tipo de antipartícula. Quando uma partícula colide com uma antipartícula do tipo correspondente, aniquilam-se mutuamente com a libertação de energia.

Átomo: Unidade fundamental da matéria comum, constituída por um núcleo minúsculo (formado por prótons e neutrões) envolvido por uma nuvem eletrónica.

Big bang: Singularidade do começo do universo.

Big crunch: Singularidade do fim do universo.

Buraco de verme: Um fino tubo de espaço-tempo que liga regiões distantes do universo. Os buracos de verme também podem ligar universos paralelos, ou universos bebês, e podem permitir viajar no tempo.

Buraco negro: Região do espaço-tempo donde nada, nem mesmo a luz, pode escapar, porque a gravidade é muito intensa (capítulo 6).

Buraco negro primevo: Buraco negro criado no início do universo.

Campo: Algo que existe através do espaço e do tempo, por oposição a uma partícula, que pode ser encontrada num ponto de cada vez.

Campo magnético: Campo responsável pelas forças magnéticas, incorporado conjuntamente com o campo elétrico no designado campo eletromagnético.

Carga elétrica: Propriedade de uma partícula através da qual ela repete (ou atrai) outras partículas que têm carga do mesmo sinal (ou de sinal contrário).

Comprimento de onda: Distância entre duas cristas ou duas cavas sucessivas de uma onda.

Condição sem fronteira: Conceito de que o universo é finito, mas não tem fronteira (no tempo imaginário).

Cone de luz: Superfície do espaço-tempo que delimita as trajetórias possíveis dos raios luminosos que se cruzam num acontecimento definido.

Conservação de energia: Lei da física que estabelece que a energia (ou o seu equivalente em massa) não pode ser criada nem destruída.

Constante cosmológica: Artífício matemático usado por Einstein para atribuir ao espaço-tempo uma tendência intrínseca para a expansão.

Coordenadas: Números que especificam a posição de um ponto no espaço e no tempo.

Cosmologia: Estudo do universo como um todo.

Cromodinâmica quântica: Teoria que descreve as interações dos quarks e glúões.

Desvio para o vermelho: Avermelhamento da luz de uma estrela que se afasta de nós devido ao efeito de Doppler.

Dimensão espacial: Qualquer das três dimensões do espaço-tempo que são espaciais, ou seja, qualquer dimensão, exceto a temporal.

Dualidade: Uma correspondência entre teorias aparentemente diferentes que leva aos mesmos resultados físicos.

Dualidade onda/partícula: Conceito em mecânica quântica de que não há distinção entre ondas e partículas: as partículas podem, por vezes, comportar-se como ondas e as ondas como partículas.

Efeito de Casimir: A tensão atrativa entre duas placas (lâminas) planas e paralelas colocadas muito perto uma da outra no vácuo. A tensão é devida a uma redução do número de partículas virtuais no espaço entre as lâminas.

Elétrão: Partícula com carga que orbita em torno de um núcleo atômico.

Energia da grande unificação: Energia acima da qual se crê que as forças eletromagnéticas, forte e fraca se tornam indistintas umas das outras.

Energia de unificação eletrofraca: Energia (cerca de 100 GeV) acima da qual a distinção entre força eletromagnética e força fraca desaparece.

Espaço-tempo: Espaço quadridimensional cujos pontos são os acontecimentos.

Espectro: Decomposição de uma onda eletromagnética (entre outros exemplos) nas frequências que a compõem.

Estado estacionário: Estado que não muda com o tempo: uma esfera que roda com velocidade constante está estacionária porque permanece idêntica em cada instante, embora se saiba que não está em repouso.

Estrela de neutrões: Estrela mantida pela repulsão decorrente da aplicação do princípio de exclusão aos neutrões.

Fase: A fase de uma onda é a posição no seu ciclo num instante dado: uma «medida» para descrever se está numa crista, numa cava ou em qualquer ponto intermédio.

Força eletromagnética: Força que se manifesta entre partículas com carga elétrica; é a segunda mais forte das quatro forças fundamentais.

Força forte: A mais forte das quatro forças fundamentais, com o menor alcance de todas. Mantém os quarks juntos no interior dos prótons e neutrões e mantém os prótons e neutrões juntos para formar átomos.

Força fraca: A segunda mais fraca das quatro forças fundamentais, com alcance pequeníssimo; afeta todas as partículas materiais, exceto as que são portadoras de força.

Fotão: Um quantum de luz.

Frequência: É o número de ciclos por segundo de uma onda.

Fusão nuclear: Processos em que dois núcleos colidem e coalescem para formar um só núcleo mais pesado.

Geodésica: Trajetória mais curta (ou mais longa) entre dois pontos.

Horizonte de acontecimentos: Fronteira de um buraco negro.

Limite de Chandrasekhar: Limite de massa de uma estrela fria estável, acima do qual a estrela colapsa para originar um buraco negro.

Massa: Quantidade de matéria existente num corpo; a sua inércia ou resistência à aceleração.

Matéria escura: Matéria nas galáxias, aglomerados de galáxias e, possivelmente, entre aglomerados que não pode ser observada diretamente, mas pode ser detetada pelos seus efeitos gravitacionais. Cerca de 90% da matéria do universo pode encontrar-se na forma da matéria escura.

Mecânica quântica: Teoria desenvolvida a partir do princípio da incerteza de Heisenberg (capítulo 4).

Neutrão: Partícula sem carga, muito semelhante ao próton, que contribui para cerca de metade do número total das partículas existentes no núcleo atómico.

Neutrino: Partícula elementar de matéria, extremamente leve (possivelmente sem massa), que é afetada apenas pela força fraca e pela gravidade.

Núcleo: Parte central do átomo, constituído apenas por prótons e neutrões, mantidos juntos por ação da força forte.

Partícula elementar: Partícula que se crê não ser subdivisível.

Partícula virtual: Em mecânica quântica, partícula que não pode ser detetada diretamente, mas cuja existência tem efeitos mensuráveis.

Peso: É a força exercida num corpo pelo campo da gravidade. É proporcional à massa do corpo, mas não se identifica com ela.

Ponte de Einstein-Rosen: Um fino tubo de espaço-tempo que liga dois buracos negros. Ver Buraco de verme.

Positrão: Antipartícula do eletrão, com carga positiva.

Princípio antrópico: Vemos o Universo tal como está porque, se fosse diferente, não estaríamos aqui para o observarmos.

Princípio de exclusão: Duas partículas de *spin* 1 idênticas não podem ter ambas (dentro de certos limites fixados pelo princípio da incerteza) a mesma posição e a mesma velocidade.

Princípio de incerteza: Não é possível ter a certeza simultânea da posição e da velocidade de uma partícula; quanto maior for a precisão com que se conhece uma, menor é a precisão com que se pode conhecer a outra.

Princípio quântico de Planck: Conceito de que a luz (ou quaisquer outras ondas clássicas) pode ser emitida ou absorvida somente em quanta discretos, cuja energia é proporcional à frequência.

Proporcionalidade: «X é proporcional a Y» significa que, quando Y é multiplicado por qualquer número, o mesmo ocorre com X. «X é inversamente proporcional a Y» significa que, quando Y é multiplicado por qualquer número, X é dividido por esse número.

Protão: Partícula com carga positiva que contribui para cerca de metade das partículas existentes no núcleo atómico.

Pulsar: Uma estrela de neutrões que gira e emite pulsos regulares de ondas de rádio.

Quantum: Unidade indivisível em que as ondas podem ser emitidas ou absorvidas.

Quark: Partícula elementar (com carga) sensível à força forte. Os protões e neutrões são compostos por três quarks cada um.

Radar: Sistema que usa impulsos de ondas de rádio para detetar a posição de objetos através da medição do tempo que um impulso simples demora a atingir o objeto a ser refletido.

Radiação de fundo de micro-ondas: Radiação resultante das elevadas temperaturas do universo primitivo, atualmente tão deslocada para o vermelho que não aparece como luz, mas como micro-ondas (ondas de rádio com comprimento de onda característico de alguns centímetros).

Radioatividade: Desintegração espontânea de um tipo de núcleo atômico originando outro.

Raios gama: Ondas eletromagnéticas de comprimento muito pequeno produzidas por decaimento radioativo ou por colisão de partículas elementares.

Relatividade geral: Teoria de Einstein baseada no conceito de as leis da física deverem ser as mesmas para todos os observadores, independentemente do seu movimento. Explica a força da gravidade em termos de curvatura de um espaço-tempo quadridimensional.

Relatividade restrita: Teoria de Einstein baseada no conceito de que as leis da física devem ser as mesmas para todos os observadores em movimento livre, isto é, uniforme e retilíneo, independentemente das suas velocidades.

Segundo-luz (ano-luz): Distância percorrida pela luz num segundo (num ano).

Singularidade: Ponto do espaço-tempo onde a curvatura se torna infinita.

Singularidade nua: Singularidade do espaço-tempo não resguardada por um horizonte.

Spin: Propriedade intrínseca das partículas elementares relacionada, mas que não se identifica, com o conceito corrente de rotação.

Tempo imaginário: Tempo referido em números imaginários.

Teorema da singularidade: Teorema que mostra que a singularidade é inevitável em certas circunstâncias, em particular que o universo começou com uma singularidade.

Teoria das cordas: Teoria física em que as partículas são descritas como ondas em cordas. As cordas têm comprimento, mas não têm qualquer outra dimensão.

Teoria da grande unificação (GUT): Teoria que unifica as forças eletromagnéticas, forte e fraca.

Zero absoluto: Temperatura mais baixa possível, à qual uma substância não contém energia calorífica.

AGRADECIMENTOS

Foram muitas as pessoas que me ajudaram a escrever este livro. Os meus colegas cientistas, sem exceções, incentivaram-me. Ao longo dos anos, os meus colaboradores foram Roger Penrose, Robert Geroch, Brandon Carter, George Ellis, Gary Gibbons, Don Page e Jim Hartle. Muito lhes devo, tal como aos meus alunos de investigação, que sempre me ajudaram quando necessário.

Um dos meus alunos, Brian Whitt, auxiliou-me muito na redação da 1ª edição deste livro. O meu editor, Peter Guzzardi, fez numerosos comentários que contribuíram para melhorar consideravelmente o livro. Além disso, relativamente à edição ilustrada, gostaria de expressar os meus agradecimentos ao pessoal da *MoonRunner Design*, que forneceu as ilustrações, e a Andrew Dunn, que me ajudou a rever o texto e a escrever as legendas. Julgo que fizeram um bom trabalho.

Não teria podido escrever este livro sem o meu sistema de comunicação. O *software*, designado como *Equalizer*, foi fornecido por Walt Waltosz, da World Plus Inc., em Lancaster, Califórnia. O meu sintetizador de fala foi fornecido por Speech Plus, de Sunnyvale, Califórnia. O sintetizador e o computador *laptop* foram montados na minha cadeira de rodas por David Mason, da Cambridge Adaptative Communication Ltd. Com este sistema posso comunicar melhor agora do que antes de ter perdido a fala.

Ao longo dos anos tive algumas secretárias e assistentes com que escrevi e revi este livro. Quanto às secretárias, estou muito grato a Judy Fella, Ann Ralph, Laura Gentry, Cheryl Billington e Sue Masey. Os meus assistentes foram Colin Williams, David Thomas, Raymond Laflamme, Nick Phillips, Andrew Dunn, Stuart Jamieson, Jonathan Brenchley, Tim Hunt, Simon Gill, Jon Rogers e Tom Kendall. Eles, as minhas enfermeiras, colegas, amigos e familiares permitiram-me ter uma vida muito cheia e continuar a investigação apesar da minha doença.

STEPHEN HAWKING