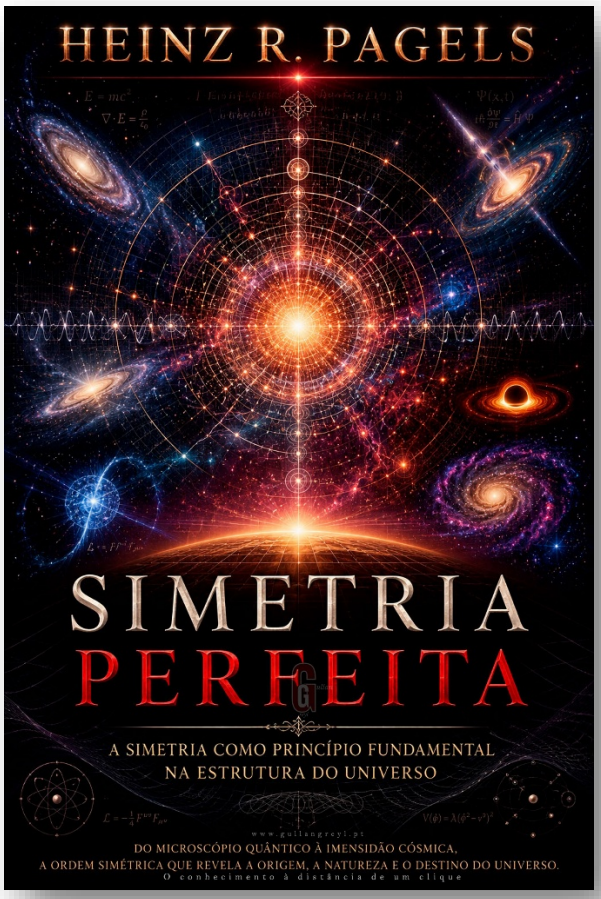
SIMETRIA PERFEITA

A Simetria como Princípio Fundamental na Estrutura do Universo



Autor: Heinz R. Pagels

Edição e Preparação Digital: Gullan Greyl

Edição original: Perfect Symmetry: The Search for the Beginning of Time, 1985.

Edição portuguesa concluída em 02-08-2023

Livroteca Gullan Greyl

Sinopse

Em Simetria Perfeita, Heinz R. Pagels leva o leitor para uma das zonas mais vertiginosas da ciência moderna: o ponto em que a física quântica do infinitamente pequeno se encontra com a cosmologia do infinitamente grande. Das estrelas e galáxias ao Big Bang, dos buracos negros às partículas elementares, o livro reconstrói a tentativa humana de compreender o universo como uma entidade física governada por leis inteligíveis — não apenas como imensidão misteriosa, mas como estrutura profunda que a razão começa a decifrar.

A viagem começa no "jardim de Herschel", esse universo dinâmico onde estrelas nascem, vivem, morrem e reciclam a matéria de que também somos feitos. A partir daí, Pagels conduz-nos para regiões cada vez mais extremas: anãs brancas, estrelas de neutrões, supernovas, buracos negros, quasares, galáxias e superaglomerados. O cosmos deixa de ser um palco imóvel e transforma-se num organismo físico em evolução, onde cada corpo celeste participa numa história maior de criação, colapso, transformação e continuidade.

Mas o centro secreto da obra está na simetria. Pagels mostra que, no universo primordial, antes da complexidade que hoje observamos, a realidade teria existido num estado de extrema simplicidade, temperatura e ordem. À medida que o universo se expandiu e arrefeceu, essa simetria perfeita quebrou-se, dando origem às partículas, forças, estruturas e diferenças que compõem o mundo atual. A diversidade cósmica nasce, assim, não da ausência de ordem, mas da sua transformação: o universo visível como remanescente assimétrico de uma perfeição inicial.

Em Simetria Perfeita, a ciência lê-se como uma epopeia da origem. A criação do universo a partir do nada, o Big Bang, a inflação cósmica, as teorias de grande unificação, a supersimetria, a gravidade, a matéria, o espaço e o tempo surgem como capítulos de uma única pergunta: que princípio profundo permite que do vazio emerja um cosmos ordenado, dinâmico e compreensível?

Mais do que uma obra de cosmologia, este livro é uma travessia pela beleza matemática do real. Pagels escreve como quem observa o universo simultaneamente com o telescópio, o acelerador de partículas e a imaginação filosófica. E, ao fazê-lo, transforma a origem do cosmos numa experiência intelectual quase iniciática: do microcosmo quântico à imensidão cósmica, tudo parece apontar para uma ordem escondida — uma simetria primordial cuja quebra deu nascimento ao universo que habitamos, interrogamos e, lentamente, começamos a compreender.

Esta edição foi preparada para formato digital, com intervenção ao nível da organização, revisão e uniformização do texto, de forma a assegurar clareza, consistência e legibilidade em ambiente de leitura eletrónica.

Foram realizadas correções pontuais sempre que necessário, respeitando integralmente o conteúdo e a estrutura da obra original.

O objetivo desta preparação é preservar a fidelidade ao texto, garantindo simultaneamente uma experiência de leitura fluida e acessível.

Gullan Greyl

Não se trata de saber mais,
mas de ver melhor

Índice

Agradecimentos 1

Prefácio 2

PARTE UM 7

O Jardim de Herschel 7

CAPÍTULO 1 8

O Jardim de Herschel 8

CAPÍTULO 2 20

O Nascimento e a Vida das Estrelas 20

CAPÍTULO 3 57

Morte das Estrelas – Necroscopia Astral 57

1. As Anãs Brancas 59

2. As Estrelas de Neutrões 65

3. Buracos Negros 79

CAPÍTULO 4 102

A Descoberta das Galáxias 102

CAPÍTULO 5 146

Radiogaláxias e Quasares 146

CAPÍTULO 6 169

– Por Que É o Universo Granuloso? – 169

CAPÍTULO 7 189

– Cosmologia Clássica – 189

PARTE DOIS 220

O Universo Infinito 220

CAPÍTULO 8 221

O Universo Primitivo 221

CAPÍTULO 9 238

Campos, *Quanta* e Simetria 238

1. Relativista 251

2. Campos 275

3. Quântica 284

4. Antipartículas 286

5. Renormalização 289

6. Campos de Padrão 295

7. Simetrias de Padrão e Quebra de Simetria 299

CAPÍTULO 10 307
O Modelo Padrão 307
1. Os Leptões 311
2. Os Quarks 317
3. Os Gluões 323
4. Gluões Coloridos e Cromodinâmica Quântica 324
5. Os Gluões Eletrofracos e o Modelo de Weinberg-Salam 331
6. Para Além do Modelo Padrão 338
CAPÍTULO 11 343
Termodinâmica e Cosmologia 343
CAPÍTULO 12 358
O Big Bang 358
1. A Singularidade 366
2. O Limiar da Quebra da Simetria Electrofraca: 10^{15} K 368
3. Hadronização: 10^{14} K 370
4. A Era dos Hadrões: 10^{14} K a 10^{12} K 371
5. A Era dos Leptões: 10^{12} K a 10^{10} K 376
6. A Era dos Fotões: 10^{10} K a 10^{3} K 381
7. A Recombinação e o Fim do Big Bang 386
8. Para Além da Síntese Electrofraca 388
PARTE TRÊS 393
Ideias Loucas 393
CAPÍTULO 13 394
Teorias do Campo Unificado 394
1. Uma Invenção Livre do Espírito 396
2. Teorias de Grande Unificação (GUTs) 400
3. O Deserto 408
4. O Decaimento do Protão e a Assimetria Matéria-Antimatéria do Universo 413
5. Para Além das GUTs Simples 420
CAPÍTULO 14 429
Monopólos Magnéticos 429
CAPÍTULO 15 452
A Unificação da Gravidade 452
1. Supersimetria e Supergravidade 459
2. A Quinta Dimensão e Para Lá Dela 467

CAPÍTULO 16 477
Antes do *Big Bang* – O Universo Inflacionário 477
CAPÍTULO 17 506
Antes da Inflação – A Origem do Universo 506
PARTE QUATRO 528
Reflexões 528
CAPÍTULO 18 529
O Computador Cósmico 529
CAPÍTULO 19 541
Ciência na Primeira Pessoa 541
Bibliografia 558

SIMETRIA PERFEITA

A Simetria como Princípio Fundamental na Estrutura do Universo

Heinz R. Pagels

Agradecimentos

Na preparação deste livro tive a sorte de ter amigos e colegas capazes de fazer crítica franca ou de dar sugestões que vieram encontrar lugar no texto. Beneficiei dos comentários de Jeremy Bernstein, John Brockman, Malcolm Diamond, John Faulkner, Randall Furlong, George Greenstein, Alan Guth, Edward Harrison, Joseph H. Hazen, Nicolas Herbert, James McCarthy, Richard Ogust, Jim Peebles, Anthony Tyler e Anthony Zee. Estou especialmente grato pela crítica pormenorizada de George Field e de Engelbert Schucking nas secções do livro que tratam de astrofísica e de cosmologia. Alice Mayhew e Catherine Shaw fizeram a maior parte do trabalho redatorial do texto e ajudaram a transformar o «meu inglês» em «verdadeiro inglês». As ilustrações imaginosas de Matthew Zimet deleitam a vista e muito contribuem para realçar o texto. Finalmente, quero agradecer aos diretores da Academia das Ciências de Nova Iorque a sua compreensão pelo meu interesse em escrever sobre ciência.

No princípio Deus criou os céus e a Terra. A Terra era informe e vazia, e havia trevas sobre a face do abismo, e o espírito de Deus movia-se sobre as águas.
GÉNESIS

Prefácio

Os livros infantis que constituíram o meu primeiro contato com o universo abriram a minha imaginação aos pensamentos de viagens à Lua, aos planetas e às estrelas. Quando, já mais velho, visitei o Planetário Fels, em Filadélfia, e o Planetário Hayden, em Nova Iorque, aquela perceção simples e centrada em mim próprio foi desfeita. O drama e a força do universo dinâmico esmagaram-me. Aprendi que uma única galáxia contém mais estrelas do que todos os seres humanos que alguma vez existiram e vi projeções de aglomerados de galáxias a moverem-se no vazio do espaço como cardumes de peixes nadando no mar. A realidade da imensidão e longevidade do universo causou-me uma espécie de «choque existencial» que abalou os fundamentos do meu ser. Tudo quanto experimentara ou conhecera parecia insignificante quando situado nesse vasto oceano do que existe.

Embora o sentimento de temor respeitoso quanto às dimensões e esplendor do universo nunca me tenha abandonado completamente, ao recordar a experiência de criança, constato que o universo me proporcionava um ecrã no qual poderia projetar os meus sentimentos acerca da imensidão da existência: esse oceano exterior espelhava o que tinha dentro de mim. Mais tarde, ao estudar física teórica nas Universidades de Princeton e de Stanford, a minha atitude em relação ao universo alterou-se. O universo tornou-se mais um enigma, desafiando-me como cientista, do que um ecrã para a projeção dos meus sentimentos, um enigma que deixava pistas dispersas e complexas para a sua solução. O universo, a despeito das suas dimensões, é uma entidade física governada pelas leis do espaço, do tempo e da matéria. Um dia (e esse dia ainda não está próximo) os físicos poderão vir a conhecer as leis que descrevem a criação do universo e a sua evolução posterior. A explicação lógica dos fundamentos da existência física estará então completa.

À medida que nos embrenhamos no estudo do universo, vale a pena lembrarmo-nos de que não há muito tempo, no início deste século, os físicos ainda estavam intrigados com as propriedades dos átomos. Eram tão pequenos (alguns cientistas eminentes duvidavam, inclusivamente, da sua existência) e comportavam-se de maneira tão estranha e incontrolável que houve quem pensasse que estavam para além da compreensão científica. No entanto, após importantes descobertas experimentais e teóricas, os físicos dos anos 20 inventaram a teoria quântica, que explicava o estranho mundo do átomo. Conceitos físicos novos e pouco familiares foram incorporados na teoria quântica, sobrevivendo até aos dias de hoje.

Analogamente, ao tentarem compreender a origem e evolução do universo, os físicos terão certamente necessidade de inventar conceitos novos e pouco familiares. Os cientistas não compreendem ainda as leis fundamentais que descrevem a origem do universo, pelo menos não tão bem como compreendem as leis que descrevem os átomos. Mas hoje em dia muitos cientistas estão excitados por tal compreensão estar a processar-se, como resultado da síntese intelectual de duas disciplinas científicas: a teoria quântica, que estabelece as leis dos mais pequenos de todos os corpos — as partículas quânticas —, e a cosmologia, que estabelece as leis que governam o maior de todos os corpos — o universo na sua globalidade.

Uma razão de peso para a crescente intimidade entre a física quântica e a cosmologia é o sucesso da teoria do *big-bang* do universo primitivo, de acordo com a qual, se imaginarmos que recuamos no tempo, veremos então o universo contrair-se e as galáxias aproximarem-se umas das outras até se fundirem num gás quente e uniforme, com todos os constituintes da matéria — as partículas quânticas — interagindo com enorme energia. Explicar as propriedades desse gás quente composto por partículas quânticas interagindo entre si é o objetivo da moderna teoria quântica. Os físicos calculam que as elevadas temperaturas e as altas energias das partículas «quânticas» presentes no universo primitivo são fisicamente únicas — as energias tornam-se tão elevadas que não podem ser reproduzidas nos laboratórios, na Terra. Daí que o único «laboratório» capaz de testar as teorias da interação entre partículas quânticas a energias altíssimas seja o próprio universo.

Uma outra razão para o aumento da intimidade entre a teoria quântica e a astronomia é o facto de os astrónomos observarem agora objetos exóticos, como estrelas de neutrões, que consistem em matéria comprimida até densidades elevadíssimas, e possivelmente buracos negros, em que a própria textura do espaço e do tempo sofre estranhas distorções. Tal como o universo primitivo, estes estranhos objetos apresentam condições físicas extremas que não

podem ser reproduzidas na Terra. Uma vez que são as propriedades do espaço, do tempo e da matéria, especialmente em condições extremas, que os físicos se esforçam por entender, estes novos objetos põem-lhes à disposição laboratórios extraterrestres onde podem verificar as leis físicas. Se tivesse de resumir o conteúdo otimista deste livro numa única frase, essa seria a seguinte: «Do microcosmos ao macrocosmos, da sua origem ao seu fim, o universo é descrito por leis físicas compreensíveis pelo espírito humano.»

Creio que os físicos compreenderão um dia, talvez não muito distante, as leis básicas da criação quântica do universo (muito provavelmente a partir do nada) tão bem como os astrofísicos compreendem hoje o interior das estrelas. O universo, cuja simples menção invoca um sentido de transcendência, será compreendido como estando sujeito a leis naturais, tal como os outros corpos materiais. Apesar da sua imensidão e idade, o universo nunca mais parecerá o mesmo.

Tal concretização do programa das ciências naturais terá um profundo impacto no pensamento humano. À medida que amadurece o conhecimento do nosso universo, o antigo sentimento de terror e maravilha face às suas dimensões e longevidade parece inapropriado, como se se tratasse de uma sensibilidade pertencente às Eras primitivas. Há milhares de anos muitos eram os que concebiam o Sol como uma presença divina; hoje em dia muita gente concebe o universo como algo que está para lá da compreensão humana. Todavia, da mesma maneira que o Sol é hoje compreendido em termos de processos astrofísicos, também o universo será, de maneira análoga, compreendido. No passado os mitos e as narrativas religiosas da criação moldaram os valores dos que acreditavam neles; da mesma forma, a cosmologia científica emergente moldará os valores daqueles que a aceitarem. Por intermédio das descobertas científicas, a ordem externa do universo influencia a nossa consciência e os nossos valores.

Este livro está dividido em quatro partes. A primeira, «O jardim de Herschel», dá ao leitor uma visão geral do universo dinâmico descoberto pelos astrónomos — as estrelas, as anãs brancas, as estrelas de neutrões, os buracos negros, o gás e a poeira interestelares, os quasares, as galáxias e a sua distribuição no espaço em aglomerados e superaglomerados de galáxias e o cosmos como um todo. Desta parte do livro o leitor poderá retirar uma ideia não só das dimensões do universo e do conhecimento que temos dos seus habitantes, mas também dos enigmas que a moderna astronomia tem de enfrentar, como seja o saber como nascem as estrelas e como evoluem as galáxias. Discutirei algumas das soluções sugeridas

para estes e outros enigmas astronómicos, para os quais esperamos encontrar uma solução final à medida que formos dispondo de mais dados de observação. O estudo do universo faz-se atualmente com satélites e radiotelescópios, um estudo a que traduz, nas palavras do astrónomo americano Edwin Hubble, um «anseio [...] mais velho do que a história».

Enquanto a primeira parte do livro descreve o universo observado no espaço, as duas partes seguintes descrevem uma exploração concetual do universo no tempo. A segunda parte, «O universo primitivo», descreve a impressionante imagem do universo quando tinha apenas alguns segundos ou minutos de vida — o *big-bang* quente, uma teoria que nasceu da aplicação das leis da física das partículas quânticas à totalidade do universo. Sem utilizar matemática complicada, descreverei o quadro concetual básico do pensamento acerca das partículas quânticas — a disciplina conhecida por «teoria quântica relativista do campo» — e como se aplica ao estudo do universo primitivo. Por mais estranho que pareça, os físicos compreendem melhor o universo quando este tinha apenas alguns segundos ou minutos de idade do que em Eras anteriores ou posteriores, uma vez que não era então mais do que um gás uniforme e bastante simples de partículas quânticas, cujas propriedades são conhecidas. Hoje em dia o universo primitivo é mais bem compreendido do que o tempo meteorológico.

Mas o próprio sucesso da teoria do *big-bang* quente dá aos físicos confiança para continuarem a explorar concetualmente o universo antes do primeiro nanossegundo (um milésimo de milionésimo de segundo) até ao próprio momento em que teve origem. A terceira parte do livro, «Ideias loucas», abandona o terreno firme explorado pela observação astronómica e pelas experiências laboratoriais de altas energias e especula acerca da natureza do universo antes do primeiro nanossegundo. Apresento as «ideias loucas» do reportório concetual dos físicos teóricos, que poderão talvez explicar a dinâmica do universo primitivo, ideias como as GUTs (teorias da grande unificação[1]), monopólos magnéticos, supersimetria e o mundo das muitas dimensões adicionais. Se estiverem corretas — e muitos físicos pensam que sim —, resultará daqui uma impressionante imagem do universo primitivo.

O universo começa num estado muito quente de extrema simplicidade e simetria e, à medida que se expande e arrefece, a sua simetria perfeita quebra-se, dando lugar ao aparecimento da complexidade que hoje observamos. O universo atual é o remanescente arrefecido e assimétrico do estado quente inicial, neste sentido muito semelhante aos complexos cristais de gelo que resultam de um gás uniforme de vapor de água. Descrevo o universo inflacionário — uma era hipotética anterior ao *big-bang* —, que pode explicar alguns

aspetos enigmáticos do universo contemporâneo, tais como a sua uniformidade e idade, e também dar uma explicação para a origem das galáxias. O penúltimo capítulo desta terceira parte do livro descreve, no que respeita à especulação, alguns modelos matemáticos recentes da origem do universo e o modo como a textura do espaço, do tempo e da matéria pode ser criada do nada absoluto. O que é que pode ter uma simetria mais perfeita do que o nada absoluto? Pela primeira vez na história os cientistas construíram modelos matemáticos que explicam a própria criação do universo a partir do nada.

Há ainda uma quarta parte, «Reflexões», que expressa as minhas opiniões e atitudes (não que as outras partes do livro não contenham muitas das minhas opiniões e tendências intelectuais enquanto físico teórico). Aqui o leitor encontrará um capítulo desenvolvendo a metáfora do universo como um computador cósmico, sendo as partículas quânticas o *hardware*, as leis da física o *software* e a evolução do universo a execução do programa. No capítulo final, chamado «Ciência na primeira pessoa», exploro os pensamentos e sentimentos que algumas pessoas tiveram acerca do significado do nosso universo tão estranhamente coerente.

Nova Iorque, Nova Iorque
Felton, Califórnia
1984

---

[1] Do inglês: grand unified theories. (*N. do T.*)

PARTE UM

O Jardim de Herschel

A mais maravilhosa e profunda experiência que um homem pode ter é o sentido do misterioso. É o princípio subjacente à religião, mas também a todo o esforço sério na arte e na ciência [...] Penso que quem nunca teve esta experiência deve estar, se não morto, pelo menos cego. O sentimento de que por detrás de qualquer coisa que possa ser experimentada existe algo que o nosso espírito não pode abranger e cuja beleza e excelência apenas nos atingem indiretamente, como débil reflexo, é a religiosidade. Neste sentido, sou religioso. Basta-me maravilhar-me com estes segredos e tentar, humildemente, captar com a minha mente uma mera imagem da sublime estrutura de tudo quanto existe.

ALBERT EINSTEIN, 1932

CAPÍTULO 1

O Jardim de Herschel

> Existem duas espécies de felicidade ou contentamento às quais nós, mortais, estamos adaptados: a primeira experimentamo-la pensando, a segunda, sentindo. A primeira é a mais pura e a menos confusa. Compreenda o homem alguma vez a espécie de ser que é, quão grande é o ser que lhe deu existência, quão transitórias são as coisas materiais, tome disso consciência sem paixão, em íntima serenidade filosófica, e eu afirmo que então será feliz, tão feliz quanto lhe é possível sê-lo.
>
> *WILLIAM HERSCHEL (de uma carta ao seu irmão Jacob)*

William Herschel, o maior astrónomo do século XVIII, começou a carreira como jovem oboísta nos Hanoverian Foot Guards, numa parte da Alemanha então sob o domínio de Jorge II de Inglaterra. Nascido em 1738, pretendia tornar-se músico e compositor profissional. Contudo, aquando da batalha de Astenbeck, esteve «tão perto do campo de ação que se encontrava ao alcance de um tiro de canhão». O pai aconselhou-o a fugir. Para evitar a incorporação no serviço militar regular foi para Inglaterra, com 19 anos de idade, acompanhado do irmão Jacob, e aí continuou a carreira de músico. Em 1766 foi nomeado organista da Capela Octagon na animada cidade de Bath, onde também tocava na orquestra da Pump Room[1].

Só aos 35 anos despertou em Herschel o interesse pela astronomia. Ajudado pela irmã, Caroline, e pelo irmão Alexander,

fabricou um ótimo telescópio de reflexão numa fundição que construiu em casa. Sem dúvida que a habilidade para os instrumentos musicais lhe foi muito útil na construção do instrumento de precisão. Com o auxílio do telescópio, descobriu um novo planeta — Úrano —, que, inicialmente, julgou ser um cometa. Desde os tempos antigos que os únicos planetas conhecidos eram os seis que são observáveis a olho nu. Ninguém tinha previsto a existência de mais um planeta, pelo que o choque desta descoberta tornou Herschel e o seu telescópio instantaneamente famosos. Sabendo o que fazer para exprimir gratidão para com o seu país adotivo, chamou ao novo planeta *Georgium Sidus* (Estrela de Jorge), em honra do rei Jorge III, nome que, posteriormente, foi mudado. Foi eleito para a Royal Society de Londres, Jorge III tornou-se seu protetor e a sua carreira na astronomia estava lançada.

A entrada de Herschel na astronomia não foi incomum — vários grandes astrónomos observacionais começaram as carreiras em profissões diferentes, apenas com um interesse secundário pela astronomia. Após a importante descoberta astronómica de um estranho e novo planeta, também Herschel não resistiu ao anseio de continuar a explorar o universo. A paixão pela ciência e a paixão pela música eram movidas pelo mesmo desejo: dar realidade à beleza de uma imagem do mundo.

As realizações de Herschel em astronomia são ainda mais notáveis quando vistas em retrospetiva; com efeito, muitas das suas observações e análises não puderam ser completamente apreciadas senão no século XX. Concluiu, por exemplo, que, devido

à velocidade finita da luz, vemos os objetos celestes distantes como eram no passado. Quando contemplamos as profundezas do universo, vemo-lo como era milhões e milhões de anos atrás, quando estava a ser emitida a luz que agora recebemos. Notavelmente, o universo encerra o registo do seu passado, à semelhança dos estratos sedimentares das rochas, que contêm o registo geológico do passado da Terra. E este facto abriu a janela à visão evolucionária do universo sustentada atualmente.

Herschel tornou-se obcecado pelos problemas da determinação da estrutura da Via Láctea e da localização da posição do nosso sol dentro desta. Estava até convencido de que algumas nebulosas eram exteriores à Via Láctea e semelhantes a ela, antecipando, assim, a teoria das galáxias como «universos-ilhas». Dado que lhe era impossível calcular as distâncias às estrelas com as técnicas de que dispunha, a sua conceção da galáxia Via Láctea estava quantitativamente errada. Muito para seu crédito como cientista, mas também para seu desapontamento pessoal, abandonou, por fim, o conceito da Via Láctea como um grande disco (o que, de facto, está correto) quando compreendeu que os seus métodos de observação eram inadequados à tarefa de determinar com precisão a forma da galáxia. Mas foi o primeiro a mostrar que as estrelas da Via Láctea não estão dispostas simetricamente em torno do Sol — um facto importante, corroborado pelas observações modernas. Destruiu, assim, para sempre a ideia do céu como uma esfera celeste a rodear o Sol. A despeito de especulações posteriores, mais nenhum progresso se deu até Harlow Shapley, astrónomo americano, publicar os seus estudos sobre a forma da Via Láctea, cerca de 140 anos mais tarde.

No dia da sua eleição para a Royal Society, Herschel recebeu do seu amigo Dr. W. Watson, Jr., uma cópia do novo catálogo de 103 nebulosas publicado por Charles Messier e Pierre Méchain. Começou imediatamente a apontar o seu maravilhoso telescópio para estes estranhos objetos, na esperança de descobrir mais alguns que pudessem não constar do catálogo.

Em vez disso, descobriu 2000 novas nebulosas e começou a elaborar a sua própria lista. Foi este o começo de um novo catálogo (ao qual o seu filho John juntou muitas mais nebulosas, que observou no hemisfério sul, alguns anos mais tarde), que constituiu a base de todos os modernos catálogos de galáxias.

Herschel descobriu também muitos sistemas de estrelas duplas — duas estrelas em órbita uma da outra — e mostrou que obedeciam à lei da gravitação de Newton. Hoje em dia sabemos que cerca de metade de todas as estrelas observáveis pertencem a tais sistemas binários. A descoberta de Herschel de que a lei de Newton se aplicava ao movimento de estrelas longínquas, e não somente ao dos planetas em torno do Sol, foi fundamental. Mostrou também que o Sol não está fixo no espaço, movendo-se em direção à estrela Lambda de Hércules — ideia revolucionária comparável à declaração de Copérnico de que a Terra rodava em torno do Sol. Tal como muitos dos seus contemporâneos, Herschel pensava que a Lua, os planetas e o Sol eram habitados (julgava existir uma superfície fria sob a atmosfera quente do Sol). Talvez ninguém antes ou depois dele tenha passado tanto tempo a olhar por um telescópio.

Herschel produziu uma grande mudança concetual na astronomia. Anteriormente a ele, as pessoas tinham uma visão mecânica e newtoniana das estrelas, julgando-as submetidas unicamente à força da gravidade. Mas Herschel, inteiramente em consonância com a nossa visão moderna, sugeriu que outros processos dinâmicos estavam a dar forma ao universo. No estilo barroco daquele tempo, escreveu acerca da possibilidade de velhas estrelas colidirem entre si para formarem novas estrelas:

> Se, porventura, não fosse demasiado perigoso continuar uma antiga suposição de um renovamento naquilo a que, figurativamente, chamo laboratórios do universo, as estrelas que formam estas extraordinárias nebulosas, devido a uma espécie de degenerescência ou desgaste natural, tendo deixado de estar aptas para os fins primitivos e tendo-se as suas forças de projétil, se algumas tiveram, retardado nas atmosferas umas das outras, poderiam, finalmente, aproximar-se e, quer em sucessão, quer devido a um enorme choque universal, unir-se para formar um corpo novo. Talvez o repentino e extraordinário clarão de uma nova estrela na cadeira da Cassiopeia em 1572 tenha sido desta natureza.

Herschel apercebeu-se da grande diversidade dos céus — mesmo no seu tempo, quando o universo observável era muito mais simples do que aquele que hoje contemplamos. Via o universo como um lugar em mudança e em evolução e dizia que examinar as estrelas era como examinar um vasto jardim, no qual algumas plantas são velhas, outras jovens, umas estão a nascer, outras a morrer. Embora não possamos ver o crescimento de uma planta, vemos, de facto, muitos exemplares dessa planta em todas as fases da sua vida, e essa observação dá-nos uma pista para

compreendermos o respetivo crescimento. De maneira análoga, o astrónomo vê uma contínua evolução no desenvolvimento das estrelas, e talvez nas galáxias e nos aglomerados de galáxias, e essa é a sua chave para a compreensão da dinâmica de mudança no universo. Herschel escreveu:

> Este método de observar os céus parece colocá-los sob uma nova espécie de luz. Eles parecem agora um jardim luxuriante, que contém enorme variedade de produções em diferentes canteiros, e pelo menos uma vantagem podemos colher desta situação: é a de podermos estender o alcance da nossa experiência a tempos imensos. Por que, e continuo a usar a similitude com o reino vegetal, não será quase a mesma coisa vivermos para testemunhar a germinação, floração, foliação, fecundação, amarelecimento, murchidão e corrupção de uma planta ou, num relance, observarmos um vasto número de exemplares escolhidos de cada uma das fases pelas quais a planta passa no decurso da sua existência?

O universo dinâmico é o jardim de Herschel. Podemos levar esta analogia ainda mais longe. Outrora os botânicos estudavam as plantas apenas como organismos isolados. Contudo, à medida que a vida das plantas se tornou mais bem compreendida, os botânicos aperceberam-se de que, longe de existir independentemente, cada planta dependia, para viver, de um sistema ecológico, de um ambiente complexo.

O mesmo acontece com os planetas, as estrelas e as galáxias. Embora os astrónomos possam estudá-los independentemente, torna-se cada vez mais claro que existe uma interação complexa entre todos os corpos celestes que observamos. Por exemplo, os

átomos dos planetas e os dos nossos corpos compreendem muitos elementos químicos pesados que foram produzidos a partir de elementos mais leves nas fornalhas nucleares das estrelas em tempos remotos. O ritmo a que nascem novas estrelas nos braços de uma galáxia espiral influencia a dinâmica de toda a galáxia, que, por sua vez, influencia a formação de estrelas. Tal como a vida num jardim, a vida no universo depende de uma relação complexa entre as partes e o todo. Para compreendermos esta relação deixemo-nos vaguear no «jardim luxuriante» de Herschel e colhamos uma panorâmica rápida do que lá existe.

Os céus fervilham com uma grande variedade de corpos celestes. Além de milhares de milhões de estrelas semelhantes ao nosso sol, os astrónomos descobriram uma imensidade de espécies diferentes de estrelas. Entre estas encontra-se Betelgeuse, a estrela «supergigante vermelha» da constelação de Oríon, tão grande que o seu raio é igual ao da órbita da Terra. Como Betelgeuse é tão grande e está relativamente tão próxima, foi a primeira estrela cujo disco foi resolvido — podemos hoje vê-la como um disco circular, e não como um ponto luminoso, usando uma técnica óptica de observação chamada *speckle interferometry*[2].

Os astrónomos descobriram também algumas estrelas no fim das suas vidas, as anãs brancas e as estrelas de neutrões. O Sol acabará por transformar-se numa gigante vermelha e depois numa anã branca, uma minúscula estrela brilhando com as últimas reservas de energia. As estrelas de massa superior à do Sol terão um destino mais dramático. Algumas explodirão numa «Supernova», libertando num único segundo o equivalente a toda a energia que o Sol terá

libertado em toda a sua existência de milhares de milhões de anos. «A nova estrela na cadeira da Cassiopeia em 1572», a que Herschel se referiu, foi a primeira observação de uma explosão de supernova feita no Ocidente. O remanescente desta explosão é uma minúscula estrela de neutrões composta por matéria comprimida até à densidade de um núcleo atómico, isto é, algumas toneladas por centímetro cúbico. Grandes quantidades da matéria dispersa por esta explosão espalham-se pelo espaço, enriquecendo o gás interestelar em elementos pesados. Esta matéria acabará por dar origem a novas estrelas, num gigantesco processo de reciclagem. Embora ninguém tenha visto uma nova estrela em formação, os astrónomos sabem que os lugares de nascimento das estrelas são as nebulosas gasosas densas, como a nebulosa de Oríon, nos braços da nossa galáxia espiralada.

Imaginemo-nos a voar para fora do sistema solar, para além da Via Láctea, e olhando para trás. Que veríamos? Primeiro, contemplaríamos dentro da forma discoide os grandes e belos braços em espiral da nossa galáxia, que contém estrelas novas (como o nosso sol) e grandes quantidades de gás e poeira interestelares. Mais além veríamos os braços encurvarem-se e abraçarem-se num «bojo central», de forma grosseiramente esférica, composto por estrelas velhas, albergando no núcleo fontes de imensa energia, provavelmente um gigantesco buraco negro. E, finalmente, olhando acima e abaixo do plano do disco, veríamos o «halo» da galáxia, pelo menos tão grande como o disco, grosseiramente esférico, composto por cerca de uma centena de «aglomerados globulares», dispersamente distribuídos, de velhas estrelas ligadas gravitacionalmente umas às outras e em órbita em

torno da própria galáxia. Aquilo que poderíamos ver na nossa galáxia seria apenas parte da história. Existem ainda outras componentes invisíveis — radiação infravermelha, raios X, campos magnéticos e partículas subatómicas. Sabemos hoje que a galáxia está rodeada por uma coroa de gás quente e que a maior parte da massa da galáxia se encontra, provavelmente, sob a forma de matéria escura, e não de estrelas visíveis ou de gás. A nossa galáxia é uma entidade dinâmica complexa que mal começámos a compreender.

Se a observássemos de uma posição ainda mais distante, veríamos que está adornada com galáxias-satélites mais pequenas — as sete «galáxias anãs» e Leão I e Leão II, outras duas pequenas galáxias — em órbita à volta dela. A juntar a estas galáxias anãs, pobres em estrelas e de forma aproximadamente esférica, veríamos, próximo da nossa galáxia, a Grande e a Pequena Nuvem de Magalhães, que são galáxias pequenas de forma irregular. A Grande Nuvem de Magalhães está a ser despedaçada pelas interações de maré gravitacional com a nossa galáxia. A prova disso é a existência da corrente de Magalhães — uma gigantesca corrente de gás que liga a nossa galáxia a essa Nuvem de Magalhães.

Abarcando um volume ainda maior, veríamos a nossa vizinha galáxia de Andrómeda, outra galáxia espiralada semelhante à Via Láctea, com o seu grupo de galáxias-satélites mais pequenas orbitando em torno dela. Existem outras galáxias no nosso grupo local, todas elas nos subúrbios de um aglomerado de galáxias em forma de disco — a «supergaláxia» da Virgem. Este aglomerado é apenas um dos muitos grupos e galáxias deste tipo. Os

aglomerados de galáxias tendem a agrupar-se em «superaglomerados» de galáxias. O universo visível contém, pelo menos, 100 mil milhões de galáxias — um número que ultrapassa a nossa compreensão corrente.

A Natureza foi generosa para com os astrónomos, pois ofereceu-lhes em abundância, para observação, diferentes estrelas e galáxias em todas as fases das respetivas vidas. Graças a esta abundância, os astrónomos podem elaborar uma imagem de um universo dinâmico, incorporando nela a vida das estrelas e a evolução das galáxias, embora nenhumas mudanças possam ser detetadas dentro dos limites de uma vida humana.

Se bem que a Natureza tenha sido generosa ao oferecer uma imensa variedade de estrelas e galáxias, foi ainda mais generosa no que toca às dimensões do espaço. Os próprios astrónomos ficam atónitos com as dimensões do universo sempre que refletem sobre o significado das distâncias que calculam. Apesar de o seu número ser elevadíssimo, as estrelas não se estorvam umas às outras devido à vastidão do espaço à sua volta. Se o Sol fosse comprimido até ficar do tamanho de uma ervilha, a sua vizinha mais próxima, a Próxima do Centauro, a companheira binária de Alfa do Centauro, ficaria afastada cerca de 145 km e a vizinha que se seguiria em proximidade, a estrela de Barnard, ficaria a 200 km de distância, deixando um enorme campo de manobra para as estrelas. Se fosse, pelo contrário, toda a nossa galáxia da Via Láctea comprimida até ficar do tamanho de uma ervilha, a vizinha mais próxima, a galáxia de Andrómeda, ficaria distanciada apenas 10 cm, o que é ainda um espaço enorme, embora as galáxias colidam de tempos a tempos,

especialmente nos aglomerados densos de galáxias, como o aglomerado de Coma, ou da Cabeleira, onde se encontram muito apinhadas.

O jardim de Herschel — o universo — é muito maior do que ele podia ter imaginado. Os novos corpos celestes exóticos recentemente descobertos ter-lhe-iam causado tanta excitação como a nós. O universo tem-se mostrado muito mais peculiar do que alguém jamais poderia ter imaginado. Levando a razão aos limites do espaço e do tempo, a descendência científica de Herschel criou uma nova visão da realidade.

Os cientistas de hoje olham o universo como um enigma, para cuja solução existem pistas dispersas, constituindo um desafio, que se crê poder vir a ser vencido um dia, talvez mais próximo do que se pensa.

A primeira parte deste livro descreve o território explorado pelos astrónomos, conduzindo à descoberta do universo moderno. Organizada capítulo por capítulo, primeiro, explorando as estrelas, depois, passando para as galáxias, aglomerados e superaglomerados de galáxias, esta descrição aborda, finalmente, a imensidade do universo como um todo. Ao lidarem com escalas de distâncias ainda maiores, estão também os astrónomos a olhar para um passado mais longínquo, o que constitui uma progressão em direção às profundezas do universo, aproximadamente a par e passo com o desenvolvimento de instrumentos cada vez mais poderosos para a exploração astronómica. Salientei as descobertas

astronómicas mais recentes, as quais serão, porém, colocadas no contexto histórico dos grandes passos que as precederam.

Observemos agora mais de perto os objetos do jardim de Herschel — as estrelas, o gás e as galáxias — com a finalidade de os conhecermos como um bom jardineiro conhece as suas plantas. Olhemos cuidadosamente, pois este maravilhoso jardim está em evolução, de tal maneira que nunca voltará a ser o mesmo.

---

[1] Célebre sala das Termas de Bath. (*N. do T.*)

[2] Termo que poderia talvez traduzir-se por interferometria de granulação. (*N. do T.*)

CAPÍTULO 2

O Nascimento e a Vida das Estrelas

> Os cientistas afirmam normalmente que as suas convicções são baseadas em observações, não em teorias [...] Nunca encontrei nenhum que pusesse em prática essa afirmação [...] A observação não é suficiente L...] a teoria desempenha uma parte importante na formação das convicções.
>
> *ARTHUR S. EDDINGTON, The Expanding Universe[1], 1933*

As estrelas nascem, vivem e morrem. Enchendo o firmamento noturno como faróis num oceano de escuridão, guiaram os nossos pensamentos ao longo de milénios até ao porto seguro da razão. Foi na tentativa de compreender o movimento das estrelas e dos planetas que o espírito humano, pela primeira vez, concebeu a ideia de lei natural. As estrelas, porém, não são apenas objetos de investigação científica. Tal como o Sol e a Lua, estão incrustadas no nosso inconsciente — sentimos a sua presença mesmo sem as vermos.

Dispostas no céu em aparente desordem, constituem um ecrã perfeitamente adequado à projeção dos nossos sentimentos. Nesta disposição viram antigos sacerdotes e poetas figuras míticas e da Natureza: as estrelas eram deuses — arquétipos da permanência num mundo transitório. Quando comparadas com a vida humana ou com a vida das nações e dos impérios, parecem viver para sempre, indiferentes às paixões da nossa existência. Porém, apesar

das imensas distâncias que delas nos separam, com exceção do nosso sol, sentimos de algum modo que o destino da humanidade lhes está profundamente ligado. Temos esperança de que a vida na Terra possa partilhar sempre da permanência das estrelas, das galáxias e do próprio universo. Se essa esperança de permanência mais não é do que a projeção nos céus do nosso moderno mito de progresso e, portanto, tal como as antigas projeções das figuras míticas, também uma ilusão, só o tempo o dirá. As estrelas, tal como os deuses, que em tempos recuados representaram, continuam a interferir com os nossos sentimentos mais profundos. Mas que são elas?

Neste capítulo iremos dar uma panorâmica geral daquilo que os astrofísicos aprenderam acerca do nascimento e vida das estrelas, dando ênfase especial às descobertas mais recentes. No capítulo seguinte descreve-se a morte espetacular das estrelas. Embora examinemos as estrelas como se fossem indivíduos, é importante ter presente que são membros de uma sociedade maior, a galáxia, que as nutre à nascença, na qual passam a vida e que recebe os seus despojos, depois da morte. As estrelas não podem existir fora das galáxias; tal separação letal da parte e do todo violaria os princípios do companheirismo cósmico.

Durante muito tempo aqueles a quem as estrelas causavam perplexidade tentaram compreender a origem da luz estelar fazendo apelo a processos físicos familiares, como, por exemplo, a combustão comum. Séculos atrás, Nicolau de Cusa e outros filósofos especularam que as estrelas mais não eram do que sóis distantes. Se houvesse outras estrelas comparáveis ao Sol, então

as quantidades de luz e calor irradiadas por todas as estrelas seriam, na verdade, muito grandes. Que poderia alimentar estas imensas irradiações? Nenhum processo jamais visto na Terra poderia explicar como ardem as estrelas. Não houve qualquer explicação possível até Einstein ter mostrado, na primeira década deste século, que a matéria e a energia se podem converter uma na outra e até os físicos experimentais terem explorado o núcleo atómico.

Inspirados por estas novas descobertas, os físicos sugeriram que essa transformação de matéria em energia consistiria na emissão espontânea de partículas quânticas pelo núcleo atómico, conhecida por radioatividade, mesmo na ausência de uma explicação pormenorizada de tal processo. Algumas destas ideias foram adotadas por Arthur S. Eddington, astrónomo inglês, no seu importante livro *The Internal Constitution of Stars*[2], publicado em 1926, onde aplicou com êxito as leis da física atómica recentemente descobertas ao interior das estrelas e esboçou uma formulação dos principais problemas com que se defrontam os astrofísicos, os cientistas que estudam a física das estrelas. O problema fundamental consistia em descobrir a fonte da energia estelar. Eddington insistiu corajosamente em que tal só poderia dever-se a processos nucleares subatómicos, dizendo que «a medição da libertação de energia subatómica é uma das observações astronómicas mais comuns, e, a menos que os argumentos deste livro sejam inteiramente falaciosos, temos um razoável conhecimento das condições de densidade e temperatura da matéria que está a libertá-la».

De acordo com os cálculos de Eddington, o centro de uma estrela como o Sol teria uma temperatura de 40.000.000 K — muito quente, na verdade. (Cálculos mais recentes apontam para uma temperatura próxima dos 14.000.000 K). Uma vez que a temperatura nas estrelas vulgares mede a energia do movimento de partículas microscópicas, temos de concluir que no centro de uma estrela os núcleos atómicos, como o núcleo do hidrogénio, constituído por um único protão, são muito energéticos e movem-se com extrema rapidez, esmagando-se uns contra os outros continuamente. Os físicos de 1926, porém, acreditavam que, se os núcleos atómicos se aproximassem uns dos outros, repelir-se-iam. Não se fundiriam para formar um núcleo mais pesado e libertar a energia nuclear desejada. Mesmo à temperatura elevada reinante no centro de uma estrela, a barreira repulsiva que impede o contato seria demasiado alta para ser superada. Todavia, Eddington continuou a insistir em que os processos nucleares eram responsáveis pela energia de uma estrela.

O impasse veio a resolver-se em 1928 com a invenção da nova teoria quântica e com a descoberta, feita por George Gamow, R. W. Gurney e E. U. Condon, do que ela implicava — que as partículas não tinham de ultrapassar a energia da barreira de repulsão, podendo passar num «túnel» por baixo dela. A energia necessária para as partículas nucleares passarem por baixo da barreira de repulsão era bastante menor do que a necessária para a ultrapassar. Agora a sugestão de Eddington podia ser posta em ação. Em 1929 os físicos Robert d' Atkinson e Fritz Hontermans mostraram como este «efeito túnel» podia explicar a produção de energia das estrelas, através de fusão nuclear.

No entanto, permaneciam desconhecidas as reações nucleares exatas que poderiam ocorrer no núcleo de uma estrela (constituído principalmente por núcleos de hidrogénio e de hélio vogando à deriva). Como podiam, em rigor, os núcleos de hidrogénio fundir-se para dar origem aos núcleos mais pesados de hélio, processo este designado por fornalha nuclear? Em 1938 Hans Bethe, nos Estados Unidos, e, independentemente, Carl Friedrich von Weizsäcker, na Alemanha, deduziram matematicamente a primeira de duas reações nucleares — o «ciclo do carbono» — que respondiam a esta questão. Mostraram como, começando unicamente com núcleos de hidrogénio e um núcleo de carbono, como agente catalítico, se podia transformar hidrogénio em hélio, libertando enormes quantidades de energia. Bethe e Charles Critchfield indicaram ainda outro caminho pelo qual o hidrogénio podia ser transformado em hélio sem necessidade de um catalisador de carbono — a «cadeia protão-protão», sugerida também por von Weizsäcker — e Bethe provou, além disso, que esta reação e a do ciclo do carbono eram as únicas possíveis. Contudo, malgrado a descoberta dos mecanismos da fornalha nuclear responsáveis pela libertação da energia das estrelas, os cientistas verificaram quão longa e árdua era toda a tarefa de explicar as propriedades observáveis das estrelas a partir deste processo.

Na década de 50 começaram a ser feitos modelos de estrelas em computador, o que forneceu um novo método, que havia de revelar as complexas consequências astrofísicas das leis físicas elementares. Os conhecimentos resultantes do estudo das explosões nucleares no laboratório científico de Los Alamos

ajudaram os astrofísicos no esforço de compreenderem os processos nucleares no seio das estrelas. Podiam agora programar num computador as equações que descreviam o interior das estrelas — a temperatura, a pressão e as complexas reações nucleares. Fizeram grandes progressos. Em 1955 Fred Hoyle e Martin Schwarzschild introduziram a inovação de usar simulações de computador da evolução de uma estrela vulgar para mostrarem como esta se dilata e transforma numa estrela gigante vermelha.

Em traços gerais, os modelos matemáticos das estrelas tratados por computador constituíram um sucesso notável; hoje em dia dispomos das bases de uma teoria sobre a formação das estrelas muito conforme com as observações astronómicas. Os astrofísicos compreendem hoje os aspetos mais importantes da evolução das estrelas, desde o nascimento até à morte. A teoria está longe de estar completa; existem lacunas, problemas e observações incompreensíveis. No entanto, os sucessos mais importantes permitem-nos ser otimistas ao ponto de acreditarmos que uma teoria completa das estrelas e da evolução estelar possa estar concluída ainda neste século. Há até quem diga que já se encontra ao nosso alcance. As estrelas são entidades muito complexas e os cientistas continuam a descobrir novas peculiaridades, mas que não passam de pormenores, comparadas com as principais características já conhecidas e compreendidas. Contudo, ninguém pode ficar satisfeito até que mesmo o comportamento estelar mais bizarro se torne claro, e isso pode levar mais tempo.

Nas últimas décadas os astrofísicos descobriram mais coisas sobre as estrelas do que nos séculos anteriores, descobertas estas

que foram impulsionadas por vários avanços científicos de importância, alguns deles já referidos. Em primeiro lugar, a tecnologia usada nas observações astronómicas deu grandes passos em frente: detetores eletrónicos de grande sensibilidade que podem «ver» objetos muito pouco luminosos; o advento dos satélites artificiais; o nascimento da astronomia de raios X e os novos telescópios ópticos, de infravermelhos e de rádio e respetivos sistemas eletrónicos associados, que melhoraram grandemente as possibilidades de observação. Em segundo lugar, o aparecimento da teoria quântica dos átomos, a compreensão teórica e experimental da física nuclear e da física dos plasmas — o estudo dos gases eletronicamente neutros das partículas carregadas —, dotaram a astrofísica moderna dos seus fundamentos teóricos. Confiantes na compreensão das leis do mundo microscópico das partículas atómicas e subatómicas, os cientistas avançaram na elaboração de modelos matemáticos de objetos macrocósmicos, como as estrelas. Em terceiro lugar, os computadores de alta velocidade permitiram resolver as equações matemáticas que descrevem as múltiplas interações que têm lugar numa estrela e possibilitaram que os astrónomos observacionais compilassem grandes quantidades de informação. Sem estes computadores seria muito difícil aos astrofísicos confrontar as teorias com as observações e aos astrónomos processar a informação proveniente dos instrumentos (astrónomos).

Segundo os astrofísicos, as estrelas são esferas de gás quente, principalmente hidrogénio e hélio, cuja coesão é assegurada pela gravidade. Uma vez que a intensidade das forças gravitacionais é proporcional à massa, quanto maior for a massa das estrelas, tanto

mais intensas serão as forças que tendem a fazer com que a estrela entre em colapso. Pode estabelecer-se uma analogia entre uma estrela e um halterofilista olímpico que transpira e geme enquanto segura o peso sobre a cabeça, lutando contra a força da gravidade. Como a gravidade tende a fazer cair o peso sobre o atleta, este, por sua vez, exerce uma força igual e oposta para impedir a queda. No caso do atleta, esta força resistente tem como origem a energia química que está a ser libertada nos seus músculos. Porém, no caso da estrela, com um peso muito maior a suportar, donde provém a força resistente?

A pressão, como a de um gás dentro de um balão, é o resultado do movimento rápido das partículas de gás em colisão umas com as outras ou com as paredes. Quanto mais rapidamente colidem, tanto maior é a pressão; quanto mais elevada é a pressão, tanto mais se expande o gás, evitando que o balão murche. A velocidade de uma partícula está relacionada com a energia do seu movimento. Então descobrir uma fonte para a pressão que se oponha ao colapso gravitacional numa estrela e descobrir a fonte de energia calorífica que produza essas rápidas colisões são um e o mesmo problema.

Num pedaço de matéria, em condições normais, os núcleos dos átomos — as minúsculas e maciças partes centrais — estão muito separados uns dos outros. Todavia, no centro de uma estrela dificilmente se pode falar de «condições normais». Com efeito, o enorme peso, devido a toda a massa da estrela a fazer pressão sobre o núcleo — equivalente a cerca de 2 milhões de toneladas sobre uma área do tamanho de uma pequena moeda —, comprime os núcleos dos átomos de hidrogénio uns contra os outros. Além de

estar submetido a uma pressão elevada, o núcleo encontra-se a uma temperatura suficientemente alta para «acender» a fornalha termonuclear que transforma hidrogénio em hélio, processo este que gera energia calorífica. Se a temperatura no núcleo for insuficiente para acender a fornalha nuclear, a estrela não viverá muito tempo — uns meros 20 milhões de anos.

À medida que uma estrela se contrai, cerca de metade da energia gravitacional libertada transforma-se em energia calorífica, a qual sustém a contração da estrela. Daqui que, a curto prazo, a contração gravitacional forneça energia calorífica. A fornalha nuclear no núcleo é crucial somente porque o calor que gera pode compensar as perdas de calor da superfície da estrela, o que retarda a contração durante muito mais tempo. Vemos, assim, que é o equilíbrio dinâmico entre a gravidade, que tenta produzir o colapso da estrela, e o calor que este colapso gera, o responsável pela estabilidade temporária da estrela.

As estrelas não são realmente estáveis; apenas o aparentam porque vivem muito mais tempo do que nós. Desde o nascimento, a partir do gás cósmico, até à morte, os núcleos estão em contínua contração. A fim de evitar um colapso total, uma estrela tem de encontrar sempre novas fontes de energia que lhe proporcionem um aumento do tempo de vida. As fontes químicas de energia podem manter uma estrela acesa durante apenas cerca de 20 milhões de anos, período longo comparado com a duração de uma vida humana, mas curto à escala cosmológica. A fornalha nuclear de hidrogénio pode alimentar uma estrela com a massa do Sol durante milhares de milhões de anos, e a combustão de outros

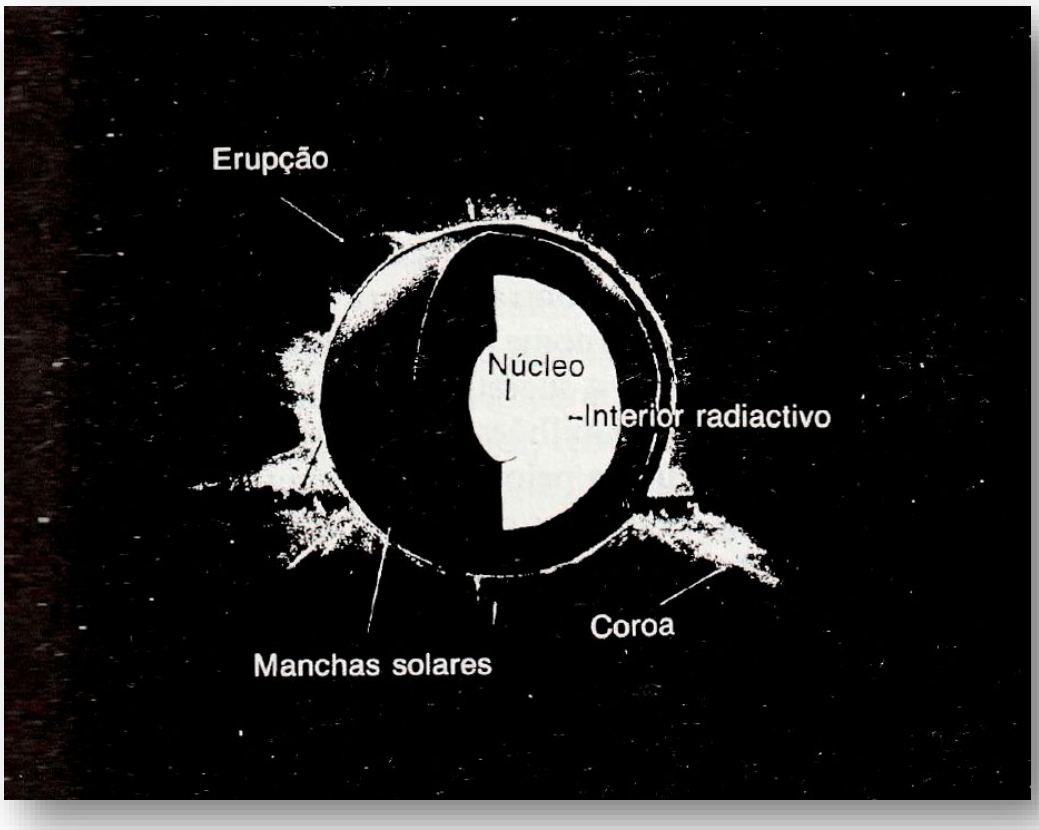
elementos, como o hélio, pode prolongar este período. Durante o período da fornalha nuclear as estrelas parecem estáveis, mas, de facto, estão a contrair-se, conquanto muito lentamente. Por fim, as estrelas devem morrer, devido à implacável compressão da gravidade e ao facto de qualquer fonte de energia ser finita.

Como podemos visualizar as estrelas? Podemos imaginá-las como consistindo numa série de camadas. Nas profundezas, no centro da estrela, existe um minúsculo núcleo cujas dimensões são apenas uma centésima parte das da estrela. O núcleo, que consiste em correntes de convecção de matéria quente, encerra o segredo da vida da estrela. As reações nucleares no núcleo não só fornecem o calor que prolonga a vida da estrela, mas também, nas últimas fases da existência desta, «cozinham» novos elementos pesados, essenciais para a formação dos planetas e da vida. Os nossos próprios corpos são feitos de material estelar.

Fora do núcleo de algumas estrelas existe uma camada gasosa que transfere a energia radiante do núcleo para as camadas de convecção exteriores. Este processo faz lembrar o aquecimento de água num fogão, com a diferença de se desenrolar ao longo de milhares de milhões de anos. A fornalha nuclear no núcleo é como o lume por baixo da panela. A água dentro da panela transfere o calor da chama para a superfície, onde a água se transforma em vapor. Na superfície de uma estrela como o nosso sol os astrónomos solares descobriram uma variedade de processos físicos complexos, parte dos quais não foi ainda compreendida. Contudo, sabemos que o Sol tem uma «atmosfera» que se estende muito para além da sua superfície e derrama matéria solar pelo espaço, por meio do vento

solar. Num certo sentido, os processos nucleares que ocorrem no interior do núcleo de uma estrela estão intimamente relacionados com regiões longínquas do espaço.

Os mesmos processos nucleares que produzem a energia e pressão que se opõem à gravidade também produzem fotões — partículas de luz. Mas as estrelas são opacas e os fotões não podem sair diretamente para fora do núcleo. Em vez disso, movem-se ao acaso no interior da estrela, numa espécie de «passeio de bêbado», colidindo com átomos de gás. Passados 10 milhões de anos, os fotões atingem, finalmente, a superfície — enfim, livres — , viajando então através do espaço interestelar e acabando por atingir os nossos olhos.



O interior de uma estrela como o nosso sol pode ser visualizado como uma série de camadas. A fornalha nuclear situa-se no minúsculo núcleo, no centro. A camada que envolve o núcleo transfere a radiação para as camadas convectivas exteriores, que, por sua vez, transportam a energia calorífica para a superfície. As manchas solares na superfície são manifestações de grandes campos magnéticos. Uma coroa solar de gás quente e brilhante, visível durante os eclipses, circunda o Sol.

Hoje compreendemos os complexos processos que se desenrolam no interior das estrelas, mas os físicos do século XIX, que desconheciam a existência das forças nucleares, tinham grande dificuldade em perceber como é que as estrelas podiam libertar tanta energia por períodos de tempo tão longos. Hermann von Helmholtz, físico alemão do século XIX, pensava que o Sol obtinha a energia a partir da contração gravitacional. Lord Kelvin, o físico inglês que deu nome à escala das temperaturas absolutas, calculou, aceitando a sugestão de Helmholtz, que a idade do Sol seria de apenas 20 milhões de anos. Este facto, esta tão curta vida do Sol calculada pelos físicos, criou um terrível problema aos geólogos e biólogos do século XIX, os quais haviam concluído, com base em fósseis encerrados em camadas de rochas sedimentares, que a Terra era muitíssimo mais velha. Como poderiam ser a Terra e a vida na Terra mais antigas do que a idade que se calculava para o Sol? Era algo que não fazia sentido.

Este conflito em torno das idades do Sol e da Terra criou uma divisão entre os físicos do século XIX, que se agruparam em «catastrofistas» — os que acreditavam que Deus intervinha periodicamente na Natureza (por exemplo, inundando a Terra) — e «uniformistas» — os que acreditavam que o mundo evolui lentamente em longos períodos de tempo, guiado por leis naturais. Lord Kelvin, que defendia vigorosamente o seu cálculo da idade do Sol, chefiou o ataque contra a visão uniformista. Partindo do

princípio de que a energia nuclear é a fonte da energia do Sol, os astrofísicos do século XX mostraram que a idade do Sol é comparável aos 4500 milhões de anos que atualmente se estimam para a idade da Terra. Constitui consequência gratificante da disputa oitocentista entre físicos, geólogos e biólogos o facto de todos eles hoje estarem de acordo no que toca à cronologia da história da Terra, medida em milhares de milhões, e não milhões de anos.

As estrelas vivem muito tempo. Segundo os astrofísicos, a duração de vida de uma estrela é, aproximadamente, proporcional ao inverso do quadrado da sua massa (ou, de modo mais geral, uma lei do inverso da potência). Daí que uma estrela com massa dez vezes superior à do Sol viva apenas um centésimo do tempo de vida estimado para o nosso sol — 10 mil milhões de anos —, isto é, uns escassos 100 milhões de anos, enquanto uma estrela com massa 90 vezes superior só pode viver 1 milhão de anos — nada, à escala do tempo cósmico —, o que pode explicar a razão por que não vemos muitas estrelas de massa elevada — desaparecem muito rapidamente. Nas estrelas que atingem massas superiores a 90 vezes a do Sol sucede que o peso esmagador aquece o núcleo até temperaturas altíssimas e causa a expulsão das camadas exteriores, reduzindo a massa da estrela. Tais estrelas tornam-se, portanto, instáveis, pelo que o valor máximo da massa de uma estrela parece ser de 90 vezes a massa do Sol.

E que se pode dizer sobre a massa mínima das estrelas, com massa muito inferior à do Sol? Tais estrelas são difíceis de observar, uma vez que não são, nem muito quentes, nem muito brilhantes.

Existe entre os astrónomos observacionais uma espécie de competição informal para descobrir a estrela menos luminosa — uma estrela que seja intrinsecamente pouco brilhante, e não apenas por estar muito distante. Durante algum tempo o recorde pertenceu à estrela VB10 da constelação Águia, a qual foi recentemente «batida» pela estrela RG0050-2722 da constelação Escultor.

Procurar a estrela menos luminosa não é um passatempo inútil, pois o conhecimento de um limite mínimo para a luminosidade estelar tem considerável importância para a teoria das estrelas. Uma vez que a luminosidade está relacionada com a massa, estrela menos luminosa tem também a menor massa — cerca de 2,3% da massa do Sol no caso da atual detentora do recorde. Abaixo de uma certa massa as estrelas não acendem as fornalhas nucleares e não podem arder. Assim, o conhecimento da massa mínima que uma estrela pode ter permitiria restringir os modelos teóricos admissíveis do processo de formação das estrelas. Sem massa suficiente o gás a partir do qual as estrelas se formam não pode atingir uma concentração suficientemente elevada para dar origem a uma estrela, sendo, assim, importante saber qual é a massa mínima necessária.

Uma interrogação que desponta ao contemplarmos objetos astronómicos de pouca massa é a seguinte: onde acabam as estrelas e começam os planetas? Júpiter, o planeta gigante do nosso sistema solar, tem uma massa de apenas 0,1% da massa do Sol — cerca da vigésima parte da massa da estrela de menor massa. Os astrofísicos acreditam que o processo de formação de planetas, mesmo dos grandes planetas, como Júpiter, é diferente

do processo de formação das estrelas. A teoria mais aceite sustenta que os planetas se formam a partir de um disco achatado de gás e detritos cósmicos que envolve uma estrela recém-nascida. Os astrofísicos suspeitam de que podem não existir objetos com massas situadas entre a massa da estrela menos luminosa e a massa dos grandes planetas, como Júpiter — ou talvez ainda os não tenham descoberto.

À medida que queimam o «combustível» nuclear, as estrelas procedem a contínuos ajustamentos, como os que os pais fazem às roupas de uma criança em crescimento. Devido a estes ajustamentos, mudanças e movimentos, as estrelas como o Sol emitem ruídos — uma sinfonia completa de sons. Qual é a origem da canção solar?

Recordemos que a luz criada pela fornalha nuclear no núcleo de uma estrela não pode sair imediatamente porque as estrelas são opacas. Consequentemente, a radiação luminosa aquece o gás das camadas exteriores, agitando-as, tal como o Sol aquece o ar num dia quente. Os gases quentes transportam calor para a superfície da estrela. Todo este movimento interior dos gases dentro de uma estrela dá origem a ondas sonoras. Estas viajam dentro da estrela (demora uma hora para que o som atravesse o interior do Sol), que atua como um gigantesco altifalante.

Alguns astrofísicos são da opinião de que no Sol esta energia acústica é lançada na coroa solar — a quentíssima alta atmosfera do astro. Outros, contestando esta posição, creem que a energia acústica vai parar à cromosfera do Sol, a camada superior, sendo a

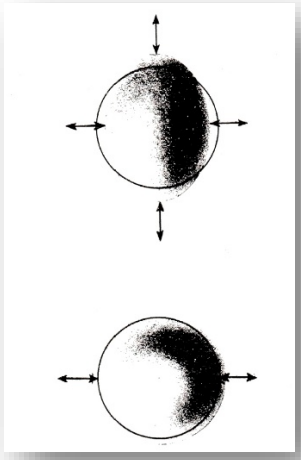
coroa, por sua vez, aquecida pelas correntes elétricas geradas pelo campo magnético solar. Dado que a coroa solar não é muito densa, não pode irradiar o conteúdo excedentário de energia e, em vez disso, expande-se, transportando para longe essa energia, como um poderoso motor a jato, ejetando gases quentes. Esta coroa solar em expansão chama-se «vento solar» e prolonga-se bastante para lá da Terra, até aos planetas exteriores. Como parte do sistema cósmico cíclico, o vento solar despeja em cada segundo centenas de milhões de toneladas de matéria solar no espaço exterior. A utilização de satélites artificiais capazes de penetrar no vento solar e transmitir para a Terra informações acerca da sua atividade possibilitou aos cientistas ouvir os rangidos, gemidos, gritos, estampidos e rufos da nossa canção solar.

O Sol não só canta, como também vibra. Observando cuidadosamente a forma do Sol, os cientistas notaram que vibra em vários modos de frequência, como uma tigela de gelatina trémula. Alguns destes modos de vibração têm apenas a duração de alguns minutos; outros demoram horas. Tais observações dos complicados movimentos exteriores do Sol deram aos cientistas algumas pistas sobre o movimento interno do núcleo, que não podem observar diretamente. Por exemplo, analisando os modos exteriores de vibração do Sol, são capazes de determinar com que rapidez o núcleo roda relativamente à camada exterior.

Como estas descobertas deixam entrever, os cientistas aprenderam muito acerca das estrelas através de um estudo cuidadoso da nossa estrela local, o Sol. Vivemos hoje na idade de ouro da astronomia solar. Quase anualmente são descobertas

novas e interessantes propriedades do Sol. Mais se aprenderá na próxima década, à medida que os satélites artificiais construídos para examinar o Sol de perto forem sendo enviados para as suas viagens inaugurais. E, no entanto, o Sol é apenas uma estrela de entre milhares de milhões de outras. Se pretendemos abranger a variedade das estrelas, os respetivos tipos e evolução, só o estudo do Sol não é suficiente. Como é que os astrónomos estudam as estrelas longínquas?

Convém verificar que quase toda a informação que os astrónomos adquirem acerca das regiões remotas do universo lhes chega via radiação eletromagnética — luz de vários comprimentos de onda, correspondendo à luz visível, ondas de rádio, radiação infravermelha e de micro-ondas, raios X e raios gama. Estudando a mensagem de luz que vem de fontes distantes, os astrónomos conseguem construir uma imagem do universo. A luz transporta uma enorme quantidade de informação, como, aliás, se torna evidente, se considerarmos em que medida o nosso conhecimento está baseado naquilo que vemos.



Modos de vibração simples do Sol. Os astrónomos solares extraem da observação destas e de outras vibrações mais complexas pistas sobre os processos que ocorrem no interior do Sol.

A luz das estrelas transporta duas importantes informações: a cor da estrela, que pode ir do avermelhado ao azulado, e a respetiva luminosidade. A cor está relacionada com a temperatura da superfície, implicando a cor azulada uma superfície muito quente e a cor avermelhada uma temperatura relativamente baixa. A luminosidade corresponde à energia total libertada pela estrela. Se fizermos um gráfico da luminosidade em função da temperatura de um grande número de estrelas, encontraremos, não um conjunto de pontos dispersos ao acaso, mas uma região estreita, uma linha, ao longo da qual se dispõem a maior parte das estrelas. Este gráfico é conhecido por «diagrama de Hertzsprung-Russell» (nomes dos seus descobridores) e a linha que corre das estrelas quentes e brilhantes até às frias e de brilho fraco chama-se «sequência principal».

Conhecer a localização de uma determinada estrela no diagrama de Hertzsprung-Russell é, provavelmente, a única informação importante de que os astrónomos dispõem acerca dela. Diz-lhes se a estrela tem muita ou pouca massa, facto que determina o respetivo destino, bem como várias outras propriedades. Como as estrelas evoluem queimando hidrogénio, movem-se lentamente ao longo da linha da sequência principal, acabando por atingir a altura das suas vidas em que ocorrem acontecimentos dramáticos e se afastam dessa linha. As estrelas que não estão na sequência principal são conhecidas por serem, em certos aspetos, especiais. São disso exemplo as gigantes vermelhas e as anãs brancas.

Embora a luminosidade e a temperatura sejam características muito importantes das estrelas, é com certeza da observação do espectro da sua luz que retiramos as informações mais pormenorizadas. A luz parece branca, mas, se for refratada por meio de um prisma, decompõe-se num espectro de cores. Sobrepostas ao contínuo do espectro, do vermelho ao azul, aparecem riscas espectrais distintas, compostas por intensas e estreitas faixas coloridas. Podem faltar, porém, algumas riscas. Estas riscas espectrais de cores específicas, ou comprimentos de onda, são devidas a transições específicas ocorridas nos átomos que emitem essa luz. A partir da disposição das cores no espectro de uma estrela, podemos deduzir que tipos de átomos — que elementos químicos — compõem uma estrela distante. Os espectros estelares são as impressões digitais das estrelas.

Auguste Comte, o teórico social francês fundador do positivismo, observou em 1825, no *Cours de philosophie positive*, que algo que nunca conheceríamos seria a composição química das estrelas. Precisamente na mesma altura em que Comte fazia esta afirmação os alemães Joseph von Fraunhofer e, mais tarde, Gustav Kirchhoff e Robert Bunsen estabeleciam os princípios da moderna espectroscopia. Algumas décadas mais tarde o estudo da composição química das estrelas tinha-se transformado numa excitante área de investigação. As observações de Comte fazem-me lembrar aquelas pessoas, nas quais incluo cientistas, que afirmam hoje que nunca conheceremos a origem do universo. Da mesma maneira que Comte pensava que nunca conheceríamos a composição das estrelas por estarem tão distantes no espaço, estas

pessoas argumentam hoje que nunca conheceremos a origem do universo por estar tão distante no tempo. Mas estão enganadas.

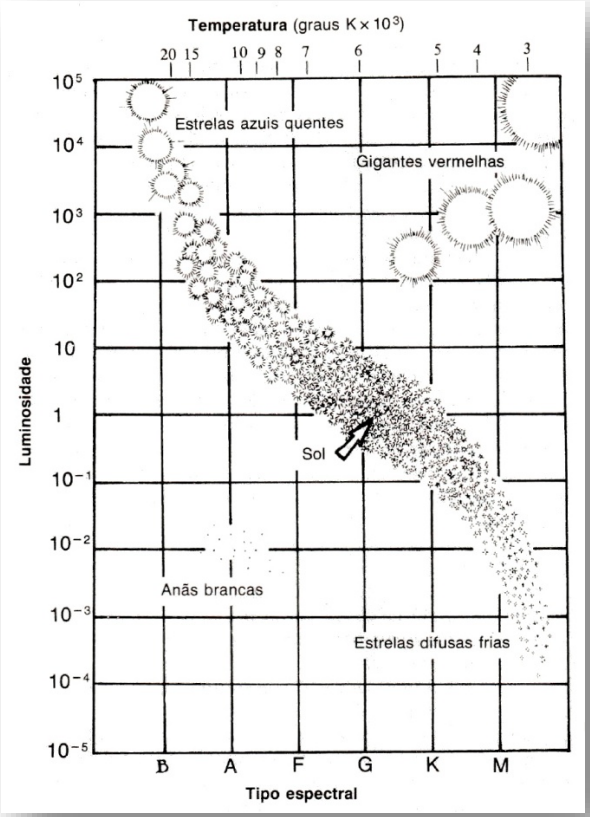


Diagrama Hertzsprung-Russell esquemático da luminosidade das estrelas em função da temperatura. A maior parte das estrelas estão na sequência principal e movem-se ao longo dela no decurso das suas vidas. A seta indica a localização do nosso sol. No canto superior direito encontram-se as estrelas gigantes vermelhas, enquanto no canto inferior esquerdo estão as anãs brancas — ambos os tipos são especiais, uma vez que não se situam na sequência principal.

Através do exame dos espectros os cientistas descobriram que as estrelas diferem na sua composição química. Todas as estrelas contêm sobretudo hidrogénio e hélio. Todavia, algumas têm uma abundância relativamente elevada — 1% a 2% — de elementos químicos mais pesados do que o hidrogénio e o hélio, enquanto outras apresentam apenas alguns vestígios — uma minúscula fração de 1% — destes elementos pesados. Esta distinção conduz

a uma classificação das estrelas em população I — aquelas que contêm elementos pesados — e população II — com uma percentagem relativamente baixa de elementos pesados.

O astrónomo germano-americano Walter Baade descobriu estas duas populações de estrelas distintas em 1942, durante a segunda Guerra Mundial, utilizando o telescópio de 2,5m de Mount Wilson, perto de Los Angeles. Curiosamente, a guerra desempenhou um papel indireto nesta descoberta. Dado que Baade era cidadão alemão, não podendo juntar-se ao esforço de guerra americano, foi deixado como único utilizador do telescópio. Mais tarde, devido à possibilidade de ataques inimigos, Los Angeles foi colocada em regime de ocultação de luzes durante a noite, o que criou excelentes condições de observação em Mount Wilson. Baade conseguiu observar muitas estrelas individuais na parte central da galáxia de Andrómeda e encontrou aí as duas populações distintas de estrelas. A população I era constituída por jovens estrelas azuladas, enquanto a população II era formada pelas estrelas vermelhas mais velhas. Fora desta parte central, nos braços em espiral da galáxia, encontrou principalmente estrelas jovens da população I, do tipo do nosso sol. Os aglomerados globulares de estrelas que habitam o halo, envolvendo a galáxia, datam da altura da formação desta; estas estrelas constituem a população II, mais antiga.

Qual é a origem das duas populações de estrelas? A maioria dos astrofísicos acreditam que somente elementos leves, como o hidrogénio e o hélio, e um pouco de deutério e lítio, mas nenhuns elementos químicos pesados, teriam sido «cozinhados» no *big-bang*, que esteve na origem do universo. Chamam a estes primeiros

elementos «elementos primordiais» para os distinguirem dos elementos que mais tarde foram sintetizados nas fornalhas nucleares dentro das estrelas ou nas explosões de supernovas. No princípio o universo consistia apenas num gás constituído pelos elementos primordiais hidrogénio e hélio. As estrelas vieram depois, quando o gás, ao condensar-se, as formou. As primeiras estrelas não podem ter tido quaisquer elementos pesados, pela simples razão de que estes não existiam. Estas estrelas «puras» — estrelas hipotéticas, porque ninguém jamais as viu — constituem a população III de estrelas. Contudo, estas estrelas primordiais realizaram um trabalho importante, dado que «cozinharam» os primeiros elementos mais pesados, como o carbono e o azoto, a partir dos elementos leves, hidrogénio e hélio, nas suas fornalhas nucleares e por transmutação. Ao morrerem em explosões espetaculares, as estrelas da população III espalharam a sua produção de elementos pesados pelo espaço interestelar. Estes raros elementos pesados acabaram por ir parar às mais velhas estrelas observadas — as estrelas da população II. Estas, por sua vez, «cozinharam» mais elementos pesados, que foram espalhados pelo espaço e foram enriquecer a população I, a mais jovem população de estrelas.

Os elementos pesados presentes no meio interestelar de gás e poeiras que impregnam o espaço entre as estrelas tendem a coalescer quando os átomos colidem, formando pequenos grãos de poeira, com as dimensões de um centésimo de milésimo do centímetro. Estes minúsculos grãos de poeira congregaram-se em nuvens escuras e nebulosas nos braços em espiral das galáxias, que são os locais de nascimento de novas estrelas. Por um grande

processo cíclico nascem novas estrelas dos despojos das que morreram. Mas como é que as estrelas nascem realmente?

É pouco provável que alguma vez assistamos ao nascimento de uma estrela. As estrelas são como animais selvagens. Podemos ver as muito jovens, mas nunca o seu nascimento, que é um acontecimento velado e secreto. As estrelas nascem em espessas nuvens de poeira e gás, nos braços em espiral das galáxias, tão espessas que a luz visível não pode penetrá-las. No entanto, e a despeito da impossibilidade de se poder observar diretamente o nascimento das estrelas, os astrónomos deram grandes passos em frente na compreensão deste complexo processo, que envolve muitos tipos de interações físicas, e por isso muitos enigmas permanecem, mas a maioria dos astrofísicos estão seguros de que uma teoria completa se encontra em marcha.

Uma galáxia é um lugar sujo. No espaço entre as estrelas existe um meio interestelar composto por gás (moléculas e átomos dispersos) e poeiras (minúsculas partículas de matéria) vogando à deriva. É muito pouco denso, não compreendendo mais de 5% da massa da galáxia, mas em certas regiões encontra-se mais concentrado. O satélite *Copernicus*, lançado em 1972 e mantido operacional até 1980, projetado para medir o espectro da luz na região do azul e do ultravioleta longínquo, luz esta que não pode penetrar na atmosfera da Terra e só pode ser detetada acima dela, revolucionou a nossa compreensão do meio interestelar. A partir das informações recolhidas por este satélite, os astrónomos ficaram a conhecer as abundâncias de vários átomos e moléculas e

respetivas temperaturas, bem como muitas outras propriedades do meio interestelar.

Este impregna o disco da galáxia. É particularmente denso nos braços da espiral delineada pelas jovens estrelas brilhantes. A relação entre os braços da espiral e as jovens estrelas brilhantes que neles se encontram é como a relação entre a galinha e o ovo: quem apareceu primeiro? Será que enormes vagas no meio interestelar varrem matéria ao acaso para os braços em espiral, criando, assim, as condições ideais para a formação de estrelas jovens? Ou será que a existência de «viveiros» estelares é uma condição prévia para a formação dos braços? Os astrónomos teóricos lutam com estas questões difíceis e concebem modelos de computador do processo de formação das estrelas com o objetivo de para elas encontrarem respostas. No entanto, e independentemente de quaisquer modelos de computador, pelas observações torna-se claro que o gás e as poeiras se concentram sob a forma de milhares de «complexos de nuvens moleculares gigantes» — enormes nuvens de gás — e que é aí que as estrelas nascem.

Até há cerca de quinze anos os astrónomos pensavam que a maior parte do gás na nossa galáxia era composto por átomos isolados, mas o uso de radiotelescópios e satélites detetores sensíveis à radiação emitida pelas moléculas permitiu concluir que 10% a 50% desse gás é molecular — átomos agrupados — e tende a concentrar-se em nuvens gigantes. Cerca de 99% deste gás é hidrogénio molecular — 2 átomos de hidrogénio agrupados —, mas foram detetadas, pelo menos, 53 outras moléculas, incluindo as de

álcool etílico, ou seja, de *vodka*. As nuvens de gás próximas do centro da nossa galáxia contêm *vodka* suficiente para encher mais de 10 000 taças do tamanho da Terra.

Uma propriedade das nuvens moleculares de gás (em contraste com as nuvens atómicas) é que são muito mais frias e mais densas. Estes complexos gigantes de nuvens moleculares são os objetos de maior massa na galáxia. Foram detetados mais de 4000, ocupando principalmente os braços em espiral. Constantemente sujeitas a um complexo processo de transformação, as nuvens parecem viver muito pouco tempo, apenas um centésimo da idade do Sol ou da Terra, especulando-se acerca do seu nascimento, idade média e morte. No entanto, parece claro que são o local de nascimento das estrelas de grande massa e vida curta e, possivelmente, também das estrelas de vida longa.

Contemplando a constelação de Oríon numa noite clara, pode distinguir-se a grande nebulosa na Espada do Caçador, não longe das três estrelas brilhantes do Cinturão. A nebulosa de Oríon é um viveiro estelar rico em fenómenos físicos complexos, um «laboratório» celeste, bem como o local onde se encontram os complexos gigantes de nuvens moleculares. No centro das enormes nuvens de poeiras e de gás em agitação encontra-se o Trapézio, um conjunto de quatro estrelas brilhantes que atuam sobre o material da nebulosa, iluminando-o, enquanto excitam o gás, fazendo-o brilhar por si próprio. Enormes ondas de choque, que talvez ponham em marcha a formação de estrelas, podem ser vistas, propagando-se nas nuvens. Formaram-se novas estrelas no

interior da nebulosa de Oríon, em locais inacessíveis à nossa observação. Como é que os astrónomos o descobriram?

Há mais de 180 anos William Herschel notou a existência de «energia radiante» para além da extremidade vermelha do espectro da luz visível. Em linguagem moderna, diríamos que descobriu que o Sol emite não só a luz visível, mas também luz com comprimento de onda na região dos infravermelhos. Hoje em dia a astronomia de infravermelhos é um novo e importante ramo da ciência astronómica, com alguns observatórios já construídos e dedicados a esse estudo, como, por exemplo, o *infra-red telescope facility* (IRTF), da NASA, o *United Kingdom infra-red telescope* (UKIRT), ambos no cume do extinto vulcão havaiano Mauna Kea, o *Wyoming infra-red telesc*ope, o *multiple mirror telescope* (MMT), no Arizona, o novo refletor de 2,12m do México e, mais recentemente, o *infra-red astronomy satellite* (IRAS)[3], de 1 t, colocado em órbita em torno da Terra, mas já fora de serviço. Estes instrumentos, sensíveis à luz infravermelha, facultam-nos, atualmente, uma nova e empolgante compreensão de constituintes do cosmos outrora invisíveis.

Ao contrário da luz visível, a radiação infravermelha pode penetrar nas espessas nuvens dos viveiros estelares, como a nebulosa de Oríon. Através da deteção de micro-ondas e radiação infravermelha, os astrónomos Eric E. Belkin e Gerry Neugebauer, do California Institute of Technology[4], descobriram em 1965 um objeto misterioso — o objeto B-N — no interior de uma das nuvens de Orion. Na década seguinte seis outros centros difusos de radiações infravermelhas foram detetados na nuvem gigante. Os

astrónomos interrogaram-se então se seriam ou não protoestrelas — corpos à beira de se tornarem verdadeiras estrelas, capazes de queimarem combustível nuclear. Parece agora que essas fontes de radiação infravermelha são devidas a estrelas com idade mais avançada do que as protoestrelas — estrelas jovens de grande massa, as quais libertam massa de uma forma extremamente energética, provocando fortíssimos ventos estelares de partículas, cuja origem é desconhecida. À distância de cerca de um mês-luz da estrela, a corrente atinge a nuvem de gás circundante, aquecendo-a e produzindo radiações na região das micro-ondas e do infravermelho. Na realidade, a nuvem de gás amplifica os sinais rádio de micro-ondas, os quais são detetados na Terra — uma bela confirmação do processo físico envolvido. Os astrónomos especialistas em radiações infravermelhas e de rádio descobriram objetos semelhantes — estrelas jovens encerradas ainda em casulos de poeira e gás — não somente na nebulosa de Oríon, mas também nas nebulosas de Ómega e de Trífido —, o que constitui talvez a comprovação mais convincente de que estas nebulosas são regiões de formação de estrelas.

Em 25 de Janeiro de 1983 foi dado um grande impulso à astronomia de radiações infravermelhas com o lançamento do IRAS, que fez um mapa completo da radiação infravermelha no céu. De sete em sete meses o telescópio varria por duas vezes o céu inteiro. A sua principal missão consistia em examinar os «pontos quentes» onde os astrónomos suspeitam terem origem as estrelas. O satélite localizou dúzias de fontes de radiação infravermelha na nebulosa de Tarântula, situada na Grande Nuvem de Magalhães, uma galáxia vizinha. Algumas destas fontes devem-se,

provavelmente, à interação de estrelas jovens com a espessa poeira que as envolve; outras poderão ser verdadeiras protoestrelas — uma possibilidade que entusiasma os astrónomos.

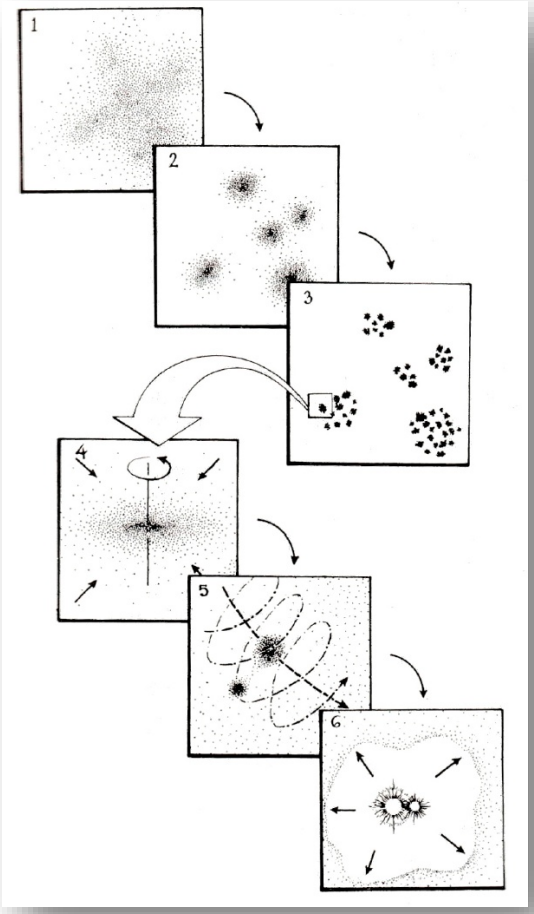
Como nasce uma estrela? Para responder a esta pergunta recorremos aos teóricos da astrofísica e aos seus modelos de computador. Estes cientistas elaboraram uma imagem da formação das estrelas em complexos gigantes de nuvens moleculares — nuvens constituídas principalmente por hidrogénio molecular e cuja massa é muito maior do que a de uma única estrela. No princípio a grande nuvem começa a fragmentar-se em pedaços mais pequenos, devido à atração gravitacional mútua de todas as partículas de gás e poeira, e a massa de cada um destes pedaços é, aproximadamente, igual à massa de uma estrela. Daqui concluímos que algumas estrelas nascem em grupos. Os astrónomos observaram vários «aglomerados abertos», grupos de estrelas jovens, deslocando-se em conjunto na mesma direção, que, provavelmente, nasceram juntas a partir da mesma concentração de poeiras e gás.

As estrelas de grande massa e vida curta surgem nas grandes nuvens de gás; outras, fugindo a estes esquemas de produção em massa, nascem individualmente a partir de uma classe diferente de nuvens escuras muito mais pequenas conhecidas por «glóbulos». Cerca de duzentos destes glóbulos, quase esféricos, ricos em moléculas grandes, foram encontrados até à distância de 15.000 anos-luz do Sol e alguns deles estão a comprimir-se, possuindo massa necessária para formar uma única estrela. Julga-se que o nosso sol foi criado a partir de um destes glóbulos. A existência de

glóbulos escuros sugere a possibilidade de haver muitos outros mecanismos de formação de estrelas em que não intervenham as nuvens gigantes de gás. É, inclusivamente, possível que diferentes mecanismos possam coexistir dentro de uma mesma nuvem gigante. Examinaremos, no entanto, o que acontece a uma massa de gás numa região de formação de estrelas.

Acompanhando a evolução de uma dessas massas, assistimos ao desenrolar de complexos processos físicos. Embora constituam apenas cerca de 1% da massa do gás, os grãos de poeira são importantes núcleos nos quais as moléculas de gás do meio se podem agrupar e moldar. A presença de poeiras pode também proteger as moléculas da intensa radiação, que poderia dissociá-las. A ubíqua força de gravidade tenta agrupar as massas de poeiras e gás, enquanto a ação dos campos magnéticos, do calor, da turbulência e da rotação nessas massas tende a dispersá-las, competindo com a gravidade. Por fim, a gravidade triunfa em todas estas disputas. No caso da formação de uma estrela, a competição é um processo longo — cerca de 10 milhões de anos. A contração gravitacional é acompanhada pelo aumento de densidade da massa, e, quando esta se torna opaca, a temperatura sobe. Quando a massa de gás atinge dimensões algumas centenas de vezes superiores às do nosso sistema solar, a temperatura é de 0ºC — o ponto de congelação da água. Ao contrair-se até ficar do tamanho do sistema solar, a temperatura eleva-se a milhares de graus Celsius — mais quente do que o ponto de fusão dos metais. Cerca de 100.000 anos depois de ter começado a contrair-se esta massa de gás cabe na órbita da Terra e a sua temperatura é de centenas

de milhares de graus. Tal objeto, que não é ainda uma estrela, recebe a designação de protoestrela.



Representação esquemática do processo de formação de uma estrela a partir da contração gravitacional de uma nuvem de gás e de poeiras. Uma nuvem gigante situada no braço em espiral da nossa galáxia fragmenta-se em pedaços, cada um dos quais continua a fragmentar-se e, animado de movimento de rotação, a contrair-se ao longo de um período de milhões de anos. A figura ilustra a formação de um sistema binário de estrelas — um resultado possível deste processo. Finalmente, as estrelas acendem-se e varrem para longe as nuvens remanescentes. Tais nascimentos estelares são acompanhados por jatos de matéria ejetada das novas estrelas, cuja origem não foi ainda compreendida.

Se acompanharmos a evolução de uma protoestrela, que dura cerca de 10 milhões de anos, veremos que continua a contrair-se e a aquecer. Finalmente, a temperatura no centro atinge os 10

milhões de graus Celsius necessários para iniciar a reação nuclear de fusão do hidrogénio, e nasce uma verdadeira estrela.

Esta imagem teórica simples do colapso de uma nuvem de gás apresenta problemas quando comparada com as observações. Os astrónomos observam indiretamente duas rápidas correntes de matéria que flui em direções opostas a partir das estrelas recém-nascidas; estas correntes são milhões de vezes mais intensas do que o vento solar com origem no Sol. A imagem de contração contínua não explica a origem destes misteriosos jatos gémeos de matéria emitidos pelas estrelas jovens. Os astrofísicos estão a aperceber-se de que o nascimento das estrelas é dinamicamente mais complexo do que anteriormente imaginaram e trabalham arduamente, inventando novos modelos e modificando os antigos.

Alguns modelos de computador do processo de formação das estrelas correntemente em voga mostram duas possíveis saídas para uma nuvem em contração, dependendo, entre outras coisas, do valor inicial do momento de rotação. Uma possibilidade é a poeira e o gás em contração transformarem-se em duas estrelas orbitando uma em torno de outra — um sistema binário. A outra é formar-se só uma estrela, sendo o momento de rotação distribuído por um sistema planetário auxiliar. Se estes modelos estiverem corretos, então metade das estrelas da nossa galáxia, uma vez que se pode observar que são binárias, não terá planetas acompanhantes, enquanto a outra metade — estrelas isoladas — tem, provavelmente, sistemas planetários. Esta poderosa conclusão — todas as estrelas isoladas têm planetas — depende de complicados pormenores do processo de transferência do momento

de rotação dentro da nuvem de gás, processo este que ainda não é bem compreendido. Alguns astrofísicos pensam que o momento de rotação é transportado por um vento estelar e que o momento de rotação dos planetas não é importante.

Apesar destas complicações, muitos astrofísicos teóricos acreditam que uma estrela isolada, como o nosso sol, deixa atrás de si, ao nascer, um disco de matéria excedentária que roda em torno dela, como a clara de um ovo estrelado envolve a gema. Foi a partir deste disco, designado por «nebulosa solar», que o sistema solar subsequentemente se formou. Contudo, os problemas matemáticos da elaboração de modelos teóricos da nebulosa solar são particularmente difíceis porque não existem dados observacionais para guiarem o autor do modelo. Nunca ninguém viu um sistema planetário em nenhum estádio de formação, exceto no último. Todavia, se acreditamos nestes modelos de nebulosa solar, então os planetas formaram-se à medida que a matéria no disco começou a agrupar-se em massas concentradas. Massas maiores acumularam ainda maior quantidade de matéria. Uma teoria sustenta que a cintura de asteroides situada entre Marte e Júpiter é uma coleção de massas que nunca se combinaram para formar um verdadeiro planeta. Ao acender-se pela primeira vez, o Sol emitiu um vento fortíssimo que varreu para longe os detritos que não se tinham agrupado em corpos de massa elevada, como planetas e luas.

Se este cenário para a origem do sistema solar está correto, então os sistemas planetários devem ser, pelo menos, tão comuns como as estrelas idênticas ao nosso sol, uma estrela isolada típica.

Aprenderemos muito mais acerca da origem do nosso sistema solar nos anos que se seguem, graças aos dados já recolhidos pelas sondas enviadas aos planetas. No entanto, não devemos esperar respostas rápidas. A despeito da ida dos homens à Lua, da recolha de rochas lunares e dos dados selenológicos, continuamos a ter dúvidas quanto à origem da Lua — para já não falar da origem do sistema solar.

O estudo teórico da nebulosa solar, o disco achatado de matéria que rodeia o sol-nascente, é apenas um exemplo específico do fenómeno astrofísico geral dos «discos de acreção». Os anéis à volta do planeta Saturno são outro exemplo de discos de acreção, o mesmo sucedendo, possivelmente, com as estrelas numa galáxia em forma de disco. Os buracos negros e as estrelas de neutrões devem também estar envolvidos por um disco de gás quente, sendo provável que emitam um sinal energético quando o gás se precipita sobre eles. É então evidente que, em determinadas circunstâncias, a matéria, sob a forma de gás ou de poeira, tende a formar um disco semiestável circundando um objeto de grande massa. O estudo matemático dos discos de acreção conduzir-nos-á, no futuro, a uma mais profunda compreensão não somente da origem do sistema solar, como também dos enigmáticos sinais que emanam do espaço longínquo.

Cerca de metade de todas as estrelas da nossa vizinhança passam a vida com uma companheira: estão agrupadas aos pares. Algumas, desdenhando tais convenções sociais estelares, são membros de agrupamentos com três ou quatro elementos. Os pares binários têm sido muito estudados. Em alguns casos, as duas

estrelas que os compõem orbitam tão perto uma da outra que os binários recebem a classificação de «binários de contato» — as estrelas, efetivamente, tocam-se, trocando grandes quantidades de massa. Durante vários anos os astrofísicos andaram intrigados sem saberem por que razão os membros dos pares binários apresentavam idades tão diferentes. Pensavam que, se tivessem nascido ao mesmo tempo, a partir de uma mesma nuvem de gás, deveriam ter a mesma idade. Todavia, se duas estrelas trocam massa durante a sua evolução, então a idade aparente pode ser bastante diferente, uma vez que adicionar ou subtrair massa a uma estrela pode alterar-lhe a idade aparente. A observação dos binários de contato, por um lado, e a elaboração de modelos, por outro, confirmaram estes mecanismos de troca.

A maior parte das estrelas, uma vez formadas, levam uma vida sem nada de especial a registar. Queimando hidrogénio nos núcleos, «cozinhando» hélio, fazendo ajustamentos, vão cantando e vibrando ao longo de milhares de milhões de anos. Só uma pequena parte das estrelas apresenta comportamento não convencional. Entre estas encontram-se as estrelas T do Tauro, estrelas bebés agitadas, com idades compreendidas entre 100.000 e 1 milhão de anos — muito pouco tempo para uma estrela —, as quais apresentam espectros de emissão complexos e muitas vezes anómalos. Estão normalmente envolvidas por gás quente e jatos luminosos de matéria, o que dificulta ainda mais a interpretação dos processos físicos. São enigmáticas, devido à diversidade de características peculiares que possuem, diversidade que, uma vez compreendida, muito ensinará aos astrofísicos acerca destas

estranhas e jovens estrelas, bem como acerca dos gases que as acompanham.

Outra classe de estrelas com interesse são as cefeides variáveis — variáveis porque o seu brilho oscila com um período que vai de três dias a várias semanas. São focos de brilho variável no céu. A Polar — a estrela do Norte — é uma cefeide variável cujo brilho varia de 10% num período de quatro dias. As cefeides são estrelas velhas que descobriram que podem libertar a energia mais eficientemente fazendo pulsar o seu brilho intrínseco. Mas a propriedade mais notável de uma cefeide reside no facto de a taxa de variação estar relacionada com o brilho de uma forma precisa.

Esta relação importante — a relação período-luminosidade — foi descoberta em 1912 por Henrietta Leavitt, do Harvard College Observatory, ao examinar chapas fotográficas da Pequena Nuvem de Magalhães enviadas para Harvard por um observatório de Arequipa, no Peru. Uma vez que as nuvens de Magalhães, galáxias irregulares situadas nos limites da Via Láctea, estão muito distantes, as suas estrelas estão todas aproximadamente à mesma distância (da mesma maneira que Paris e Lião estão aproximadamente à mesma distância de São Francisco). Como sabemos que todas estas estrelas estão à mesma distância de nós, podemos concluir que a luminosidade *aparente* (o brilho das estrelas observado através de um telescópio) é proporcional à luminosidade *intrínseca* (a quantidade total de energia, em forma de luz, que sai da estrela). Leavitt notou que os períodos das cefeides variáveis da Pequena Nuvem de Magalhães estavam relacionados com a luminosidade aparente, portanto também com

a luminosidade intrínseca. O que descobriu foi que, quanto mais brilhante for a estrela, maior será o período de pulsação, regularidade esta que ficou conhecida como a relação período-luminosidade. Calibrando esta relação com o auxílio das cefeides vizinhas, cujas distâncias absolutas conhecemos, podemos obter uma relação entre o período observável e a luminosidade aparente, por um lado, e a distância até à estrela, por outro. A importância das cefeides variáveis consiste em que fornecem aos astrónomos um método de determinarem a distância às estrelas longínquas e até a distância a outras galáxias nas quais possamos observar outras cefeides. As cefeides de brilho variável são a régua com que se mede o tamanho do universo.

Quem poderia imaginar, um século atrás, o que hoje sabemos acerca do nascimento e da vida das estrelas? Temos vivido numa idade de ouro da astrofísica, na qual compreendemos pela primeira vez os processos fundamentais da vida das estrelas. Continuam, no entanto, a existir problemas, que não são problemas de princípio, mas antes problemas de complexidade, que continuarão a ocupar os astrofísicos nas décadas vindouras. Ajudados por novos instrumentos e computadores, os astrofísicos continuarão a construir modelos de estrelas, confrontando teorias com observações cada vez mais aperfeiçoadas.

Os povos antigos adoraram o Sol como fonte de vida. No futuro, à medida que cresce o nosso conhecimento dos sistemas estelares, talvez concluamos que a vida planetária é parte da evolução de um sistema estelar. As estrelas criaram as condições para a vida, e, por esse facto, estamos-lhes vinculados. Contudo, as estrelas não são

seres eternos e independentes; são a prole da galáxia. A sua vida, e portanto a nossa, está interligada com processos galáticos que ocorrem a escalas de tempo incompreensíveis para a perspetiva humana. No entanto, vista em perspetiva temporal alargada, nenhuma parte do universo é verdadeiramente independente do todo. Não será possível que, da mesma maneira que o meio-ambiente, na Terra, deu forma à vida, em escalas de tempo de milhões de anos, o meio-ambiente do universo dê forma ao futuro da vida em escalas de tempo de milhares de milhões de anos? A vida pode vir a descobrir que todo o universo se torna o palco da sua existência. Estará o nosso destino a ser decidido nas estrelas? Ou será que esta visão estelar do nosso futuro não passa de uma ilusão incessantemente refletida nos espelhos da nossa mente?

---

[1] Literalmente, O Universo em Expansão. (*N. do T.)*

[2] Literalmente, A Constituição Interna das Estrelas. (*N. do T.*)

[3] Satélite de astronomia de infravermelhos. (*N. do T.*)

[4] Instituto de Tecnologia da Califórnia. (*N. do T.*)

CAPÍTULO 3

Morte das Estrelas – Necroscopia Astral

> Este «estremecimento perante o belo», este facto incrível de uma descoberta motivada pela busca da beleza em matemática encontrar réplica exata na Natureza, leva-me a dizer que a beleza é aquilo a que o espírito humano reage no mais profundo de si mesmo.
>
> *SUBRAHMANYAN CHANDRASEKHAR*

As estrelas são uma imagem da eternidade. Parecem fixas, imutáveis, indestrutíveis, e foi assim que os astrónomos as imaginaram durante séculos. Esta imagem da eternidade das estrelas alterou-se, porém, drasticamente no espaço de uma vida humana. Tal como os seres vivos, as estrelas morrem. A sua substância transforma-se, os elementos espalham-se pela galáxia, os restos de algumas delas ficam fechados em túmulos celestes e em tal segurança que permanecem fora do alcance do espaço e do tempo.

Compreender o destino das estrelas faz parte de um enigma mais geral da física, chamado o «problema do estado final», que consiste em determinar qual o destino final de todas as coisas, sejam elas quais forem, se esperarmos o tempo suficiente. Muitos dos objetos materiais que nos cercam parecem ter uma qualidade de permanência — os mares, as montanhas, a atmosfera da Terra. No entanto, se esperarmos o suficiente, todos se transformarão. Até

os átomos a partir dos quais esses objetos são feitos estão destinados à extinção e à aniquilação. Qual é o fim último das coisas? Qual o destino da nossa galáxia e do universo? Os físicos especulam acerca da resposta a esta pergunta e chegam a uma variedade de conclusões — o assunto não está encerrado. No entanto, podemos chegar a algumas conclusões concretas se examinarmos o destino das estrelas. Aqui descobrimos novos processos físicos tão surpreendentes que até os cientistas que primeiro os compreenderam só relutantemente aceitam as conclusões do raciocínio, pois a força esmagadora da gravidade que acompanha a morte violenta das estrelas cria condições físicas que desafiam o nosso próprio entendimento das leis da Natureza. Acompanhemos uma estrela na sua agonia de morte.

Após milhares de milhões de anos a queimar hidrogénio e a convertê-lo em resíduos de hélio, acaba-se o combustível de hidrogénio no núcleo da estrela -— uma crise de energia que sela o seu destino. Recordemos que a fornalha nuclear alimenta as elevadas temperaturas que permitem à estrela resistir ao colapso gravitacional. Quando o fogo se apaga, a estrela entra de novo em colapso. Os astrofísicos anteveem três possíveis destinos para as estrelas que entram em colapso: tornam-se anãs brancas, ou estrelas de neutrões, ou buracos negros. É sobretudo da massa total da estrela que depende o respetivo destino futuro. As estrelas com massa inferior a 1,4 vezes a massa do Sol tornam-se anãs brancas — estrelas minúsculas feitas de matéria milhares de vezes mais densa do que a matéria vulgar. As estrelas de maior massa sofrem uma explosão de supernova, dando origem a uma estrela de neutrões — basicamente um núcleo atómico gigante, do

tamanho de uma cidade. As estrelas cujo núcleo tem uma massa superior a cerca de duas vezes a massa do Sol transformam-se, em princípio, num buraco negro — objeto no qual o próprio espaço fica «virado do avesso». Como chegaram os astrónomos a estas bizarras conclusões acerca da morte das estrelas? Que provas temos da existência de tão estranhos objetos? Examinemo-los um a um.

1. As Anãs Brancas

Cale-se. Não diga disparates

ARTHUR S. EDDINGTON

Nunca ninguém, nos tempos modernos, viu o colapso de uma estrela da nossa galáxia; os astrofísicos podem, porém, elaborar uma imagem dos anos finais de uma estrela através do uso de modelos de computador que reproduzam as leis da física nuclear e da termodinâmica. Uma vez transformado em hélio todo o hidrogénio no núcleo de uma estrela, perde-se o precioso equilíbrio da estrela e o minúsculo núcleo (com apenas cerca de um centésimo das dimensões da estrela) começa a ser comprimido pela pressão imensa das camadas de gás exteriores. Esta compressão gravitacional aquece o núcleo, que, por sua vez, aquece as camadas exteriores, fazendo com que o hidrogénio dessas camadas (que não está ainda esgotado) arda vigorosamente. Através de um complexo processo de transferência de energia, as camadas exteriores expandem-se, fazendo dilatar a superfície da estrela. A estrela fica assim «inchada», com milhares de vezes o volume inicial, e torna-se vermelha — consequência do facto de as camadas exteriores, ao

expandirem-se, terem arrefecido, tornando-se o gás a temperaturas mais baixas vermelho, em vez de branco. Tal estrela — um quente e minúsculo núcleo cercado por uma enorme quantidade envolvente de gás quente — chama-se uma gigante vermelha. São exemplos as estrelas vermelhas Pólux e Arcturo.

O núcleo continua a ser comprimido até uma densidade de milhares de vezes a densidade do núcleo numa estrela normal, até que a temperatura do centro (que vai subindo) atinge 100 milhões de Kelvin. A esta alta temperatura inicia-se um novo processo de fornalha. Os núcleos de hélio no centro do núcleo fundem-se para formar o elemento carbono, mais pesado, e neste processo toda a estrela sofre uma redução do seu enorme volume, atingindo um novo equilíbrio. Graças a esta nova energia proveniente da transformação do hélio em carbono, o núcleo dá à estrela um novo alento. Mas não por muito tempo.

O que acontece a seguir depende principalmente da massa total da estrela. Ou as camadas exteriores da estrela têm massa suficiente para continuarem a comprimir o núcleo e a fazer subir a temperatura, ou não. As estrelas de massa elevada transformam-se em estrelas de neutrões ou em buracos negros. Nas estrelas que têm relativamente pouca massa, como o nosso sol, as camadas exteriores não têm o peso suficiente para continuarem a comprimir o núcleo. Em vez disso, o calor intenso gerado pelo hélio que arde no núcleo ejeta para o espaço interestelar as camadas exteriores da estrela. Os farrapos de hidrogénio resultantes são chamados «nebulosas planetárias», das quais foram encontradas cerca de um milhar na nossa galáxia. A expressão «nebulosa planetária» é uma

designação incorreta dada pelos primeiros astrónomos que observaram as nebulosas e as julgaram semelhantes a planetas. Estas nebulosas nada têm a ver com os planetas; são o remanescente do invólucro exterior das estrelas moribundas.

Os farrapos de gás acabam por dispersar-se no espaço interestelar e tudo o que fica da estrela é o seu núcleo nu, com dimensões aproximadamente iguais às da Terra — uma estrela anã branca. Através dos evos, as anãs brancas perdem energia, mudando, sucessivamente, da cor branca para a amarela, desta para a castanha e, finalmente, para o preto. A transição de uma anã branca para uma anã preta demora tanto tempo que não deve existir ainda nenhuma anã preta na nossa galáxia. E quanto às anãs brancas? Existem ou não?

A história das anãs brancas começa em 1844 no observatório de Königsberg (na altura pertencente à Prússia, hoje parte da União Soviética), quando Friedrich W. Bessel viu oscilar a imagem de Sírio, a estrela mais brilhante dos nossos céus. Que poderia causar a oscilação de uma estrela? Bessel concluiu que Sírio estava acompanhada por outra estrela, escura e de grande massa, que, à medida que orbitava, atraía Sírio, produzindo um movimento oscilante na respetiva posição no céu. Bessel não viu essa estrela escura, mas o seu palpite provou ser correto quando dezanove anos mais tarde Alvan Clark, construtor de telescópios americano, localizou a ténue companheira de Sírio ao testar uma nova lente de 45 cm. Alvan Clark, membro de uma distinta família americana de construtores de telescópios, construiu mais tarde as grandes lentes

de refração para os primeiros telescópios dos observatórios de Lick e de Yerkes.

No entanto, havia qualquer coisa de estranho na companheira de Sírio. Em 1910 Henry Norris Russell, o codescobridor do diagrama de Hertzsprung-Russell, notou que esta estrela não se enquadrava na sequência principal, e essa exceção causou-lhe grande preocupação. Talvez não existisse a correlação que tinha descoberto entre o brilho da superfície e a densidade das estrelas. Pediu então ao astrónomo Edward Pickering para lhe obter o espectro da companheira de Sírio. Russell registou o seguinte:

> Caracteristicamente, Pickering enviou uma nota ao gabinete do observatório e em breve obteve a resposta [...] de que o espectro desta estrela era A. Eu sabia o suficiente, mesmo nesses dias «pré-históricos», para concluir rapidamente que existia uma grande incoerência entre aquilo a que nessa altura poderíamos chamar valores «possíveis» do brilho da superfície e da densidade. Devo ter mostrado que não estava somente perplexo, mas também desanimado com esta exceção àquilo que parecia ser uma norma segura sobre as características estelares, mas Pickering sorriu-me e disse: «São precisamente estas exceções que conduzem ao avanço do nosso conhecimento.»

Nos sete anos seguintes foram descobertas mais duas estrelas que, tal como esta, constituíam exceções à regra.

De um modo geral, as estrelas de brilho ténue (esta era verdadeiramente de brilho fraco, com apenas $1/400$ da luminosidade do Sol) deveriam ter cor vermelha, enquanto a companheira de

Sírio era, em vez disso, de um branco incandescente. A única explicação para o brilho ténue era a sua extrema pequenez. Mas, sendo tão pequena, não deveria ter massa suficiente para influenciar o movimento de uma estrela pesada como a Sírio. Havia uma saída para este enigma, que consistia em admitir que a companheira de Sírio era, na verdade, muito pequena, mas feita de matéria 3000 vezes mais densa do que a matéria das estrelas vulgares. Todavia, esta solução para o enigma soava a disparate. Com efeito, nas primeiras décadas deste século não se tinha conhecimento da existência de nenhuma espécie de matéria tão densa. Refletindo sobre esta estranha mensagem da ténue companheira de Sírio, o astrónomo britânico Sir Arthur Eddington disse em 1927:

> Que resposta se pode dar a esta mensagem? A resposta que a maior parte de nós deu em 1914 foi: cale-se, não diga disparates.

A mensagem era realmente um disparate, se interpretada em termos da física newtoniana. A solução do enigma da companheira de Sírio teve de aguardar até ao aparecimento da teoria quântica dos átomos, em 1927, e até ao trabalho de um indiano de 19 anos, Subrahmanyan Chandrasekhar, realizado em 1930. Partindo dos trabalhos anteriores de Ralph H. Fowler, na Inglaterra, que mostrou que uma estrela entra em colapso quando acaba o combustível nuclear, Chandrasekhar viu aquilo a que conduzia o colapso: uma nova forma de matéria superdensa, tão densa que 1 cm^3 pesaria 600 kg. Como podemos imaginar semelhante matéria?

Fowler tinha feito uso do «princípio de exclusão», descoberto em 1925 pelo físico quântico Wolfgang Pauli, segundo o qual os eletrões (pequenas partículas carregadas eletricamente enxameando em torno do núcleo atómico) não podem colocar-se uns em cima dos outros — excluem-se uns aos outros e, se tentarmos empurrar dois eletrões para o mesmo estado, repelir-se-ão mutuamente. Esta força de repulsão não é devida ao facto de as cargas elétricas dos eletrões se repelirem por serem do mesmo sinal; é uma espécie inteiramente nova de força de repulsão, muito mais intensa do que a força elétrica. Chamada «interação de troca», só pode ser compreendida com base na teoria quântica e não tem analogia na física clássica. É a sua existência a nível atómico que evita que as nuvens de eletrões que rodeiam os átomos entrem em colapso.

Se imaginarmos um gás de eletrões sujeito a pressão, a interação repulsiva de troca entre os eletrões estabelece uma «pressão de Fermi» resistente para contrariar esta compressão. Contudo, é necessário pressionar o gás fortemente antes que se possa sentir esta pressão de Fermi resistente, a qual só entra em jogo quando os eletrões são empurrados para tão perto uns dos outros que as suas ondas associadas começam a sobrepor-se. Tais condições existem no interior das estrelas. Aquilo que Chandrasekhar compreendeu foi que a teoria da relatividade restrita implicava que a pressão eletrónica de Fermi, nascida do estranho mundo da teoria quântica, resistiria ao colapso gravitacional e estabilizaria a estrela, desde que a respetiva massa total não fosse muito grande.

Chandrasekhar calculou que seria este o caso das estrelas com massa inferior a 1,4 vezes a massa do Sol — uma massa crítica conhecida por «limite de Chandrasekhar». Em algumas destas estrelas a densidade da matéria para a qual se estabelece o equilíbrio entre a gravidade e a pressão de Fermi é de 600 kg/cm^3, valor que explicava exatamente o comportamento da companheira de Sírio. Esta estrela, uma anã branca, foi outrora uma estrela normal, mas a certa altura o combustível hidrogénio do núcleo esgotou-se, tendo a estrela estabilizado de novo graças à pressão de Fermi. Os astrónomos detetaram até hoje mais de 300 anãs brancas.

Algumas anãs brancas, como a companheira de Sírio, são componentes de um sistema estelar binário, cujo segundo membro é uma estrela normal. A anã pode orbitar muito perto da estrela normal e absorver gás dela. Este gás, composto, na sua maior parte, por hidrogénio, cai na anã, começa a acumular-se e, após um intervalo de tempo suficiente, atinge uma quantidade crítica. Então, ao fundir-se em hélio, o hidrogénio explode todo de uma vez na superfície da anã, como se de milhares de bombas de hidrogénio se tratasse. Foram observadas centenas destas explosões de «novas» — confirmação adicional das insólitas propriedades das anãs brancas.

2. As Estrelas de Neutrões

> Num dia chi-chhou do quinto mês do primeiro ano do reinado de Chi-Ho [4 de Julho de 1054], uma estrela visitante apareceu a sudeste de Thien-K'uan, medindo alguns centímetros. Passado mais de um ano, desapareceu.
>
> *TOKTAGU, Crónicas da Dinastia Sung*

Destino mais espetacular aguarda as estrelas com massa superior ao limite de Chandrasekhar. Depois de o hélio começar a arder no núcleo, as camadas exteriores têm massa suficiente para continuar a exercer pressão sobre o núcleo, de maneira que este continua a comprimir-se e, portanto, a aquecer. A temperatura sobe tanto que se iniciam novos processos de fornalha nuclear. O núcleo de carbono arde violentamente, cozinhando com rapidez até elementos mais pesados. O interior da estrela velha rapidamente se assemelha ao de uma cebola, apressando várias camadas distintas. Na parte de fora do núcleo encontram-se os elementos mais leves — hidrogénio e hélio; nas camadas intermédias, o carbono e o hélio, e, à medida que se penetra mais fundo, as camadas vão contendo elementos cada vez mais pesados — magnésio, silício, enxofre, e assim sucessivamente, até ao ferro, o elemento mais pesado que as estrelas produzem por processos de fornalha nuclear.

O núcleo da estrela é composto principalmente por ferro, que não é o elemento mais pesado existente, mas tem a propriedade especial de não «arder» na fornalha nuclear. O ferro é a cinza final da fornalha nuclear — não existe processo de extrair energia dos núcleos de ferro através da sua fusão. Não é claro o que acontece em seguida. Contudo, segundo parece, logo que tenham sido sintetizadas no núcleo quantidades suficientemente grandes de ferro, a fornalha nuclear pára, a pressão que evita o colapso gravitacional da estrela desaparece abruptamente e esta sofre um colapso catastrófico. A imensa massa, impedida de entrar em colapso em direção ao centro da estrela durante milhares de milhões de anos, sofre agora esse colapso no espaço de segundos.

Essa energia — equivalente à produção total de energia da estrela durante toda a vida anterior, de milhares de milhões de anos — é libertada em poucos segundos e a explosão é tão brilhante como mil milhões de sóis. Se uma das estrelas nossas vizinhas sofresse uma tal explosão de supernova (nenhuma está a isso fadada, exceto talvez Sírio), então apareceria no céu um segundo sol, tão brilhante como o nosso, que nos assaria vivos.

Que acontece durante a explosão de supernova? Uma vez que ninguém viu uma supernova na nossa galáxia desde 1604, quando explodiu a estrela de Kepler, na constelação da Serpente, não tivemos oportunidade de observar nenhuma de perto com os instrumentos modernos[1], razão por que os astrónomos estabeleceram uma «vigilância supernova» (*supernova watch*), de maneira que, se tal explosão ocorrer na nossa galáxia (como tudo leva a crer), muitos observatórios apontarão imediatamente para lá os seus instrumentos. Foram observadas cerca de quatrocentas supernovas em galáxias distantes (algumas tão brilhantes que ofuscaram toda a galáxia durante várias semanas). E, na ausência de observações pormenorizadas de uma supernova na nossa galáxia, os astrofísicos desenvolveram complexos programas de computador que reproduzem cada microssegundo do colapso e explosão subsequente. Seria interessante comparar tais modelos de computador com observações de uma supernova próxima logo que uma seja vista.

Quando uma estrela de grande massa sofre um colapso, criam-se condições extremas. A temperatura e a pressão tornam-se elevadíssimas, tão grandes que depois do colapso até elementos

mais pesados do que o ferro são produzidos por transmutações nucleares no seio da matéria que explodiu. Todos os elementos metálicos a que damos tanto valor — níquel, prata, ouro e urânio — foram criados nessas explosões de supernovas e lançados no espaço, tendo alguns deles vindo a fazer parte de novas estrelas. Os astrofísicos calculam as abundâncias relativas de cerca de 90 elementos que são produzidos nas estrelas e nas explosões de supernovas, as quais estão em notável acordo com as observadas na Natureza, fornecendo, assim, alguma confirmação de que os modelos estão razoavelmente corretos. Um elemento pesado, em particular o tecnécio, que é radioativo, tem um período de semidesintegração de 200.000 anos, ou seja, suficientemente curto para que já tenha decaído todo há muito na Terra. Mas as riscas espectrais do tecnécio podem ainda ser observadas nas estrelas gigantes vermelhas — prova direta de que as estrelas criam novos elementos.

Os modelos pormenorizados do que acontece depois do colapso diferem entre si. Num desses modelos, idealizado por Hans Bethe, da Universidade de Cornell, o material da parte exterior da estrela ressalta do núcleo em colapso para o espaço. Sterling Colgate, do Laboratório Nacional de Los Alamos, desenvolveu um modelo diferente, no qual uma erupção de neutrinos energéticos — partículas subatómicas criadas pelas reações nucleares no núcleo em colapso — varre literalmente as camadas exteriores da estrela para longe com um vento de neutrinos oriundo do núcleo. Ambos os mecanismos — ressalto e vento de neutrinos — expelem o invólucro exterior.

No entanto, todos os modelos preveem que o núcleo, que é o remanescente da supernova, se transforme num novo estado da matéria — uma estrela de neutrões. Objetos deste tipo foram postulados teoricamente já em 1933 pelos astrofísicos Fritz Zwicky e Walter Baade e, independentemente, por Lev Landau, físico soviético. Estes cientistas pretendiam ir além do trabalho de Chandrasekhar. Mas o que é uma estrela de neutrões?

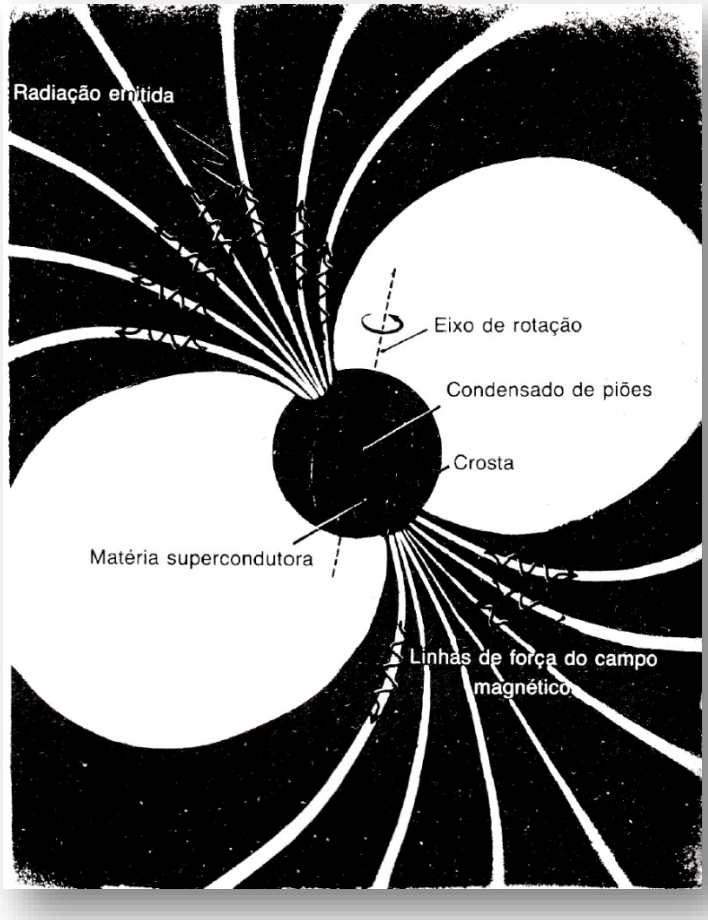
Numa anã branca é a pressão de Fermi dos eletrões que resiste à pressão da gravidade. Contudo, se a gravidade for suficientemente intensa — como é o caso nas estrelas que explodem em supernovas —, os eletrões são efetivamente comprimidos contra os protões (partícula encontrada nos núcleos atómicos), que transformam em neutrões (outra componente dos núcleos). Os neutrões, tal como os eletrões, obedecem também ao princípio de exclusão de Pauli — não se podem colocar dois neutrões no mesmo estado um em cima do outro. É a pressão neutrónica de Fermi resultante que resiste à força da gravidade e estabiliza a estrela de neutrões.

As estrelas de neutrões não são objetos vulgares e as suas propriedades sobressaltam a imaginação. O termo «estrela» é um pouco enganador, uma vez que estes objetos não são verdadeiras estrelas. 1 cm^3 da matéria nuclear de uma estrela de neutrões pesa 600 milhões de toneladas. São esferas com cerca de 12 km de diâmetro — as dimensões de uma cidade. Mas ninguém irá visitar uma estrela de neutrões nem passear à volta dela; são lugares a evitar.

Usando as leis físicas das partículas elementares, uma equipa de físicos teóricos dirigida por Malvin Ruderman, da Universidade de Colúmbia, construiu modelos de estrelas de neutrões que sugerem que existe uma espécie de crosta ferrosa na superfície da estrela, extremamente lisa e com a espessura aproximada de 1 m. Nesta crosta é possível que existam minúsculas «montanhas» com a altura de apenas 0,0001 cm. No entanto, para «trepar» a uma dessas montanhas seria necessária a energia consumida numa grande cidade durante um ano inteiro, devido à intensíssima força de gravidade. A maior parte do interior de uma estrela de neutrões, por baixo da crosta, é composta por núcleos e outras partículas subnucleares comprimidas à densidade nuclear, formando um «cristal» sólido de matéria nuclear. Os físicos, porém, suspeitam de que, alguns quilómetros abaixo da superfície, a matéria adquire ainda outra propriedade — torna-se supercondutora, isto é, conduz a eletricidade sem qualquer resistência. Correntes elétricas intensíssimas podem, assim, fluir sem perdas no interior das estrelas de neutrões e produzir os correspondentes campos magnéticos intensíssimos, que desempenham um papel importante na produção dos sinais emitidos pelas estrelas de neutrões.

Alguns físicos especulam que o centro de uma estrela de neutrões consiste num condensado de piões — um novo estado da matéria. Os piões, partículas subnucleares observadas em laboratórios dotados de aceleradores, podem ser imaginados como a cola que mantém coeso o núcleo atómico. Sob condições extremas, como as que reinam no núcleo de uma estrela de neutrões, os piões condensam-se para formar uma espécie de gás capaz de suportar um peso enorme. As condições extremas

reinantes no interior de uma estrela de neutrões conduzem os físicos até às fronteiras dos seus conhecimentos de física subnuclear. Alguns suspeitam de que o núcleo é formado pelos quarks constituintes das partículas nucleares. Mesmo sendo as características ainda objeto de debate, a maioria dos físicos estão entusiasmados com a ideia de as estrelas de neutrões poderem ser uma espécie de «laboratório natural» para testar as novas ideias relativas ao mundo subnuclear.



Um pulsar, ou estrela de neutrões, em rotação, com cerca de 10 km de diâmetro, é acompanhado por um forte campo magnético que drapeja em redor. As partículas carregadas eletricamente que seguem as linhas de força do campo magnético produzem um feixe de radiação que varre o espaço «à maneira da luz de um farol» e que pode ser detetado na Terra. A superfície da estrela é, provavelmente, constituída por uma crosta de núcleos de ferro. Por baixo encontra-se uma espécie de «cristal» de outros núcleos atómicos e partículas subnucleares. À medida que se caminha para o interior, a matéria torna-se supercondutora — conduz correntes

elétricas sem resistência. O centro da estrela de neutrões consiste, possivelmente, num «condensado de piões», composto por partículas subnucleares.

Tudo isto é teoria. Existem realmente estrelas de neutrões? Certamente que sim; foram descobertas por acaso em 1967. Eis parte da história.

Em Cambridge, na Inglaterra, Antony Hewish dirigia uma equipa que projetava um radiotelescópio — um campo de 1,80 ha coberto com 2048 antenas alongadas, destinado a ser usado na identificação de quasares distantes através das cintilações. Hewish teve a sorte de ter na equipa Jocelyn Bell-Burnell, uma estudante graduada de 24 anos, a qual, ao examinar a saída de dados da antena que varria o céu à medida que a Terra rodava, observou uma «pequena arranhadela», um sinal de rádio distinto vindo de um determinado ponto do céu. Teria sido bastante fácil desprezar tal sinal como sendo apenas ruído de fundo. A saída de dados da antena estava a ser registada num rolo de papel e a «pequena arranhadela» consistia apenas em curtos saltos no traço que cobria centenas de metros de papel, examinados centímetro a centímetro por Bell.

Um mês mais tarde viu de novo o sinal e tratou de analisar a «arranhadela» em pormenor. Viu que consistia em sinais periódicos, com cerca de um segundo de duração. Quando comunicou este resultado a Hewish, seu orientador de tese, este respondeu: «Oh, isso explica tudo: devem ser feitos pelo homem.» Hewish pensou a princípio que o sinal estava a ser captado de uma fonte local, como, por exemplo, a ignição defeituosa de um automóvel. Contudo, depois de afastar todas essas possibilidades,

verificou o período do sinal de Bell e descobriu, para seu espanto, que este marcava o tempo certo com precisão de uma parte em 10 milhões. Nenhum sinal de nenhuma fonte terrestre poderia ter esta precisão. O entusiasmo começou.

A equipa de Hewish considerou então a louca possibilidade de o sinal estar a ser enviado por uma civilização extraterrestre, tentando talvez comunicar com outras sociedades. Todavia, Bell cedo descobriu outra fonte pulsante. Pouco depois foram detetadas mais quatro em diferentes partes do céu, o que tornava a ideia de uma civilização extraterrestre pouco provável. Tornava-se claro que uma nova espécie de objeto astronómico tinha sido descoberta. Ao anunciar esta descoberta em 1967, o grupo de Cambridge sugeriu que os pulsares, como então foram chamados, podiam ser estrelas de neutrões — os remanescentes do colapso das estrelas que explodiam, previstos teoricamente em 1933 por Fritz Zwicky e Walter Baade, suposição que veio a revelar-se correta. Conhecemos, atualmente, mais de 300 destas estrelas de neutrões, cada uma delas constituindo o remanescente pulsante de uma explosão de supernova.

Que faz com que as estrelas de neutrões pulsem tão rapidamente? F. Pocini, astrofísico italiano, e Thomas Gold, astrofísico da Universidade de Cornell, propuseram em 1968 uma explicação elementar. As estrelas grandes rodam e têm campos magnéticos de forma semelhante à do campo magnético da Terra, que é conhecido como um campo de «dípolo». Se uma estrela grande entra em colapso até se tornar uma minúscula estrela de neutrões, como acontece numa supernova, então a velocidade de

rotação da estrela aumenta enormemente, da mesma maneira que a velocidade de rotação de uma patinadora do gelo aumenta quando ela recolhe os braços e as pernas. O campo magnético associado à estrela entra em colapso juntamente com ela e aumenta enormemente. Mas o eixo de rotação e o eixo do campo magnético (determinado pelos polos norte e sul da estrela de neutrões) não coincidem necessariamente. Assim, à medida que a estrela de neutrões roda rapidamente, o campo magnético, cujo eixo é diferente, drapeja excentricamente à volta dela. As partículas eletricamente carregadas que se encontram na vizinhança da estrela de neutrões precipitam-se sobre ela, produzindo um feixe de radiação que roda com ela, como o feixe de luz de um farol. Este «efeito de farol» é o responsável pelos sinais de rádio vistos pela primeira vez por Bell. A frequência da pulsação do sinal de rádio corresponde exatamente à rotação extremamente rápida da estrela de neutrões.

A maior parte destas estrelas rodam várias vezes num segundo. No entanto, recentemente, detetou-se uma que roda com a incrível frequência de 640 vezes por segundo! É difícil imaginar uma esfera com as dimensões de uma cidade a rodar tão rapidamente. Este pulsar com o período de milissegundos era, provavelmente, membro de um sistema de estrelas binário que adquiriu esta rapidíssima rotação ao consumir completamente a estrela companheira.

Enquanto os físicos teóricos elaboravam modelos destes objetos bizarros, os astrónomos observacionais não estiveram parados. Talvez a mais dramática confirmação de que os pulsares são

remanescentes das supernovas tenha vindo das observações ópticas de Dan Taylor, John Cocke e Michael Disney. Estes cientistas estudaram a estrela de Baade, no centro da nebulosa do Caranguejo, constituída pelos restos difusos de uma supernova observada e registada pelo historiador chinês Toktagu em 1054 e posteriormente identificada como um pulsar com uma frequência de 30 vezes por segundo. Taylor, Cocke e Disney decidiram examinar mais de perto a luz emitida pelo pulsar do Caranguejo (em vez do sinal de rádio), usando um telescópio convencional, no observatório Steward, na Universidade do Arizona. Acoplaram-lhe um sistema eletrónico de sincronização que efetivamente fazia «piscar» o aparelho detetor do telescópio com a frequência do pulsar. Se o «piscar» ocorresse sincronizadamente com a fase de emissão do pulsar, então a luz seria detetada; se ocorresse com a fase de não emissão, nenhuma luz seria detetada. Desta maneira, esperavam verificar experimentalmente a ideia de que o pulsar do Caranguejo pulsava em luz visível com a mesma frequência dos sinais de rádio.

Se, na realidade, a estrela piscasse, apareceria um sinal que percorreria uma sequência de pequenos pontos verdes no mostrador para o qual os astrónomos olhavam.

Na noite de 15 de Janeiro de 1969 decidiram fazer nova tentativa (depois de algumas dificuldades preliminares). Por acaso o assistente tinha deixado ligado um gravador, pelo que o momento da descoberta ficou registado. A voz de Disney (inglês, pronúncia inconfundível) fez-se ouvir: «Temos aqui um raio de um impulso.»

Cocke respondeu: «Hei!», e alguns momentos depois, incapaz de se conter, exclamou: «Oh! Não me diga que está a pensar que é isso! Não pode ser!»

Efetuadas mais experiências e verificações para terem a certeza de que o efeito que estavam a ver era real, e não um sinal qualquer gerado pela própria aparelhagem, Disney disse: «Meu Deus, venha cá vê-lo!», e ambos riram. «Este é um momento histórico!»

«Humm», advertiu Cocke, «espero que seja um momento histórico.» Mas, passados alguns dias, a descoberta de que o pulsar do Caranguejo «piscava» na região do visível foi confirmada por outros observatórios. A ideia bizarra de que uma estrela de neutrões é o remanescente de uma supernova obteve, assim, a mais impressionante confirmação.

Os astrónomos descobriram desde então centenas de pulsares na nossa galáxia, e é possível que existam milhões deles — os remanescentes de estrelas outrora ardentes. Uma ideia oriunda da periferia da física teórica foi arrastada para o centro das observações astronómicas.

A partir de 1969 os astrónomos notaram a existência de «avarias», ou seja, acelerações rápidas, na rotação de algumas estrelas de neutrões. Para espanto de alguns, essas avarias repetiam-se. Qual seria a causa? Para as explicar os astrofísicos elaboraram modelos matemáticos pormenorizados. Segundo algumas teorias, as avarias são devidas à abertura de fendas na

crosta sólida — tremores fazem com que a estrela de neutrões encolha, ainda que ligeiramente, e, como sucede com a patinadora do gelo ao recolher os braços, rode mais rapidamente. Originados por instabilidades complexas no estranho interior da estrela, dão aos cientistas uma oportunidade de testarem os modelos teóricos da estrutura destes corpos celestes.

Uma das mais belas confirmações das propriedades das estrelas de neutrões veio da observação dos «emissores intensos de raios X» (*X-ray bursters*). Nos últimos anos da década de 70 os cientistas projetaram vários satélites com capacidade para detetar raios X, encontrando, assim, muitas fontes de raios X que apresentavam misteriosas variações periódicas de intensidade. As flutuações na intensidade de alguns destes emissores tinham uma frequência medida em segundos; outros aumentavam-na enormemente a intervalos regulares de algumas horas. Que poderia causar este comportamento tão preciso e periódico?

Hoje em dia o consenso é que as erupções de raios X, cuja energia iguala em alguns segundos à emitida pelo Sol em duas semanas, são devidas a estrelas de neutrões que orbitam em torno de estrelas vulgares — um sistema estelar binário pouco quente.

Expliquemos como ocorrem as erupções periódicas, breves e longas.

Se a estrela de neutrões é jovem, possui um campo magnético característico muito intenso junto aos seus polos norte e sul. Ao ser

atraído gravitacionalmente da companheira próxima da estrela de neutrões, o hidrogénio gasoso ionizado é arrastado na direção dos polos pelo forte campo magnético — uma corrente constante de gás quente que emite raios X e pulsa com a frequência rotacional da estrela de neutrões, ou seja, em alguns segundos.

Por outro lado, se a estrela de neutrões é velha, o campo magnético é muito mais fraco. Então acontece algo de muito notável. O hidrogénio gasoso arrancado à estrela companheira distribui-se por toda a superfície da estrela, em vez de se concentrar somente nos polos. Este material deposita-se na superfície e não emite raios X até se ter acumulado, no espaço de algumas horas, uma quantidade crítica. Explode então de uma só vez num «relâmpago termonuclear» espetacular, produzindo uma erupção de raios X que pode ser vista na Terra. O processo continua a repetir-se periodicamente — uma confirmação impressionante da interligação entre o trabalho teórico sobre as estrelas de neutrões e as observações baseadas na astronomia de raios X.

Face aos resultados sensacionais obtidos com os raios X, os astrónomos começaram a estudar a luz a energias ainda mais altas do que as dos raios X, construindo detetores espectrais de raios gama e enviando-os para os céus em balões e satélites. Os raios gama, uma forma de luz de muito alta energia, fornecem novas e importantes informações acerca das estrelas de neutrões ou dos buracos negros onde se desenvolvam processos energéticos que permitam emiti-los.

Os raios gama provêm dos núcleos atómicos ou da aniquilação eletrão-positrão, sendo, assim, independentes do estado químico da matéria. Facultam mais um conjunto de «impressões digitais» pormenorizadas que ajudam a identificar os processos físicos complexos que rodeiam os objetos compactos exóticos. Embora ainda na sua infância, a astronomia dos raios gama progride rapidamente.

A ideia das estrelas de neutrões nasceu na imaginação dos teóricos, que persistiam em perguntar: «Qual é o estado final da matéria? Que acontece quando as estrelas de grande massa não puderem suportar por mais tempo o respetivo peso?» Para responderem a tais interrogações recorreram ao saber crescente da física atómica e nuclear — saber este confirmado em experiências laboratoriais pormenorizadas —, e, no caso das anãs brancas e das estrelas de neutrões, as respostas revelaram-se corretas. Contudo, os físicos teóricos sabiam que para estrelas de massa ainda maior, até a formação de estrelas de neutrões ou de anãs brancas é insuficiente para suster o colapso. Tanto quanto era do seu conhecimento, não havia nada que pudesse deter a contração gravitacional. No entanto, alguma coisa *tinha* de acontecer. Segundo os teóricos, o que acontece é que se forma um buraco negro — num certo sentido, espaço e tempo entram em colapso para criar um objeto do qual nem a própria luz pode escapar.

3. Buracos Negros

A própria força da gravidade de um tal corpo faria com que toda a luz por ele emitida a ele regressasse.

JOHN MITCHELL, reitor de Thornhill, no Yorkshire, 1784

[...] que a força de atração de um corpo celeste poderia ser tão grande que a luz não pudesse fluir dele.

MARQUÊS DE LAPLACE, 1798

Ao contrário do que se passa com as anãs brancas e as estrelas de neutrões, não existem provas universalmente aceites sobre a existência dos buracos negros, embora muitos astrónomos estejam convencidos de que devem existir e alguns acreditem até que já foram encontrados. Os buracos negros são uma previsão da teoria da relatividade geral de Einstein — a moderna teoria da gravidade. A teoria da relatividade geral foi testada experimentalmente com grande êxito. No entanto, alguns críticos dirão que as verificações experimentais, apesar do sucesso, foram todas realizadas para campos gravitacionais fracos e que a teoria nunca foi verificada para campos gravitacionais muito intensos — do tipo que se poderá encontrar nos buracos negros. Todavia, se a teoria se aplicar realmente a esses campos fortes, então teremos de concluir que os buracos negros são uma consequência inelutável do colapso de estrelas de grande massa.

A descrição adequada das propriedades dos buracos negros é dada pela teoria da relatividade geral de Einstein, que especifica a curvatura do espaço associada aos campos gravitacionais. A teoria da relatividade geral, altamente matemática, não é facilmente apreendida pelo leigo. No entanto, a imagem física do mundo que descreve — o mundo do espaço geométrico curvo — pode ser compreendida sem dificuldade exagerada. Como é que os cientistas visualizam o espaço curvo?

Imagine-se o disparo de raios laser através do espaço vazio com o fim de determinar a geometria do espaço. No espaço plano vulgar propagar-se-iam segundo linhas retas, de modo que raios paralelos jamais se encontrariam.

Suponhamos que medíamos os raios com grande precisão e descobríamos que não viajam em linhas retas, mas segundo curvas suaves. Concluiríamos, a seguir, que este encurvar dos raios é devido à curvatura intrínseca do espaço, da mesma maneira que um avião a jato, viajando entre duas cidades do Globo muito distantes, segue uma trajetória curva devido à curvatura da superfície da Terra. Disparando raios laser em todas as direções, podemos determinar a curvatura deste espaço tridimensional. Isto é análogo a fazer rolar uma pequena esfera ao longo de uma superfície curva. Examinando as trajetórias descritas pela esfera, podemos achar a superfície curva. Usando a luz das estrelas distantes ou o radar a partir da Terra (que viaja da mesma maneira que os raios de luz), os cientistas mostraram que o espaço se afasta da forma plana junto dos corpos de massa elevada, como o Sol. A ideia central da relatividade geral é que a curvatura do espaço e a sua influência sobre o movimento das partículas ou dos raios de luz são completamente equivalentes à gravidade. A intensa gravidade do Sol produz um desvio pequeno, mas mensurável, no trajeto de um raio de luz.

Imagine-se agora uma situação extrema na qual toda a massa do Sol fosse comprimida até ficar com o raio de alguns quilómetros. A gravidade e a curvatura do espaço junto deste sol compacto seriam enormes. Se enviássemos um raio de luz na direção deste

objeto, nunca regressaria. A curvatura do espaço torna-se tão grande que até a luz é apanhada pela gravidade. Um raio de luz rasante a um objeto deste tipo descreve, na realidade, uma órbita em espiral convergindo para ele. Uma vez que a luz não pode deixar tal objeto, ele «aparece-nos» como um buraco negro no espaço.

A fronteira de um buraco negro, que não é uma superfície material verdadeira, mas simplesmente uma fronteira matemática que limita o espaço do qual nenhuma luz escapa, chama-se «horizonte de acontecimentos»; qualquer acontecimento que ocorra dentro desta fronteira nunca pode ser observado fora dela. O horizonte de acontecimentos é uma porta com um só sentido — pode entrar-se, mas jamais se pode sair.

Como será cair num buraco negro? Se o buraco negro for pequeno, como o que se forma a partir de uma estrela que entrou em colapso, então estaremos em grandes dificuldades. As forças de maré junto a um buraco negro são enormes e o gradiente gravitacional numa distância de apenas alguns centímetros é suficiente para destruir tudo.

O tempo sofre também estranhas distorções. Um observador que não tivesse sido destruído e caísse no centro de um buraco negro poderia ver o tempo a desacelerar. Mas o observador em queda nunca poderia comunicar a estranha experiência a um amigo que estivesse fora. O observador exterior, ao ver o amigo cair no buraco negro sem o poder socorrer, constataria a sua longa demora para cruzar o horizonte de acontecimentos.

À medida que a massa dos buracos negros aumenta, tornam-se também maiores e menos densos. Se existirem buracos negros com a massa de biliões de sóis, poder-se-á atravessar o horizonte de acontecimentos sem sofrer qualquer efeito nocivo. Porém, mesmo assim, continuará a não haver saída e decorrerão apenas alguns minutos até acontecer um desastre. Encontrar-se-á uma singularidade espaço-tempo — um ponto no qual a densidade de energia se torna infinita — que se crê existir exatamente no centro do buraco. Podem até imaginar-se buracos negros maiores. É concebível que o universo inteiro esteja em vias de se transformar num gigantesco buraco negro, e nós vivemos dentro dele — um universo que um dia parará de se expandir e entrará em colapso sobre si próprio.

Foi só nos anos 60 e 70 que os físicos teóricos e os astrofísicos se aperceberam das insólitas propriedades dos buracos negros, apesar de essas propriedades poderem ter sido deduzidas diretamente da teoria da gravitação de Einstein de 1915. Esta demora de meio século deveu-se, em parte, ao facto de não haver provas observacionais isentas de ambiguidade da existência de buracos negros que forçassem os cientistas a considerar esta possibilidade. Por vezes, os físicos necessitam de que lhes abram os olhos com uma nova descoberta para começarem a levar a sério as próprias teorias. Outra causa para a demora foi o facto de a conclusão de que os buracos negros existiam exigir que a teoria da relatividade fosse explorada até limites extremos e haver muitas dúvidas de que fosse aplicável em tais circunstâncias extremas. Curiosamente, a possibilidade de existência de buracos negros foi reforçada pela impressionante descoberta das estrelas de neutrões.

Se pode existir algo tão invulgar como uma estrela de neutrões, por que razão não pode existir também um buraco negro? Hoje em dia os buracos negros transitaram da orla especulativa da teoria física para o centro dessa teoria. Não passa um ano sem que os teóricos evoquem os buracos negros para explicarem mais uma observação astronómica peculiar.

Um dos primeiros pioneiros do desenvolvimento da teoria dos buracos negros foi o astrónomo alemão Karl Schwarzschild, que no Inverno de 1915 descobriu uma solução simples, mas exata, das equações de Einstein da relatividade geral, a qual descrevia a curvatura do espaço à volta de uma massa esfericamente simétrica como o Sol — uma realização que impressionou Einstein. Schwarzschild descobriu que para uma massa suficientemente compacta existia um raio finito (hoje chamado «raio de Schwarzschild» — o raio do horizonte de acontecimentos) no qual as ondas de luz emitidas teriam um comprimento de onda infinitamente grande, o que equivalia a dizer que a luz não podia escapar.

O passo importante seguinte foi dado em 1939 por J. Robert Oppenheimer em dois artigos, o primeiro escrito em colaboração com George M. Volkoff e o segundo em colaboração com Hartland Snyder, ambos estudantes graduados, motivados, não tanto pelo trabalho de Schwarzschild, como pelo desejo de compreenderem o que acontece quando uma estrela entra em colapso. O primeiro artigo examinava as consequências do colapso gravitacional de uma estrela numa supernova. Mas Oppenheimer e Volkoff reconheceram que, sob certas condições particulares, a pressão da gravidade

impede até a formação de uma estrela de neutrões, o que susteria o colapso, e «a estrela continuaria a contrair-se indefinidamente, nunca alcançando o equilíbrio».

No segundo artigo Oppenheimer e Snyder foram mais longe, analisando os pormenores de um colapso interminável. Aperceberam-se de que um observador no interior de uma estrela em colapso nunca poderia comunicar a um observador exterior qual seria o seu destino. Nenhuma radiação escapa de tal objeto. «A estrela», comentaram, «tende, assim, a isolar-se totalmente de um observador distante; só o seu campo gravitacional persiste.» Embora se tenha demonstrado de forma convincente ser a existência de objetos tão estranhos uma consequência lógica da teoria da relatividade geral, ninguém, incluindo Oppenheimer e os seus colaboradores, foi ao ponto de sugerir que existissem realmente no universo, tendo sido relegados para o reino das curiosidades matemáticas, onde permaneceram por várias décadas.

Em 1963 Roy Kerr, neo-zelandês, descobriu uma solução matemática exata para a equação de Einstein que descreve um buraco negro em rotação. Este resultado notável ia mais longe do que a primitiva solução de Schwarzschild, a qual descreve somente massas que não estão em rotação. O trabalho matemático de Kerr implicava que era possível extrair energia de um buraco negro em rotação. Ao extrair-se energia, a velocidade de rotação abrandaria. A solução de Kerr estimulou muito o interesse por estas «curiosidades matemáticas». S. Chandrasekhar, que tanto contribuiu para a astronomia moderna, fez notar o seguinte:

> Em toda a minha vida científica, que se estende por mais de 45 anos, a experiência que mais me abalou foi compreender que uma solução exata das equações da relatividade geral de Einstein, descoberta pelo matemático neo-zelandês Roy Kerr, constitui a *representação absolutamente exata* de incontáveis buracos negros de grande massa que povoam o universo. Este «estremecimento perante o maravilhoso», este facto incrível de uma descoberta motivada pela busca da beleza em matemática encontrar réplica exata na Natureza, leva-me a dizer que a beleza é aquilo a que o espírito humano reage no mais profundo de si mesmo.

John A. Wheeler, o visionário da física teórica americana, fomentou ainda mais o interesse por estes objetos bizarros e teve a coragem intelectual de instar os cientistas a levá-los a sério. Dirigindo-se à Academia Nacional das Ciências em 1973, por ocasião do 500º aniversário do nascimento de Copérnico, Wheeler usou o termo «buraco negro», que tinha inventado num artigo de 1968 e que perdurou desde então. Aos cientistas ali reunidos disse que a ideia do colapso gravitacional total se apresentava aos físicos como um grande desafio.

Para o enfrentarem os teóricos aprofundaram desde então o nosso conhecimento dos buracos negros. Muito desse trabalho foi inspirado por Stephen Hawking, um brilhante físico inglês da Universidade de Cambridge. Já famoso pelos trabalhos matemáticos que anteriormente realizara, aplicou os seus talentos à investigação dos buracos negros. Ele e Jacob Bekenstein, físico israelita, descobriram independentemente uma relação surpreendente entre os buracos negros e a entropia — a medida termodinâmica do grau de desordem de um sistema físico. Pode

perguntar-se o que é que os buracos negros, consequência da teoria da relatividade, têm a ver com a entropia, uma propriedade termodinâmica. No entanto, existe uma profunda ligação.

Para explicar esta ligação tenho de dizer algo acerca da entropia. Os sistemas altamente ordenados, como um cristal, com os átomos cuidadosamente arrumados, têm baixa entropia, enquanto os sistemas altamente desordenados, como os gases, com os átomos vogando à deriva caoticamente, têm entropia elevada. A segunda lei da termodinâmica estabelece que a entropia de um sistema físico fechado nunca diminui — as coisas podem tornar-se mais caóticas, mas nunca menos. Como consequência, a informação acerca da estrutura pormenorizada de um sistema físico tende sempre a deteriorar-se; de facto, a perda de tal informação (convenientemente definida) é precisamente proporcional ao aumento da entropia de um sistema físico.

Podemos compreender agora como é que a entropia está relacionada com os buracos negros. Se um corpo cair num buraco negro, perde-se para sempre — não há processo, para quem esteja de fora, de o poder recuperar. Em particular, a informação perde-se para sempre quando caem corpos num buraco negro, e a perda de tal informação acarreta um aumento da entropia do buraco negro. Aquilo que Hawking e Bekenstein mostraram foi que a entropia de um buraco negro era proporcional à área da superfície do seu horizonte de acontecimentos — a fronteira do buraco. Uma vez que, de acordo com a segunda lei da termodinâmica, a entropia, ou aumenta sempre, ou mantém-se constante, parece que os buracos negros têm sempre de aumentar a área de superfície e,

portanto, tornam-se sempre cada vez maiores. Tudo indica não haver maneira de nos livrarmos de um buraco negro. Todavia, esta conclusão não é verdadeira. Facto assinalável, um buraco negro entregue a si mesmo acabará por evaporar-se em radiação. Como podemos compreender isto?

No seguimento dos estudos da termodinâmica dos buracos negros, Hawking constatou que era possível atribuir a um buraco negro uma temperatura inversamente proporcional ao seu raio. Compreendeu ainda que qualquer objeto cuja temperatura é não nula deve emitir radiação, tal como o carvão quente emite luz vermelha. Mas toda a ideia de um buraco negro se centra no facto de nada poder escapar dele, incluindo a radiação. Parece então existir um enigma: como podem os buracos negros irradiar?

Para espanto dos outros cientistas, Hawking desvendou este enigma em 1974, descobrindo o meio pelo qual os buracos negros irradiam precisamente aquilo que a sua temperatura exige. O argumento pode ser simplificado da seguinte maneira: embora seja verdade que nenhuma radiação interior ao horizonte de acontecimentos — a superfície do buraco — pode escapar, o mesmo não se passa com a radiação que lhe é imediatamente exterior. Hawking argumentou que o intenso campo gravitacional imediatamente exterior à superfície do buraco podia criar espontaneamente pares partícula-antipartícula. As teorias quânticas do campo das partículas elementares introduzem processos análogos, os quais foram já observados em laboratório. Segundo Hawking, uma das partículas do par criado aí dentro do buraco — perde-se para sempre —, enquanto a outra se escapa e

pode aniquilar-se ao encontrar uma antipartícula também em fuga, produzindo radiação pura. A radiação que se escapa, hoje chamada «radiação de Hawking», pode ser calculada para os grandes buracos negros que se possam formar a partir de estrelas que entraram em colapso e a sua intensidade é baixíssima. No entanto, os microburacos negros são «quentes» e irradiam a massa rapidamente, numa espetacular erupção de radiação de Hawking. Os microburacos negros eventualmente criados na altura do *big-bang* poderiam estar a explodir apenas agora. As erupções reveladoras têm sido procuradas, mas nunca foram vistas. Talvez existam hoje somente buracos negros grandes, muito estáveis, e todos os microburacos negros se tenham evaporado há muito tempo.

Com todos estes desenvolvimentos teóricos fascinantes, os buracos negros transitaram das «curiosidades matemáticas» para o centro da astronomia especulativa. A maior parte dos físicos pensam que eles devem existir, mas não há ainda dados observacionais que confirmem este preconceito teórico sem margem para dúvidas. Claro está, nunca poderemos «ver» um buraco negro, porque não há nada para ver. Como podem então ser detetados os buracos negros?

Os buracos negros têm assinaturas observáveis, das quais a mais importante é a enorme quantidade de energia radiante (que *não* é a radiação de Hawking) oriunda da vizinhança de um buraco negro, radiação que provém da matéria que lá cai. Um buraco negro é como um aspirador cósmico, suga tudo o que se encontra na vizinhança. Orbitando perto de uma estrela vulgar, absorve,

portanto, matéria desta. Ao precipitar-se no buraco negro, a matéria acelerada emite muita radiação, incluindo raios X, que pode ser detetada na Terra. Os buracos negros são hoje comummente invocados como possíveis explicações para a observação de enormes erupções de energia originárias de um pequeno volume do espaço. Só um buraco negro parece capaz de explicar tal libertação de energia.

Há duas situações em que os buracos negros, caso existam, poderiam ser observados. A primeira é quando em órbita em torno de uma estrela vulgar, cuja matéria consomem, emitindo, assim, radiação. A segunda é no núcleo das galáxias, onde grandes buracos negros podem estar a destruir estrelas inteiras. Examinemos a informação que tem sido facultada por uma área científica inteiramente nova, a astronomia dos raios X, um dos mais empolgantes avanços da astronomia nas últimas décadas.

Os raios X são uma forma de luz, mas com um comprimento de onda bastante mais pequeno do que o da luz visível. Há todas as razões para esperar que o espaço esteja cheio de raios X, tal como o está de luz visível. Infelizmente para os astrónomos, os raios X não penetram na atmosfera da Terra e durante anos a existência de raios X cósmicos foi objeto de mera conjetura. Nos princípios da década de 60, voos de foguetões e de balões acima da atmosfera ofereceram-nos o primeiro vislumbre do céu dos raios X. Facto espantoso, foram detetadas fontes de raios X, entre elas a situada no coração da nebulosa do Caranguejo, que sabemos hoje albergar uma estrela de neutrões. Apesar de estes voos de foguetões terem possibilitado observações com não mais de cerca de cinco minutos

de duração, tal foi suficiente para permitir concluir que estavam a desenrolar-se mudanças dramáticas nas poucas fontes de raios X detetadas.

Encorajados por estes resultados, cientistas de Cambridge, Massachusetts, que tinham organizado um grupo de investigação, conhecido por American Science and Engineering, propuseram que a NASA colocasse na órbita da Terra um pequeno satélite que pudesse melhorar a deteção dos raios X no céu. Dirigido por Riccardo Giacconi, o grupo de cientistas lançou um satélite, chamado *Uhuru* (palavra swahili que significa «liberdade»), de uma plataforma *offshore* de pesquisa de petróleo junto da costa do Quénia no dia 12 de Dezembro de 1970, dia da independência deste país. O custo deste projeto, incluindo o satélite, o veículo de lançamento e a carga, foi de cerca de 13 milhões de dólares. Nunca, em tempos recentes, foi tanto conhecimento astronómico adquirido por quantia tão modesta.

O *Uhuru*, que explorou apenas uma parte do céu noturno, localizou 125 fontes de raios X diferentes nos primeiros 70 dias de observação. Os céus estavam repletos de fontes de raios X! Era como se tivesse sido levantado um véu que cobrira os céus. Se pudéssemos orbitar por cima da atmosfera e estivéssemos equipados com óculos especiais, de maneira a podermos ver os raios X, um céu inteiramente novo e peculiar teria aparecido. Os astrofísicos, ainda a «digerir» a informação enviada pelo *Uhuru*, lançaram posteriormente outros satélites, como o observatório *Einstein*, o SAS-3, o *Hakucho*, satélite de raios X japonês, e vários outros. Conhecem-se, atualmente, mais de 300 fontes de raios X

diferentes, que fornecem indicações importantes acerca da natureza do universo.

Muitas destas fontes de raios X são pulsares, facilmente identificáveis, dada a extrema regularidade dos respetivos sinais, devido ao seu movimento de rotação. A localização de cerca de uma dúzia destes pulsares de raios X foi feita com tanta precisão que os astrónomos ópticos podem voltar os telescópios para essas fontes e identificar o corpo visível correspondente. As intensidades dos raios X e das ondas de rádio emitidas por esses pulsares aumentam, por vezes, de um fator superior à 1000. Suspeita-se de que tal acontece quando o «ponto quente» da estrela de neutrões (o polo magnético norte ou sul, locais onde a matéria atinge a estrela com maior facilidade) está virado para a Terra, de tal maneira que sofremos um impacto direto do feixe de raios X e de ondas de rádio.

Existem outras fontes de raios X que não se enquadram em nenhuma categoria bem determinada. Algumas são sistemas binários — um pulsar orbitando em torno de uma estrela vulgar. O compacto e maciço pulsar arranca matéria à companheira, processo este que resulta na emissão de raios X. Pensa-se que a mais potente fonte de raios X, a SCO X-1, seja um sistema binário e que o pulsar (ou talvez uma anã branca) orbite muito próximo de uma estrela vulgar. Grandes quantidades de gás são arrastadas para o interior do pulsar, sendo elas as responsáveis pelas grandes erupções de raios X. A SCO X-1 é uma fonte de raios X tão poderosa que pode perturbar as comunicações rádio de longa distância na Terra, ao sobrevoar-nos no céu noturno.

Tornou-se claro desde os primeiros dias que numa fonte de raios X, chamada Cisne X-1 (Cyg X-1) não se observavam pulsações regulares, razão por que não podia ser identificada como um pulsar. O que os astrónomos viam eram grandes flutuações de intensidade com durações entre um vigésimo de segundo e dez segundos e pequenos sinais periódicos que duravam apenas alguns segundos. Face a estes estranhos resultados, voltaram a atenção para este objeto estranho e encontraram um corpo visível correspondente — uma estrela catalogada como HDE 226868 —, que teria, pelo menos, 30 vezes a massa do Sol. A maioria concluiu que os raios X eram devidos a um sistema binário formado por uma companheira de grande massa numa órbita de pequeno raio em torno da estrela HDE 226868, talvez apenas a um terço da distância da órbita de Mercúrio ao nosso sol, distorcendo, assim, a forma da estrela e arrancando-lhe enormes quantidades de gás. A análise dos sinais sugeriu que a companheira era de massa extremamente elevada e muito compacta, tão compacta que não podia ser uma estrela de neutrões; tinha de ser um buraco negro.

Em Novembro de 1982 observações feitas no observatório interamericano de Cerro Tololo, no Chile, sugeriram a existência de um segundo candidato a buraco negro, chamado LMC X-3, na Grande Nuvem de Magalhães, o qual descreve uma órbita de um pequeno raio em torno de uma estrela normal no período de 1,7 dias, engolindo gás e emitindo raios X. Através da medição da velocidade da estrela visível, determinou-se que a companheira tem uma massa, aproximadamente, 10 vezes maior do que a massa do Sol, demasiado elevada para ser uma estrela de neutrões. O buraco negro acabará por engolir completamente a companheira, ficando

isolado e completamente invisível. É possível que existam muitos buracos negros como este, os quais, tendo devorado completamente as companheiras, vagueiam agora pela galáxia, sombrios e solitários.

A conclusão radical de que a Cyg X-1 e o LMC X-3 são buracos negros não é partilhada por todos os astrónomos. Alguns pensam que uma estrela de neutrões poderia produzir os sinais de raios X observados. Deve dizer-se que os dados da Cyg X-1 e do LMC X-3, conquanto consistentes com a hipótese de se tratar de buracos negros, não provam que o sejam. É, pois, provável que a questão da existência de buracos negros permaneça em aberto por mais algum tempo.

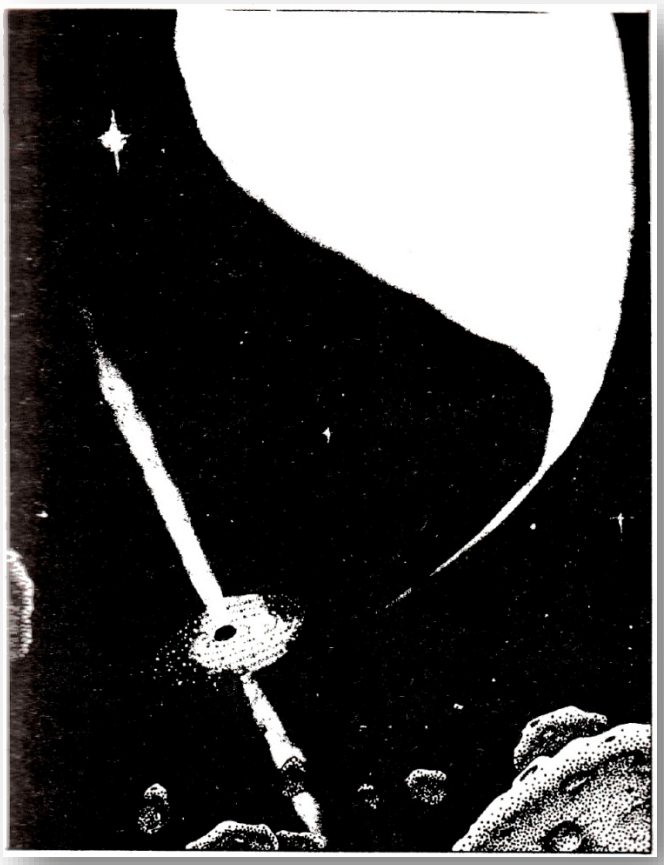
Como poderá a questão ser esclarecida? Os astrofísicos teóricos estão, presentemente, a elaborar modelos matemáticos de buracos negros envoltos em gás, numa tentativa de descobrir as «assinaturas» características dos buracos negros — sinais que só podem ser atribuídos a estes objetos. Tais «assinaturas» poderiam ser sinais emitidos pelo gás quente ao ser sugado. De acordo com um destes modelos, o gás que rodeia um buraco negro deveria formar um «disco de acreção» semelhante ao dos anéis em volta do planeta Saturno, onde poderiam aparecer pontos quentes — fontes intensas de raios X. Estes pontos quentes orbitariam em torno do buraco e do respetivo disco num milésimo de segundo e deveriam emitir sinais com essa frequência até serem engolidos por ele. Tais discos de acreção podem também formar-se à volta dos pulsares, sendo as suas propriedades semelhantes às dos que

circundam os buracos negros, o que torna difícil a distinção entre os dois.

Observar tais sinais regulares provenientes dos discos de acreção foi a missão do HEAO-1 (*high energy astronomy observatory*)[2], um satélite de raios X lançado em 1977. Todavia, não foram registados quaisquer sinais com a frequência de 1 milissegundo vindos da Cyg X-1 ou de qualquer outro candidato a buraco negro em sistemas binários. Contudo, outras fontes de raios X existem, chamadas emissores intensos, que emitem sinais periódicos de raios X, que alguns cientistas pensam poderem ser indicativos de buracos negros. No entanto, esta explicação das erupções de raios X não goza hoje de grande apoio. Pelo contrário, sabe-se agora que são devidas a estrelas de neutrões orbitando em torno de estrelas normais. As provas da existência de buracos negros em sistemas binários permanecem inconsistentes.

Os buracos negros que se julga existirem em certos sistemas binários são insignificâncias, comparados com os monstros que se crê encontrarem-se nos núcleos das galáxias — buracos negros de massa muito elevada, com milhões de vezes a massa do Sol. Sabe-se que há qualquer coisa nos núcleos das galáxias que liberta enormes quantidades de energia, possivelmente buracos negros gigantes. Quem não aceita a realidade dos buracos negros refere-se ao objeto ou processo em termos mais neutrais, como «a máquina», «o monstro» ou «o primeiro motor». Mas ninguém duvida de que existe qualquer coisa de incomum no centro das galáxias.

O núcleo galáctico mais próximo é, como não podia deixar de ser, o da nossa própria galáxia — a Via Láctea. Não podemos vê-lo diretamente, por estar obscurecido por poeiras e gás. No entanto, as ondas de rádio e as radiações infravermelhas e de raios X que penetram a poeira e o gás podem dizer-nos qualquer coisa acerca dele. Sabemos hoje que existe uma concentração maciça de estrelas no núcleo galáctico. Se o nosso sistema solar estivesse situado lá, o nosso céu noturno estaria iluminado pelo brilho de uma centena de luas cheias, vindo de todas as estrelas no núcleo. Foram detetadas várias fontes de radiação infravermelha, que, provavelmente, serão enormes estrelas gigantes vermelhas. As observações do gás ionizado sugerem a existência de uma única fonte ionizante, que pode ser identificada com uma intensa e compacta fonte de rádio — presumivelmente, um buraco negro —, situada mesmo no centro da galáxia, num lugar conhecido por Sagitário A Ocidental. De acordo com o modelo de buraco negro, um destes objetos, situado numa posição central e de massa muito elevada, aquece o gás envolvente, que é a fonte de raios X e radiação ionizada. O buraco negro, em si, é bastante pequeno — tem as dimensões do sistema solar — e comporta-se razoavelmente bem. Mas está cercado de violência.



Um buraco negro em órbita em torno de uma estrela companheira e sugando matéria dela. A matéria capturada — gás quente — forma um disco de acreção à volta do buraco antes de ser engolida por ele. Segundo alguns modelos teóricos, a energia radiante é ejetada pelo sistema do disco de acreção do buraco negro ao longo do eixo de rotação.

Em Dezembro de 1983 os astrónomos de Cal Tech exploraram a região do Sagitário A Ocidental, utilizando as 27 antenas de rádio do Very Large Array, em Socorro, Novo México, e obtiveram uma imagem de alta resolução. Descobriram três braços de gás, movendo-se em rápido redemoinho, centrado (possivelmente caindo) no ponto onde se crê estar situado o buraco negro. Se esta interpretação for correta, então, como disse um dos astrónomos de Cal Tech, «é a primeira vez que vemos realmente matéria a cair num buraco negro». Em 1984 os astrónomos descobriram ainda

que o núcleo estava envolto em intensos campos magnéticos — uma descoberta que os deixou perplexos.

Outros indícios da existência de buracos negros nas galáxias provêm, não da nossa galáxia, mas sim da investigação de galáxias extremamente ativas e presumivelmente jovens (a Via Láctea é uma galáxia razoavelmente madura). Cresce a convicção de que os quasares — objetos astronómicos altamente energéticos e situados a enormíssimas distâncias — são um estádio primitivo do desenvolvimento das galáxias. Poderão os quasares ser uma manifestação da dinâmica dos buracos negros? Parece não haver mais nada capaz de explicar a prodigiosa produção de energia destes objetos. Os quasares poderiam ser alimentados por gás absorvido, algum do qual seria posteriormente devolvido ao espaço, formando uma «tocha cósmica» — uma estrutura em jato que faz lembrar alguns corpos celestes que, por vezes, se observam. Todas as galáxias ativas, como as violentas Seyferts, os objetos BL do Lagarto, os quasares, todos podem ser diferentes estádios evolucionários do desenvolvimento das galáxias, estádios nos quais «a máquina» dos núcleos ainda não atingiu a serena meia-idade exemplificada pela nossa galáxia.

O papel dos buracos negros de massa muito elevada nas galáxias ativas é, evidentemente, especulação. Todavia, se os astrónomos tivessem de fazer apostas quanto ao melhor local para os situarem — localização central e massa muito elevada —, muitos apostariam no núcleo da galáxia M 87, uma galáxia elíptica gigante que desempenha um papel muito importante na coesão gravitacional da «supergaláxia» da Virgem, que é composta por mais de 1000

galáxias. O núcleo da M 87, bastante ativo, como o revelam as emissões de rádio e de raios X, talvez seja um velho quasar.

O interesse pela M 87 motivou em 1978 o exame observacional do respetivo núcleo por intermédio da mais recente tecnologia eletrónica capaz de amplificar sinais luminosos. Os astrónomos descobriram não só um pico na intensidade da luz no núcleo da M 87, como também que a velocidade do movimento das estrelas aumentava abruptamente à medida que diminuía a sua distância relativamente ao centro da galáxia, como se fossem atraídas para um objeto de grande massa aí situado. Concluiu-se que estas observações estão em perfeita concordância com a existência de um buraco negro de grande massa — 5 mil milhões de vezes a massa do Sol — no centro da galáxia M 87, ainda que tal facto não pudesse ser definitivamente comprovado. Com o advento do telescópio espacial, a lançar em 1986, a M 87 será observada com uma precisão 20 vezes maior e talvez aprendamos mais sobre ela.

Recentemente, um grupo de astrónomos europeus empreendeu uma extensa e pormenorizada análise dos dados obtidos pelo *international ultraviolet explorer*, um satélite lançado em 1978, concluindo que uma das mais brilhantes das energéticas galáxias de Seyfert, a NGC 4151, continha um buraco negro com cerca de 100 milhões de vezes a massa do Sol. Por sorte, o satélite estava a observar esta galáxia quando o núcleo se acendeu subitamente. As nuvens de gás em órbita em torno do núcleo foram também ativadas sequencialmente, e esta informação permitiu aos investigadores determinar a velocidade e distância das nuvens ao centro. Uma vez conhecidos estes elementos, tornou-se um mero

exercício de física newtoniana calcular a massa do objeto central, assim se estabelecendo a presença de um buraco negro.

Os buracos negros percorreram um longo caminho desde os tempos em que eram curiosidades matemáticas da imaginação científica. Hoje são correntemente invocados pelos teóricos para explicar quase todas as novas observações astronómicas que carecem de uma poderosa fonte de energia numa pequena região do espaço. No entanto, não sabemos ao certo se existem. Como nota Edwin Turner, astrofísico de Princeton, «os buracos negros são uma ideia central da teoria das galáxias ativas e para quem tem inclinações empíricas a observação direta é essencial». Não seria maravilhoso poder obter uma prova definitiva da existência dos buracos negros? Talvez alguma prova incontestável esteja em vésperas de aparecer. Ou talvez não. É possível que nenhuma prova definitiva seja encontrada, mas que a opinião científica se incline gradualmente a favor ou contra a ideia de buracos negros. De certa maneira, seria uma pena que tal acontecesse. Ideias tão impressionantes e ricas em consequências como as que implicam a existência de buracos negros merecem, ou confirmação por observação, impressionante e rica em consequências, ou rejeição conclusiva. Gostaríamos que a história dos buracos negros chegasse a uma conclusão estrondosa, e não lamurienta.

Numa análise retrospetiva do presente século, podemos constatar quão longe os astrofísicos avançaram na compreensão das estrelas — como são criadas, como vivem e como morrem. Conhecem-se hoje as principais características da evolução estelar. E as estrelas não só parecem hoje menos misteriosas, como até

parecem mais amáveis. Como nós, também elas nascem, vivem e morrem e, tal como nós, sobrevivem numa sociedade mais vasta — a galáxia.

O estudo das galáxias é um ramo da astronomia que, ao contrário do que acontece com a astrofísica, está ainda na infância. Não existe ainda uma compreensão sistemática da evolução das galáxias, uma compreensão que nos leve desde as sementes primordiais das quais elas nasceram no *big-bang* até à sua magnificência atual. Isso será para o futuro. Mas muito se aprendeu já acerca das galáxias, principalmente nas últimas décadas, e é para este assunto que agora nos voltamos.

---

[1] Em 23 de Fevereiro de 1987 deu-se uma explosão de supernova na Grande nuvem de Magalhães. (*N. do T.*)

[2] Observatório astronómico de altas energias. (*N. do T.*)

CAPÍTULO 4

A Descoberta das Galáxias

> Posso, pois, terminar como comecei. Da nossa casa, na Terra, contemplamos as distâncias e esforçamo-nos por imaginar a espécie de mundo no qual nascemos. Hoje penetrámos já nas profundezas do espaço. Conhecemos intimamente a nossa vizinhança imediata. Contudo, à medida que a distância aumenta, o nosso conhecimento esbate-se [...] até que no último horizonte indistinto procuramos por entre nebulosos erros de observação pontos de referência que tenham mais algum significado. A busca continuará. O anseio é mais velho do que a história. Nunca será satisfeito e nunca será eliminado.
>
> *EDWIN HUBBLE (no seu último artigo científico)*

O facto de o céu noturno aparecer repleto de estrelas contribui para a ilusão de que o espaço imenso do universo deve estar também uniformemente repleto de estrelas. Esta ilusão é tão persuasiva que foi só neste século que os astrónomos puderam provar definitivamente que as estrelas são parte das galáxias — os «universos-ilhas» — e que estas são os principais habitantes do cosmos.

Se pudéssemos sair da Via Láctea, veríamos que ela é um imenso disco em espiral, cujos braços difusos se enrolam em torno de um bojo central de estrelas, dentro do qual se esconde o misterioso núcleo da galáxia. As fronteiras desse bojo estão assinaladas por um anel de nuvens espessas e granuladas de hidrogénio molecular. Se olhássemos mais de perto, veríamos que os braços são

delineados por estrelas azuis brilhantes e contêm grandes quantidades de poeiras e gás concentrados em nebulosas produtoras de estrelas. O nosso próprio sol está situado no rebordo interior de um desses braços, o braço de Oríon: uma estrela entre as centenas de milhares de milhões que constituem a galáxia.

Ainda que o disco em espiral e o bojo central sejam as características mais salientes da nossa galáxia, se olharmos para o espaço que a rodeia — o halo —, veremos que contém «aglomerados globulares» de estrelas, cada um deles um sistema independente constituído por 50.000 a 1 milhão de estrelas ligadas gravitacionalmente, formando um agregado grosseiramente esférico. Alguns astrofísicos pensam que nos centros dos aglomerados globulares existem pequenos buracos negros com massas de 100 vezes a do Sol. Os astrónomos descobriram algumas centenas destes aglomerados de estrelas (cerca de uma centena no halo e outra centena no disco), que se dispõem simetricamente em torno da galáxia e tendem a concentrar-se junto do centro. Cada aglomerado globular é, de facto, um satélite da galáxia Via Láctea. Os astrofísicos aprenderam que a maioria das estrelas que compõem os aglomerados globulares são muito velhas, datando da época da formação da própria galáxia. Os aglomerados globulares são, provavelmente, lembranças de coisas passadas e constituem uma pista, por enquanto não compreendida, para a origem da própria galáxia.

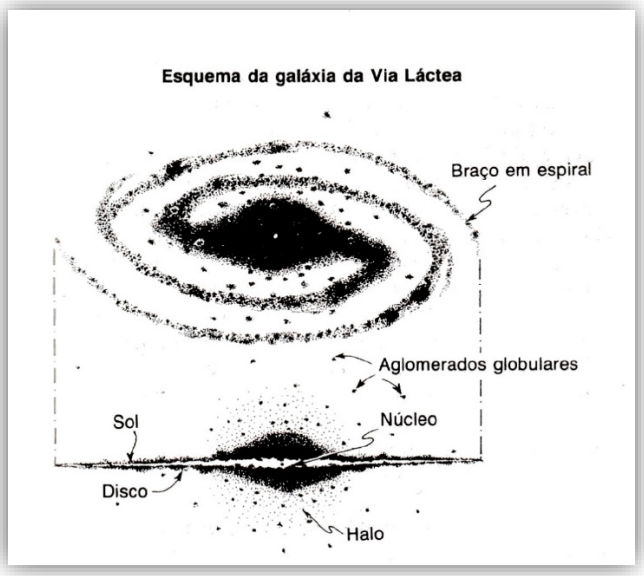
A «senhora» desta vasta galáxia — disco, bojo central e halo de aglomerados globulares — é a gravidade, a única força que governa o movimento de estrelas muito separadas. A lei da gravidade,

descoberta por Newton, é bastante simples: a força entre duas massas é sempre atrativa, sendo diretamente proporcional ao produto dessas massas e inversamente proporcional ao quadrado da distância entre elas. No entanto, e apesar desta simplicidade, a lei da gravidade tem de explicar as complexas configurações de milhares de milhões de estrelas descrevendo trajetórias tortuosas, bem como o facto de a galáxia ter mantido a forma ao longo de milhares de milhões de anos, sem entrar em colapso e sem se desfazer em pedaços. Os físicos matemáticos e os astrofísicos que estudam o problema da estrutura da galáxia estão conscientes de que tentar seguir o movimento de apenas algumas centenas de estrelas obedecendo à lei da gravidade atinge os limites da capacidade dos nossos maiores computadores. Para tratar um número de estrelas adequado à descrição das galáxias é necessário usar métodos estatísticos, que obtêm médias sobre grandes quantidades de estrelas. Assim, apesar da simplicidade da lei da gravidade, os físicos ainda consideram difícil determinar matematicamente propriedades como a forma e a estabilidade da galáxia. Mas sabem que o carácter atrativo e de longo alcance da gravidade garante que um sistema dinâmico de estrelas não possa ser absolutamente estável para sempre. Uma galáxia deve mudar e evoluir, talvez até drasticamente.

A imagem padrão da galáxia que os astrónomos conservaram durante décadas — um bojo central, uma espiral em disco e um halo de aglomerados globulares — começou a mudar drasticamente nos princípios da década de 70. Os estudos matemáticos de Donald Lyndem-Bell, da Universidade de Cambridge, e de um grupo da Universidade de Princeton, que incluía Jeremiah P. Ostriker, P. J. E.

Peebles e Amos Yahil, mostraram que o disco de uma galáxia em espiral não seria dinamicamente estável a menos que a galáxia estivesse rodeada por um halo extenso e de massa elevada constituído por matéria negra. Se esta ideia estiver correta, então quase toda a massa da galáxia — até 90% — reside, não no gás e nas estrelas visíveis, mas numa nova componente, o halo invisível.

O trabalho de J. Einasto e seus colaboradores no observatório de Tartu, na Estónia, trouxe apoio observacional, e não teórico, à hipótese da existência de um halo invisível e de massa elevada. Estudando o movimento da nossa galáxia relativamente às galáxias próximas, descobriram que ela se movia com bastante rapidez, concluindo então que deveria ter muito mais massa para, assim, poder estar ligada gravitacionalmente ao sistema das galáxias vizinhas. Esta circunstância levou-os a sugerir que a massa que faltava estaria contida num halo de grande massa.



Representação grosseira da nossa galáxia, a Via Láctea, quando vista de cima e de lado. O diâmetro é de cerca de 100 000 anos-luz. Esta ilustração mostra as principais componentes visíveis da galáxia — o bojo central de estrelas envolvendo o núcleo,

os braços em espiral e o halo dos aglomerados globulares, pequenos sistemas estelares independentes concentrados perto do centro da galáxia. Conquanto revele as principais componentes visíveis, esta ilustração não mostra a coroa de gás quente que envolve a galáxia, os campos magnéticos no disco ou a matéria escura que se estende bem para além do bordo visível da galáxia. A dinâmica das galáxias começa agora a ser compreendida. O centro da galáxia pode albergar um grande buraco negro.

O indício mais surpreendente de que existiria um halo invisível surgiu quando foi medida a velocidade do gás que orbita em torno das galáxias, longe dos bordos visíveis. Se toda a massa de uma galáxia estivesse concentrada nas estrelas visíveis, então a velocidade do gás em órbita deveria diminuir à medida que a sua distância relativamente à galáxia aumentasse, da mesma maneira que a velocidade de um planeta em órbita em torno do Sol é tanto menor quanto mais longe se encontra. Em vez disso os astrónomos Vera C. Rubin, W. Kent Ford, Jr., e Norbert Thonnard, da Carnegie Institution, em Washington, D.C., tiveram a surpresa de descobrir que a velocidade do gás em órbita não diminui, mas permanece constante, indicando, assim, que a maior parte da massa da galáxia não está contida na região visível, mas se estende muito para além dela, um efeito, aliás, já reconhecido por Martin Schwarzschild, Leon Mestel e outros na década de 50. A maioria dos astrónomos estão hoje convencidos de que as galáxias têm halos invisíveis de massa elevada e de que a distribuição das fontes luminosas numa galáxia não é indicativa da distribuição de massa. Os mais populares candidatos ao estatuto de matéria escura do halo são as novas partículas quânticas imaginadas pelos físicos teóricos.

Quando os novos detetores de luz ultravioleta foram colocados em satélites orbitais, como o *internationrnal ultraviolet explorer*, os astrónomos do fim da década de 70 confirmaram a conjetura de

1956 do físico de Princeton Lyman Spitzer, Jr... segundo a qual a nossa galáxia está envolvida por uma coroa de gás quente que se estende para cima e para baixo do disco. Esta coroa, que absorve a luz ultravioleta das estrelas distantes, podendo, portanto, ser detetada, não está relacionada com o hipotético halo de matéria invisível. É evidente que o disco da galáxia, onde se encontram todas as estrelas, expele explosivamente gás quente para o espaço por cima e por baixo do disco, em correntes gigantescas. Uma vez no espaço, o gás quente arrefece, perde velocidade e volta a cair no disco galáctico — um ciclo denominado «fonte galática». A energia da fonte galática parece provir das explosões de supernovas das estrelas do disco. Tais coroas de gás quente aparecem também envolvendo outras galáxias.

Na última década não só mudou a imagem que se tinha da «arquitetura» em grande escala da nossa galáxia, como mudaram também os pontos de vista dos astrónomos que estudam os pormenores da sua estrutura interna. Os habitantes de maior massa da nossa galáxia não são estrelas, mas sim os complexos gigantes de nuvens moleculares que se concentram nos braços da espiral, cuja existência era anteriormente desconhecida. Estes complexos de nuvens são lugares de formação de estrelas e sede de complicados processos físicos que desempenham um papel importante na evolução da nossa galáxia. Recentes observações e estudos do núcleo da nossa galáxia sugerem também que alberga um objeto compacto de enorme massa, talvez um buraco negro. Como estas descobertas inesperadas indicam, os cientistas estão a começar agora a entender o que se passa realmente numa galáxia. Muito há ainda para aprender.

O disco da galáxia Via Láctea tem cerca de 100 000 anos-luz de diâmetro. Se nos afastarmos dela cerca de 1 milhão de anos-luz e olharmos para trás, veremos que não está só — está rodeada de um grupo de galáxias satélites mais pequenas. Entre as mais conspícuas contam-se a Grande e a Pequena Nuvem de Magalhães (visíveis do hemisfério sul da Terra), galáxias de forma irregular perto da fronteira da Via Láctea que foram despedaçadas e distorcidas pela interação gravitacional de maré com a nossa galáxia. Mais longe ainda estão as galáxias satélites anãs — Dragão, Escultor, Sextante, Pégaso, Forno, Ursas Maior e Menor, Carina, ou Quilha, Leão I e Leão II —, todas elas sistemas estelares galáticos pequenos e dispersos, quando comparados com a Via Láctea. Haverá, provavelmente, outras pequenas galáxias satélites como estas na nossa vizinhança, as quais têm passado despercebidas, por serem demasiado ténues para serem detetadas, todas dentro de um raio de 1 milhão de anos-luz do centro da Via Láctea.

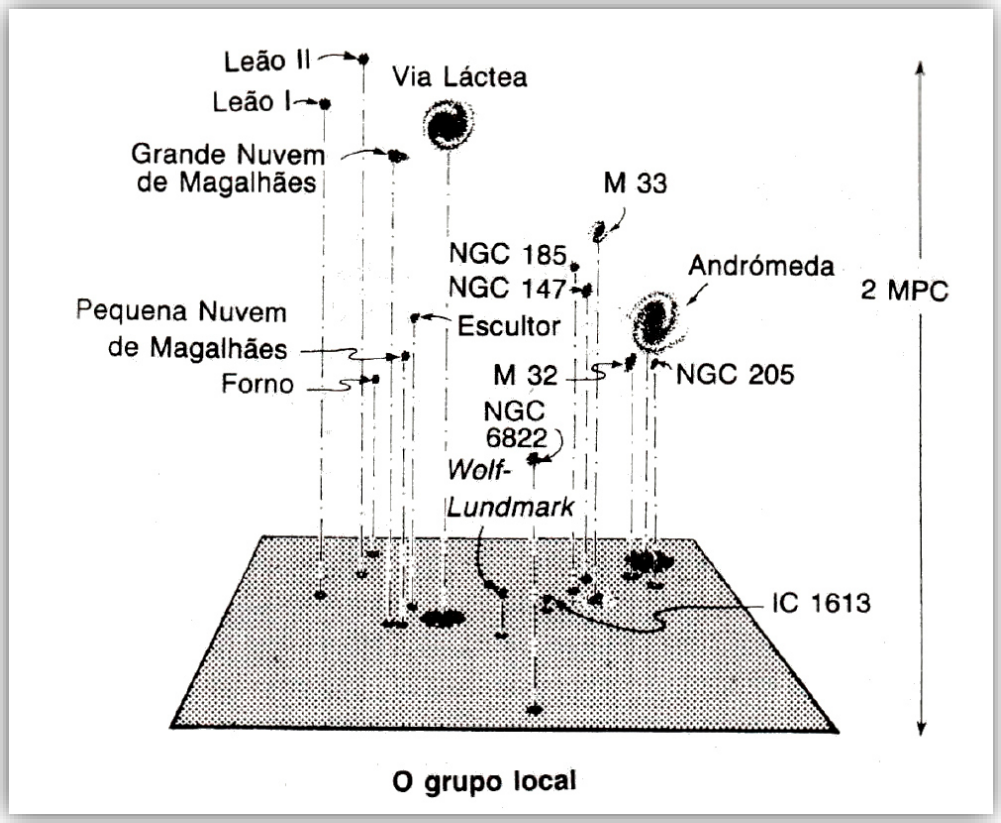
A pouco mais de 2 milhões de anos-luz encontra-se outra grande galáxia em espiral, comparável à Via Láctea, a nossa irmã, a galáxia de Andrómeda. Olhando para a galáxia de Andrómeda a partir da Terra, podemos ter uma boa ideia do aspeto que a nossa galáxia deve ter vista de longe, uma vez que as duas galáxias são muito semelhantes. A galáxia de Andrómeda, como a nossa, está rodeada de galáxias satélites mais pequenas, das quais as mais proeminentes são as duas galáxias elípticas companheiras que orbitam em torno dela. Todo o sistema de galáxias da Andrómeda se move em direção ao sistema da Via Láctea à velocidade de 90 km por segundo, pelo que daqui a milhares de milhões de anos as duas galáxias poderão colidir. Tal colisão não será a catástrofe que

se possa imaginar, dado que as galáxias são, na maior parte, espaço vazio. As duas galáxias irão passar uma através da outra, mas o efeito sobre o gás interestelar das duas galáxias irá distorcer a forma de cada uma delas, indicação de que uma colisão teve realmente lugar.

As galáxias da Via Láctea e de Andrómeda, juntamente com as suas satélites associadas e algumas galáxias desgarradas, cerca de 17 ou 18 galáxias no total, constituem, coletivamente, aquilo a que os astrónomos chamam «o grupo local». É este o nosso «lar» dentro do conjunto das galáxias que povoam o universo — o nosso cantinho do cosmos. Quando os astrónomos querem fazer observações pormenorizadas das galáxias, examinam o «grupo local», porque estas são as galáxias mais próximas.

Se examinarmos regiões mais vastas do espaço para lá do grupo local, encontraremos mais e mais galáxias; o seu número parece não ter fim. No entanto, pode notar-se que as galáxias não estão espalhadas ao acaso no espaço, mas, pelo contrário, tendem a juntar-se em grupos, formados por algumas centenas de grandes galáxias, acompanhadas, provavelmente, por alguns milhares de galáxias mais pequenas. O mais próximo destes aglomerados de galáxias é a parte central do aglomerado da Virgem, afastada de nós entre 30 e 60 milhões de anos-luz e composta por centenas de grandes galáxias em espiral. Mais longe, a uma distância entre 200 e 400 milhões de anos-luz, encontra-se o aglomerado de Coma, ou Cabeleira, albergando, pelo menos, 1300 grandes galáxias. O universo está povoado por estes aglomerados de galáxias. Todavia, os aglomerados de galáxias não são os maiores conjuntos: a

hierarquia do agrupamento em aglomerados vai até aos superaglomerados.



O grupo local de galáxias. Mapa tridimensional da nossa região do universo que mostra a nossa galáxia Via Láctea e os seus vizinhos. Existem cerca de 100 mil milhões de galáxias no universo observável.

Aglomerados de galáxias como os de Coma, ou Cabeleira, e da Virgem e vários outros aglomerados mais pequenos tendem a congregar-se em tais superaglomerados — gigantescos aglomerados de aglomerados. O nosso grupo local de galáxias, por exemplo, é parte de um superaglomerado que também engloba o aglomerado da Virgem. Tais superaglomerados de galáxias são os maiores objetos bem definidos que têm nome humano — facto registado em *The Guinness Book of World Records*[1]:

> Recorde epónimo. O maior objeto que detém um nome humano é o superaglomerado de galáxias conhecido por Abell 7, do nome do astrónomo Dr. George O. Abell, da Universidade da Califórnia. O grupo de aglomerados tem uma dimensão linear estimada em 300 000 anos-luz e a sua descoberta foi anunciada em 1961.

As galáxias são os habitantes primeiros e visíveis do universo. Estas grandes ilhas de estrelas estão agrupadas segundo uma hierarquia, que consiste em galáxias, aglomerados de galáxias e superaglomerados. Por que razão se organiza o universo desta maneira peculiar? Por que razão não estão as estrelas, ou mesmo as galáxias, uniformemente distribuídas pelo espaço? Alguns astrónomos especulam que a hierarquia dos aglomerados representa um desenvolvimento evolutivo do universo. Todas as estruturas do universo são instáveis, mudam e evoluem, ainda que muito lentamente, segundo os padrões humanos. De acordo com esta imagem, as galáxias primordiais formaram-se pouco após a origem do próprio universo e atravessaram uma série de estádios evolucionários. As galáxias estão hoje já amadurecidas e a Era da grande evolução galática já terminou. Estamos agora numa nova Era, em que as galáxias dentro dos aglomerados se movem ao encontro umas das outras, talvez para formar aglomerados e superaglomerados mais compactos.

Uma série de interrogações ainda mais profundas do que as suscitadas pela distribuição peculiar das galáxias no espaço diz respeito a saber por que é que as galáxias existem, qual a sua origem e quais os pormenores da sua evolução subsequente. Pouco se sabe das respostas a tais interrogações. Walter Baade, um dos

grandes astrónomos deste século, fez notar que a nossa compreensão atual das galáxias é tão incompleta como o era a das estrelas no princípio deste século, antes de entendermos a libertação de energia da fornalha nuclear. Não conhecemos ainda os mecanismos energéticos fundamentais responsáveis pela evolução das galáxias. No entanto, não obstante a nossa ignorância no que toca à sua dinâmica, é notável quanto progrediu o nosso conhecimento das galáxias neste século.

As conquistas alcançadas pela astronomia moderna são devidas a dois desenvolvimentos de grande importância: primeiro, o advento de novos instrumentos — telescópios de alta resolução, tecnologia óptica, mecanismos de *timing* e, mais recentemente, radiotelescópios, computadores e observatórios orbitais; segundo, o novo conhecimento que adquirimos na Terra acerca das propriedades da matéria.

A astronomia é uma ciência observacional, e não experimental. Não é possível alterar experimentalmente os objetos astronómicos para ver como se modificam fisicamente — as estrelas estão muito longe e são muito grandes para que possamos fazê-lo. Contudo, é possível realizar experiências aqui na Terra e extrair informações acerca das propriedades da matéria — o comportamento dos átomos, das moléculas, da luz e dos gases. Uma vez que estas propriedades são universais, até a matéria das estrelas e galáxias distantes deve obedecer às leis da Natureza descobertas na Terra e as observações dos astrónomos podem ser interpretadas pelos astrofísicos em termos de processos físicos familiares. As experiências terrestres, complementadas por modelos teóricos

computadorizados, caminham lado a lado com as observações astronómicas — uma combinação poderosa que abre o cosmos à investigação racional.

A história da descoberta das galáxias exteriores e da distribuição das estrelas na nossa Via Láctea é parte da grande história científica deste século, que está longe de ter acabado. Somos como crianças que, tendo apreendido o ambiente da própria casa e do jardim, viram que havia outros edifícios para lá da cerca e que existe um mundo muito mais vasto, por enquanto desconhecido, situado para além das vizinhanças. Virá o dia em que saberemos até como é que o ambiente foi criado. Para começar esta história, recuemos alguns séculos, até aos primeiros filósofos e astrónomos cujas especulações e observações abriram caminho à moderna conceção das galáxias.

Thomas Wright, um inglês de Durham, registou os seus pensamentos cosmológicos em *An Original Theory*, or New Hypothesis of the Universe[2], publicado em 1750. Do nosso ponto de vista moderno, representam uma mistura de teologia e astronomia: via o universo como uma relação divina. Supunha que a Via Láctea aparecia como uma faixa no céu por ser uma camada achatada de estrelas. Um dos seus modelos representava-a como uma «mó», enquanto noutros dispôs as estrelas concentricamente à volta de um centro «sobrenatural». Especulou que a Via Láctea seria apenas um entre muitos sistemas estelares existentes no universo — uma ideia que antecipava as observações do século XX. Rejeitou mais tarde muitas das especulações que publicara, mas não antes que um artigo de jornal sobre o seu trabalho, embora

equívoco, estimulasse o interesse do jovem Immanuel Kant, na longínqua Königsberg, por questões de cosmologia.

Em 1755, com 31 anos de idade, Kant publicou a *História Natural Universal e Teoria dos Céus*, na qual ensaiava o desenvolvimento de uma cosmologia que tivesse em conta a nova física newtoniana: o movimento das estrelas na Via Láctea tinha de obedecer à lei da gravitação de Newton. Salientou também que as diferentes espécies de nebulosas então observadas podiam requerer explicações diferentes, o que hoje sabemos ser verdade. Algumas destas nebulosas eram, como ele conjeturou, grandes sistemas de estrelas exteriores. Kant escreveu:

> É de longe mais natural e inteligível olhar [as nebulosas] como sendo, não enormes estrelas isoladas, mas sistemas de muitas estrelas, cuja distância as apresenta tão juntas que a sua luz, sendo individualmente impercetível, nos atinge, devido ao seu imenso número, com uma uniforme e pálida claridade [...] O que está em perfeita harmonia com a ideia de que estas figuras elípticas são precisamente universos, ou seja [...] vias lácteas.

Infelizmente para o jovem Kant, não só este importante trabalho foi publicado anonimamente, como os editores entraram em falência e o livro ficou fechado num armazém. As ideias cosmológicas de Kant só lograram ampla circulação com o despontar da sua fama como filósofo. Em 1761 o astrónomo L. H. Lambert publicou algumas ideias às quais tinha chegado independentemente, mas que eram semelhantes às de Kant, o qual sentiu que tinha de defender a sua prioridade. Em 1791 as observações de William Herschel sobre a rotação do anel interior de

Saturno — observações que apoiavam algumas das ideias de Kant — instigaram-no novamente a defender a sua prioridade. Contudo, muitas delas, especialmente a sugestão de que algumas nebulosas eram sistemas de estrelas exteriores, não puderam ser comprovadas senão no século XX, tendo a natureza das nebulosas permanecido, entretanto, objeto de especulação.

William Herschel, o pai da moderna astronomia observacional, dedicou a maior parte dos últimos anos de vida a determinar a forma do nosso sistema estelar. Foi o primeiro a mostrar que as estrelas não estão dispostas simetricamente à volta do Sol. Concluiu também que a galáxia é uma estrutura com dois lóbulos, existindo um intervalo na distribuição das estrelas, que hoje sabemos ser devido a uma faixa escura de nuvens de poeira e gás que se estendem desde a constelação do Cisne até às latitudes meridionais. Herschel, tal como Newton antes dele, calculou as distâncias até às estrelas partindo do princípio de que eram tão brilhantes como o Sol. Este pressuposto não era correto, como o próprio Herschel veio a compreender, acabando por desesperar de alguma vez poder determinar a estrutura precisa do nosso sistema estelar, a Via Láctea.

Durante o século XIX pouco progresso foi feito no sentido da compreensão da estrutura da Via Láctea, o que não é surpreendente, pois é extremamente difícil determinar a estrutura de qualquer coisa se estivermos *dentro dela*, em vez de a olharmos de fora. No século XIX não era sequer claro que existisse um «lado de fora» para o nosso sistema estelar. Até às primeiras décadas deste século a maior parte dos astrónomos pensavam que o

universo consistia num único sistema estelar que preenchia o espaço infinito.

Pesem embora as muitas conceções erradas acerca da estrutura global do universo, os astrónomos dos séculos XVIII e XIX fizeram várias observações importantes, entre as quais as primeiras medições diretas das distâncias até às estrelas próximas, que foram de importância crucial na determinação da forma da nossa galáxia. Como se pode medir a distância para um objeto longínquo, como uma estrela, sem lá ir? Os astrónomos usaram o método da paralaxe, que é facilmente ilustrado se olharmos para um dos nossos dedos a uma distância de cerca de 30cm. Olhemos primeiro para o dedo só com um olho e depois só com o outro. O dedo muda de posição em relação ao que se vê ao fundo. Conhecendo o ângulo subtendido pelo salto e o comprimento da «linha de base» entre os olhos, pode calcular-se geometricamente a distância dos olhos ao dedo. Foi usado este mesmo método da paralaxe para medir a distância até às estrelas. Usando como linha de base o diâmetro da órbita da Terra em torno do Sol, fazem-se duas observações da mesma estrela com um intervalo de seis meses. O ângulo medido é o «salto» aparente da posição da estrela contra o fundo fixo das estrelas ainda mais distantes. A observação da paralaxe confirmaria também definitivamente o modelo solar de Copérnico — a Terra move-se em torno do Sol, e não o Sol em torno da Terra.

Nos últimos anos da década de 1830 os novos telescópios eram já suficientemente aperfeiçoados para que a paralaxe fosse visível. Entre 1838 e 1839 Friedrich Wilhelm Bessel mediu a distância à estrela 61 do Cisne, Friedrich Georg Wilhelm Struve, a distância à

Alfa da Lira, na constelação de Vega, e Thomas Henderson, a distância à Alfa do Centauro. Os cálculos mostraram que todas estas distâncias eram de, pelo menos, vários anos-luz — mesmo as estrelas mais próximas estavam muito afastadas. Tornou-se claro, a partir de outras medições, que estas e outras estrelas na vizinhança do Sol se moviam a velocidades de dezenas ou mesmo centenas de quilómetros por segundo e que o Sol fazia parte de um sistema estelar dinâmico.

As mais misteriosas de todas as observações foram talvez as das nebulosas — uma nova classe de objetos astronómicos, distintos das estrelas, planetas e cometas, cuja relevância permaneceu obscura até ao presente século. 103 destas nebulosas, manchas indistintas entre as estrelas, foram catalogadas pelo astrónomo francês Charles Messier em 1784, cujo colega Pierre Méchain acrescentou mais seis em 1786. Messier era um caçador de cometas (os cometas foram uma obsessão dos astrónomos do século XVIII) e compilou um catálogo famoso «[...] para que os astrónomos não confundam estas mesmas nebulosas com os cometas que começam a brilhar». Posteriormente, William Herschel publicou um catálogo de milhares de nebulosas, trabalho que foi continuado pelo filho John, o qual publicou em 1864 o *General Catalogue*, contendo uma lista de 5000 nebulosas e aglomerados estelares. John Herschel passou vários anos na cidade do Cabo, observando os céus do hemisfério sul, especialmente as nuvens de Magalhães, acerca das quais observou que «as nebulosas devem ser encaradas como sistemas *sui generis*, que não têm similares no nosso hemisfério [norte]». Hoje olhamos as nuvens de Magalhães como galáxias independentes.

Ampliando o trabalho dos Herschels, J. L. E. Dreyer publicou em 1890 o *New General Catalogue*; o sistema de classificação destes trabalhos ainda hoje é usado. Por exemplo, a galáxia Andrómeda é classificada como M 31 (M para significar Messier) e um dos seus satélites é NGC 205 (NGC = New General Catalogue). Algumas das nebulosas classificadas por Messier e por Dreyer foram identificadas como sendo galáxias muito afastadas da nossa. Todavia, nem todas as nebulosas são galáxias exteriores, tendo o facto de existirem espécies muito diferentes de nebulosas causado grande confusão. Sucede que muitas das nebulosas classificadas pelos primeiros astrónomos fazem, na realidade, parte da nossa galáxia — «os pontos quentes», regiões nos braços em espiral da nossa galáxia, em que gás e poeira se acumulam para formar novas estrelas, como as nebulosas de Oríon e de Trífido. Um grupo completamente diferente é o das «nebulosas planetárias» — envolventes gasosos das estrelas velhas que foram expelidos para o espaço, mas que se parecem com os discos dos planetas (daqui a designação incorreta), e objetos, como a nebulosa do Caranguejo, que são, na realidade, remanescentes de uma supernova a vaguear no espaço e que em breve se dispersarão no meio interestelar. Outros objetos astronómicos indistintos, originalmente classificados como nebulosas, acabaram por ser identificados como sendo os aglomerados globulares de estrelas distribuídos pelo nosso halo galático. Com tantos objetos astronómicos diferentes parecendo nebulosas, não é de estranhar que até há pouco persistisse tanta confusão quanto ao que eram ou não eram.

Na década de 1840 William Parsons, 3º conde de Rosse, abandonou o Parlamento para se dedicar à sua paixão pela

astrologia e usou a fortuna pessoal para construir um enorme telescópio de reflexão. Este «leviatão» excedia as possibilidades tecnológicas daquele tempo e, embora funcionasse, não era tão útil como os telescópios mais pequenos. Contudo, Rosse, apontando o monstro para os céus, observou as nebulosas com grande pormenor e foi o primeiro a verificar que muitas delas tinham uma estrutura em espiral. Os seus apontamentos científicos estão cheios de desenhos de espirais. Mas não tinha razões para acreditar que se situassem fora do nosso sistema estelar. De facto, até ao presente século não havia razões para que alguém acreditasse que o nosso sistema estelar tivesse um «lado de fora».

Depois das descobertas de Rosse a fotografia tornou-se cada vez mais importante na exploração das nebulosas. James Edward Keeler, um americano do Observatório Lick, na Califórnia, continuou os trabalhos de fotografia de nebulosas realizados pelo astrónomo inglês Isaac Roberts, que foi o primeiro a mostrar, a partir de fotografias que resolviam os bordos da galáxia de Andrómeda, que ela era semelhante a outras nebulosas em espiral. Keeler fotografou mais de 100.000 galáxias, a maioria delas em espiral. O universo estava ricamente povoado destes estranhos e maravilhosos objetos.

O começo deste século marcou um ponto de viragem na tecnologia astronómica. Os maiores telescópios de refração viáveis (que usam grandes lentes como elemento de focagem) foram construídos nos Observatórios Lick e Yerkes no fim do século XIX. Para que a astronomia observacional progredisse tornava-se necessário que os novos telescópios fossem refletores, com um

único e imenso espelho como elemento principal de recolha de luz e de focagem. A tecnologia para construir tais espelhos começava agora a despontar e os Estados Unidos, que viriam a dominar a astronomia observacional nos 50 anos seguintes, começaram a tomar a dianteira sobre a Europa.

Um único indivíduo notável, George Ellery Hale, talentoso investigador do MIT[3], tornou-se a força impulsionadora da nova astronomia americana. Ainda jovem, Hale, aparentemente livre das contradições quanto aos objetivos a atingir que afligem a maior parte dos jovens, impressionou os superiores com a sua maturidade. Pediu dinheiro aos americanos ricos para a construção de novos telescópios, selecionou os locais para a instalação, supervisionou a construção e encorajou os melhores observadores a utilizá-los. Promoveu também a integridade da astrofísica como disciplina científica de direito próprio, distinta da astronomia e da física.

Sob a direção de Hale, apareceram telescópios que permitiram revelar a verdadeira identidade das nebulosas. Por que razão era tão difícil compreender que as nebulosas em espiral eram exteriores ao nosso sistema estelar? Primeiro, porque estavam extremamente distantes, tão distantes que não puderam ser vistas como um sistema de estrelas individualizadas (mas sim como gás) até à instalação do telescópio de reflexão *Hooker* de 2,5m de Mount Wilson, construído por Hale em 1919. Segundo, porque as nebulosas em espiral não estavam uniformemente distribuídas no céu. Existia uma «zona de ausência» — uma parte dos céus na qual as espirais não apareciam. Sabemos hoje que essa zona

corresponde ao disco da nossa galáxia e que a poeira no disco obscurece qualquer espiral distante. Mas nessa altura o facto de não se verem espirais na região da Via Láctea pareceu a muitos astrónomos uma indicação de que as nebulosas em espiral estavam de algum modo correlacionadas com a estrutura do nosso sistema estelar, não lhe sendo exteriores.

Estes enigmas relativos às nebulosas em espiral e à forma do nosso sistema estelar só foram desvendados na década de 20. O astrónomo americano Harlow Shapley contribuiu, talvez mais do que qualquer outro indivíduo, para a determinação da estrutura correta da nossa galáxia: um disco de estrelas, com o nosso sol junto do bordo, envolvido pelo halo dos aglomerados globulares. Como chegou Shapley a esta notável imagem da galáxia, que, na altura em que a propôs, parecia tão peculiar?

A pista que Shapley seguiu para decifrar o enigma da forma da nossa galáxia encontra-se na distribuição dos aglomerados globulares — essas esferas de estrelas espalhadas dentro do halo. Antes de Shapley (trabalho realizado em 1918) os astrónomos sabiam que a distribuição destes aglomerados de estrelas não é simétrica — a maior parte concentra-se num dos hemisférios da esfera celeste. Cerca de um terço dos aglomerados globulares encontra-se na constelação do Sagitário, que ocupa apenas 2% do céu. Supondo que o centro da nossa galáxia coincidia com o centro da distribuição dos aglomerados globulares, Shapley mostrou, usando o novo telescópio de Mount Wilson, que o centro da nossa galáxia se encontra na direção de Sagitário. Utilizando a relação período-luminosidade de Leavitt para as cefeides nos aglomerados

globulares, calculou que a distância para o centro da galáxia era realmente muito grande. O ponto de vista de Shapley quanto à geometria da galáxia recebeu confirmação posterior pelo astrónomo holandês Jan Oort, que mostrou que as estrelas se movem em torno do centro galático.

Se Shapley estiver certo, estamos no bordo da nossa galáxia. Contudo, se for assim, então outras questões se levantam. Por que razão o centro da galáxia, com aquela concentração de estrelas, não brilha mais intensamente na constelação de Sagitário? Por que razão as estrelas que vemos no céu noturno estão distribuídas tão uniformemente? A resposta a estas interrogações é que a nossa galáxia — tal como todas as galáxias em espiral — contém grandes quantidades de gás e poeiras que obscurecem a luz brilhante das estrelas do centro da galáxia. De facto, o gás e as poeiras impedem a passagem da luz vinda de todas as estrelas, exceto daquelas (milhares) que estão na vizinhança do nosso sol e podem ser vistas a olho nu, as quais apresentam uma distribuição uniforme no céu.

Durante muito tempo a poeira na nossa galáxia criou grandes dificuldades de observação. A poeira galática não só obscurece as estrelas distantes, como também avermelha a respetiva luz. Este avermelhar tornava difícil aos astrónomos identificar a cor correta da estrela, uma informação que é útil para estimar a distância a que esta se encontra de nós. Só em 1930, depois das cuidadosas observações de Robert Trümpler, astrónomo suíço do Observatório Lick, é que a complexidade associada à poeira e ao gás e o problema de determinar a geometria da nossa galáxia foram, finalmente, esclarecidos. Estudando grupos de estrelas em aglomerados

abertos, Trümpler demonstrou que na nossa galáxia a intensidade da luz das estrelas se reduz a metade por cada 3000 anos-luz percorridos dentro do disco. Utilizando este e outros factos, os astrónomos puderam «corrigir» as suas observações das estrelas distantes e estabelecer escalas de distâncias precisas.

Embora Shapley tivesse razão acerca da estrutura da nossa galáxia, a Via Láctea, e defendesse vigorosamente a sua posição, insistia, no entanto, ironicamente, em que as nebulosas em espiral eram parte da nossa galáxia, e não exteriores a ela. Walter Baade argumentou contra a visão de Shapley em 1920, citando as provas fornecidas por uma fotografia de longa exposição, tirada com um telescópio de 150cm, que parecia resolver a espiral M 33 em estrelas separadas. A ser verdade, tal significaria que a M 33 estava muito longe e não era parte da nossa galáxia. Shapley pensava que as imagens na fotografia não eram estrelas, mas sim gás. As divergências entre Shapley e alguns colegas atingiram o auge num famoso debate travado com o astrónomo Heber Doust Curtis na Academia Nacional das Ciências em 1922.

Os argumentos de Shapley dependiam de algumas medições muito precisas da rotação das nebulosas em espiral feitas pelo astrónomo Adrian van Maanen. Se as espirais estivessem a rodar tão rapidamente como as medições de van Maanen indicavam e se se admitisse que estavam muito longe (como o oponente de Shapley afirmava), então as estrelas na espiral teriam de estar a mover-se a velocidades comparáveis à velocidade da luz, se não maiores até. Isto era impossível, porque nada se move mais rapidamente do que a luz. Logo, ou as espirais eram objetos

próximos, como Shapley afirmava, ou as medições de van Maanen estavam muito erradas. Apesar do extremo cuidado que van Maanen teve, as suas medições da rotação das galáxias estavam no limite da tecnologia daquele tempo e provou-se estarem erradas. As galáxias em espiral rodam realmente, mas muito mais lentamente do que van Maanen calculara. Assim, Shapley estava também errado, embora só o viesse a saber muito mais tarde.

Curtis, o seu oponente no debate, defendia o ponto de vista contrário, isto é, que as espirais estavam muito longe e representavam galáxias distintas, ponto de vista que hoje aceitamos.

Centrou os argumentos na M 31, a nebulosa da Andrómeda, e argumentou que não podia ser uma nuvem de gás na nossa galáxia. Todavia, as provas de Curtis eram, na realidade, um pouco fracas. Fez notar que existia um número considerável de novas — presumivelmente explosões de estrelas — na nebulosa da Andrómeda, ainda que a luminosidade observada, ao contrário da das novas vulgares, fosse bastante variável. A prova de que estas eram verdadeiras novas, como as observadas na nossa galáxia, não era, segundo Shapley, convincente; não havia razão para acreditar que a nebulosa da Andrómeda não fosse mais do que uma nuvem de gás local.

O debate relativo à natureza extragalática das nebulosas em espiral apenas foi resolvido através do trabalho cuidadoso de Edwin Hubble e de vários outros que se dedicaram ao estudo das galáxias. Hubble seguiu a pista do astrónomo John C. Duncan, que, ao

procurar novas na galáxia de Andrómeda, localizara uma estrela variável e cintilante que era, provavelmente, uma cefeide. Encontrou mais 40 cefeides em Andrómeda e muitas mais nas outras galáxias. Pela medição da taxa de cintilação, estabeleceu a luminosidade absoluta e, comparando-a com a luminosidade aparente, pôde determinar a distância a que se encontravam, uma distância tão grande que as galáxias estavam decerto muito para além do sistema da Via Láctea. Nas palavras do célebre astrónomo Allan Sandage, as observações de Hubble «provaram, sem margem para qualquer dúvida, que as nebulosas eram galáxias exteriores, com dimensões comparáveis às da nossa, o que abriu a última fronteira da astronomia e estabeleceu pela primeira vez o correto valor concetual do universo. As galáxias são a unidade da matéria que define a estrutura granular do universo.» Depois do trabalho de Hubble os outros astrónomos rapidamente se convenceram de que estas nebulosas eram gigantescos e longínquos sistemas-ilhas de estrelas semelhantes ao nosso.

Edwin Hubble viveu no reino das galáxias. Como muitos outros grandes astrónomos, começou por trabalhar numa disciplina sem qualquer relação com a astronomia: foi estudante de Direito Romano e de Inglês na Universidade de Oxford. Após um ano de exercício, decidiu «trocar as leis pela astronomia». Veio a utilizar o telescópio refletor Hooker de 2,5m instalado em Mount Wilson, na Califórnia — um excelente par homem-máquina —, onde realizou o seu grande trabalho. Como tantos astrónomos observacionais e cientistas experimentais, desdenhava o pensamento especulativo, preferindo, em seu lugar, um ceticismo crítico. Ele comentou:

> Os investigadores tentam satisfazer a curiosidade e estão acostumados a utilizar qualquer meio razoável que possa ajudá-los a atingir o fugidio objetivo. Uma das poucas características universais é um ceticismo saudável relativamente às especulações não verificadas. Estas são olhadas como assuntos de conversa até que possam ser concebidos testes de verificação. Só então atingem a dignidade de tópicos para investigação.

Através deste trabalho e do dos seus colaboradores, as galáxias, outrora objeto de meras «especulações não verificadas», atingem agora «a dignidade de tópicos para investigação».

Nos últimos anos da década de 20 tornou-se claro que o universo estava organizado em galáxias de grande variedade de tamanhos e formas, cada uma formada por um número de estrelas que podia variar dos milhares de milhões aos biliões. Tanto quanto se podia ver, entre as galáxias encontrava-se o negro espaço vazio. As galáxias eram ilhas de luz num mar de escuridão vasto e sem fim. A razão pela qual o universo se organizou desta maneira, em vez de se constituir numa população uniforme de estrelas, permanece um mistério até aos nossos dias, mistério que, uma vez desvendado, fará luz sobre o problema da origem do universo, dado ser claro para os astrónomos de hoje que as galáxias são quase tão velhas como o próprio universo.

O trabalho de Hubble sobre as galáxias conduziu-o mais tarde a dar um importante contributo para a compreensão da expansão do universo, problema sobre o qual muitos tinham especulado antes dele. Tal como outras grandes descobertas, a da expansão do universo não se fez toda de uma só vez. O caminho foi preparado

durante os anos que vão de 1912 a 1923 pelo astrónomo americano Vesto Slipher, que levou a efeito cuidadosas medições da mudança de cor da luz emitida pelas galáxias próximas. Este astrónomo descobriu que a maior parte das galáxias tinha luz que se desviava em direção ao vermelho.

Sabemos que a luz emitida por um objeto que se afasta de nós se desvia para frequências mais baixas, que correspondem à cor vermelha, tal como o apito de um comboio parece mais grave quando este se afasta — o chamado «efeito Doppler». Quer isto dizer que a interpretação mais simples do «desvio para o vermelho» de Slipher era que a maior parte das galáxias se afasta de nós — estranha conclusão, se estivéssemos convencidos de que as galáxias se movem ao acaso no espaço do universo. Contudo, as observações de Slipher ajustavam-se lindamente a um modelo cosmológico de 1917, baseado na nova teoria da relatividade geral de Einstein, inventado pelo astrónomo holandês Willem de Sitter, que admitia a possibilidade de o espaço do universo se expandir, levando consigo as galáxias, de maneira que todas as galáxias se afastavam umas das outras. Durante algum tempo as observações de Slipher foram conhecidas por «efeito de Sitter». Carl Wirtz, astrónomo alemão, usou este efeito de Sitter em 1923. Combinando-o com estimativas grosseiras das distâncias às galáxias baseadas nas dimensões aparentes destas, propôs uma lei, relacionando a velocidade com a distância, segundo a qual a velocidade de uma galáxia (que pode ser determinada a partir do desvio para o vermelho) é proporcional à distância que nos separa dela, antecipando, assim, a lei de Hubble.

Hubble aplicou as suas grandes capacidades de observador e medidor do universo ao problema do desvio para o vermelho das galáxias. Através de meticulosas medições das distâncias, utilizando as cefeides e a extensão que M. L. Humason fez do trabalho de Slipher acerca do desvio para o vermelho, pôde demonstrar claramente que, quanto mais distante se encontra a galáxia, tanto maior é a sua velocidade de afastamento, relação conhecida por «lei de Hubble». Desde os tempos de Hubble os astrónomos alargaram as observações a objetos que se encontram a distâncias mais de dez vezes maiores, continuando sempre a verificar-se a lei de Hubble, dentro dos limites dos erros de medição.

A lei de Hubble implica que o universo se expande uniformemente. Que quer isto dizer? Dado verificar-se que todas as galáxias distantes se afastam de nós, poderia pensar-se que a nossa galáxia está, de certo modo, no meio da expansão e que ocupamos o centro de uma explosão cósmica. Isto é um mal-entendido. A propriedade de uma explosão uniforme, como a que a lei de Hubble implica, é que, seja qual for a galáxia, todas as outras se afastam dela; não temos uma posição privilegiada.

Outro mal-entendido frequente acerca da expansão do universo é que as galáxias se movem no espaço da mesma maneira que um peixe nada na água de um rio. A interpretação dada à expansão do universo pela teoria da relatividade geral de Einstein é que as galáxias se movem com o espaço, da mesma maneira que um pedaço de madeira se move com a corrente de um rio. Pode imaginar-se que é o próprio espaço do universo que se expande

como uma folha de borracha a alongar-se; as galáxias acompanham este movimento.

Segundo Einstein, o espaço não é uma entidade fixa e imutável, um conceito que existia nas nossas cabeças. A geometria e as propriedades do espaço físico devem ser determinadas instrumentalmente por aquilo que podemos medir. Mas como pode alguém medir a geometria do espaço do universo inteiro? Temos de usar as próprias galáxias como «marcos» que definem esse espaço. Se o fizermos, descobriremos que o próprio espaço deve estar em expansão, uma vez que as galáxias se afastam. A lei de Hubble, de grande significado cósmico, é, assim, a primeira lei empírica da estrutura do universo como um todo. Todos os modelos matemáticos ulteriores do universo deverão ter em conta esta lei.

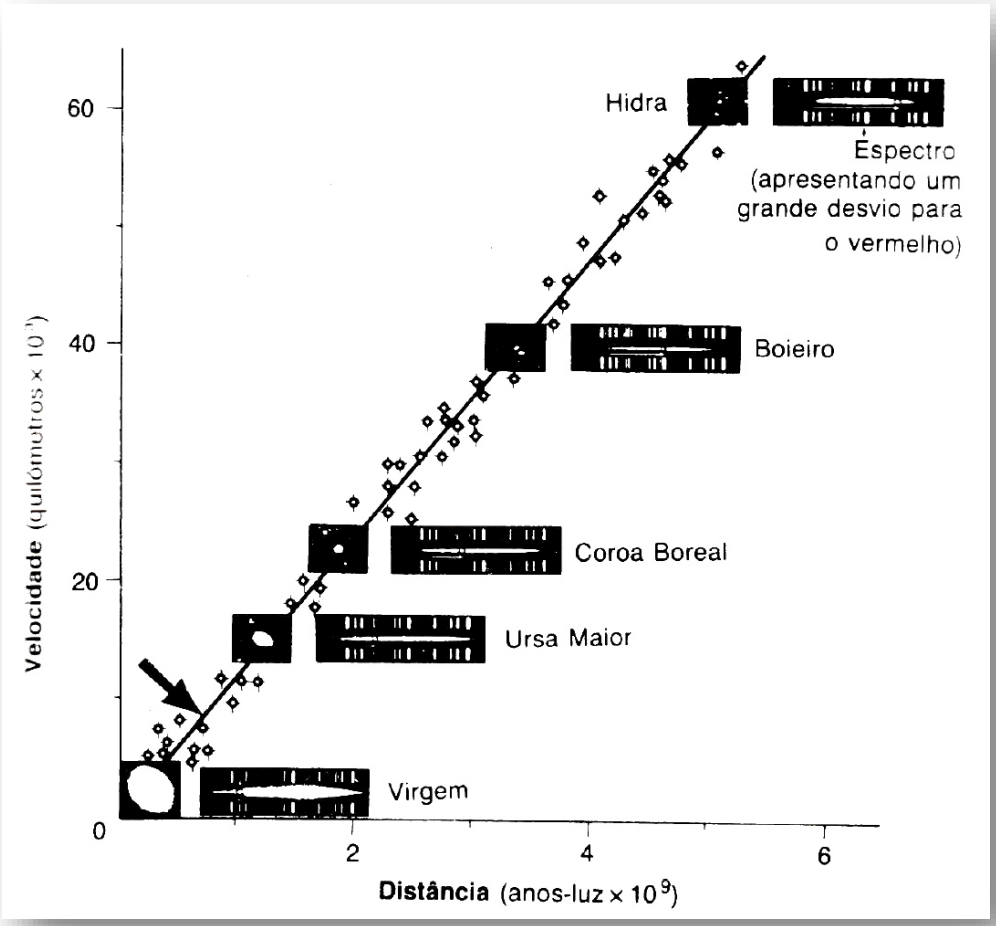
Hubble dedicou muitos anos da sua vida à classificação das galáxias, definindo com precisão os diferentes tipos e formas. As galáxias podem agrupar-se em dois grandes grupos: elípticas e em espiral. As galáxias elípticas têm formas que vão de arranjos quase esféricos de estrelas até discos achatados. Como exemplos de galáxias elípticas podem citar-se as duas galáxias satélites da galáxia de Andrómeda. As galáxias elípticas têm uma grande variedade de massas e de tamanhos e, curiosamente, quase não contêm gás ou poeiras, mas apenas estrelas. Compreendem mais de 60% das galáxias e a maior parte delas são pequenas galáxias «anãs».

Pelo contrário, as galáxias em espiral são discos de estrelas com um bojo central e um halo, cuja massa visível varia entre 10 e

algumas centenas de milhares de milhões de massas solares, albergando grandes quantidades de gás e poeiras. Estão subdivididas em espirais normais (como a nossa galáxia e a Andrómeda) e espirais em forma de barra, apresentando duas barras ou jatos que emergem da região central, as quais terminam em braços em espiral (não existem espirais em forma de barra no nosso grupo local, mas cerca de um terço de todas as espirais têm forma de barra). Tanto as espirais normais como as espirais em forma de barra podem ser classificadas consoante o grau de estreitamento dos braços em torno do bojo central.

O sistema de classificação de Hubble é representado pelo famoso «diagrama de diapasão». Na haste do diagrama está a galáxia lenticular (em forma de lente), como uma espiral, mas sem os braços nem o gás. Nenhuma galáxia lenticular tinha sido ainda observada quando Hubble propôs o seu sistema de classificação, mas posteriormente muitas foram encontradas.

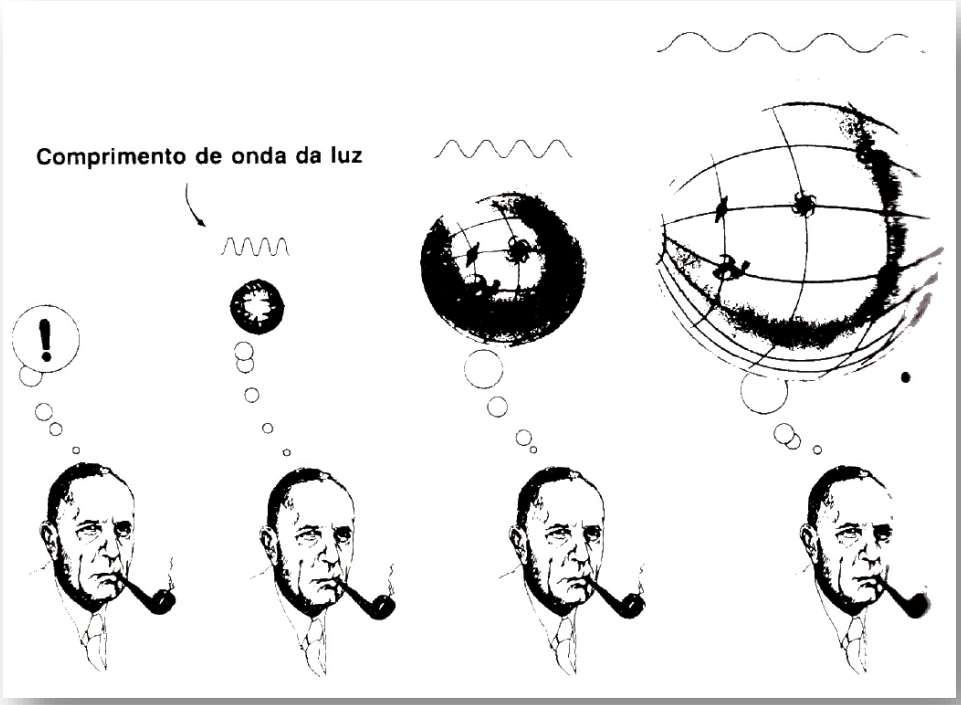
Algumas galáxias não encaixavam neste esquema — as galáxias irregulares, como as nuvens de Magalhães perto da nossa Via Láctea, a que Walter Baade chamou «o caixote do lixo do sistema de Hubble». As galáxias irregulares encontram-se normalmente junto de uma galáxia vulgar. Talvez tenham sido elas próprias galáxias vulgares, cuja forma foi distorcida pelas interações gravitacionais de maré com as suas vizinhas. Baade, que usou o sistema de classificação de Hubble durante 30 anos, observou que, «embora tenha procurado obstinadamente sistemas que não encaixassem nele, o número [...] é tão pequeno que posso contá-los pelos dedos da minha mão».



Os desvios para o vermelho galático (proporcional à velocidade da galáxia) em função da distância que nos separa da galáxia. A relação linear entre estas grandezas é conhecida por lei de Hubble. Mostra-se o tamanho aparente da galáxia, bem como as linhas espectrais de certas galáxias, com indicação do desvio para o vermelho. A seta indica até onde foram as medições de Hubble no estudo original. Desde então a lei tem-se conservado válida (dentro dos erros das medições) até distâncias 10 vezes maiores do que as investigadas originalmente. Os pontos do gráfico são dados experimentais autênticos.

Qual é, porém, o significado do sistema de classificação das galáxias? Muitos dos astrónomos, incluindo Hubble, que primeiro refletiram sobre esta questão pensavam que os diferentes tipos de galáxias representavam diferentes estádios da evolução galática. Julgava-se que as galáxias evoluíam de elípticas para espiraladas — da esquerda para a direita no diagrama de diapasão. Houve alturas em que alguns astrónomos inverteram esta suposição e

pensaram que a evolução ia das espiraladas para as elípticas — da direita para a esquerda no diagrama.



Edwin Hubble contempla o universo em expansão. A lei de Hubble trouxe grande credibilidade às especulações anteriores dos astrónomos e cosmólogos segundo as quais o espaço do universo estaria realmente em expansão. Aqui está ilustrada uma analogia bidimensional do universo FWR fechado. À medida que o espaço se expande, as galáxias acompanham a expansão geral — o «deslocamento de Hubble». O comprimento de onda de uma onda de luz deixada pelo *big bang* aumenta também à medida que o espaço se expande, o que implica que tal radiação perde energia no universo e, portanto, a temperatura do banho de radiação luminosa também diminui. A temperatura da radiação de fundo é, atualmente, de uns meros 2,7K.

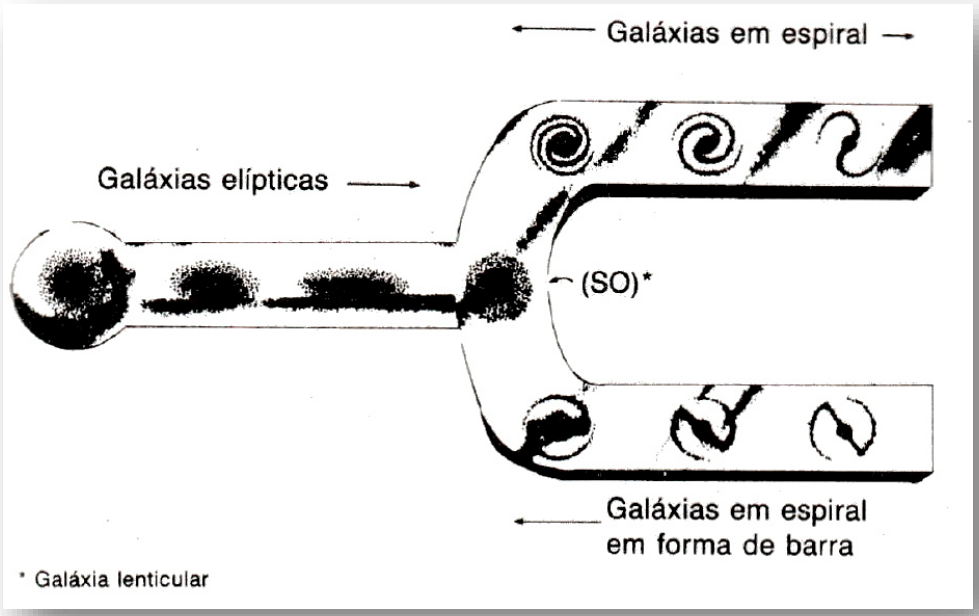


Diagrama de diapasão de Hubble para a classificação das galáxias. Na cabeça do diapasão estão as galáxias elípticas, classificadas segundo a respetiva excentricidade. Na junção da forquilha com o suporte vê-se a SO, ou *galáxia lenticular*. Os dois dentes da forquilha são as galáxias em espiral e as galáxias em espiral em forma de barra. Os astrónomos chegaram a pensar que esta classificação representava uma sequência evolutiva, mas hoje creem que as diferentes espécies de galáxias são uma consequência dos diferentes ambientes e condições em que se deu a sua formação.

Atualmente, a maior parte dos astrónomos acreditam que o sistema de classificação tem pouco a ver com a evolução galática e que todas as galáxias começaram a existir aproximadamente ao mesmo tempo. Sabe-se hoje que todas as galáxias contêm estrelas com cerca de 10 mil milhões de anos — forte indício de que todas as galáxias têm, pelo menos, esta idade, portanto, aproximadamente, a mesma idade que o universo. Do ponto de vista moderno, os diferentes tipos de galáxias são análogos às diferentes raças humanas, que também não constituem uma sequência evolutiva, representando, em vez disso, diferentes respostas a ambientes físicos diversos. As diferentes espécies de galáxias da classificação de Hubble refletem, provavelmente, diferenças na massa total e nas velocidades de rotação das galáxias

individuais — condições físicas estabelecidas aquando da sua criação. Todavia, mesmo hoje ninguém sabe o porquê das diferentes formas das galáxias, nem por que têm aquelas dimensões, nem tão-pouco por que formam aglomerados e superaglomerados. Tais enigmas, bem como o enigma, com eles relacionado, da origem e evolução das galáxias, estão na fronteira da pesquisa atual; voltaremos a tratá-los num capítulo posterior.

A que distância estão as galáxias? Hubble pôde, através das observações do brilho pulsante das cefeides, determinar a distância às galáxias mais próximas, mas a distância às galáxias mais longínquas, nas quais já não podiam ver-se as brilhantes cefeides, exigia outros métodos. Um deles consiste em supor que as estrelas mais brilhantes das diferentes galáxias têm todas a mesma luminosidade. Usando-as como «velas padrão», os astrónomos ficaram a dispor de outra unidade de medida. Para verificar esta hipótese pode medir-se também o tamanho aparente de nuvens de gás muito luminosas ou o brilho das supernovas em galáxias distantes, de modo a obter o valor da distância por um processo independente. Tais métodos (todos eles dão, aproximadamente, os mesmos resultados) funcionam razoavelmente até distâncias de 100 milhões de anos-luz. Para lá desta distância os astrónomos são obrigados a usar o brilho aparente das próprias galáxias como «vela padrão», visto não ser possível resolver os corpos que as constituem. Mas este processo pode ser arriscado, pois, à medida que os astrónomos olham para lugares cada vez mais distantes no espaço, estão também a ver as galáxias como eram há muito tempo. Se as galáxias evoluem, então talvez a quantidade de luz emitida tenha sido maior no passado, e não igual à das galáxias

mais velhas que vemos perto. O método poderia levar-nos, assim, a estimativas erróneas das distâncias.

Um exemplo de quanto as nossas medições de distâncias astronómicas dependem das hipóteses acerca da luminosidade dos objetos distantes é a descoberta de Walter Baade, durante a Segunda Guerra Mundial, de duas populações de estrelas. Eis como o trabalho de Baade implicava ser o universo duas vezes maior do que anteriormente se pensava.

Depois do trabalho de Hubble sobre as cefeides variáveis nas galáxias próximas pensou-se estarem determinadas as distâncias até às galáxias vizinhas. Contudo, com base nestas distâncias, os astrónomos podiam também estabelecer as dimensões absolutas das outras galáxias. Parecia que a nossa galáxia era cerca de duas vezes maior do que as outras galáxias espiraladas — uma completa anomalia nesta classe, que desapareceu depois de Baade ter mostrado em 1952 que as cefeides usadas por Shapley para determinar a distância aos aglomerados globulares, portanto as dimensões da Via Láctea, eram as velhas estrelas da população 11, enquanto que as cefeides observadas por Hubble na galáxia de Andrómeda eram as jovens estrelas brilhantes da população 1 nos braços da espiral. As velhas e jovens cefeides das duas populações de estrelas eram fisicamente diferentes e, não surpreendentemente, tinham relações período-luminosidade diferentes. As distâncias às galáxias estimadas por Hubble, que não tinham em conta a distinção entre as duas populações de estrelas, revelaram-se demasiado pequenas. Depois do importante trabalho de Baade a escala de distâncias do universo foi multiplicada por um

fator de 2 e, em consequência disso, a nossa galáxia deixou de ser anormalmente grande, tornando-se uma galáxia espiralada de tamanho normal. Existe ordem nos céus.

É difícil estimar com confiança as distâncias às galáxias mais longínquas. No entanto, é extremamente importante conhecer essas distâncias para que possamos traçar o mapa da estrutura do universo em grande escala. Até que o façamos, o nosso mapa atual do universo, ou mesmo o mapa do interior da nossa própria galáxia, pode ser comparado aos mapas do Novo Mundo, ou da Ásia, desenhados pelos cartógrafos com base nos relatos feitos pelos primeiros navegadores europeus — uma mistura de observações e de palpites. O uso de novos instrumentos, como o telescópio espacial, possibilitará aos astrónomos elaborar mapas do cosmos com maior rigor.

Se quisermos estudar os pormenores da estrutura interna de uma galáxia, não existe melhor local para começar do que em casa, na nossa galáxia, a Via Láctea. De que é então feita a nossa galáxia e, por inerência, de que são feitas as outras galáxias espiraladas como ela? As duas principais componentes visíveis são as estrelas e o meio interestelar, que, embora constitua apenas 5% da massa das estrelas, é muito importante na determinação da dinâmica da galáxia. O meio interestelar compõe-se de gás, poeiras, raios cósmicos (partículas quânticas de alta energia) e campos magnéticos. Todas estas componentes — adicionadas talvez a outras componentes desconhecidas, como a matéria escura no halo — constituem a nossa galáxia e interagem de maneira complicada e ainda não completamente entendida. A galáxia é um sistema rico,

em permanente transformação química, evoluindo dinamicamente, tão complexo como o clima terrestre. Examinemos estas componentes e vejamos o que sabem os astrónomos.

As estrelas, que são as componentes mais proeminentes, podem ser classificadas segundo o seu tipo espectral, temperatura de superfície, luminosidade, massa, composição química e idade — uma variedade de características que estão inter-relacionadas e são razoavelmente bem compreendidas. Na nossa galáxia as estrelas estão densamente concentradas junto do núcleo e fracamente concentradas no disco. As estrelas mais velhas encontram-se no disco, enquanto as mais novas, como o nosso sol, tendem a permanecer junto do plano médio da galáxia. Todas as estrelas do disco se movem em trajetórias complicadas, ao mesmo tempo que oscilam para cima e para baixo do plano médio, rodando num determinado sentido em torno do centro. Os astrónomos estimam que a velocidade de rotação do nosso sol em torno do centro da galáxia deve ser entre 200km e 300km por segundo. Algumas estrelas de alta velocidade «voam» bastante para fora do disco e passam a maior parte do tempo no halo. Sabe-se que estas estrelas solitárias do halo são bastante velhas, tendo, provavelmente, escapado dos aglomerados globulares que povoam o halo, ou talvez sejam as estrelas originais a partir das quais se formaram os próprios aglomerados globulares.

O primeiro indício sólido da existência de um meio interestelar surgiu nos anos 20. Os astrónomos foram capazes de observar o meio graças à sua absorção característica da luz das estrelas e concluíram que havia nuvens de átomos de cálcio ionizado e de

átomos neutros de sódio flutuando entre as estrelas. Todavia, pensaram que, sendo as estrelas feitas principalmente de hidrogénio, se existisse um gás interestelar, deveria conter também grandes quantidades de hidrogénio gasoso neutro. O problema de verificar esta hipótese residia no facto de, ao contrário dos átomos de cálcio e de sódio, que absorvem a luz, mesmo frios, os átomos de hidrogénio só emitirem luz visível quando a altas temperaturas. Uma vez que os átomos de hidrogénio nas nuvens de gás interestelar estariam realmente a uma temperatura muito baixa, nunca poderiam ser vistos, por muito abundantes que fossem.

O processo de observar as nuvens de hidrogénio frio foi descoberto pelo astrónomo holandês H. C. van de Hulst durante a Segunda Guerra Mundial. Jan Oort, director do observatório de Leiden, tinha lido os artigos do americano Grote Seber, radioastrónomo pioneiro, e comunicou-os ao seu grupo em 1944. Oort salientou a importância da deteção do hidrogénio neutro, e van de Hulst pegou neste assunto. Sabia que o hidrogénio neutro consiste num único eletrão em órbita em torno de um único protão. O protão e o eletrão têm ambos *spin* ½ — podemos imaginá-los como pequenos piões em rotação —, pelo que o hidrogénio neutro e frio pode ter o eletrão e o protão a rodar no mesmo sentido ou em sentidos opostos. Existe uma pequena diferença de energia entre estas duas configurações de *spin*: e, quando o átomo de hidrogénio descreve um salto quântico do estado de mais alta para o de mais baixa energia, emite radiação. Pode calcular-se o comprimento de onda desta radiação, que é de 21cm, precisamente o comprimento suscetível de ser detetado por um radiotelescópio convenientemente projetado.

Quando, depois da guerra, os astrónomos construíram radiotelescópios, foram vistas pela primeira vez as enormes nuvens de hidrogénio neutro no disco, especialmente nos braços da galáxia. A radiação de 21cm abriu um novo caminho para fazer o mapa da galáxia, confirmando a conclusão anterior de que a galáxia era uma espiral. Usando os novos satélites de raios X, bem como os telescópios de infravermelhos e de rádio, os astrónomos tiveram nos últimos 15 anos a surpresa de descobrir que o meio interestelar era composto, não por nuvens de forma grosseiramente esférica e espalhadas entre as estrelas, como anteriormente tinham pensado, mas sim por enormes «cascas» concêntricas de gás em expansão, que podem ter sido produzidas por um fortíssimo vento estelar proveniente de uma estrela quente ou de uma supernova. Os radioastrónomos encontraram também indícios de bastante gás fora do plano da galáxia, formando uma coroa galática.

Além do gás, a nossa galáxia contém ainda poeiras — pequenas partículas de matéria com talvez um centésimo de milésimo de centímetro de tamanho. Não existe consenso no que toca à composição química desta poeira granular: podem ser silicatos, grafite ou cristais de gelo sujos. A poeira absorve a luz das estrelas (o azul mais do que o vermelho), de modo que estas, vistas através da poeira, parecem mais ténues. É importante conhecer as propriedades de absorção das poeiras se se pretender medir distâncias a partes diferentes da galáxia.

Além de absorverem a luz das estrelas, os grãos de poeira também a dispersam. Quando se observa a passagem da luz através da poeira, verifica-se que fica muitas vezes ligeiramente

polarizada — as oscilações da onda dão-se segundo orientações preferenciais, e não aleatórias, o que significa que os grãos de poeira alongados que absorvem ou dispersam a luz devem também estar orientados. A explicação mais plausível para esta orientação uniforme é a existência de um campo magnético galático, penetrando as nuvens de poeira. Como se fossem pequenas agulhas de bússola em número incontável, os grãos alinham-se segundo o campo magnético galático. Foi a partir da luz polarizada das estrelas que os astrónomos aprenderam que o meio estelar possui um campo magnético.

Provas adicionais a favor dos campos magnéticos galáticos provêm da observação dos raios cósmicos — partículas elementares de alta energia que atingem a Terra, vindas uniformemente de todas as direções do espaço. Estas partículas cósmicas podem ser produzidas a seguir a uma explosão de supernova ou, como parece mais provável, a partir da aceleração de material interestelar (as abundâncias relativas de elementos nos raios cósmicos estão de acordo com as do meio estelar). Mas, tendo as partículas cósmicas energia tão elevada, deveriam «voar» para fora das galáxias após terem sido produzidas e só deveríamos poder observar muito poucas. Em vez disso, detetamos grandes quantidades de raios cósmicos. Qualquer coisa deve suster estas partículas de alta energia no disco da galáxia, sendo a única explicação possível a existência de um campo magnético galático, uma vez que os campos magnéticos podem «engarrafar» partículas eletricamente carregadas e é isto exatamente o que o campo magnético galático faz.

Além de traçar mapas das nuvens de hidrogénio interestelar, os radioastrónomos descobriram, juntamente com os astrónomos de infravermelhos, a existência de moléculas interestelares que emitem radiação próxima dos famosos 21cm de comprimento de onda do hidrogénio. Estas descobertas assinalaram a maturidade dos novos métodos de radioespetroscopia na parte final da década de 60 e na década de 70, métodos estes que permitiram estudar as riscas espectrais de rádio, que revelam a presença de moléculas exatamente da mesma maneira que as riscas do espectro visível indicam a presença de átomos. Deste modo, foram descobertas, entre outras, as moléculas de monóxido de carbono e de aldeído fórmico. Contudo, os astrónomos ficaram completamente surpreendidos com a descoberta, em nuvens densas, de grandes moléculas orgânicas, como as de álcool etílico e de cianoacetileno. Que faziam nas profundezas do espaço essas pouco mais de 50 moléculas, diferentes e tão grandes? As moléculas grandes formam-se a partir de outras mais simples e só se conservam em condições bastante especiais de temperatura, presentes nas quentes nuvens de gás em contração que formam os complexos gigantes de nuvens moleculares descobertas na década de 70 e que são quase certamente as precursoras de novas estrelas. Assim, a existência destas grandes moléculas forneceu mais algumas pistas para o conhecimento do complexo processo de formação das estrelas.

As estrelas nascem nas densas nuvens de gás que se encontram nos braços em espiral das galáxias. Mas qual é a origem de tais braços? Se os braços em espiral tivessem sido sempre formados pelas mesmas estrelas rodando juntas no disco, as estrelas junto

do centro rodariam mais rapidamente do que as estrelas de fora. Se esta imagem estivesse correta, então os braços em espiral enrolar-se-iam em torno da galáxia dentro do tempo que levam a executar algumas rotações galáticas, isto é, algumas centenas de milhões de anos. Mas não é isto o que se observa — a forma dos braços mantém-se inalterada à medida que a galáxia roda.

Para resolver este «enigma do enrolamento» o astrónomo sueco Bertil Lindblad sugeriu em 1941 uma nova teoria de ondas de densidade, que foi mais tarde desenvolvida, na década de 60, pelo matemático americano Chia Chiao Lin e seus colaboradores. De acordo com ela, uma onda de densidade de estrelas em forma de espiral é gravitacionalmente autossustentada. Novas estrelas e poeiras estão continuamente a entrar (e a sair) na onda de densidade em espiral, da mesma maneira que as partículas de água numa onda oceânica de forma constante estão continuamente a mudar. Embora esta teoria resolva o enigma do enrolamento, levanta outros problemas — não explica por que razão os braços não se desvanecem após algumas rotações da galáxia.

Numa teoria alternativa, desenvolvida por Alar Toomre, a onda de densidade em forma de espiral é gerada por complexas interações gravitacionais de maré com galáxias vizinhas. Demonstrou-se, através de simulações em computador de duas galáxias passando perto ou através uma da outra, que as interações de maré entre as galáxias podem dar origem à estrutura em espiral. Esta imagem sugere que o grande número de espirais que vemos hoje se deve a uma elevada taxa de interações dinâmicas entre as galáxias ao longo da sua história natural.

Outro modelo teórico dos braços da espiral, apresentado por Phillip Seiden, incorpora as novas ideias acerca da formação de estrelas nas nuvens de poeiras moleculares. Quando as estrelas de grande massa se formam nas nuvens densas, geram ondas de choque, que, por sua vez, fazem com que as poeiras se concentrem mais, engendrando a formação de mais estrelas.

À medida que as ondas de formação das estrelas se deslocam para a periferia, afastando-se do centro da galáxia, a rotação desta transforma-as nos braços em espiral. Quando introduzido num computador, este modelo de «formação estocástica das estrelas» gera imagens impressionantes, que parecem, exatamente, galáxias espiraladas reais.

Entretanto, nenhum destes modelos teóricos para explicar a origem dos braços em espiral é universalmente aceite. Passou já mais de um século desde que Lord Rosse viu pela primeira vez as muitas nebulosas em espiral, e, no entanto, como Alar Toomre recentemente resumiu, «parece hoje claro que a estrutura espiralada das galáxias é um problema complexo, sem resposta única e clara». As interações do campo magnético galático, raios cósmicos, poeiras, gás e estrelas, todas sujeitas à ação da gravidade, são muito complicadas. Porém, a dinâmica dessas interações, que está a ser deslindada pelos astrónomos de hoje, é a principal responsável pela estrutura da nossa galáxia e pelos processos que dão origem ao nascimento das estrelas.

Todavia, ao mesmo tempo que se esforçam por compreender os processos que observam na nossa galáxia, os astrónomos

continuam a embrenhar-se cada vez mais nas profundezas do espaço e a descobrir novos corpos celestes — radiogaláxias, quasares, talvez até buracos negros. Ultrapassaram em muito as especulações setecentistas acerca das nebulosas, embora permaneçam empolgantes interrogações de fundo — acerca da origem e evolução das galáxias, da sua dinâmica interna e processos geradores, da natureza do misterioso núcleo das galáxias —, mas o facto de estas interrogações provocantes permanecerem sem resposta garante que a astronomia galática permanecerá um campo aberto e atraente nas décadas vindouras.

As galáxias — a mais importante estrutura do nosso universo — foram descobertas na primeira metade deste século. Os instrumentos que dominaram este período foram os grandes telescópios ópticos, utilizados em conjugação com a moderna tecnologia fotográfica. Durante muito tempo parecia ser esta a única maneira de fazer astronomia. Mas durante e depois da Segunda Guerra Mundial os cientistas desenvolveram a tecnologia do radar e, consequentemente, os radiotelescópios (já em funcionamento antes da guerra), que abriram outra janela sobre o universo.

A guerra marcou um ponto de viragem nas ciências, resultando não só no aparecimento de novas técnicas astronómicas, como também numa reorganização social dos laboratórios científicos. Antes da guerra os tempos de observação nos grandes telescópios eram atribuídos pelo director do observatório, normalmente um ditador benevolente com um estilo muito próprio. Depois da guerra a tomada de decisões mudou para um método mais complexo e

democrático, um sistema abertamente competitivo, o que destruiu a vida sossegada no alto da montanha, com o telescópio. Como notou Jesse L. Greenstein, astrónomo de renome e membro do pessoal dos observatórios de Mount Wilson e de Mount Palomar desde 1948, «o mundo mudou. As personagens da peça mudaram. Em vez dos cavalheiros, passámos a ter o brilhante e agressivo jovem génio interessado em tudo, sem se importar com quem pisa, muito ansioso por fazer a descoberta do dia, da semana ou do que quer que seja. É uma perda e é um ganho.» Nos últimos anos da década de 40 as publicações dos astrónomos aumentaram, a carga administrativa dos grandes observatórios cresceu e o conhecimento humano do universo expandiu-se.

A consciência humana estendeu-se até às profundezas do espaço, ao núcleo das estrelas e ao princípio do tempo. Os cientistas fizeram perguntas novas, repetiram as antigas, perguntas estas que pela primeira vez adquiriram conteúdo empírico. No próximo capítulo continuaremos a exploração do reino das galáxias, especialmente das revelações dos radiotelescópios, instrumentos que estão a abrir um universo novo para a ciência astronómica.

---

[1] As edições mais recentes do *Guinness Book* atribuem o recorde ao próprio universo, com o nome de Friedmann. (*N. do T.*)

[2] Literalmente, Uma Teoria Original, ou Nova Hipótese do Universo. (N. do T.)

[3] Massachusetts Institute of Technology, ou Instituto de Tecnologia de Massachusetts. (N. do T.)

CAPÍTULO 5

Radiogaláxias e Quasares

Para sondar as profundezas do universo eram necessários instrumentos maiores e mais sensíveis. A construção de grandes radiotelescópios é uma história de visão científica e de luta, competição e controvérsia, persistência e diplomacia.

J. STANLEY HEY

George Ellery Hale dedicou os últimos anos da sua vida à construção do maior dos telescópios de reflexão, cujo coração era um imenso espelho de precisão com 5m de diâmetro, que concentrava a luz e podia resolver corpos celestes muito ténues com exposições de duração razoável. Um segundo telescópio *Schmidt*, um novo dispositivo com uma lente especial, colocada antes do espelho, que podia abarcar uma área enorme com distorção mínima junto dos bordos, trabalhava acoplado ao primeiro, constituindo, assim, o conjunto, um poderoso par de instrumentos para explorar os céus. Hale ainda viu o grande espelho de 5m chegar à Califórnia pouco antes da sua morte, em 1938.

Após a Segunda Guerra Mundial, o telescópio entrou em funcionamento no observatório de Mount Palomar, na Califórnia, e foi durante várias décadas o maior instrumento óptico do mundo, tendo sido chamado, com toda a propriedade, «o telescópio Hale». Hubble propôs-se usá-lo para continuar o trabalho de contagem das

galáxias e esperava conseguir os melhores resultados fazendo observações durante os períodos de ocultação da Lua. Numa tarde de 1948, em Pasadena, Califórnia, teve lugar, numa residência particular, uma reunião dos principais astrónomos para decidirem sobre a distribuição dos tempos de observação. Nela Hubble ouviu os colegas Richard Tollman, Walter Baade, Ira Bowen e outros explicarem-lhe que a contagem das galáxias não ia produzir qualquer informação útil acerca da expansão do universo. A proposta de Hubble, cujo prestígio e presença tanto tinham pesado na obtenção de fundos para a construção do telescópio, não seria concretizada. Embora profundamente desiludido, Hubble aceitou a decisão como um cavalheiro.

Enquanto os astrónomos observacionais, de posse do novo telescópio, davam grandes passos no sentido da exploração do universo, abria-se o novo campo da radioastronomia. Em 1931 Karl Jansky, dos Bell Telephone Laboratories, detetara pela primeira vez ruídos rádio, «este sussurrar vindo das profundezas do universo», como lhes chamou um locutor de rádio. Oito anos mais tarde o engenheiro de telecomunicações Grote Reber tornou-se o primeiro radioastrónomo, ao construir, recorrendo às suas poupanças, um refletor de rádio, em forma de disco, com cerca de 10m de diâmetro, nas traseiras da sua casa, em Wheaton, Illinois, um projeto que, nos seus aspetos gerais, fazia lembrar a construção do primeiro telescópio de reflexão de Herschel em Bath, Inglaterra. O radiotelescópio de Reber tinha suficiente capacidade de deteção direcional de sinais para lhe permitir determinar que os sinais mais fortes provinham do centro da Via Láctea e de duas fontes rádio nas

constelações do Cisne e de Cassiopeia, respetivamente, tendo ele traçado então o primeiro e grosseiro mapa rádio da nossa galáxia.

Os astrónomos, porém, prestaram pouca atenção às observações de Reber. Os seus interesses pareciam tão diferentes daqueles dos engenheiros de telecomunicações que as grandes potencialidades da radioastronomia só mais tarde foram compreendidas, especialmente nos Estados Unidos. Os engenheiros de telecomunicações estavam principalmente interessados nas ondas de rádio que podiam ser refletidas na atmosfera para transmissão de sinais a longa distância, e não nas ondas que Reber estudava e que penetravam na atmosfera, vindas de fora. Os astrónomos, por seu lado, não estavam nada interessados em ondas de rádio. Reber fazia notar em 1948 que, quando tentava obter apoio para a construção de um grande radiotelescópio, «[...] a atitude da maior parte das pessoas era considerarem-me inofensivo, pensarem que, se não mostrassem qualquer interesse pelo que dizia, me iria embora sem causar problemas».

Alguns astrónomos pensavam que a radioastronomia, devido ao facto de os comprimentos de onda da radiação envolvida serem muito grandes, não podia competir com o poder de resolução da astronomia observacional, que dispunha dos comprimentos de onda menores da luz visível. Mas não se aperceberam de que, se vários radiotelescópios fossem colocados bastante distanciados, a abertura efetiva e, consequentemente, o poder de resolução poderiam ser grandemente aumentados. Contudo, a razão principal pela qual estas primeiras iniciativas não eram seguidas devia-se ao facto de ninguém conseguir conceber mecanismos responsáveis

pela emissão dos sinais rádio que não fossem gás ionizado e este praticamente só existir em certos pontos da galáxia, não tendo, assim, interesse especial. Acabou, porém, por verificar-se que o universo é ricamente povoado por fontes radioastronómicas cujos mecanismos de emissão ninguém podia prever na década de 40.

Os grandes avanços na radioastronomia ocorreram depois da segunda Guerra Mundial, quando os cientistas ganharam experiência no uso da nova tecnologia eletrónica, de que é exemplo o radar. Uma nova estirpe de cientistas de telecomunicações, radar e astronomia envolveu-se então ativamente na construção de grandes «pratos» de rádio, «plantações de antenas» e equipamento eletrónico auxiliar necessário para desenvolver a nova astronomia. Embora cientistas como Jansky e Reber, nos Estados Unidos, tenham feito o trabalho pioneiro antes da guerra, foram ingleses, holandeses e australianos que, seguindo-lhes os passos, fizeram a maior parte das primeiras grandes descobertas. As condições meteorológicas quase sempre desfavoráveis da Inglaterra não afetavam as ondas de rádio, ao contrário do que se passava com a luz normal das estrelas. Na Universidade de Cambridge, então o centro da ciência na Inglaterra, J. A. Ratcliff, uma figura destacada na investigação de rádio e ionosférica, encorajava jovens engenheiros de telecomunicações com capacidade inventiva, como Martin Ryle, a começarem a trabalhar em radioastronomia. Ryle ligou primeiro vários pequenos «pratos» de rádio para obter a capacidade de resolução de um único, grande. Um outro grupo, dirigido por Bernard Lovell, em Manchester, construiu o radiotelescópio de Jodrell Bank — com 76,2m de diâmetro, era o primeiro dos grandes «pratos». Na Austrália dispositivos de rádio

de outro tipo, as chamadas «cruzes de Mill» (cada uma delas consistindo numa antena colocada no chão, com a forma de uma enorme cruz), perscrutavam os céus do Sul. A radioastronomia progredia rapidamente.

Embora os radioastrónomos tivessem tido algum sucesso na década de 40, a grande recompensa para os seus trabalhos chegou nos anos 1950-1951. As grandes descobertas que fizeram nesses dois anos marcaram um ponto de viragem e estabeleceram, finalmente, a radioastronomia como uma nova e importante disciplina científica, que facultava o conhecimento profundo da natureza física do universo.

Durante estes dois anos concluíram que a intensa fonte rádio Cassiopeia A era, na realidade, uma remanescente de uma supernova — um anel de gás radioemissor —, que assinalava o local onde uma estrela havia explodido por volta do ano 1700 a. C. e que Cisne A, outra fonte rádio, revelou ser uma galáxia dupla peculiar, cuja emissão rádio é um milhão de vezes mais intensa do que a de uma galáxia normal. Nestes anos surgiu também o primeiro mapa rádio de uma galáxia exterior à nossa, a de Andrómeda. Como se estas descobertas não fossem suficientes, três grupos de cientistas, um em Harvard, outro na Holanda e um terceiro na Austrália, detetaram as radiações de 21cm do hidrogénio neutro, já previstas por H. C. van de Hulst em 1945, descobrindo, assim, um novo método de fazer o mapa da nossa galáxia. Uma vez que os desvios Doppler nos comprimentos de onda de rádio — essencialmente os desvios para o vermelho e para o azul — das nuvens móveis de hidrogénio neutro podiam ser medidos ainda com maior precisão do

que os desvios Doppler da luz visível, os mapas rádio acabaram por fornecer os maiores pormenores possíveis sobre a estrutura em espiral da nossa galáxia.

Enquanto estavam a ser feitas estas descobertas, os físicos teóricos suecos H. Alfven e N. Herlofson sugeriram pela primeira vez em 1950 que a emissão de rádio de fontes discretas provinha da radiação de sincrotrão — uma classe de ondas eletromagnéticas de alta energia emitidas pelos eletrões ao girarem à volta das linhas de força de campos magnéticos —, fornecendo, assim, uma explicação para muitas das fontes rádio. Todas as partículas carregadas eletricamente irradiam ondas eletromagnéticas se estiverem animadas de movimento helicoidal — precisamente o que se passa com os energéticos eletrões, movendo-se rapidamente num campo magnético. De facto, hoje em dia a radiação de sincrotrão pode ser observada nos laboratórios equipados com aceleradores de partículas, chamados sincrotrões (daqui o nome), sendo utilizada para uma grande variedade de fins terapêuticos. Desta forma, o mecanismo de produção das ondas de rádio observadas pelos radioastrónomos é agora bem compreendido.

Acerca dos importantes progressos registados em 1950-1951, o radioastrónomo inglês J. Stanley Hey, descobridor das emissões de rádio das explosões solares, fez notar o seguinte:

> Era já claro que a radioastronomia podia dar um grande contributo à astrofísica. A fusão da astronomia óptica com a radioastronomia já começou [...] Espera-nos muito trabalho de observação contínua [...] Para sondar as profundezas do universo eram necessários instrumentos

maiores e mais sensíveis. A construção de grandes radiotelescópios é uma história de visão científica e de luta, competição e controvérsia, persistência e diplomacia.

Continuando o trabalho do passado, uma nova geração de astrónomos veio à liça com novos e poderosos instrumentos. A sua formação também mudou, refletindo as novas necessidades. Na década que se seguiu à Segunda Guerra Mundial somente uma pequena minoria tinha formação em física; a formação dos restantes era em astronomia tradicional. Hoje ninguém seriamente interessado em estudar astronomia pode prescindir do estudo prévio da física — nuclear, das partículas elementares, do plasma, especialmente da física atómica — e 35% dos astrónomos com o grau de doutor obtiveram o doutoramento em Física.

A razão do êxito da radioastronomia também se tornou clara para todos os cientistas. Dado que as ondas de rádio são menos energéticas do que as vulgares ondas de luz, a sua produção através dos processos naturais que ocorrem nos objetos astronómicos é fácil; donde se segue que o universo regurgita de fontes rádio, como acontece com as de ondas de luz. Uma segunda razão para o sucesso da radioastronomia é a incrível eficiência de deteção dos grandes radiotelescópios. Como o detetor dos sinais rádio, no ponto focal do telescópio, pode ser arrefecido artificialmente até temperaturas muito baixas, o ruído de fundo que interfere com o sinal pode ser limitado a um valor muito pequeno; logo, a banda de frequências de rádio torna-se a parte mais sensível de todo o espectro eletromagnético. Tal sensibilidade provocou a admiração de cientistas como Robert Otto Frisch, que, certa vez, comentou:

> Quando o observatório de radioastronomia Mullard foi aqui inaugurado [Cambridge, Inglaterra] em 1958, cada convidado, no lugar que devia ocupar, encontrou, durante o almoço, um pequeno cartão branco com estas palavras escritas no verso: «Ao voltar este cartão empregou mais energia do que a que jamais foi captada, vinda do espaço, por todos os radiotelescópios.»

Não só ninguém podia ter previsto que tal sensibilidade seria possível de conseguir, como também ninguém tinha pensado que os radioastrónomos pudessem obter uma resolução de um milésimo de segundo de arco — equivalente a poder discriminar uma pequena moeda a uma distância de 4000km. O poder de resolução dos telescópios está relacionado com as dimensões do detetor. Ao contrário dos telescópios ópticos, os radiotelescópios podem ter os detetores espalhados por uma área considerável da superfície da Terra, sendo posteriormente possível combinar os vários sinais individuais. Com um poder resolvente tão incrível deixa de ser surpresa o facto de os radioastrónomos poderem discriminar os objetos mais distantes do universo.

Nas décadas que se seguiram aos grandes êxitos dos primeiros anos da década de 50 a radioastronomia sofreu uma expansão em duas áreas: por um lado, os cientistas construíram grandes refletores de prato único; por outro, desenvolveram interferómetros de base longa, consistindo em múltiplos detetores de rádio, separados entre si, que misturavam os sinais vindos da mesma fonte. Os Estados Unidos entraram neste novo campo de atividade, tal como a União Soviética, encontrando-se hoje a maior parte dos radiotelescópios nos Estados Unidos. O mais recente dos grandes radiotelescópios é o Very Large Array[1] (VLA), perto de Socorro,

Novo México, que consiste em 27 radiotelescópios ligados entre si, cada um com um prato de 25m de diâmetro. A linha de base do VLA é de cerca de 27km. Uma outra técnica, chamada *very-long-baseline* interferometry (VLBI)[2], consiste em ligar radiotelescópios separados em continentes diferentes, sendo a resolução superior à de qualquer método. Estes novos instrumentos exploram as radiogaláxias e colocam pela primeira vez os radioastrónomos em pé de igualdade com os seus colegas da astronomia óptica.

O trabalho dos radioastrónomos e dos astrónomos ópticos é muitas vezes complementar, como sucedeu no caso do levantamento cartográfico das superfícies visíveis da Lua e dos planetas, no exame da cromosfera solar e na observação das estrelas com erupções. Todavia, muitas das grandes descobertas só podiam ter acontecido graças à radioastronomia, como, por exemplo, a distribuição galática do hidrogénio neutro, o levantamento cartográfico dos campos magnéticos das galáxias, a descoberta das radiogaláxias, dos quasares, dos pulsares, das riscas de emissão das moléculas, do plasma interplanetário e das cinturas de radiação de Júpiter. Tais descobertas deram origem a uma nova visão do universo, que revelou a presença de processos dinâmicos complexos. Todo um universo anteriormente invisível deixara de o ser.

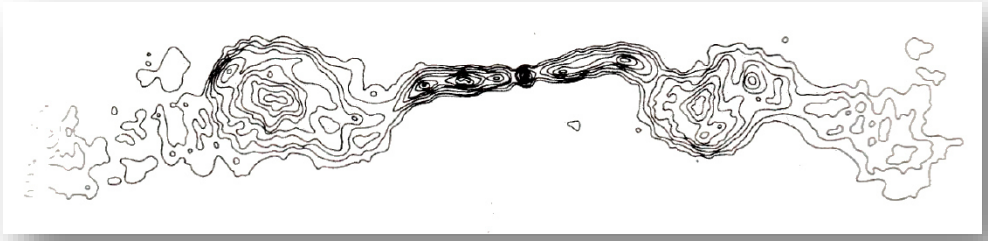
Os radioastrónomos descobriram centenas de fontes rádio nos céus, que, em muitos casos, podiam ser identificadas com corpos visíveis. Algumas eram fontes locais, na nossa galáxia, como a nuvem de gás associada à nebulosa do Caranguejo — os restos de uma supernova. Mas as mais intensas de todas as fontes rádio

identificadas opticamente, com intensidades de milhares, se não milhões, de vezes a saída de rádio da nossa galáxia, são as componentes de uma nova espécie de galáxias — as radiogaláxias. O que se observa são objetos caóticos, galáxias elípticas gigantes com regiões centrais muito luminosas, ou longos jatos de matéria ejetada dos núcleos. São, sem sombra de dúvida, galáxias muito ativas. No entanto, a maior parte das radiogaláxias que cospem radiação rádio não podem ser «vistas» opticamente — estão muito longe e têm um brilho muito fraco no visível. Contudo, estas galáxias «invisíveis» podem fornecer pistas valiosas para a solução do enigma da evolução das galáxias visíveis.

Com a construção de radiotelescópios de alta resolução tornou-se claro que as fontes rádio extragaláticas se agrupam em duas categorias: «compactas» e «não compactas». São, na sua maioria, duplas: a radiação rádio de uma única fonte provém de duas componentes distintas. Nas compactas as duas componentes estão apenas separadas por cerca de uma centena de anos-luz — uma distância relativamente pequena, se nos lembrarmos de que a nossa galáxia tem um diâmetro mil vezes maior. Algumas, porém, são «não compactas»: ocupam uma enorme região do espaço e as duas componentes, em forma de lóbulos, estão separadas por milhares, se não milhões, de anos-luz. Mas o que produz os lóbulos?

As regiões emissoras de rádio e em forma de lóbulos devem-se a longos «jatos cósmicos» de plasma (gás ionizado) que partem de um núcleo central. Os jatos, aos quais Philip Morrison, do MIT, chamou *fountains of radio stuff*[3], seriam algumas vezes desviados e torcer-se-iam no espaço, terminando em pontos quentes — os

lóbulos de intensa emissão rádio. Outras vezes apenas um jato poderia ser visto. Descobriu-se que os grandes jatos estavam normalmente associados a fontes duplas de baixa intensidade: as radiogaláxias fracas ejetam dois jatos em sentidos opostos, enquanto as poderosas ejetam, no máximo, um.



Representação gráfica da intensidade rádio de uma radiogaláxia ativa (no centro) lançando «fontes de radiação rádio» para o espaço. Dois jatos de matéria emergem de um objeto central, acabando por perder energia e inchar para formar dois grandes lóbulos.

Todas as provas apontam no sentido de os jatos serem correntes de gás expelidas do centro da radiogaláxia, como a água da agulheta de uma mangueira. O jato de gás supersónico, ao atravessar o meio interestelar da radiogaláxia, encontra o gás ténue do meio intergalático, que o desacelera, dando, inclusivamente, origem a uma frente de choque quente. As regiões em forma de lóbulo são os pontos quentes onde os jatos abrandam e a energia se acumula. Depois de a matéria do jato ter desacelerado, este volta a fluir para a galáxia, fazendo, assim, inchar os grandes lóbulos vistos pelos radioastrónomos. A corrente de jato, defletida e torcida, é apenas um gás de eletrões e outras partículas carregadas, serpenteando através do meio intergalático.

O principal enigma das radiogaláxias consiste em saber, em primeiro lugar, donde provêm estes eletrões e o que dá origem aos jatos no núcleo da galáxia. Nenhuma das ideias usuais da física está apta a explicar a criação de todos estes eletrões energéticos. Os astrofísicos têm especulado acerca de possíveis mecanismos da origem dos jatos, pelo que abordarei mais adiante algumas dessas ideias. Antes, porém, de mergulhar nestas especulações descreverei aquilo que para muitas pessoas é ainda uma descoberta mais surpreendente do que as radiogaláxias: a descoberta dos quasares, uma classe nova de objetos astronómicos. Tal como as radiogaláxias, os quasares foram primeiro descobertos pelos radioastrónomos, mas muito do estudo pormenorizado destes novos corpos celestes foi feito por astrónomos ópticos — prova de uma frutuosa colaboração entre os dois ramos da astronomia. Mas que são os quasares?

A maior parte das cerca de cem radiogaláxias detetadas até aos primeiros anos da década de 60 tinham aparência indistinta. As que foram identificadas opticamente eram, sem margem para dúvida, galáxias elípticas gigantes. Contudo, algumas das fontes rádio visíveis não eram indistintas e tinham a aparência de estrelas, emitindo uma invulgar quantidade de radiação ultravioleta. Pareciam diferentes e, na realidade, eram-no.

A descoberta aconteceu em 1963, quando Maarten Schmidt, dos observatórios Hale, identificou pela primeira vez riscas de emissão no espectro de luz de um desses objetos semelhantes a estrelas — 3C 273. As linhas espectrais estavam desviadas para o vermelho na incrível percentagem de 16%. Tal desvio para o vermelho, se fosse

devido à expansão do universo, significava que o objeto estava a cerca de 2 mil milhões de anos-luz — mesmo muito, muito longe. A estes objetos com aparência de estrelas foi dado o nome de quasares, «fontes rádio quase estelares»; classificados hoje entre os mais remotos de todos os corpos celestes, emitem imensas quantidades de energia. Identificados os quasares, Allan Sandage descobriu com o telescópio de 5m de Hale, em Mount Palomar, muitos quasares silenciosos na banda rádio, constatando que os quasares emitem muito mais luz ultravioleta do que as estrelas vulgares, o que permitia a sua deteção óptica no meio da panóplia de fontes de luz.

Para termos uma ideia das quantidades de energia libertadas por um quasar imaginemos que uma galáxia tenha as dimensões de uma sala. Um quasar não será maior, nesta escala, do que um pequeno grão de pó, dificilmente visível. No entanto, um único quasar produz 100 vezes a energia irradiada por todos os milhares de milhões de estrelas da nossa galáxia. Podemos agora apreciar quão incrível é a produção energética de um minúsculo quasar, produção esta que se mantém durante cerca de 10 milhões de anos, dando origem a uma quantidade total de energia equivalente à conversão total da massa de 100 milhões de sóis.

O que complica ainda mais o enigma dos quasares é que as suas regiões ativas são muito pequenas, com dimensões não muito maiores do que cerca de uma semana-luz, ou seja, cerca de 10 vezes o tamanho do sistema solar. Como conseguem os astrónomos saber o tamanho de um objeto tão distante? Uma vez que nada viaja com velocidade superior à da luz, para que haja uma flutuação

coerente na intensidade da luz numa determinada região do espaço as dimensões dessa região não podem ser superiores à distância percorrida por um sinal de luz enquanto dura essa flutuação. Observa-se que a intensidade luminosa dos quasares varia com períodos que chegam a ser de uma semana, o que quer dizer que a fonte deve ter dimensões inferiores a uma semana-luz.

Tal como as radiogaláxias, os quasares têm regiões emissoras de rádio compactas e não compactas. Os menos brilhantes assemelham-se ao núcleo das radiogaláxias mais brilhantes, parecendo as radiogaláxias menos ativas, galáxias vulgares. Esta sequência contínua de atividade galática sugere uma sequência evolutiva em que os quasares, que teriam sido em muito maior número no passado, evoluem para radiogaláxias, que, quando «sossegam», se transformam em galáxias vulgares — uma evolução do núcleo galático no sentido de a atividade diminuir ou se tornar esporádica. No entanto, esta possível sequência evolutiva, embora intrigante, continua ainda a ser objeto de debate entre os astrónomos. Pode estar completamente errada ou ser apenas uma parte da história da evolução galática. E, todavia, parece atraente.

Que mecanismos físicos poderão ser responsáveis pela energia dos quasares ou das radiogaláxias e produzir os jatos que se observam? Para responder a esta pergunta temos de abandonar o reino das observações e valer-nos da imaginação especulativa dos teóricos da astrofísica, os quais pensam que o núcleo da galáxia alberga uma «máquina monstruosa», que pode envolver um ou mais buracos negros ou objetos compactos similares. Edwin Salpeter, da Universidade de Cornell, e Yakob B. Zel'dovich, físico

soviético, estão entre aqueles que partilham desta opinião, teorizando que numa das fases iniciais da formação das galáxias, quando estas se compunham essencialmente de estrelas vulgares e de gás, se deu a explosão de uma ou mais estrelas no núcleo denso da galáxia, as quais se converteram em buracos negros. Então, como canibais, devorando a população circundante, os buracos negros consumiram milhões dos seus abundantíssimos vizinhos, elevando-se a sua massa até um valor da ordem de 100 milhões de massas solares, processo este que durou alguns milhões de anos. Embora os buracos negros não possam irradiar, o mesmo não se passa com o gás, que é aspirado para o seu interior. Os físicos estimam que um décimo da massa sugada para dentro de um buraco negro é convertida em energia radiante, sendo esta radiação o que vemos como sendo um quasar. Se um buraco negro consome cerca de mil milhões de estrelas com o tamanho do Sol, então a energia radiante é equivalente a 100 milhões de estrelas — a quantidade exata de que precisamos para explicar a produção de energia do minúsculo quasar.

Os buracos negros têm igualmente aceitação entre os teóricos que tentam explicar o comportamento das radiogaláxias. De acordo com os seus modelos, o gás carregado eletricamente, rodopiando em torno do buraco, forma um dínamo elétrico que projeta dois jatos de partículas de alta energia em sentidos opostos a partir do centro. Os buracos negros podem rodar e os jatos que emergem segundo os respetivos eixos podem ser alimentados pela energia da rotação. Seriam então estas as *fountains of radio stuff* vistas pelos radioastrónomos.

Outra especulação que se faz é a de que o núcleo das radiogaláxias contém dois buracos negros, orbitando perto um do outro, ou uma estrela de massa muito elevada em órbita em torno de um buraco negro, ou até duas estrelas de massas muito elevadas. Os jatos teriam origem nas interações dos dois objetos de massa muito elevada, orbitando no núcleo da galáxia.

Jeremiah Ostriker, da Universidade de Princeton, e Scott Tremaine, atualmente no MIT, formularam outro modelo intrigante, no qual o buraco negro de uma grande galáxia «canibal» consome uma galáxia «missionária» mais pequena, despedaçando as partes exteriores desta e capturando na sua órbita o buraco negro da missionária. Parte do atrativo desta ideia reside no facto de a maior parte das radiogaláxias terem existido no passado longínquo, quando as galáxias estavam mais perto umas das outras e o «canibalismo» teria sido mais frequente. Além disso, esta hipótese permite explicar algumas peculiaridades dos jatos observados. Mas existem inúmeros modelos teóricos para explicar as radiogaláxias, sendo difícil determinar qual deles está correto, se é que algum o está.

Suponhamos que os teóricos têm razão e que existem buracos negros nos núcleos dos quasares e das radiogaláxias. Que acontecerá depois de o buraco negro ter terminado o seu trabalho? Uma vez despojados os núcleos galáticos da maior parte das estrelas e do gás que continham, o buraco negro «sossega», os jatos de rádio atrofiam-se e a galáxia torna-se normal, como a maior parte das que povoam a nossa vizinhança cósmica. Provavelmente, os buracos negros nos centros das galáxias

espiraladas são muito mais pequenos do que os que existem nas elípticas, dado que a rotação angular das estrelas em espiral as mantém fora da perigosa região do núcleo.

Se estas ideias estiverem corretas, então pode realmente existir um buraco negro com a massa de alguns milhões de sóis no centro da nossa galáxia, como conjeturaram há mais de uma década os astrofísicos ingleses Donald Lynden-Bell e Martin Rees. A melhor prova em abono desta ideia provém de recentes observações rádio do centro galático que revelam a existência de uma fonte rádio compacta e não térmica (presumivelmente, um buraco negro) envolvida por braços espiralados de gás, provenientes talvez de estrelas despedaçadas. Os sinais rádio que os astrónomos veem são precisamente do género dos que se espera serem emitidos pelo gás que cai num buraco negro. Os raios X e gama com origem no centro galático e que resultam da aniquilação eletrão-positrão, um processo que poderia ter lugar à entrada de um buraco negro, constituem provas adicionais.

Até que se disponha de mais informação, este modelo do núcleo galático, fazendo apelo a um buraco negro, permanecerá na zona crepuscular da física especulativa. Felizmente, será com certeza possível recolher muito mais dados na próxima década, parte dos quais virão do Very Large Array, em Socorro, Novo México, que até agora apenas examinou uma fração das fontes rádio que pode detetar, das quais apenas algumas em alta resolução. O uso deste importante instrumento, embora ainda no começo, já revelou novas peculiaridades do núcleo galático. O programa VLBI, que utiliza simultaneamente radiotelescópios situados em diferentes

continentes, será também intensificado. O telescópio espacial, com 20 vezes o poder de resolução dos telescópios equivalentes baseados em terra, irá dizer-nos muito acerca dos quasares, pelo que poderemos vir a descobrir muitos mais jatos visíveis. Uma questão interessante, cuja resposta poderá ser dada pelo telescópio espacial, é a de saber se os quasares estão ou não agrupados em aglomerados e superaglomerados, como as galáxias.

Foram descobertas recentemente «lentes gravitacionais», constituídas por galáxias inteiras interpostas entre a Terra e um quasar distante. A galáxia, em si, pode ser invisível, mas a sua massa imensa encurva o espaço circundante, fazendo com que a luz do quasar distante se encurve à volta dela e na Terra observemos imagens múltiplas do quasar. Os astrónomos procuram mais lentes gravitacionais (uma foi já encontrada por pesquisa direta), pois podem não só facultar novos métodos de estabelecer escalas de distâncias aos quasares distantes e novos dados acerca da estrutura do universo em larga escala, como também tornar-se «mesas ópticas» numa escala cósmica.

Debatem-se atualmente os méritos da tecnologia presente e futura, como sejam os telescópios de espelhos múltiplos, que, em vez de um espelho grande, possuem vários mais pequenos, cujas imagens separadas são combinadas eletronicamente para formar uma só. Um projeto diferente, ainda em fase de estudo, é o de um telescópio com um espelho segmentado de 15m (em contraste com o velho telescópio de 5m de Hale), que se assemelhará a um favo de mel polido, em que cada uma das 60 partes hexagonais é controlada individualmente, de maneira a ter-se um foco comum e

uma só imagem. Seria um telescópio nacional. A Universidade da Califórnia planeia construir (com fundos privados) um telescópio de espelho segmentado de 10m, com 37 peças hexagonais, cada uma com cerca de 2m de largura, e cujo alcance será duas vezes superior ao de qualquer telescópio atualmente existente. Discute-se igualmente o uso de espelhos deformáveis, que utilizam novos sistemas de geração de imagens. Muitas destas inovações projetadas são o resultado da competição amigável entre construtores de telescópios e têm por fim tornar os telescópios eficientes, baratos e aptos a perscrutar fundo no cosmos. Novas componentes do cosmos ser-nos-ão certamente reveladas por estes telescópios supergigantes, planeados para a década de 90, instrumentos que hoje parecem pertencer ainda ao domínio da ficção científica.

Muitos destes novos instrumentos serão utilizados no estudo das galáxias e quasares muito distantes, situados na zona a que Allan Sandage chama a «orla do mundo». Tal estudo é interessante porque poderá permitir responder a algumas das questões cruciais com que se defronta o astrónomo moderno: que processos controlam a evolução das galáxias e qual é a estrutura do universo em grande escala? Contudo, malgrado os avanços tecnológicos que os sistemas de deteção registaram até hoje, os astrónomos sentem-se frustrados na procura de respostas para estas perguntas, pois os dados que obtêm não são facilmente interpretáveis. Por exemplo, algumas galáxias distantes têm cores inconsistentes com os seus desvios para o vermelho — indicação de que não compreendemos verdadeiramente a composição destas galáxias. Um novo problema surge ao reconhecermos que a determinação da distância a galáxias

longínquas necessita de «luzeiros padrão» — galáxias cuja luminosidade se mantém aproximadamente constante ao longo do tempo. Mas, se as galáxias estão a evoluir, a luminosidade pode ter sido muito maior no passado e os «luzeiros padrão» podem nem sequer existir. O problema da evolução das galáxias e o da estrutura do universo em larga escala encontram-se, assim, interligados.

Outra questão à qual os novos instrumentos poderão trazer alguma luz é a da determinação da idade das galáxias, que não é de fácil resposta, uma vez que, se existiam galáxias como a nossa há alguns milhares de milhões de anos, então deveriam ser demasiado pequenas e pouco brilhantes para poderem ser vistas com os telescópios ópticos atuais. Mas os quasares existiam decerto e talvez fossem precisamente os núcleos luminosos de novas galáxias, a partir das quais evoluíram as galáxias normais. Quando o telescópio espacial entrar em funcionamento, deveremos ser capazes de pôr à prova esta hipótese. Porém, por agora, e assumindo que as galáxias vulgares evoluíram a partir dos quasares, a questão da idade das galáxias identifica-se então com a idade dos quasares.

O mais velho (portanto, o mais distante) dos corpos celestes conhecidos é um quasar descoberto por meio de um dos novos telescópios gigantes, o radio-observatório de Parkes, na Nova Gales do Sul, na Austrália, catalogado como *Parkes 2000-330*. Afasta-se de nós a uma velocidade de cerca de 90% da velocidade da luz, pelo que a luz que vemos dele teve origem quando o universo tinha apenas um quinto da idade atual. Observa-se que a maior parte dos quasares estão situados numa época que vai de um terço a um

quinto da atual idade do universo — a Era dos quasares. Os astrónomos inspecionaram Eras mais recuadas, mas apenas encontraram, se é que encontraram, alguns quasares. Os mais antigos podem estar obscurecidos pelas poeiras, mas, segundo parece, os quasares «acenderam-se» e depois «apagaram-se» mais ou menos todos na mesma altura da história do universo. Se as galáxias tiveram origem em quasares, então, de acordo com algumas estimativas, estes quasares «acenderam-se» pela primeira vez há cerca de 10 a 11 mil milhões de anos — uma escala de tempo aceitável, comparada com a idade das estrelas mais velhas, e próxima da idade do próprio universo.

Qual o aspeto das galáxias antes da Era dos quasares? Os astrónomos procuram «galáxias primordiais» com instrumentos novos e muito sensíveis. Algumas podem envolver os quasares, mas são demasiado ténues para poderem ser vistas. Acredita-se terem existido entre os primeiros 100 milhões de anos e os primeiros milhares de milhões de anos depois do *big bang*. A descoberta de tais galáxias seria como encontrar um fóssil nunca antes visto. Serão estruturas difusas com fraco brilho de superfície, ou serão azuis, indicando uma elevada taxa de formação de estrelas? Que aspeto teria então o universo? Ninguém sabe ao certo — há pouca ou quase nenhuma informação sobre este período primordial e a especulação está aberta à imaginação quase livre.

Percorremos um longo caminho desde a visão de Hubble das galáxias como universos-ilhas isolados. A teoria das galáxias e as observações abriram uma nova visão sobre a sua evolução. Da mesma forma que muitas estrelas dentro de uma galáxia formam

grupos e criam um ambiente rico, também as galáxias tendem a agrupar-se em aglomerados e engendram um ambiente dentro do universo, interagindo de forma complexa. As galáxias elípticas são encontradas mais frequentemente nas regiões de densidade elevada dos aglomerados galáticos, enquanto as galáxias espiraladas existem relativamente isoladas ou nas regiões de baixa densidade destes aglomerados. Há cada vez mais provas de que as galáxias situadas próximo do centro dos aglomerados foram anteriormente galáxias espiraladas, que, de uma forma ou outra, foram despojadas do seu gás, de tal maneira que a formação de estrelas em tais galáxias parou. O ambiente galático é, assim, importante para a determinação da estrutura, forma e evolução de uma galáxia.

As estrelas evoluem, bem como as galáxias e os aglomerados de galáxias. Os mecanismos desta evolução começaram só agora a ser entendidos. Grandes obstáculos se levantam no caminho que leva ao conhecimento. Talvez nem sequer saibamos quais são as componentes materiais primárias do universo. Formas invisíveis de matéria — buracos negros à escala macroscópica, novas partículas quânticas à escala microscópica — podem ser dados importantes para a compreensão destes processos.

Mas alarguemos agora a nossa visão do cosmos para lá das galáxias, das radiogaláxias e dos quasares. Imaginemos que as galáxias são somente partículas espalhadas, como areia, pelo espaço do cosmos. Estarão distribuídas ao acaso ou existirá uma estrutura — uma nova pista para a descoberta da ordem cósmica? Tais questões conduzem-nos ao problema da determinação da

estrutura do universo em grande escala, um tópico para o qual nos voltamos em seguida.

---

[1] Literalmente, «conjunto muito grande». (*N. do T.*)

[2] Literalmente, «interferometria de linha de base muito longa». (*N. do T.*)

[3] O pitoresco da expressão original perde-se na tradução por «fontes de emissão rádio». (*N. do T.*)

CAPÍTULO 6

– Por Que É o Universo Granuloso? –

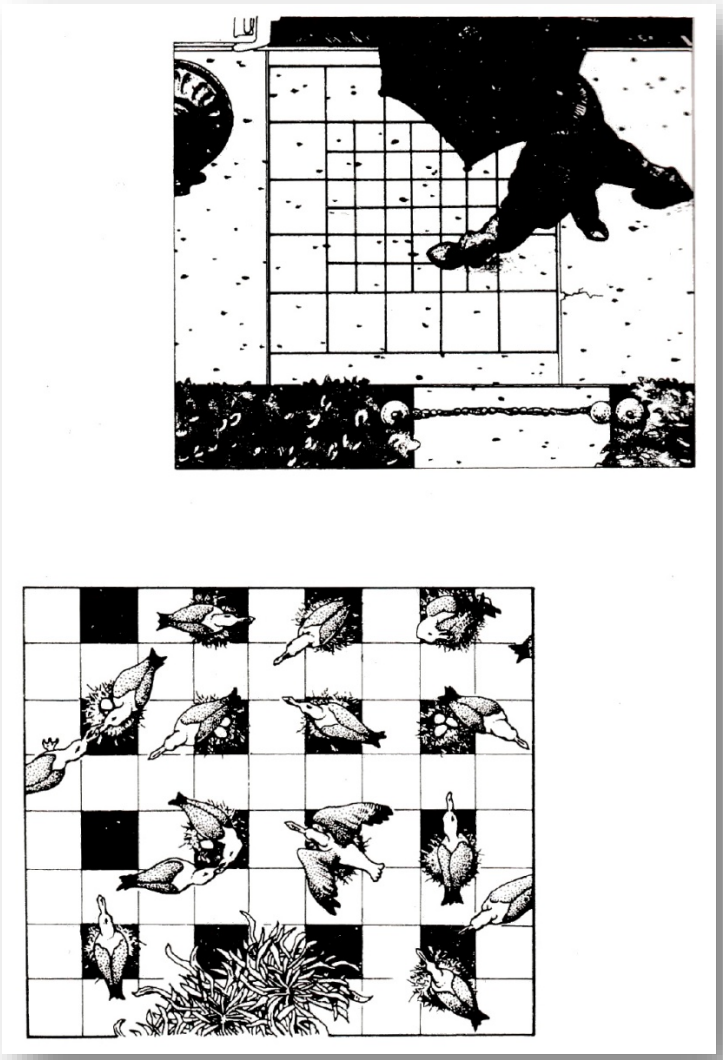
> Existem [...] grandes diferenças de opinião no que toca à formação das galáxias. Isto não pode ser considerado como um sinal de perturbação, dado que só recentemente o assunto progrediu até ao ponto em que é possível formular posições que parecem defensáveis.
>
> *P. J. E. PEEBLES, 1984*

Já alguma vez observou o começo do cair da chuva sobre um pavimento? Inicialmente, alguns pingos humedecem o chão; depois, mais pingos continuam a cair, até que todo o pavimento fica molhado. Todavia, antes que isto aconteça, pode observar-se que os pingos de chuva formam um arranjo bidimensional aleatório sobre o pavimento — o sítio onde cai um pingo é independente do facto de terem ou não caído previamente outros pingos no mesmo sítio.

Outro exemplo de um arranjo aleatório são os hábitos de nidificação de certas aves marinhas numa ilha. Cada ave tem o respetivo lugar; se observarmos do alto a zona dos ninhos, as aves parecem estar colocadas ao acaso segundo um arranjo bidimensional. Mas é uma aleatoriedade de tipo diferente do que acontece com os pingos de chuva. As aves territoriais parecem não gostar umas das outras. Imaginemos que colocamos uma grade sobre o chão, sendo cada quadrado da grade suficientemente

grande para que lá caiba uma ave. Então, dado que cada ave ocupará um determinado quadrado na grade, a probabilidade de encontrar uma outra ave num quadrado adjacente é muito baixa. As aves repelem-se umas às outras, enquanto os pingos de chuva são indiferentes à presença uns dos outros. Os arranjos, mesmo aleatórios, podem ter propriedades que forneçam informações importantes.

Imaginemos também que as galáxias exteriores sejam apenas partículas espalhadas pelo espaço do universo. Examinando esta grande paisagem, poderíamos perguntar: estão as galáxias espalhadas ao acaso, como pingos de chuva, ou existe arranjo estruturado? Os astrónomos interrogaram-se pela primeira vez acerca da organização cósmica na década de 20, quando se tornou claro que as galáxias são vastos sistemas estelares exteriores, com diâmetro de milhares de anos-luz e semelhantes à nossa Via Láctea. Constatou-se que as galáxias tendem a juntar-se, como nos aglomerados — o primeiro indício de que havia uma certa espécie de estrutura orgânica entre as galáxias. Nos primeiros anos da década de 50 o astrónomo Gérard de Vaucouleurs foi o primeiro a estudar a concentração local das galáxias, a que chamou «superaglomerado local», do qual a Via Láctea é membro. A partir do exame da aglomeração das galáxias nasceu um novo campo de investigação — o estudo da estrutura do universo em grande escala.



Diferentes tipos de acaso caracterizam a queda dos pingos de chuva sobre o pavimento e a nidificação das aves marinhas. As distribuições são diferentes porque os pingos de chuva atingem sítios no chão, independentemente do facto de terem ou não caído previamente outros pingos no mesmo sítio, enquanto as aves marinhas guardam, pelo menos, uma «distância de bicada» entre os ninhos. As correlações da distribuição aleatória das galáxias do universo poderão fornecer pistas quanto à sua evolução e origem.

Na década de 50 Donald Shane e Carl Wirtanen, no Observatório Lick, na Califórnia, começaram a pesquisar mais de um milhão de galáxias — uma quantidade de informação quase indigerível. P. James E. Peebles e os seus colaboradores, na Universidade de Princeton, pegaram nas informações de Shane-Wirtanen e processaram-nas para obterem um mapa bidimensional do céu, que

se transformou no popular cartaz *Um Milhão de Galáxias*, o qual nos dá uma imagem razoável da distribuição das galáxias — uma visão geral do cosmos.

Olhando para este quadro, podem ver-se densos aglomerados de galáxias, formando nós, e entre eles o espaço vazio. Há quem veja nisso provas da existência de uma fina estrutura filamentar. Se admitirmos que a técnica de amostragem é rigorosa, parece claro, a partir destes mapas, que existe realmente um certo tipo de estrutura em grande escala da distribuição das galáxias à escala das dezenas de milhões de anos-luz. Não estão espalhadas como os pingos de chuva. Mas a escalas de distância muito grandes, da ordem dos milhares de milhões de anos-luz — aquilo que veríamos se olhássemos para o mapa através de um vidro fosco —, o universo tem a aparência de um local regular e sem estrutura. O único problema estrutural a essas escalas de distância é este: qual é a geometria global do espaço? Aberta ou fechada? Analisaremos esta questão no capítulo seguinte, dedicado à cosmologia.

Mesmo de posse de informações mais desenvolvidas do que dispunham na década de 50, os astrónomos continuam a debater problemas relacionados com a estrutura do universo em grande escala. As duas posições extremas neste debate são ocupadas por aqueles a quem chamaremos «hierarquistas» e «filamentarianos». Os hierarquistas veem uma contínua «aglomeração hierárquica» de galáxias, abrangendo desde distâncias relativamente pequenas, com alguns diâmetros galáticos, até às distâncias máximas. Os pequenos grupos de galáxias combinam-se para formar grupos maiores e estes, por sua vez, combinam-se para formar grupos

ainda maiores — uma hierarquia contínua de aglomerados. Os filamentarianos, pelo contrário, veem provas convincentes da existência de filamentos cósmicos, consistindo em espessas «cordas» de galáxias, à escala dos superaglomerados, com grandes espaços vazios entre elas. A distinção entre estas duas posições extremas nem sempre é completamente definida, uma vez que os desvios da aglomeração regular podem parecer filamentos e os espessos e difusos filamentos parecem parte de um conjunto hierárquico regular de aglomerados.

A análise recente de dados parece vir apoiar os filamentarianos, aceitando a maior parte dos astrónomos esta posição. No entanto, outros mantêm que os filamentos aparecem devido a métodos de amostragem tendenciosos — escolhem-se as galáxias que corroboram o nosso ponto de vista. É um pouco como ler nas folhas de chá — olhamos para um arranjo aleatório e vemos nele aquilo que queremos. Precisamos apenas de nos lembrar do debate sobre os canais de Marte, na viragem do século — que eram também estruturas lineares e filamentosas —, para apreciarmos o problema. Na ausência de informações pormenorizadas, os olhos tendem a acrescentar entre os pontos linhas imaginárias, as quais, na realidade, não estão presentes.

A ênfase da conceção de aglomeração hierárquica está na suavidade da transição da aglomeração desde as pequenas escalas cósmicas até às maiores — podem encontrar-se aglomerados de todos os tamanhos. Podemos comparar isto com a distribuição de pessoas num *cocktail* muito grande. Algumas pessoas permanecem sozinhas, mas a maior parte reúne-se em grupos ou aglomerados

de tamanhos variáveis para conversar ou para ouvir o que se diz. Alguns grupos podem ter poucas pessoas, outros muitas mais. Alguns grupos podem fundir-se. As pessoas não estão distribuídas puramente ao acaso, que é a situação que se pode esperar se centenas de bolas de ténis forem espalhadas no chão de um compartimento vazio. As bolas de ténis não se aglomeram da mesma maneira que as pessoas.

Esta conceção da distribuição das galáxias é contestada pelos filamentarianos, que pensam existir ainda mais estrutura na distribuição das galáxias e que a transição de aglomeração não é hierarquicamente suave. Veem provas da existência de uma estrutura filamentar, uma malha de cordas cósmicas ou superfícies no espaço, à volta da qual as galáxias tendem a congregar-se. Noutras regiões do espaço, entre os filamentos, devem existir grandes espaços vazios. Se fôssemos ao *cocktail* e desenhássemos uma vasta rede de linhas curvas no chão ao acaso e depois disséssemos às pessoas que tinham de permanecer sobre ou perto de uma linha, a distribuição resultante das pessoas no salão formaria filamentos, o que estaria mais próximo da visão que os filamentarianos têm da distribuição das galáxias.

Que dizem as observações? Novos dados, compilados a partir da pesquisa do milhão de galáxias de Shane-Wirtanen, vêm em apoio dos filamentarianos. O astrónomo estoniano Jan Einasto começou a examinar na década de 70 mapas tridimensionais das galáxias, em vez de mapas bidimensionais. Isto, é claro, requeria a determinação não só da posição das galáxias individuais no céu, mas também das suas distâncias à Terra, através da medição dos desvios para o

vermelho e da aplicação da lei de Hubble. Einasto afirmou ter encontrado provas da existência de filamentos e de espaços vazios nos pequenos setores do céu que examinou. Depois, em 1981, usando técnicas de amostragem similares, foi encontrado um grande «buraco no espaço», uma região com 200 milhões de anos-luz de diâmetro, centrada num ponto a 500 milhões de anos-luz para além da constelação de Boieiro, na nossa galáxia. Esta descoberta parecia vir em apoio dos filamentarianos, que argumentavam que entre os filamentos tinham de existir grandes espaços vazios.

Uma inspeção do desvio para o vermelho feita a 2400 galáxias compreendidas num raio de 400 milhões de anos-luz da Terra foi completada por Marc Davis e colaboradores, no Centro de Astrofísica de Harvard-Smithsonian, em 1981. Um estudante do liceu local construiu um modelo tridimensional dos dados de Davis, usando bolas de medula suspensas de cordéis para representar as galáxias. Neste modelo vê-se claramente que existem espaços vazios e filamentos. Diz Davis: «Penso que temos agora informações suficientes para convencer toda a gente de que existe uma estrutura filamentosa aberta.» Teóricos que, como Edwin Turner, fazem análises estatísticas pormenorizadas da distribuição das galáxias concordariam com esta posição. Turner encontrou provas da existência de cordas cósmicas e de folhas (sheets). É oportuno dizer que a maioria dos astrónomos estão convencidos da existência de filamentos e superfícies. No entanto, outros há que não consideram estas provas convincentes, afirmando que estes arranjos estão somente nos olhos do observador e que a análise estatística e as amostragens são tendenciosas. O próprio Davis

admite que o tamanho da «grelha» usada no processamento computadorizado dos dados pode alterar as conclusões.

Por que razão estão os astrónomos tão interessados em estabelecer a estrutura do universo em grande escala — um interesse que vai para além do desejo de conhecer a estrutura por si só? Sabem que a estrutura do universo em grande escala fornece uma pista importante para a natureza deste e prende-se com duas importantes questões, relacionadas uma com a outra: a existência de matéria escura e a origem e evolução das galáxias. Estão altas as apostas para uma correta leitura das folhas de chá do universo.

Os astrónomos e os físicos sabem que é pouco provável que a atração gravitacional entre as galáxias individuais, atuando durante um longo período de tempo, possa ter produzido, por si só, a distribuição das galáxias que observamos. Alguma coisa mais tem de estar presente. Um candidato provável é «a massa que falta» na forma de matéria escura — matéria escondida dentro ou entre as galáxias e que não pode ser vista pelos telescópios.

Uma prova indireta, mas boa, da existência de tal matéria escura foi encontrada na década de 30 em estudos de aglomerados de galáxias pelos astrónomos Fritz Zwicky e Sinclair Smith, do Cal Tech. Concluíram que as galáxias num aglomerado se moviam demasiado depressa para permitirem que o aglomerado se mantivesse coeso, dado que a atração gravitacional mútua de todas as galáxias era insuficiente para ligar um sistema cujas partes se moviam tão depressa. Era a primeira boa prova de que os aglomerados de galáxias deviam conter algum tipo de matéria

escura que fornecesse a massa adicional para os conservar coesos. Estudos subsequentes de outros aglomerados de galáxias revelaram as mesmas características.

Mais recentemente tem havido provas acumuladas da existência de matéria escura mesmo à escala dos superaglomerados de galáxias. Em 1982 R. Brent Tully, da Universidade de Hawaii, e J. Richard Fisher, do National Radio Astronomy Observatory[1], em Green Bank, West Virginia, completaram uma investigação de nove anos durante a qual mediram as posições e os desvios para o vermelho da luz de cerca de 2200 galáxias no superaglomerado da Virgem. No centro deste superaglomerado de galáxias situa-se o próprio aglomerado da Virgem, que contém cerca de 20% das galáxias do superaglomerado. Um livro segurado na mão à distância de um braço estendido cobri-lo-ia no céu noturno. No centro do aglomerado da Virgem situa-se a M 87, uma galáxia elíptica gigante, com um remoinho gravitacional de outras galáxias, estrelas e gás quente, emitindo raios X e movendo-se à sua volta. Tully e Fisher descobriram que, envolvendo o centro da Virgem, se encontra um imenso disco de galáxias, com, aproximadamente, 35 milhões de anos-luz de diâmetro e cerca de 6 milhões de anos-luz de espessura. À volta do disco estão nuvens de galáxias, com a forma de charutos, apontadas para o núcleo — uma configuração impressionante e, de certa maneira, floral de milhares de galáxias. O disco observado não pode ser mantido só pela gravidade das galáxias que contém, as quais se movem com rapidez excessiva para que tal seja possível. De novo a matéria escura, distribuída em forma de disco à escala dos superaglomerados, poderia ser o agente responsável.

A matéria escura não só está presente à escala de distâncias das dimensões dos aglomerados e superaglomerados, como também é possível que seja a principal componente material das galáxias individuais. Como descrevi em capítulo precedente, ao longo da última década tem sido medida a velocidade do hidrogénio neutro e das estrelas isoladas orbitando bem para além da fronteira visível das galáxias espiraladas. Se toda a massa das galáxias estivesse nas estrelas — caso em que a densidade da massa baixaria até zero na fronteira da galáxia visível — , então a velocidade do gás em órbita ou das estrelas para além da fronteira deveria baixar até zero com o aumento da distância à galáxia. Em vez disso, a velocidade do gás em órbita permanece constante ou até aumenta para lá do bordo da galáxia, o que sugere a existência de matéria escura escondida na galáxia, que se estende bem para além da parte visível. Possivelmente, a parte visível da galáxia — todas as estrelas e o gás — não ultrapassa uns meros 10% da matéria total; o resto é um halo gigantesco e invisível que penetra profundamente no espaço, tão profundamente que pode fundir-se com os halos semelhantes das galáxias vizinhas.



Flor da Virgem. Uma pesquisa rádio de 2200 galáxias no superaglomerado da Virgem revelou esta figura composta. No meio encontra-se o aglomerado da Virgem, em cujo centro, por sua vez, está a galáxia elíptica gigante M 87. Envolvendo o aglomerado central de galáxias, encontra-se uma distribuição de galáxias em forma de disco com cerca de 35 milhões de anos-luz de diâmetro e várias «nuvens» de galáxias apontadas para o centro. Aglomerados de galáxias semelhantes encontram-se por todo o universo.

Que *é* a matéria escura? Os astrónomos não sabem dar uma resposta simples. No entanto, podem dar razões pelas quais alguns dos candidatos óbvios não servem. Foi considerada a possibilidade de poder ser gás, grãos de poeira, hidrogénio gelado, ou estrelas que entraram em colapso, como os buracos negros, ou planetas do tamanho de Júpiter. Todas estas possibilidades, por razões várias, foram afastadas. Com a possível exceção das anãs castanhas, nada no reportório usual da matéria parece servir.

Os físicos teóricos, que têm sempre resposta para tudo, conjeturaram que partículas quânticas eletricamente neutras

criadas no *big bang*, tais como neutrinos, gravitinos, fotinos e axiões, constituem a matéria escura. Esta ideia, inicialmente considerada rebuscada, ganhou apoios nos últimos anos, uma vez que parece lançar também alguma luz sobre o problema da origem das galáxias. Assinalavelmente, os maiores objetos que observamos — galáxias e aglomerados de galáxias — podem não somente ser estruturados pelos mais pequenos que conhecemos — as partículas quânticas —, como, na realidade, dever-lhe a existência.

É possível que tal matéria escura, sob a forma de um fluido de partículas quânticas invisíveis, perpasse o cosmos, constituindo 90% da massa total. As galáxias seriam então apenas uma pequena fração da densidade de massa total do universo, limitar-se-iam a passear-se no seio do fluido escuro. A matéria escura de partículas quânticas, se é que existe, pode congregar-se gravitacionalmente em conjuntos gigantescos, com o tamanho de galáxias ou de aglomerados, cujas dimensões dependem das propriedades das próprias partículas quânticas. Então as galáxias seriam como partículas luminosas cuja distribuição revelaria os conjuntos, filamentos e aglomerados do fluido cósmico de matéria escura. Há quem não partilhe desta opinião e pense que não deve haver relação entre a distribuição da matéria escura e da matéria visível.

Ao procurar respostas para estas questões, os astrónomos tentam igualmente resolver o enigma da origem e evolução das galáxias. A estrutura do universo em grande escala, revelada pela distribuição e movimento das galáxias, é, presumivelmente, uma remanescência do *big bang*. Recuemos, pois, no tempo até ao *big*

bang e perguntemos como é que as galáxias foram originadas a partir da bola de fogo primordial.

Esta também não é uma pergunta de fácil resposta. Os cientistas sabem que a bola de fogo primordial era muito homogénea e regular porque a radiação de fundo de micro-ondas, também ela uma remanescência do *big bang* e o mais antigo ente que podemos ver, está distribuída com uniformidade direcional (desde que eliminemos o efeito do movimento da Terra relativamente à radiação de fundo) superior a uma parte em 10.000. No entanto, ocorreram decerto pequenos desvios desta homogeneidade em grande escala, pequenas «concentrações» que se desenvolveram para dar galáxias. Olhar desta maneira para o universo é como contemplar a superfície de um lago, que, na globalidade, é regular, mas, quando olhado de perto, apresenta ondulações e até ondas. Mas por que razão é o universo granuloso?

Na realidade, estamos a fazer duas perguntas e é melhor separá-las. A primeira questão é esta: qual foi a origem das minúsculas flutuações de densidade de massa — os desvios de homogeneidade — que acabaram por se transformar em galáxias? A segunda questão é a seguinte: dado que essas minúsculas flutuações existiram na bola de fogo do *big bang*, como cresceram e se desenvolveram elas então até formarem as primeiras galáxias? Examinemos estas questões uma a uma.

Os físicos que estudam o *big bang* pensam que as minúsculas flutuações iniciais de densidade começaram como flutuações quânticas — uma componente necessária de qualquer sistema

físico, incluindo o gás das partículas quânticas que constituía a bola de fogo. Não existe nenhum modo de nos desembaraçarmos delas. A única questão é esta: como se explica que as flutuações quânticas tenham tido a amplitude adequada — suficientemente grandes para virem a transformar-se em galáxias, mas não tão grandes que pudessem destruir a homogeneidade global do universo? Para responder a esta questão há que examinar matematicamente o universo no momento em que estas flutuações apareceram, antes do primeiro milionésimo de bilionésimo de segundo, um intervalo de tempo em cuja descrição os físicos se sentem muito pouco confiantes.

Alguns sustentam que a estrutura da distribuição das galáxias que hoje vemos é uma consequência das propriedades inerentes a essas flutuações quânticas iniciais, da mesma maneira que um carvalho é uma consequência da bolota, da qual cresceu. Assim, a estrutura em grande escala do universo fornece uma espécie de janela sobre esses primeiros tempos. Outros, contestando esta opinião, defendem com igual confiança — que não é muito grande em qualquer dos dois casos — que a estrutura do universo em grande escala é uma consequência do crescimento posterior dessas flutuações e é independente das propriedades específicas iniciais das flutuações. Tais disputas são parte da vanguarda das discussões correntes.

Suponhamos que os físicos podiam explicar a origem das flutuações iniciais de densidade de massa. Poderiam então mostrar como cada flutuação cresce, porque tal «concentração» tem apenas duas forças simples atuando sobre ela. A primeira é a energia do

movimento das partículas de gás, proporcional à temperatura do gás, que tenta desfazer o grânulo. A segunda é a atração gravitacional mútua de todas as partículas, que tenta contraí-lo. Sabendo o valor da temperatura e da massa das partículas no gás primordial, os físicos podem calcular o crescimento e as dimensões finais dos grânulos de densidade de massa. Concluiu-se que, nas condições de temperatura e de massa adequadas, mais e mais matéria envolvente era varrida para dentro desses grânulos pela atração gravitacional, fazendo-os expandir-se à medida que decorria a explosão do *big bang*. Cerca de 100.000 ou um milhão de anos mais tarde o *big bang* tinha terminado e os grânulos de matéria gasosa povoavam o universo. Começou então a processar-se a formação das verdadeiras galáxias.

O cenário para a formação das galáxias depende crucialmente das dimensões e da massa dos grânulos que povoavam o universo nestes primeiros tempos. Alguns cientistas pensam que os primeiros grânulos tinham a massa e as dimensões que os aglomerados e superaglomerados de galáxias têm hoje — eram grânulos muito grandes. Subsequentemente, estes grandes grânulos do tamanho de aglomerados fragmentaram-se em grânulos do tamanho de galáxias.

Outros pensam que os primeiros grânulos começaram por ter as dimensões e a massa das galáxias de hoje — eram grânulos relativamente pequenos, que foram o embrião dessas galáxias. Segundo esta imagem, depois de as galáxias se formarem, começaram então a agrupar-se para formar os aglomerados e

superaglomerados que hoje observamos. O astrofísico P. James E. Peebles argumenta, em defesa deste ponto de vista:

> Se os aglomerados se formaram primeiro, então por que razão não estão todas as galáxias agrupadas em aglomerados? A maior parte das galáxias são como a nossa — estão nos bordos em agrupamentos abertos, começando agora a coalescer.

A diferença entre estas duas opiniões pode ser resumida como a diferença entre «primeiro, aglomerados, depois, galáxias» e «primeiro, galáxias, depois, aglomerados». Examinemos agora estas alternativas.

A opinião de que os aglomerados e os superaglomerados se formaram primeiro, como gigantescos grânulos de matéria que, subsequentemente, se fragmentaram em galáxias, começou com o modelo «panqueca» de Yakob B. Zel'dovich, do Instituto de Matemática Aplicada de Moscovo. No princípio da década de 70 Zel'dovich e os seus colaboradores mostraram que, se existiu uma concentração inicial de matéria gasosa, o seu colapso, devido à própria gravidade, não teria sido esfericamente simétrico, da maneira que o é um balão ao esvaziar-se. Pelo contrário, é mais provável que o grânulo entre em colapso até ficar uma prancha chata — uma panqueca de matéria. Na altura em que foi proposto o modelo da panqueca não existiam, contudo, partículas quânticas conhecidas que tivessem as propriedades adequadas para formar as panquecas.

Este modelo da panqueca foi revitalizado em 1980, quando os físicos propuseram que os neutrinos — minúsculas partículas quânticas — tivessem uma pequena massa e fossem um candidato viável para a matéria escura. Zel'dovich argumentou que as flutuações de densidade num gás de neutrinos com uma certa massa, impregnando o universo, cresceriam, formando grânulos com o tamanho de superaglomerados na bola de fogo primordial, os quais entrariam em colapso para acabar em panquecas de neutrinos. Mais tarde os átomos de hidrogénio, formados cerca de 300.000 anos depois do *big bang*, seriam atraídos gravitacionalmente e para as panquecas de neutrinos preexistentes, formando, assim, um gás denso e quente. Este protoaglomerado de hidrogénio gasoso fragmentar-se-ia muito mais tarde em grânulos do tamanho de galáxias, através de uma variedade de processos físicos complicados. As galáxias seriam, assim, neste modelo de panqueca, atores relativamente recém-chegados ao palco cósmico.

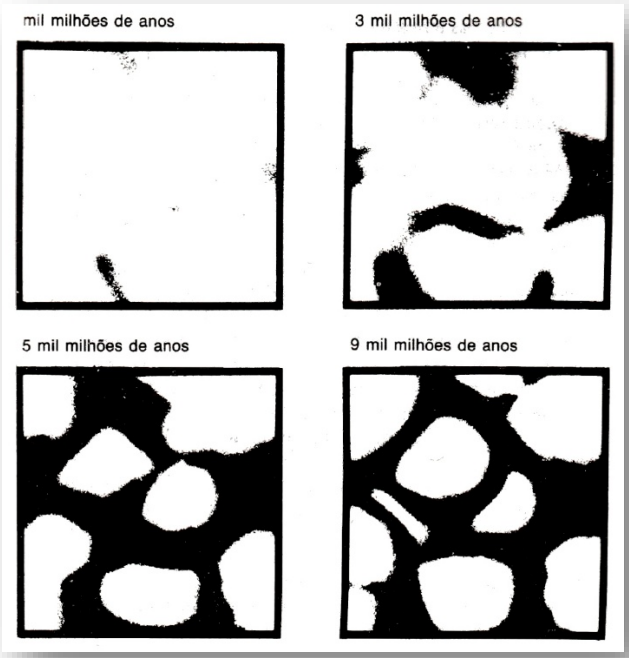
Esta conclusão acaba por tornar-se um problema para o modelo da panqueca. As galáxias têm, aproximadamente, a mesma idade que o próprio universo — são velhas, e não jovens. Um outro problema consiste em que as experiências recentes com partículas quânticas indicaram que, pelo menos, um dos neutrinos não tem massa suficiente para fazer triunfar o cenário dos grânulos em forma de panqueca. Estes problemas têm esmorecido o entusiasmo pelo modelo da panqueca.

A alternativa ao modelo do protoaglomerado em forma de panqueca é a possibilidade de as galáxias se terem formado

primeiro e só recentemente terem começado a agrupar-se em aglomerados maiores. Se a matéria escura se formou dentro dos grânulos mais pequenos do tamanho de galáxias, então deve consistir em partículas quânticas que ainda nem sequer foram detetadas. É pouco provável que estas partículas imaginadas — os candidatos oriundos da teoria das partículas são os gravitinos, fotinos e axiões — venham a ser detetadas diretamente em experiências terrestres, pois interagem muito fracamente com a restante matéria. A única maneira de as «ver» é como a matéria escura, que se conjeturou existir nas galáxias, o que a muitos parece ser batota — inventar uma partícula, que de outra maneira seria indetetável, para realizar a tarefa da formação das galáxias.

Mas, evidentemente, existe matéria escura nos halos das galáxias, que poderá muito bem ser um gás dessas partículas quânticas imaginadas, as quais poderiam formar grânulos de matéria escura no universo primitivo, com as dimensões exatas para gerar subsequentemente as galáxias. Os grânulos de matéria escura com as dimensões de uma galáxia congregar-se-iam gravitacionalmente mais tarde, levando as galáxias consigo, para dar aglomerados e superaglomerados, criando filamentos e buracos.

Qual dos pontos de vista — «aglomerados primeiro, galáxias depois», ou «galáxias primeiro, aglomerados depois» — é o correto? As cabeças pensantes não estão de acordo. Todavia, tal desacordo é sinal de progresso, porque, tal como Peebles fez notar, «só recentemente o assunto progrediu até ao ponto em que é possível tomar posições que parecem defensáveis».



Estas simulações de computador da formação de panquecas ilustram uma teoria da evolução das galáxias. A matéria escura (possivelmente, neutrinos com massa, se é que existem) entra em colapso gravitacional, formando grânulos com a massa de superaglomerados de galáxias. O colapso não é esfericamente simétrico, formando, em vez disso, folhas, ou panquecas de grande massa, que mais tarde se fragmentam em galáxias. Este modelo representa apenas uma tentativa de explicar a distribuição das galáxias que presentemente se observa.

Para resolver esta questão são necessários mais dados observacionais. O telescópio espacial começará a sondar o universo primitivo para além da Era dos quasares, aguardando-se que apareçam novas informações sobre a evolução inicial das galáxias. Talvez o enigma de saber se foram primeiro as galáxias que se formaram e depois os aglomerados, ou primeiro os aglomerados e depois as galáxias, possa ser resolvido.

No entanto, por mais que se recue no tempo com um telescópio óptico e por mais poderosos que sejam os instrumentos, não se pode recuar no tempo para além dos primeiros 300.000 anos. O universo anterior a esse instante é opaco. Não há esperança de se

poderem ver as flutuações iniciais na bola de fogo primordial que veio a dar origem às galáxias e que estão fora de vista, escondidas no *big bang*. Contudo, a sua estrutura pode ser a pista para toda a evolução galática subsequente e para a ordem do universo em grande escala.

Não sabemos por que é granuloso o universo — com pouca granulosidade certamente, mas muito importante. As galáxias, as estrelas, os planetas e a própria humanidade são parte dessa granulosidade. Os físicos teóricos, guiados pela experiência e pela observação, extrapolaram as teorias quânticas correntes para energias comparáveis às existentes antes dos primeiros microssegundos do *big bang*. Aqui especulam acerca dessas sementes quânticas iniciais, a partir das quais cresceram as galáxias — especulações que estão no limite da investigação corrente. Referir-me-ei ao seu trabalho num capítulo posterior.

Fritz Zwicky, que ajudou a descobrir os grandes aglomerados de galáxias, notou uma vez que elas eram «o último degrau no estudo do universo como um todo». O estudo da totalidade do universo chama-se «cosmologia» e representa o ponto terminal do pensamento humano acerca da ordem em grande escala do espaço e do tempo. Voltemo-nos agora para este assunto.

---

[1] Observatório Radioastronómico Nacional. (*N. do T.*)

CAPÍTULO 7

– Cosmologia Clássica –

A estrutura do mundo tem o centro em toda a parte e a circunferência em parte alguma.

CARDEAL NICOLAU DE CUSA (século XV)

Nas *Lectures on Ethics* o filósofo Ludwig Wittgenstein descreve uma experiência peculiar, dizendo:

> Penso que a melhor maneira de a descrever é dizer que, quando a tenho, *me maravilho com a existência do universo.* E estou então inclinado a usar frases como «é extraordinário que qualquer coisa exista» ou «é extraordinário que o mundo exista».

O filósofo existencialista Martin Heidegger vê estas experiências como refletindo o problema fundamental da metafísica: «Por que existe o ser? Por que não antes o nada?» Esta questão articula o enigma posto pelo facto da existência: quem precisa da complexidade do mundo em face da simplicidade do nada absoluto?

Refletindo sobre esta questão, descubro que provoca emoções ambivalentes. Sinto ansiedade quando imagino o insondável abismo silencioso do nada que poderia ter sido e experimento uma completa admiração quando tomo conhecimento do facto

misterioso, embora simples, de que o universo existe. Não sei a resposta à pergunta do filósofo, nem estou convencido de que tenha alguma resposta simples. Mas é o tipo de questão que nos faz pensar na estranheza da realidade, que nos leva ao limiar da loucura e nos diz para «acordarmos». A estranheza da existência concreta é um sentimento que nunca nos abandona completamente, por mais familiar ou confortável que o mundo se torne para nós. E ainda bem que assim é. Esse estranho sentimento, provocado pelo nosso reconhecimento da realidade da existência, rejuvenesce a nossa admiração e curiosidade acerca do universo, recriando a criança em cada um de nós. E essa admiração é o princípio da ciência.

Quando era criança, olhava muitas vezes para o céu estrelado, mas não via o que os astrónomos nos dizem lá existir. O que via era uma abóbada celeste, uma gigantesca tampa esférica negra, cobrindo a superfície da Terra. Nesta tampa opaca havia orifícios que revelavam o fogo celestial do outro lado — orifícios que apareciam como estrelas. Lentamente, à medida que a noite avançava, a abóbada, com os orifícios brilhantes e bruxuleantes, movia-se através do céu. Quase podia vê-la mover-se. Em noites claras de luar julgava ver o luar refletido pela abóbada negra, como se fosse a luz de um ecrã de cinema refletida pelas paredes da sala de projeção. O Sol e a Lua eram mantidos nas suas posições por outras esferas transparentes que os transportavam pelo céu.

Era esta a minha teoria infantil do universo. Não me lembro de que alguém me tivesse falado desta conceção dos céus, mas depois de ler Aristóteles, o filósofo grego da antiguidade, e Ptolomeu, o

astrónomo alexandrino, reconheci a minha cosmologia na deles. Aristóteles via o universo como um sistema de 56 esferas transportando os corpos celestes. Ptolomeu desenvolveu este sistema, relacionando-o com observações quantitativas. A cosmologia ptolomaica, apesar dos seus complicados ciclos e epiciclos, ajustava-se ao senso comum tão completamente que dominou a cosmologia ocidental durante um milhar de anos, até ao tempo de Copérnico. E isto não é nenhum produto da ignorância. Esquecemos hoje em dia que são a física de Aristóteles e a cosmologia de Ptolomeu que correspondem ao senso comum, e não a física de Newton ou a cosmologia de Copérnico, que são já grandes abstrações da nossa experiência quotidiana.

A palavra «teoria» deriva do grego «contemplar» — uma teoria é uma imagem da realidade. A minha teoria infantil dos céus é um exemplo. Embora não existam esferas celestes, como pensava, esta «teoria» trouxe à minha experiência uma coerência satisfatória, que não era parte da mera perceção. É este o grande poder das teorias enquanto imagens da realidade — organizam a nossa experiência de novas maneiras e tornam inteligível a complexidade das nossas perceções. Mas estaremos a abusar desse poder das teorias se confundirmos a nossa imagem da realidade com a própria realidade, isto é, se confundirmos o mapa com o terreno. Nunca devemos esquecer — como quase sempre fazemos na prática — que as nossas teorias são apenas mapas da realidade. Todas as teorias vindas das ciências naturais, a despeito da sua coerência, poder de previsão e profundidade de visão, mais não são do que maneiras de descrever a realidade material, e não a realidade em si mesma. Por exemplo, de acordo com a teoria de Newton, as leis do movimento

são equações diferenciais. No entanto, os planetas, ao moverem-se nas suas órbitas em torno do Sol, não estão a resolver as equações diferenciais que lhes dizem como mover-se, nem tampouco se encontram ligados a esferas celestes. Estão, simplesmente, a mover-se. Fomos nós que inventámos a teoria, resolvemos as equações diferenciais e vemos se a nossa imagem corresponde ou não à realidade. Os planetas não têm tais problemas.

As teorias físicas e as cosmologias podem ser pensadas como mapas da realidade. Mas isto não implica que tais mapas sejam invenções arbitrárias. Alguns são de longe melhores e mais rigorosos do que outros, cingindo melhor o território da realidade. Podemos comparar mapas diferentes e verificar quais são os melhores. Por exemplo, o «mapa» aristotélico dos céus, tal como o imaginava na minha meninice, falha logo que começamos a examinar o universo mais de perto. O mapa aristotélico é substituído pelo newtoniano, este, por sua vez, substituído pelo einsteiniano. À medida que os cientistas mais aprendem sobre a realidade através da observação e da experimentação, os mapas mudam, acomodando-se às novas descobertas. Por vezes, as descobertas experimentais são tão invulgares que não existe maneira de descrever o novo território com os mapas existentes. Então as próprias regras de fazer mapas têm de ser mudadas, como foi feito no princípio deste século com o advento da relatividade e da teoria quântica. Contudo, continuamos a fazer mapas e a realidade continua a existir. É importante ter isto em mente quando discutirmos cosmologia. Todas as cosmologias são modelos do universo, não o próprio universo.

Segundo o célebre cosmólogo inglês Herman Bondi, «a cosmologia é o campo do pensamento que trata da estrutura e da história do universo como um todo». Mas como podemos tratar o universo como um todo se ele, por definição, inclui todos os corpos físicos? Embora possamos olhar para as galáxias distantes «do exterior», não podemos olhar para o universo do exterior, porque não existe o exterior do universo. Faz sentido falar acerca das posições relativas das estrelas e das galáxias no espaço, mas não tem significado falar acerca da posição do universo no espaço. É o espaço do universo infinito ou finito? Se é finito, onde está o seu limite? Teve o universo uma origem no tempo? Se for assim, que é que existia antes do universo? Como acabará o universo? É claro que, ao considerarmos todo o universo no espaço e no tempo, estamos a considerar uma entidade de uma espécie completamente diferente, e não apenas outro objeto astronómico maior. Consequentemente, deveremos talvez invocar conceitos novos e pouco familiares se quisermos elaborar um quadro para o nosso pensamento e tentar responder às questões que levantámos.

Há séculos que os pensadores refletem sobre estas questões. O que é, porém, notável é que só nas últimas décadas a cosmologia — outrora uma área do pensamento meramente especulativa — se tornou uma ciência empírica. Vários modelos cosmológicos com os quais se tentou responder às perguntas cósmicas estão hoje a ser submetidos ao teste da observação. Por exemplo, o velho modelo estático do universo, que sustentava ser o universo infinitamente velho e essencialmente inalterável no tempo, pode hoje ser afastado (ou, pelo menos, tornado extremamente pouco provável)

pelas observações astronómicas. A capacidade de excluir modelos bem definidos é a marca da ciência empírica amadurecida.

Dois importantes avanços transformaram a cosmologia de ciência especulativa em ciência empírica. O primeiro avanço, no campo teórico, foi a criação por Einstein da teoria da relatividade geral, uma teoria abrangente do espaço, tempo e matéria, que forneceu um novo esquema conceptual para pensar acerca do universo como um todo. O segundo avanço que trouxe a cosmologia até à sua forma moderna foi a instalação de novos e poderosos instrumentos astronómicos — os grandes refletores e os radiotelescópios. A teoria de Einstein não seleciona uma cosmologia específica ou uma estrutura do universo; fornece o esquema, não os pormenores. Decidir qual a verdadeira estrutura do universo no espaço e no tempo exige, como sempre, observações pormenorizadas e para estas são necessários novos instrumentos.

Durante as primeiras décadas deste século, à medida que os astrónomos penetravam mais fundo no espaço, continuavam a ver uma hierarquia de estruturas cada vez maiores: das estrelas às galáxias, destas aos aglomerados de galáxias, tudo em expansão com o universo. Todavia, nas últimas décadas os astrónomos exploraram a estrutura do próprio universo e descobriram que esta estrutura hierárquica de concentrações cada vez maiores acaba por parar. Às maiores escalas de distâncias, de centenas de milhões de anos-luz, começamos a ver a homogeneidade do universo, que parece ser uma textura global do cosmos, e não apenas uma propriedade local da nossa região do espaço. Pela primeira vez estamos a observar propriedades espaciais do universo como um

todo. O estudo deste espaço em grande escala, regular e homogéneo, da sua evolução no tempo e de como influencia a matéria que contém, constitui o verdadeiro domínio científico da cosmologia contemporânea.

Examinemos agora mais pormenorizadamente estes dois principais fios condutores da cosmologia: o teórico e o observacional, começando pela teoria moderna do espaço e do tempo.

Esqueçamos de momento tudo acerca do espaço físico real e tentemos imaginar espaços tridimensionais puros e vazios. Suponhamos que nos encontramos num foguetão mergulhado no espaço profundo e antes de pormos a trabalhar o motor do foguetão deixamos a flutuar no espaço um dispositivo laser para indicar a nossa posição. Este dispositivo emite um raio de luz, que seguimos fielmente, nunca voltando para trás. Decorrido um certo tempo veríamos o dispositivo laser reaparecer à nossa frente. Pareceria que tínhamos descrito uma circunferência. A seguir tentamos uma direção diferente, mas acontece a mesma coisa. Este espaço não é, claramente, um espaço vulgar, no qual, se nos afastarmos sempre de um ponto, nunca mais a ele voltaremos. Este é um exemplo de um espaço não euclidiano e, embora pareça estranho, é matematicamente possível.

A descrição matemática de tais espaços «curvos» foi inicialmente apresentada, com toda a generalidade, pelo grande matemático oitocentista alemão Bernhard Riemann. Pensamos normalmente que o espaço físico vazio é plano, de maneira que, se usássemos

raios laser para formar os lados de triângulos, cubos e outras figuras geométricas, estas obedeceriam aos teoremas da geometria euclidiana. Se partíssemos num foguetão segundo uma linha reta, nunca mais voltaríamos. Contudo, o trabalho de Riemann generalizou a ideia de espaço de forma a incluir também a possibilidade da geometria não euclidiana, na qual o espaço não é plano, mas possui uma curvatura. Seria como generalizar espaços bidimensionais de forma a incluir não somente o espaço plano de uma folha de papel, como também superfícies curvas, como a de uma pêra. Riemann mostrou como a curvatura geométrica do espaço não euclidiano era completamente descrita por um objeto matemático chamado «tensor de curvatura». Em princípio, usando raios laser no espaço tridimensional e medindo ângulos e distâncias com eles, podemos determinar o tensor de curvatura de Riemann em cada ponto do espaço.

A beleza e o poder do trabalho geométrico de Riemann estão em ter estabelecido sobre sólidos alicerces matemáticos a descrição de curvas arbitrariamente complicadas em qualquer número de dimensões espaciais. Podemos facilmente visualizar a maior parte dos espaços curvos a duas dimensões, como a superfície de uma esfera ou de um donut. Mas já não somos capazes de visualizar espaços tridimensionais curvos. No entanto, os métodos matemáticos de Riemann mostram-nos como tratar tais espaços — a matemática pode guiar-nos mesmo até onde a imaginação visual já não chega.

A matemática dos espaços curvos e suas geometrias não euclidianas associadas foi bastante bem compreendida pelos

matemáticos das primeiras décadas do século XX. Mas estes progressos parecem situar-se no ramo das matemáticas «puras» e o impacto destas ideias no mundo físico real só tomou forma com a invenção da teoria da relatividade geral de Albert Einstein, em 1915-1916.

A teoria da relatividade geral nasceu da teoria da relatividade restrita de Einstein, de 1905, que estabelecia uma nova cinemática para a física: as leis de transformação que relacionavam as medições do espaço e do tempo feitas por um observador com as feitas por outro movendo-se com velocidade constante em relação ao primeiro — os chamados «observadores inerciais». A cinemática de Einstein aproximava-se das leis da cinemática newtoniana para velocidades relativas pequenas, diferindo, no entanto, consideravelmente quando a velocidade relativa dos observadores era próxima da velocidade da luz. Na cinemática newtoniana as leis de transformação para as medições do espaço e do tempo estão separadas; a transformação do tempo é independente da posição e velocidade relativa dos dois observadores inerciais. Assim, o tempo pode ter um significado universal «absoluto» para todos os «observadores inerciais». Todavia, na cinemática de Einstein as medições do espaço e do tempo tornam-se intimamente ligadas; o tempo não é absoluto, mas relativo a um observador inercial particular.

Einstein, raciocinando a partir de dois postulados gerais — o movimento uniforme absoluto de um observador inercial é indetetável e a velocidade da luz é uma constante absoluta —,

deduziu corretamente as novas leis das transformações do espaço-tempo.

Foi, porém, o matemático Hermann Minkowski que deu a interpretação geométrica dessas transformações, mostrando que, se não encarássemos as três dimensões do espaço e a dimensão do tempo como entidades separadas, antes, pelo contrário, as juntássemos num espaço-tempo a quatro dimensões, então as novas transformações de Einstein poderiam simplesmente ser vistas como correspondendo a rotações levadas a cabo neste espaço-tempo a quatro dimensões. Foi uma enorme simplificação, que criou uma nova perspetiva do espaço e do tempo. Como Minkowski comentou em 1908, «daqui para a frente o espaço em si e o tempo em si estão condenados a desaparecer nas sombras, apenas uma espécie de união entre os dois conservará uma realidade independente».

Inicialmente, Einstein não atribuiu importância à interpretação de Minkowski, pensando que teria apenas um significado matemático formal e nenhum significado físico real. Mudou, porém, de opinião, adotando completamente o ponto de vista quadridimensional, quando mais tarde se dedicou a criar a teoria da relatividade geral.

Ao tentar generalizar o princípio da sua primeira teoria da relatividade restrita de forma a incluir a gravidade, Einstein intuía já que para generalizar o princípio da relatividade seria necessário considerar geometrias não euclidianas. Com o auxílio do amigo e antigo colega Marcel Grossmann, que era matemático, Einstein

estudou a geometria de Riemann, na qual aprendeu a linguagem matemática adequada para expressar as suas intuições físicas acerca da gravidade e da geometria. Após mais de uma década de luta e de frustração, que teria feito desistir um homem de menor estatura, Einstein conseguiu, finalmente, obter um conjunto de equações — as equações de Einstein — que exprimiam a curvatura do espaço-tempo a quatro dimensões em função da matéria existente nesse espaço. As equações de Einstein realçavam conceitos como invariância, simetria e geometria, que tiveram um profundo impacto no desenvolvimento posterior da física. A gravidade foi reduzida à geometria. A matéria, o espaço e o tempo tornaram-se unificados. O espaço não euclidiano deixou de ser uma curiosidade intelectual para se transformar numa descrição correta da realidade.

As equações de Einstein, que previam um encurvamento do espaço-tempo na vizinhança do Sol, foram apoiadas por observações alguns anos depois. Recentemente, foram verificadas experimentalmente com maior precisão através do uso de satélites artificiais e de métodos de determinação de distâncias por meio de radar, tendo sido maravilhosamente confirmadas. Mas, como Einstein compreendeu enquanto inventava a teoria, as suas equações descrevem não somente a curvatura do espaço-tempo na vizinhança do Sol, mas também a curvatura do espaço-tempo do universo inteiro, constituindo o quadro conceptual da cosmologia moderna.

Einstein foi o primeiro a aplicar as equações aos problemas da cosmologia. Como a maior parte dos físicos daquele tempo,

acreditava que o universo era estático e imutável (isto antes da descoberta da expansão do universo) e, portanto, pressupôs tal facto quando procurou soluções para as suas equações. Resolveu-as então e chegou à conclusão, aparentemente absurda, de que o universo mudava. Decidiu então alterar as equações, acrescentando-lhes um termo novo, o «termo cosmológico», encontrando, assim, em 1917, um espaço estático e curvo, cheio de um gás uniforme de matéria com pressão nula, que era a solução que procurava. O universo não se movia. Se se tivesse ficado pelas suas equações iniciais sem o «termo cosmológico» e tivesse procurado soluções que evoluíssem com o tempo, Einstein poderia ter antecipado a descoberta do universo dinâmico e em expansão.

No mesmo ano de 1917 o astrónomo holandês Willem de Sitter encontrou ainda outra solução para as equações de Einstein com o termo cosmológico, mas sem matéria — um universo vazio. A solução de Willem de Sitter, para as equações de Einstein podia ser interpretada como um espaço em expansão, idêntico à superfície esticada de uma bola de borracha.

Existiam, portanto, dois modelos cosmológicos baseados nas equações de Einstein: a cosmologia de Einstein, que concebia um espaço estático cheio de matéria, e a cosmologia de Willem de Sitter, que concebia um espaço em expansão sem matéria. A solução de Willem de Sitter, uma vez que representava um universo vazio, mas em expansão, parecia a muita gente perfeitamente absurda. Era opinião comum na altura que as equações de Einstein não tinham feito muita luz sobre os problemas cosmológicos.

Em 1922-1924, porém, o meteorologista soviético Alexander Friedmann, mais tarde professor de Matemática na Universidade de Leninegrado, encontrou as soluções dinâmicas para as equações originais de Einstein (sem o termo cosmológico), que hoje acreditamos descreverem corretamente a cosmologia. O trabalho de Friedmann, infelizmente, foi completamente ignorado, embora os seus artigos tenham sido publicados numa importante revista científica e Einstein tivesse tomado conhecimento deles.

A cosmologia só adquiriu o seu moderno quadro conceptual quando em 1927 Georges Lemaître, «o pai do *big bang*», redescobriu independentemente as mesmas equações. Mas o trabalho de Lemaître foi também ignorado até o prestigiado astrónomo Arthur Eddington sublinhar a sua importância em 1930.

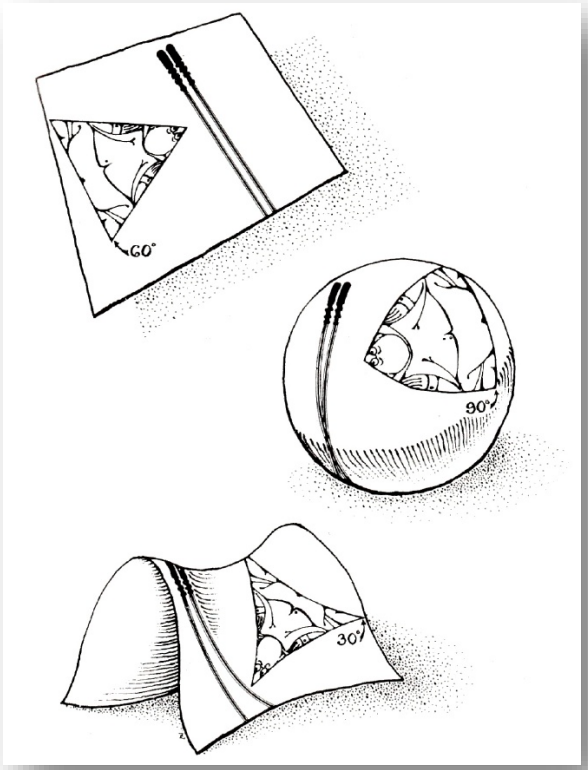
As contribuições subsequentes foram de natureza matemática. Howard P. Robertson e Arthur Walker, dois matemáticos, mostraram que as soluções de Friedmann eram, de facto, as mais gerais para as equações de Einstein, desde que se supusesse que o universo era espacialmente homogéneo e isotrópico. Demonstraram ainda que, neste caso, o espaço-tempo a quatro dimensões podia ser separado num espaço curvo tridimensional e num tempo único comum a todos os observadores «co-móveis». Hoje os modelos cosmológicos modernos baseados nestas soluções de Friedmann-Robertson-Walker são muitas vezes referidos como cosmologias FRW. Antes de os examinarmos em pormenor, vejamos por que razão são tão atraentes do ponto de vista moderno. Tal tem a ver com uma notável propriedade do universo observável.

O universo parece ser uma estrutura complicada, com planetas, estrelas e galáxias. No entanto, estes sistemas organizados surgem a escalas de distância reduzida relativamente à escala do próprio universo. À medida que observamos escalas cada vez maiores, o universo torna-se mais homogéneo — o aspeto granuloso tende a diminuir. É como olhar para a superfície da Terra, que apresenta inúmeras estruturas e grânulos, da altitude do voo de um avião a jato — vemo-la bastante regular. Tais observações levam-nos a admitir que o universo, quando visto a distâncias muito grandes, é simultaneamente homogéneo — parece o mesmo independentemente da posição de cada um dentro dele — e isotrópico — parece o mesmo em todas as direções. Um espaço que é isotrópico para todos os observadores, e não unicamente para um, é também homogéneo.

Uma prova adicional em apoio destas hipóteses veio da deteção em 1965 da radiação de micro-ondas de fundo — o calor deixado pelo *big bang*, que é o ente mais antigo até agora detetado no universo. Dentro dos limites dos erros de observação, esta radiação de fundo tem uma distribuição perfeitamente isotrópica, indicando que o universo era extremamente isotrópico já na altura do *big bang*.

Muita gente considera a homogeneidade e a isotropia do universo racionalmente satisfatórias, porque implicam que nenhum local do universo é privilegiado ou especial em qualquer aspeto que seja. A alternativa seria supor que há locais privilegiados, e então teríamos de perguntar por que razão um lugar era privilegiado e não outro qualquer. Mas nem sequer é preciso fazer essa pergunta se, como

disse Einstein, «todos os lugares do universo são semelhantes». Esta atraente ideia foi distinguida com um princípio, o «princípio cosmológico», como foi chamado pelo cosmólogo Edward Milne em 1933.



Análogos bidimensionais dos três possíveis espaços FRW tridimensionais, homogéneos e isotrópicos. Aqui um cientista bidimensional explora estes espaços. No desenho de cima — espaço plano e infinito — os raios laser paralelos nunca se encontram e a soma dos ângulos de um triângulo é de 180º. No desenho do meio está representado o espaço finito de curvatura positiva constante, no qual os raios laser paralelos convergem e se encontram e onde a soma dos ângulos de um triângulo é superior a 180º. No desenho de baixo — espaço hiperbólico infinito de curvatura negativa constante — os raios laser paralelos divergem e a soma dos ângulos de um triângulo é inferior a 180º. O espaço bidimensional de curvatura negativa constante (diferentemente dos outros dois espaços) não pode ser mergulhado num espaço tridimensional, como se tentou fazer nestes desenhos. Por esta razão, não parece ser homogéneo e isotrópico e tem um lugar especial -— o ponto de sela.

O princípio cosmológico remonta a um famoso aforisma de Nicolau de Cusa, filósofo e teólogo do século XV:

A estrutura do mundo tem o centro em toda a parte e a circunferência em parte alguma.

No seu modelo de sistema solar Copérnico retirou a Terra do centro, que, assim, deixou de ser um planeta privilegiado. Shapley mostrou, séculos mais tarde, que também o Sol não está num lugar privilegiado: estamos longe do centro da nossa galáxia. Hoje em dia sabemos até que a nossa galáxia não ocupa nenhum lugar especial entre os milhões de galáxias observados. Parece não *existir* nenhum lugar «especial».

Contudo, o princípio cosmológico, enquanto proposição científica, pode estar errado. Por exemplo, o universo inteiro e todas as galáxias nele existentes podem estar a rodar. O universo teria então um eixo de rotação, uma direção preferida, e não seria isotrópico. O princípio cosmológico e a homogeneidade e isotropia do universo, que ele implica, estão sujeitos a refutação dentro do universo observado. Atualmente, a maior parte dos dados são, porém, a seu favor.

As cosmologias FRW incorporam o princípio cosmológico, razão por que são tão atraentes. Admitindo, em conformidade com o princípio cosmológico, que o espaço tridimensional é homogéneo e isotrópico, Robertson e Walker demonstraram matematicamente que só poderiam existir três espaços geométricos com essas propriedades. Dois destes correspondem às soluções das equações de Einstein, que Friedmann tinha já encontrado, o que não é surpreendente. Os três espaços eram o espaço plano de curvatura zero (que Friedmann não tinha encontrado), o espaço esférico de

curvatura positiva constante e o espaço hiperbólico de curvatura negativa constante. No espaço plano raios laser paralelos nunca se encontram: é um espaço aberto de volume infinito. No espaço esférico raios laser paralelos convergem: é um espaço fechado com volume finito; neste espaço podemos voar em linha reta e voltar ao ponto de partida. No espaço hiperbólico raios laser paralelos divergem: o espaço é aberto e tem volume infinito.

Utilizando as equações de Einstein para analisar estes espaços, conclui-se que a curvatura varia no tempo. No espaço plano, para o qual a curvatura espacial é zero, a escala relativa das medidas de espaço e tempo varia. Destas soluções dinâmicas das equações de Einstein podemos apenas concluir que o universo não pode ser estável — deve evoluir, expandindo-se ou contraindo-se — e que o espaço do universo está a dilatar-se.

Estas soluções anteciparam, assim, a lei de Hubble, que acarreta a expansão do universo. De facto, se as galáxias estão colocadas num espaço em expansão, então afastar-se-ão também — o chamado «deslocamento de Hubble». A descoberta de Hubble trouxe, assim, grande apoio observacional às cosmologias FRW e ao universo dinâmico.

Não resolve, no entanto, o problema de saber em qual dos três espaços, plano, esférico ou hiperbólico, vivemos. Responder a esta pergunta é muito difícil. Porém, o destino do universo gira em torno desta resposta, porque as geometrias plana e hiperbólica podem corresponder a universos abertos, que continuam a expandir-se para sempre, enquanto o universo esférico e fechado acaba por

parar de expandir-se e começa a contrair-se: tem um tempo de vida finito. Uma grande variedade de métodos de observação foi proposta para tentar resolver o problema, mas nenhum alcançou grande êxito. Mencionarei dois deles.

O primeiro método consiste em tentar observar e medir a taxa de desaceleração da expansão do universo. Os dados de Hubble eram consistentes com uma taxa de expansão constante, mas as três cosmologias FRW implicam que, embora muito lentamente, a expansão do universo está a desacelerar-se. Medindo essa desaceleração com rigor, poder-se-ia determinar qual das três geometrias FRW se aplica ao mundo real e descobrir se o universo se expandirá para sempre ou se acabará por entrar em colapso. A medição da desaceleração — o desvio em reação à taxa de expansão constante no tempo — requer observações cuidadosas das galáxias mais distantes e que se verifique se há ou não qualquer variação na constante de proporcionalidade de Hubble entre a velocidade de recessão das galáxias distantes (medida a partir do desvio para o vermelho) e a distância (medida por outros meios).

Infelizmente, estimar distâncias a galáxias muito distantes é extremamente difícil. Esta dificuldade é, em parte, devida ao facto de as galáxias estarem, provavelmente, em evolução, alterando a luminosidade de modo desconhecido, pelo que utilizar os «luzeiros padrão» para medir distâncias através da luminosidade se torna pouco fiável. Consequentemente, não se pode determinar a variação da «constante de Hubble» ao longo do tempo nem achar a taxa de desaceleração que revelaria ser o universo aberto ou fechado.

Um segundo método para determinar se o universo está destinado a entrar em colapso consiste em estabelecer previamente a densidade média de massa do universo — a densidade da matéria do universo, se pudéssemos distribuí-la uniformemente. De acordo com as cosmologias FRW, existe um parâmetro crucial chamado Ω, a razão entre a densidade média de matéria observada no universo e a «densidade crítica» de cerca de 10^{-29} por centímetro cúbico, aproximadamente 10 átomos de hidrogénio por metro cúbico. Portanto, se conhecêssemos a densidade média da matéria, poderíamos também saber o valor de Ω. Se Ω tiver valor inferior a 1, então a geometria tem de ser hiperbólica; se for igual a 1, plana; se for superior a 1, esférica. Conhecendo o valor de Ω, saberíamos também se o universo entrará ou não em colapso.

O problema de determinar a densidade média da matéria reside no facto de a matéria do universo consistir tanto em matéria visível, como as estrelas e as galáxias, como em matéria invisível, como a matéria escura dos buracos negros ou das partículas quânticas microscópicas. As partes visíveis e luminosas das galáxias dão um valor de Ω de cerca de um centésimo. Só se pode medir diretamente a densidade da matéria visível. Se admitirmos que 90% da massa de uma galáxia são matéria escura, isto implicará um valor de Ω de cerca de um décimo. E no conjunto dos aglomerados de galáxias os astrónomos encontram um valor do parâmetro Ω compreendido entre um e dois décimos. Se isto esgotasse a questão, concluiríamos que vivemos num universo aberto e hiperbólico. Infelizmente, não é possível chegar a nenhuma conclusão simples, dada a possibilidade de haver mais matéria escura. Como vimos, existem fortes indícios da existência de tal matéria. De facto, a

componente material dominante do universo pode muito bem ser a matéria escura e a componente visível, as galáxias e as estrelas, apenas uma parte insignificante da massa total do universo.

A conclusão é que hoje não temos maneira segura de saber se o universo é aberto ou fechado, o que preocupa muita gente, ansiosa por conhecer o destino do universo. Todavia, na minha opinião, essas pessoas estão preocupadas com um falso problema. Os astrónomos sabem já que o parâmetro cósmico crucial, Ω, é superior a um décimo e inferior a dois décimos — uma gama de valores relativamente próximos da unidade. Por que razão é isto assim? A grandeza Ω poderia ter um valor qualquer, poderia ser 50, ou um milésimo. O enigma real consiste em saber por que razão tem um valor tão próximo de 1. Por que razão estamos colocados na fronteira entre o universo fechado e o universo aberto? É este o problema real, que iremos ponderar em futuros capítulos dedicados ao universo primitivo.

Embora não saibamos abordar questões de pormenor acerca da geometria do universo, fizemos notáveis progressos na compreensão da sua evolução ao longo do tempo. A maior parte dos cientistas mantêm hoje que o universo evoluiu a partir de um gás denso e quente de partículas quânticas que depois se expandiram rapidamente — uma explosão chamada «*big bang* quente». Tudo no universo é remanescente dessa explosão. Mas tal consenso acerca da evolução do universo nem sempre existiu.

Até há pouco tempo os cientistas estavam divididos entre os que, no espírito da cosmologia original de Einstein-de Sitter, acreditavam

que o universo se encontrava num estado estacionário, que se estendia desde o passado infinito até ao futuro infinito, e os que, de acordo com a cosmologia de Friedmann-Lemaître, acreditavam que o universo fora extremamente diferente no passado e tivera uma origem definida. É difícil imaginar dois pontos de vista mais opostos. Esta oposição foi extremamente importante para o nascimento da cosmologia como ciência empírica. A necessidade de esclarecer a questão não só promoveu a pesquisa de dados cosmológicos significativos, como também levou os proponentes de cada solução a efetuar cálculos complexos para a defender, cálculos esses que no fim provaram ser mais valiosos do que a própria solução que defendiam. Examinemos a dialética, não muito subtil, entre os defensores do *big bang* e os adeptos do «estado estacionário».

Georges Lemaître é justamente considerado o «pai do *big bang*», por já na década de 30 ter feito notar que as equações de Einstein implicavam que o universo deveria ter começado num estado de matéria muito densa — um «átomo primordial», como lhe chamou. Todavia, a visão moderna do *big bang* começou quando o físico George Gamow, tendo aceitado o desafio de Arthur Eddington de encontrar um local mais quente do que o centro de uma estrela, iniciou o estudo do universo primitivo. Como Lemaître, percebeu que, se recuássemos no tempo, o universo contrair-se-ia e a matéria dentro dele seria comprimida, tornando-se muito quente, mais quente do que no interior de uma estrela, o que implicava que os núcleos atómicos poderiam ser sintetizados no *big bang*, da mesma maneira que o são nas estrelas. Ajudado por dois estudantes, Ralph Alpher e Robert Herman, calculou então a

maneira como os núcleos atómicos poderiam ser «cozinhados» no *big bang* quente a partir do hidrogénio, o mais simples dos núcleos.

Contudo, a sua ideia de que a maior parte dos elementos mais pesados eram sintetizados no *big bang* acabou por se revelar errada (os elementos pesados são sintetizados no interior das estrelas ou nas supernovas). No entanto, por estarem a enfrentar o problema do universo primitivo, começaram também a debruçar-se sobre o calor deixado pelo *big bang*. Concluíram que o calor da explosão do *big bang* devia existir ainda, porque, ao contrário do que se passa com o calor de um fogo ou de uma estrela, não tem lugar para onde possa escapar-se — não existe um «lado de fora» do universo. O calor manifestar-se-ia como um banho de radiação de baixa temperatura, impregnando o universo inteiro. Mais ainda, a temperatura podia ser estimada. Alpher e Herman resumiram a sua profética conclusão em 1948: «A temperatura no universo, atualmente, é de cerca de 5º K.» Esta afirmação, baseada na teoria do *big bang*, deve ser comparada com as afirmações de A. A. Penzias e R. W. Wilson, que resumiram os resultados das observações feitas dezassete anos mais tarde:

> As medições da temperatura efetiva do pico de ruído [...] a 4080 MHz apontam para um valor de cerca de 3,5º K mais alto do que o esperado. O excesso de temperatura é, dentro dos limites das nossas observações, isotrópico.

Esta observação direta da radiação do *big bang* — pelo menos é esta a interpretação mais simples — era a prova concludente da teoria do *big bang*.

Mas antes de se dispor desses dados muitos cientistas consideraram muito atrativa a teoria do estado estacionário, inventada por Herman Bondi, Thomas Gold e Fred Hoyle em 1948. A ideia base era a de que, à medida que o universo se expande, nova matéria estaria contínua e espontaneamente a ser criada no espaço aberto entre as galáxias. Esta nova matéria viria a formar novas estrelas e galáxias. Os autores do modelo mostraram que a taxa de criação contínua de matéria no vazio do espaço era tão pequena que não havia conflito com qualquer observação. Com base neste raciocínio concluíram que, a despeito da expansão observada do universo, este continuaria a ter praticamente o mesmo aspeto ao longo de grandes períodos de tempo. A densidade média de galáxias permanece a mesma no passado distante ou no futuro longínquo, porque novas galáxias estão continuamente a ser criadas. O universo, segundo este modelo, é uniforme não só no espaço, como também no tempo; existe para sempre, imutável. Com uma hipótese radical, o problema da origem do universo podia ser resolvido: não tinha origem. A qualidade do eternamente imutável caracteriza esta cosmologia, uma imutabilidade expressa pelo autor do Ecclesiastes: «O que foi será; aquilo que foi feito é aquilo que será feito; não existe nada de novo debaixo do Sol.»

Apesar dos seus atrativos, a cosmologia do estado estacionário coloca alguns problemas. Primeiro, nunca foi claro por que razão a taxa de criação de matéria era exatamente a necessária para produzir a expansão — era uma suposição gratuita. A criação contínua de matéria exigida pelo modelo violava também as equações de Einstein, que teriam de ser modificadas. Estes problemas poderiam ser afastados pelos adeptos do modelo. Mas

um problema mais sério era a afirmação anterior de Gamow segundo o qual os elementos pesados foram todos sintetizados no *big bang* (*big-bang* foi, originalmente, um termo de menosprezo inventado por Fred Hoyle). Segundo os defensores do estado estacionário, não teria havido nenhum *big bang*; logo, todos os elementos deveriam ter sido «cozinhados» dentro das estrelas. Provar isto era um grande desafio. Apesar de a cosmologia do estado estacionário parecer hoje estar errada, foram os seus proponentes que, com o fim de a defenderem, fizeram os primeiros cálculos corretos da síntese dos elementos pesados nas estrelas. Esses cálculos foram feitos, inicialmente, por Hoyle e, mais tarde, em 1957, por William Fowler e Margaret e Geoffrey Burbidge e, independentemente, por Alastair Cameron.

O toque de finados da cosmologia do estado estacionário começou a soar em 1961, quando Martin Ryle e Peter Scheuer publicaram a contagem das fontes rádio, cuja densidade aumenta à medida que nos embrenhamos no espaço. Uma vez que regiões remotas correspondem a um passado remoto, estas descobertas contradiziam claramente o modelo do estado estacionário, que obrigava o universo a ser imutável no tempo. Houve, no entanto, quem pensasse que as contagens eram ambíguas, nascendo uma controvérsia entre Hoyle e Ryle.

O golpe final contra a cosmologia do estado estacionário foi desferido em 1965 pela deteção por Penzias e Wilson da radiação de micro-ondas de fundo, o calor deixado pelo *big bang*. Já em 1941 se sabia que algumas moléculas do espaço interestelar estavam ativadas e em 1965 George Field sugeriu que tal era devido à

radiação. Os engenheiros de telecomunicações estavam também perturbados pela existência de um ruído persistente nas suas antenas, cuja causa desconheciam. Retrospetivamente, sabemos hoje que estes efeitos eram devidos à radiação cósmica de fundo.

Nos laboratórios Bell, em Holmdel, Nova Jérsia, Penzias e Wilson tinham uma grande antena em forma de trompa acoplada a um dispositivo recetor de rádio, arrefecido quase até ao zero absoluto e, por isso, muito sensível à radiação de baixa temperatura. Por mais que tentassem, não conseguiam eliminar um nível irredutível de ruído no recetor, que correspondia a uma radiação de fundo com uma temperatura de cerca de 3 K.

Entretanto, não muito longe dali, na Universidade de Princeton, Robert H. Dicke, um físico experimental, tinha refletido acerca de um universo primitivo quente e sugerira que havia um processo de detetar a radiação por ele deixada. Idêntica sugestão tinha também sido feita em 1964 por dois soviéticos, Igor Novikov e Andrei Doroshkevich. Dicke convenceu dois colegas, P. G. Roll e D. T. Wilkinson, a construir um pequeno detetor de micro-ondas e a usá-lo. Peebles, em Princeton, ao corrente da sugestão de Dicke, mas desconhecendo o anterior trabalho de Alpher e Herman, estava a refazer os cálculos da temperatura da radiação de fundo. Por via indireta, Penzias e Wilson tomaram conhecimento do trabalho de Peebles e souberam por Dicke que tinham, provavelmente, detetado a radiação do *big bang*. O artigo de Penzias e de Wilson, relatando as suas medições da temperatura da radiação de fundo, foi, porém, precedido de um trabalho de Dicke, Peebles, Roll e Wilkinson que dava a interpretação cosmológica dessa

temperatura. Quando Penzias e Wilson viram o trabalho mencionado nos jornais, perceberam que tinham feito uma grande descoberta.

Antes os colegas experimentadores de Peebles em Princeton tinham já montado uma antena em forma de trompa no telhado de um edifício do campus com o fim expresso de detetar a radiação, que mediram como tendo um comprimento de onda mais curto, o mesmo acontecendo depois com Penzias e Wilson. Os dados adicionais que recolheram harmonizavam-se perfeitamente com a distribuição de comprimentos de onda que se esperava para a radiação do *big bang*.

A descoberta dos quasares reforçou também a ideia do *big bang*. Maarten Schmidt, que realizou importante trabalho de observação sobre quasares, afirmou na década de 70 que há 2 mil milhões de anos existiam mais 1000 a 10.000 quasares. Se os quasares estivessem extremamente distantes, como a maioria dos astrónomos sustentava, a observação de Schmidt de que tinham existido mais quasares no passado constituiria prova claramente a favor de um universo mutável e em evolução e contra o modelo do estado estacionário.

Depois destas emocionantes descobertas os apoiantes da cosmologia do estado estacionário passaram a constituir apenas uma pequena fração dos cientistas. Como notou Dennis Sciama, anterior apoiante do estado estacionário, «considerada juntamente com as contagens das fontes rádio e os desvios para o vermelho dos quasares, a radiação de fundo em excesso cria dificuldades

muito grandes para a teoria do estado estacionário». Triunfava o modelo do *big bang*; o modelo do estado estacionário tornou-se uma peça de museu.

Hoje, posta de parte a cosmologia do estado estacionário, esquece-se muitas vezes a motivação original que levou à criação do modelo. Primeiro, a simplicidade lógica e conceptual do modelo eliminava elegantemente o problema inquietante da origem do universo. Em segundo lugar, porque em 1948 surgiu um sério problema com a ideia do *big bang*: a idade do universo, estimada a partir da taxa de expansão, era inferior à idade do sistema solar — uma discrepância absurda. O modelo do estado estacionário, pelo contrário, não apresentava tal dificuldade. Mas em 1952 os trabalhos de Walter Baade em astronomia levaram a um reajustamento das dimensões do universo por um fator de 2, aumentando, assim, a idade do universo numa proporção comparável, o que eliminava a discrepância. A recalibração por Baade das escalas de distância e de tempo do universo tornou o *big bang* um modelo possível.

Com os resultados das observações apoiando tão fortemente a cosmologia do *big bang*, os físicos teóricos voltaram com renovada confiança ao cálculo dos elementos pesados que teriam sido produzidos na exploração inicial. Nessa altura era claro que a maior parte dos elementos pesados eram produzidos nas estrelas à custa da fornalha nuclear, mas alguns, distinguidos pela expressão «elementos primordiais», deviam ter sido produzidos em grandes quantidades no calor intenso do *big bang*. Destes, o hélio é o mais abundante.

A maior parte da matéria visível do universo é composta por hidrogénio, mas cerca de 27% são hélio. Todos os outros elementos constituem apenas uma pequena percentagem. Já em 1964, antes da descoberta da radiação de fundo, Peebles, em Princeton, e Yakob Zel'dovich, na União Soviética, independentemente um do outro e desconhecendo o trabalho anterior de Gamow, Alpher e Herman, tinham iniciado os seus cálculos da abundância de hélio e da temperatura do universo. Hoyle (a despeito da oposição à ideia do *big bang*) e o seu colaborador R. J. Taylor mostraram no mesmo ano que estrelas que ardessem durante toda a idade do universo podiam converter somente 2% ou 3% do hidrogénio em hélio. Consequentemente, as estrelas não podiam ter produzido hélio suficiente para explicar a abundância atual deste elemento. Estimaram que a quantidade de hélio produzido nas primeiras fases do *big bang* fosse de 36%. Posteriormente, outros físicos, entre os quais Hoyle, William Fowler e Robert Wagoner, fizeram mais cálculos cuidadosos da síntese do hélio a partir do hidrogénio no *big bang* e obtiveram um número de cerca de 25%, precisamente de acordo com o que era observado. Os resultados destes cálculos dependiam sensivelmente dos pormenores da explosão do *big bang* — a taxa de expansão do universo e as propriedades das partículas quânticas interagindo no gás primordial —, de modo que o facto de a abundância do hélio observada ter sido corretamente prevista deve ser encarado como um grande triunfo do modelo do *big bang*.

Ainda mais notável foi a previsão da abundância relativa de deutério, cujo núcleo atómico consiste num único protão e num único neutrão. O deutério não pode ser produzido nas estrelas e sobreviver (exceto, possivelmente, nas supernovas), mas os

astrónomos observaram uma minúscula quantidade de deutério — entre cerca de um centésimo e um milésimo de 1% — no universo. A única explicação viável era a de que todo o deutério é primordial, foi produzido no *big bang*. A abundância observada de deutério é facilmente explicada quando os físicos fazem cálculos partindo desta suposição.

Os cálculos das abundâncias relativas de elementos leves primordiais, como o hélio, o deutério e o lítio, dependem das propriedades do universo quando apenas tinha alguns minutos de idade. O facto de estes cálculos preverem com tanta precisão as propriedades do universo que presentemente observamos — propriedades que de outra maneira não têm uma explicação coerente — não pode ser acidental. Por esta razão, muitos cientistas creem que compreendem realmente o estado do universo quando tinha apenas alguns minutos de idade. Curiosamente, sentem que entendem melhor o estado do universo no período compreendido entre os primeiros segundos e as primeiras centenas de milhares de anos do que antes ou depois deste intervalo, por o universo ser relativamente simples durante aquele período, essencialmente um gás de partículas cujas interações são conhecidas. As complicações ainda não começaram.

A formação das galáxias começou só depois das primeiras centenas de milhares de anos. Embora as leis da física que descrevem as interações das partículas individuais na formação das galáxias sejam compreendidas, a complexidade dos processos físicos torna difícil dizer qual delas desempenha o papel dominante. Compreendemos o universo contemporâneo dos últimos milhares

de milhões de anos razoavelmente bem, principalmente porque as observações podem guiar-nos na elaboração de modelos de estrelas e galáxias. Falta o «elo» correspondente ao período que se estende desde cerca de algumas centenas de milhares de anos depois do *big bang*, quando se iniciou a formação das primeiras galáxias ou estrelas por contração do gás de hidrogénio, até cerca de 1 a 2 milhares de milhões de anos depois do *big bang*, quando as galáxias já haviam sido formadas. Nos próximos anos novos instrumentos astronómicos fornecerão pistas para este misterioso período. Com um pouco de sorte na obtenção de novos dados, e com discernimento e persistência, a origem das galáxias tornar-se-á tão bem compreendida como a origem das estrelas o é atualmente.

Para períodos de tempo anteriores aos primeiros milissegundos do universo não é evidente que conheçamos sequer as leis quânticas específicas que governam as interações. É um período de tempo que está na fronteira da atual investigação. Aqui os problemas da cosmologia e da física quântica fundem-se num problema comum.

Para estudar o universo primitivo os cosmólogos juntam as suas forças às dos teóricos das partículas quânticas. Descobriram que o universo era tão intensamente quente e as partículas tão energéticas que as propriedades quânticas da matéria se tornavam extremamente importantes. O universo primitivo pode ser imaginado como um poderoso acelerador de partículas, um novo campo de ensaios para os mais avançados modelos das partículas quânticas. Dedicarei a parte seguinte deste livro a descrever esta excitante área de investigação.

Mas, por mais que recuemos no tempo, existe sempre alguma coisa — um gás de partículas quânticas. Que acontecerá quando nos aproximarmos da origem mesma do universo? Como podemos sequer pensar acerca disso? Donde proveio a matéria? Será que o tempo tinha algum significado? O universo começou no caos ou na simplicidade? Que aconteceu antes de o universo ter sido criado? Os físicos teóricos que especulam sobre estas questões profundas não têm respostas simples, nem únicas, tendo, no entanto, começado a abordá-las de uma maneira racional e matemática, o que representa um progresso em relação às atitudes precedentes. É notável que os conceitos atuais da teoria do campo quântico e da teoria da relatividade, embora, provavelmente, bastante incompletos, facultem um primeiro quadro conceptual para a reflexão acerca do profundo e difícil problema da origem mesma do universo. Não temos de nos render à nossa ignorância, nem desistir da tentativa de compreender racionalmente a Natureza, nem abandonar-nos a uma ilusão egoísta.

A tentativa de compreender a origem do universo é o maior desafio colocado às ciências físicas. Munidos de novos conceitos, os cientistas preparam-se para enfrentar este desafio, sabendo, todavia, que o sucesso pode estar muito distante. No entanto, quando a origem do universo for compreendida, abrirá uma nova visão da realidade, situada no limite da nossa imaginação, uma visão englobante, bela, maravilhosa e impregnada do mistério da existência. Será o nosso legado intelectual aos nossos descendentes e o nosso tributo aos heróis da ciência que iniciaram esta grande aventura do espírito humano, sem que nunca a vissem terminada.

PARTE DOIS

O Universo Infinito

É tarefa para os futuros investigadores conhecer todas as propriedades do universo a partir das leis da física fundamental dos campos quânticos.

A. D. DOLGOV E YAKOB B. ZEL'DOVICH, 1981

CAPÍTULO 8

O Universo Primitivo

Parece, pois, que devemos rejeitar a ideia de um universo permanente e imutável e que devemos assumir que as propriedades fundamentais que caracterizam o universo como o conhecemos hoje são o resultado direto de uma espécie de desenvolvimento evolutivo que deve ter começado há alguns milhares de milhões de anos atrás [...] *Partindo desta base, o problema da cosmologia científica pode ser formulado como uma tentativa para reconstruir o processo evolutivo que vem desde a simplicidade dos primeiros dias da criação até à presente e imensa complexidade do universo que nos rodeia.*

GEORGE GAMOW, 1951

Vannevar Bush, o célebre político da ciência, comparou no seu ensaio *Os Construtores* a descoberta científica e a investigação com o trabalho de extração numa pedreira e posterior colocação de pedras para construir um edifício. As pedras utilizadas são variadas e todo o esforço parece altamente desorganizado, sem nenhum arquiteto ou mestre de pedreira vigiando o projeto. Não existe plano geral para o progresso científico. Bush escreve:

Nestas circunstâncias, não é de forma alguma estranho que os trabalhadores sigam caminhos, por vezes, erráticos. Uns ficam bastante satisfeitos se lhes derem algumas ferramentas para cavar, desenterrando estranhos blocos, empilhando-os à vista dos colegas, aparentemente sem se importarem se encaixam ou não num lugar qualquer [...] Outros não cavam, passando, em vez disso, todo o tempo

> a discutir sobre o arranjo exato de uma cornicho ou de um pilar. Outros ainda passam os dias a tentar demolir um bloco, ou dois, que um rival colocou no lugar. Outros, finalmente, nem cavam, nem discutem, mas acompanham o grupo, esgravatando aqui e ali e apreciando o panorama. Alguns sentam-se e dão conselhos; outros limitam-se a sentar-se.

O sentido que se retira desta metáfora é que, a despeito da atividade ao acaso, o edifício da ciência está a ser construído e pode um dia apresentar-se como uma estrutura acabada.

Contudo, se, por outro lado, compararmos o desenvolvimento da ciência com a evolução da vida na Terra, então estaremos a compará-lo com um processo que nunca estará acabado. A investigação científica faz lembrar a evolução na reação às condições ambientais mutáveis, no oportunismo, na atenção aos pormenores, especialmente na cegueira peculiar no que respeita ao rumo a seguir. Do ponto de vista de algumas pessoas, a cultura humana, da qual a ciência é uma parte, não é mais do que a continuação da evolução da vida no reino dos símbolos e das ideias e estas, tal como as espécies, procuram sobreviver face às mudanças ambientais.

Examinando a evolução das ciências naturais, somos surpreendidos pela simbiose entre a astronomia e a física, cada uma delas uma disciplina distinta e ao mesmo tempo completada com a contribuição da outra. A astronomia, sem dúvida a ciência mais antiga, teve início há milénios com a observação cuidadosa dos céus, uma atividade que continua até aos dias de hoje, em que o trabalho dos astrónomos observacionais é complementado pelo dos

astrofísicos, que elaboram modelos matemáticos pormenorizados com a finalidade de explicarem as observações.

A física é uma ciência relativamente jovem, que se ocupa da descoberta das leis da matéria, do espaço e do tempo, nada cuidando se essas leis se aplicam ao movimento das estrelas ou às moléculas dos nossos corpos. As leis da física descobertas na Terra podem, num certo sentido, saltar para os céus e ser aplicadas às mais distantes galáxias e quasares; as leis descobertas hoje aplicam-se tanto ao passado distante como ao futuro. A lei física é universal — é um facto nunca desmentido por nenhuma observação. E, devido a esta universalidade, a astronomia, que estuda os céus, e a física, lentamente elaborada a partir da nossa experiência na Terra, desenvolveram um relacionamento estreito e íntimo.

O facto de o universo inteiro ser governado por leis naturais simples é notável, profundo e, simultaneamente, absurdo. Como pode a vasta variedade da Natureza, a multidão das coisas e dos processos, obedecer a algumas leis universais, poucas e simples? Isaac Newton descobriu a resposta. Na sua formulação da mecânica vimos pela primeira vez uma separação conceptual clara entre as «condições iniciais» de um sistema físico e «as leis do movimento». Se nos forem dadas as condições iniciais de um sistema físico, como as posições e os momentos de milhares de milhões de partículas, condições estas que podem ser arbitrariamente complicadas, então as leis do movimento determinam precisamente a evolução subsequente deste sistema no tempo. O mundo foi, assim, dividido em duas componentes: as condições iniciais, que representam o

complicado estado do mundo, e as leis universais simples, que determinam a evolução subsequente deste. Poucas vezes teve uma ideia tão vastas e profundas consequências.

Newton e os seus sucessores aplicaram as leis do movimento e da gravidade com crescente sucesso aos movimentos da Lua, dos planetas e dos cometas. O domínio da lei da gravitação universal foi até alargado às estrelas distantes pelos estudos de William Herschel sobre os sistemas estelares binários. O sucesso que a física newtoniana conheceu ao esclarecer o movimento dos céus fez com que os físicos acreditassem que os métodos matemáticos baseados em leis naturais universais se tornariam o mais poderoso instrumento conceptual de investigação do cosmos. Essa confiança foi amplamente justificada em 1864 pela descoberta de um novo planeta, Neptuno, feita com base em estudos matemáticos das perturbações na órbita de Úrano, que permitiram prever a sua localização.

Uma maravilhosa confirmação posterior da universalidade das leis naturais e da unidade do universo foi o desenvolvimento da espectroscopia estelar visual realizado por William Huggins, bem como as fotografias do espectro de absorção de hidrogénio da estrela Alfa da Lira (Vega) obtidas por Henry Draper em 1 de Agosto de 1872. Huggins, astrónomo amador inglês, foi inspirado pelas descobertas espectroscópicas de Kirchhoff e pela primeira observação por Fraunhofer de um espectro de absorção estelar em 1823, vindo a observar muitos espectros estelares. Até à introdução da fotografia este trabalho era mais arte do que uma ciência. Draper, deão da escola médica da Universidade de Nova Iorque e

membro de uma distinta família americana de homens de ciência, passou muitos anos a desenhar e a construir, pacientemente, um telescópio de reflexão de 71cm no seu observatório de Hastings-on-Hudson e a desenvolver as recentes técnicas fotográficas de que precisava para fotografar os espectros das estrelas ténues. Ninguém o tinha feito antes. Estas grandes conquistas de Huggins e Draper estabeleceram para sempre que a matéria de uma estrela é composta pelos mesmos átomos que encontramos na Terra. A astrofísica tornou--se uma ciência adulta.

A simbiose da astronomia e da física, já frutuosa para ambas no período dominado pela física clássica newtoniana, floresceu em pleno durante o século XX, com o advento da nova teoria quântica dos átomos em 1927. Derrubando o anterior sistema newtoniano, a teoria quântica apresentava-se como uma nova mecânica, radical nas implicações que trazia à realidade, de alcance geral e estranha, mas sublimemente coerente. Os físicos iniciaram a aplicação da nova teoria, com sucesso, à química, à física nuclear, à matéria sólida e ao mundo subnuclear das partículas elementares. Sabiam que os novos conceitos quânticos não se aplicavam somente aos átomos terrestres; eram universais e aplicavam-se também às estrelas distantes.

A nova mecânica quântica atingiu alguns dos seus grandes triunfos ao esclarecer os enigmas não solucionados da astrofísica. Chandrasekhar, recorrendo às regras da nova teoria quântica, foi o primeiro a entender a bizarra matéria que compõe a densa companheira de Sírio, inaugurando a moderna teoria das anãs brancas. Hans Bethe, Carl von Weizsäcker e muitos outros, que

desenvolveram a teoria da fornalha nuclear no interior das estrelas, foram os pioneiros da moderna astrofísica. As estrelas de neutrões, ou pulsares, são feitas de matéria que nem sequer existe na Terra; para as compreender é preciso estudar as partículas subnucleares descobertas nos laboratórios de física das altas energias.

Por vezes, como físico teórico, gosto de provocar os meus amigos astrofísicos perguntando por que razão têm tanta dificuldade em compreender as propriedades das estrelas — como nascem, evoluem, se consomem e explodem — quando as leis fundamentais da física estão completamente compreendidas. Mas isto é um pouco como perguntar a um biólogo como explica as propriedades de uma célula começando pelas regras da química quântica — a complexidade da tarefa é proibitiva, uma dificuldade essencial inerente à própria natureza do objeto de investigação.

Esta complexidade nasce devido a uma espécie de «desacopulamento causal» entre os diferentes níveis de organização encontrados à medida que nos movemos do microcosmos para o macrocosmos. Por exemplo, para compreender a química temos de compreender as regras a que obedecem os eletrões de valência nas regiões exteriores do átomo. Os pormenores do núcleo atómico — os quarks dentro dos protões e neutrões — estão «causalmente desacopulados» das propriedades químicas do átomo. Outro exemplo, retirado da biologia molecular, deste «desacopulamento causal» é o facto de a função biológica das proteínas estar desacopulada da maneira como foram codificadas no material genético. Em ciência abundam exemplos deste «desacopulamento causal» — uma separação importante entre os

níveis materiais da Natureza que aparece refletida no estabelecimento de disciplinas científicas separadas.

Vemos, assim, que uma coisa é conhecer as leis fundamentais da física microscópica e outra bem diferente é intuir as suas implicações para o sistema macroscópico que pretendemos compreender. Embora o conhecimento das leis físicas fundamentais possa ajudar os astrofísicos, a sua capacidade de fazer as perguntas adequadas, por exemplo, que processos são importantes, quais podem ser ignorados e quais são as características cruciais de um sistema particular, é, por si só, um talento criativo.

Ninguém duvida hoje de que as leis da física, descobertas a partir de experiências realizadas na Terra, fornecerão os fundamentos de uma teoria completa das estrelas. No entanto, mesmo em presença de tal sucesso, torna-se claro que a relação tradicional entre a física e a astronomia não pode continuar como no passado. A relação deve, de facto, tornar-se mais íntima, pois os astrónomos observacionais estão agora a explorar os quasares, os núcleos das galáxias e o *big bang*, cada um deles caracterizado por processos de tal intensidade que as experiências físicas terrestres não se lhes podem comparar. Daí que para testarem as suas teorias de alta energia os físicos só possam ser guiados pelas observações astronómicas. O universo inteiro torna-se agora o único campo de ensaios para as leis da física. Nada menos servirá, porque é o universo nas suas condições mais extremas que os físicos se esforçam por compreender.

A necessidade de uma relação mais íntima entre a astronomia e a física torna-se acentuadamente evidente se examinarmos a teoria do *big bang* da origem do universo, de acordo com a qual, se recuarmos no tempo, a temperatura do universo aumentará, possivelmente, sem limite. A temperatura é uma medida de energia do movimento das partículas, neste caso das partículas quânticas. Quais são então as leis físicas que governam a interação das partículas quânticas a estas energias elevadíssimas?

Ninguém sabe responder a esta interrogação com segurança. Seria excitante se os físicos pudessem verificar as teorias criando, durante uma fração de segundo, num laboratório dotado de acelerador, as condições reinantes no início do *big bang*. Os maiores aceleradores que existem fazem-nos efetivamente recuar à épocas em que o universo tinha apenas um milésimo de milionésimo de segundo de idade — uma conquista imensa. Mas as respostas às perguntas importantes que os físicos hoje fazem acerca do universo dependem das propriedades que apresentava *antes* de ter esse milésimo de milionésimo de segundo de idade. Tais momentos iniciais estão para além do alcance de qualquer experiência prática com aceleradores. Logo, para explorarem o universo primitivo e responderem às questões sobre a sua origem, os físicos e outros cientistas adotam uma atitude diferente. Em vez de procurarem pistas nos aceleradores de altas energias, olham para o «grande acelerador no céu», o *big bang*, e para as suas consequências. O universo inteiro torna-se uma «experiência» para dedução das leis últimas da matéria. Esta nova ciência integra o conhecimento dos corpos mais pequenos que conhecemos — os quanta — e o do maior — o cosmos.

Como procedem os cosmólogos e os físicos quânticos para investigarem o universo primitivo? Começam por utilizar as teorias quânticas relativistas das partículas quânticas que explicam as experiências de alta energia feitas em laboratórios dotados de aceleradores. Pondo de lado a precaução, extrapolam aqueles modelos teóricos até às energias elevadíssimas do universo primitivo. Utilizam então esses modelos para deduzirem algumas propriedades notáveis do universo: por que se compõe de matéria, e não de quantidades iguais de matéria e antimatéria, ou por que existiram na bola de fogo primordial minúsculas flutuações quânticas que acabam por dar origem às galáxias e aos superaglomerados.

Nalguns aspetos, esta tarefa de fazer modelos matemáticos do universo primitivo faz lembrar a tarefa mais antiga de fazer modelos do interior das estrelas com base na física nuclear. Ninguém pode entrar no interior de uma estrela para verificar diretamente estes modelos, tal como ninguém pode recuar até à bola de fogo do *big bang* para confirmar os modelos das partículas quânticas de alta energia. Todavia, com essa analogia cessa a semelhança dessas duas tarefas; os contrastes são significativos. Para começar, existem imensas estrelas com propriedades diferentes e em diferentes fases de evolução, fornecendo aos astrofísicos um colossal volume de informações que constrangem enormemente os modelos matemáticos. Em contraste com a multidão de estrelas, existe um único universo primitivo que não pode ser observado diretamente. Esta Era primordial deixou alguns fósseis espalhados, pistas contemporâneas das suas propriedades — as galáxias, a sua distribuição no espaço, a radiação de fundo de micro-ondas, a

abundância relativa de elementos químicos. Mas a diferença mais profunda entre os modelos das estrelas e os modelos do universo primitivo consiste em que os físicos exploraram já experimentalmente leis da física nuclear aplicáveis ao interior das estrelas, ao passo que não é possível realizar nenhuma experiência terrena para verificar as leis que se aplicam ao universo primitivo.

A tarefa dos cosmólogos que estudam o universo primitivo parece impossível. É como se fosse pedido aos astrofísicos que, em vez de empregarem as leis conhecidas da física nuclear, as deduzissem a partir das propriedades observadas das estrelas! Existe, no entanto, esperança; esta reside no facto de o universo primitivo poder ser um objeto mais simples do que o interior de uma estrela. Os físicos acreditam que nesses quentes e primeiríssimos tempos as interações das partículas quânticas eram muito mais simétricas; esperam que as complexidades desapareçam e a situação física se torne manejável. Se assim aconteceu, então os cosmólogos têm boas hipóteses de compreender o universo primitivo. Permanece em aberto a questão de saber se tal não passa de um vão desejo, mas os atuais preconceitos teóricos militam em seu favor.

Vale a pena recordar que, se alguém tivesse sugerido há 40 anos que os físicos poderiam vir a conhecer o estado do universo com apenas três minutos de idade, ter-se-iam rido dele. Contudo, os dados retirados da observação, como sejam a abundância relativa de elementos químicos e a temperatura da radiação de fundo de micro-ondas, fornecem novas e poderosas informações, cuja existência não podia ter sido prevista nessa altura. Da mesma

maneira, talvez surjam na próxima década novos dados cosmológicos que se revelem grandes auxiliares na resolução do enigma da origem do universo.

Como é que os cientistas chegam a este ponto nas suas investigações? Que foi que os levou a pensar acerca destes problemas desta nova maneira?

A primeira sugestão moderna acerca da origem do universo remonta ao tempo em que George Lemaître postulou o «átomo inicial» no princípio do universo, «átomo» semelhante à bola de fogo do *big bang*. Segundo ele, o «átomo» explodira em fragmentos, e estes em fragmentos menores, até o universo se tornar o que é hoje. Em 1951 escreveu:

> A evolução do mundo pode ser comparada a uma girândola de fogo-de-artifício que acaba de consumir-se: ficaram alguns pavios avermelhados, cinzas e fumo. Postados sobre um monte de cinza arrefecida, testemunhamos o lento arrefecimento dos sóis e tentamos lembrar-nos do brilho passado da origem dos mundos.

Embora Lemaître, «o pai da teoria do *big bang*», tenha dado o primeiro passo, a moderna versão do *big bang* deve-se a George Gamow e aos seus dois pupilos Ralph Alpher e Robert Herman, os quais efetuaram em 1948 cálculos da síntese dos elementos químicos no *big bang* primordial e, ao fazê-lo, trouxeram a ideia do *big bang* do reino da especulação para o reino da ciência das observações. Alpher e Herman estimaram que a temperatura do universo é hoje de 5K acima do zero absoluto — uma estimativa

confirmada quando esta temperatura foi medida com o valor de 2,7K cerca de 18 anos depois por Penzias e Wilson. Esta observação direta da temperatura da radiação de fundo desempenhou um importante papel para convencer a maior parte dos cientistas de que a ideia do *big bang* estava correta.

O programa de Gamow para calcular a abundância relativa de elementos químicos criados no *big bang* permaneceu esquecido durante 16 anos, mas foi reiniciado em 1964 com os trabalhos de Peebles, em Princeton, e de Zel'dovich, na União Soviética, ambos ignorando o trabalho anterior de Gamow, Alpher e Herman, e de Hoyle e R. J. Tayler, na Inglaterra. Em 1967 William Fowler, Fred Hoyle, Robert Wagoner e outros investigadores posteriores fizeram apurados cálculos de computador para demonstrar que as quantidades exatas de hélio e deutério — elementos que, nas quantidades observadas no universo, não podiam ter sido originados no interior das estrelas — podiam provir do *big bang*. Estes cálculos, cujo sucesso dependia das propriedades do universo quando tinha apenas alguns segundos ou minutos de idade, deram à maior parte dos físicos a confiança de que o estado do universo nessa altura poderia ser bastante bem conhecido.

Alguns físicos quiseram remontar a tempos mais antigos, recuando mais do que alguns segundos. John A. Wheeler, então na Universidade de Princeton, deslocando-se a vários locais do país, qual conferencista itinerante, continuou a lembrar aos colegas que o universo primitivo «põe a física em confronto com a maior das suas crises». Na União Soviética Zel'dovich e colaboradores tornaram-se os heróis pioneiros de toda esta área, sublinhando que

o campo final de prova para a teoria das partículas quânticas é o universo primitivo e levando a cabo muitos cálculos para apoiar este ponto de vista. Nos Estados Unidos, relativistas, astrónomos e físicos, reconhecendo a importância crescente da teoria da relatividade, organizaram em 1963 a I Conferência do Texas sobre Astrofísica Relativista, uma série de conferências que continuam até aos dias de hoje a incluir a discussão de investigações sobre o universo primitivo.

Todavia, enquanto se davam estes desenvolvimentos, a maior parte dos físicos de partículas quânticas do princípio da década de 70 não estavam muito interessados no *big bang*. Como fez notar Steve Weinberg, «era extraordinariamente difícil [para os físicos] levar a sério *qualquer* teoria sobre o universo primitivo». As atenções estavam, sim, viradas para os novos e excitantes modelos das partículas elementares, tais como o modelo de quarks da matéria subnuclear e a ideia de que as interações eletromagnéticas e fracas, normalmente consideradas distintas, não eram mais do que manifestações de um único «campo unificado». Com tão empolgantes ideias para explorar, não admira que o problema do universo primitivo não interessasse aos físicos de partículas. Ironicamente, o próprio êxito destas novas teorias de campo aplicadas às experiências da física das altas energias levou os físicos a quererem alongar a comparação das teorias com as observações. E aqui, por força das circunstâncias, foram levados a examinar as consequências das fórmulas quando aplicadas ao *big bang*.

Heinrich Hertz, físico alemão do século XIX, escreveu certa vez:

> Não podemos escapar à sensação de que estas fórmulas matemáticas têm uma existência independente e uma inteligência própria, mostrando ser mais sábias do que nós e até do que os próprios descobridores e que podemos obter mais delas do que, inicialmente, lhes foi lá posto.

Este poder que as equações da física têm de iluminar o desconhecido foi sensacionalmente exemplificado pela aplicação das novas teorias das partículas quânticas ao universo primitivo. Weinberg, fazendo-se eco do pensamento de Hertz, escreveu:

> Na física é muitas vezes assim: o nosso erro não é tomarmos as nossas teorias muito a sério, mas sim não as tomarmos suficientemente a sério. É sempre difícil verificar que estes números e equações, com que trabalhamos às secretárias, têm qualquer coisa a ver com o mundo real.

Weinberg, um dos primeiros teóricos das partículas quânticas a tomar a sério o universo primitivo, persuadiu muitos colegas a agir semelhantemente através do seu influente livro *Os Três Primeiros Minutos*. Como Zel'dovich e o seu colaborador A. D. Doglov fizeram notar num recente artigo, «muitos físicos devem o contato com a cosmologia moderna ao livro *Os Três Primeiros Minutos*, de Steven Weinberg (1977)».

Alguns físicos, porém, desempenhando o papel de críticos conservadores, pensam que os teóricos foram longe demais ao explorarem no papel o universo primitivo. Extrapolar a partir de teorias que funcionam em domínios de energia relativamente baixa, como aqueles a que funcionam os aceleradores terrestres, para domínios de energias elevadíssimas é uma empresa duvidosa. Os

críticos podem muito bem ter razão. No entanto, a particularidade digna de nota nestes esforços recentes para compreender a origem do universo não consiste em saber se essas descobertas estão certas ou erradas, o que é, sem dúvida, muito importante, mas em compreender que pela primeira vez o problema da origem do universo está a ser discutido em termos racionais e matemáticos. O advento da teoria quântica relativista do campo unificado introduziu novos conceitos no reportório dos teóricos, conceitos que podem vir a poder explicar a origem do universo. Não admira, pois, que os físicos teóricos queiram prosseguir. Se atingirem o objetivo, e isso pode demorar algum tempo, então podemos considerar este facto como uma das grandes proezas intelectuais da ciência. Até esse dia podemos juntar-nos todos e observar as explorações, frustrações e sucessos dos seus esforços.

Contudo, por mais fundo que mergulhemos na explosão do *big bang*, a matéria está sempre presente. Como poderemos então entender o exato momento da origem? Donde provém a matéria do universo? As leis da física deixarão de ser válidas? Deveremos então desistir e adotar uma atitude mística? Certos físicos, abandonando todas as precauções, avançam, tentando captar a faísca donde tudo proveio. Invocam «ideias loucas», novos conceitos, que não contradizem nenhuma experiência, nem, por outro lado, têm qualquer resultado experimental que os apoie, envolvendo espaços a cinco ou mais dimensões, «supersimetrias», e «teorias de grande unificação» (GUTs). Só alguns físicos confiam que tais ideias, ainda não testadas, estejam completamente certas. No entanto, se delas nascer uma imagem racional da origem do cosmos, isso poderá significar o fim da ciência física.

É irónico que para compreender o que de maior conhecemos — o universo inteiro — seja necessário dominar as leis que governam as mais pequenas entidades — as partículas quânticas. Tais ironias abundam em ciência. Tal como a maior parte das pessoas, que pensam e sentem, os cientistas são sensíveis ao mistério da existência. Contudo, ironicamente, à medida que exploram o universo, mais e melhor o compreendem como sujeito às leis físicas, como qualquer outra entidade material, embora, é claro, como entidade única, cuja elucidação requer novos conceitos. O mistério da existência está menos no universo material observado do que na nossa capacidade para o compreender. E em devido tempo também se poderá tornar menos misterioso.

Há quem objete contra a abordagem puramente científica e muitas vezes reducionista, apelando, em vez disso, para a sua intuição, profundamente. sentida, segundo a qual a totalidade do universo constitui uma unidade especial, com uma lei própria, sentindo que a abordagem reducionista nega uma certa ordem vital do ser. Todavia, esta atitude preconcebida faz fechar o espírito e perder o encanto do avanço das descobertas científicas, através das quais está a criar-se uma nova visão do mundo, que irá influenciar profundamente todos os aspetos da nossa cultura.

Durante milénios os seres humanos criaram sistemas culturais simbólicos que respondiam à necessidade de se sentirem ligados à totalidade do que existe. Contudo, tais visões holistas do universo, embora sirvam um fim importante, são, do ponto de vista das ciências naturais, heuristicamente estéreis e empiricamente vazias. O que a ciência está hoje a criar é uma nova imagem do universo e

do lugar da humanidade dentro dele, uma imagem isenta de qualquer disparidade entre holismo e reducionismo, uma imagem para a qual tal distinção cessa de ser relevante. Podemos até chegar a ver o universo dentro de um grão de areia.

Na primeira parte deste livro, «O jardim de Herschel», foram descritos os habitantes do nosso universo, as estrelas e as galáxias, um universo descoberto principalmente pelas observações dos astrónomos. Nesta parte, «O universo primitivo», deixamos o terreno seguro das observações astronómicas e entramos no mundo dos conceitos da física moderna. Embora estes conceitos estejam baseados em experiências, a ênfase, nesta parte, é colocada um pouco diferentemente, dado que o nosso tópico — o universo primitivo — se situa fora do alcance da observação direta. Em vez de observações diretas, são os conceitos físicos e a realidade que implicam que agora nos guiam através do universo primitivo.

A fim de compreendermos o modo como os físicos veem o universo primitivo, nos próximos capítulos iremos fazer uma excursão à «teoria quântica relativista dos campos» — a linguagem da física moderna — e ao mundo das partículas subatómicas. Armados com estes conhecimentos e com alguns de termodinâmica, estaremos então preparados para recuar até ao momento do *big bang* e mostrar a influência que as propriedades das partículas tiveram sobre aquele espantoso acontecimento.

CAPÍTULO 9

Campos, *Quanta* e Simetria

Conhecemos já muitas leis da Natureza e esperamos descobrir ainda mais. Ninguém pode prever qual será a próxima lei a ser descoberta. No entanto, existe uma estrutura nas leis da Natureza, à qual chamamos «leis de invariância», que, em alguns casos, é tão rica em consequências que leis da Natureza houve que foram adivinhadas com base no postulado de pertencerem a essa estrutura de invariância.

EUGENE P. WIGNER

Werner Heisenberg, aluno do físico Arnold Sommerfeld em Munique, Alemanha, doutorou-se em 1924. Contudo, antes de lhe ter sido concedido o grau de doutor em Físico-matemática teve de submeter-se a um exame oral, normalmente considerado uma formalidade ritual, mas, de facto, a última oportunidade que os físicos tinham de negar a um estudante a entrada na sua tribo profissional. Existia nesse tempo uma certa animosidade entre Sommerfeld e outro professor. E, tal como muitas vezes acontecia em tais disputas, um professor tentava embaraçar o outro, atormentando os respetivos alunos. Durante o exame oral foi pedido a Heisenberg pelo antagonista de Sommerfeld que calculasse o poder de resolução de um microscópio — cálculo elementar se se souber um pouco de óptica. Heisenberg, que era um génio, não soube fazer os cálculos porque desconhecia as propriedades físicas em causa, o que muito embaraço causou a Sommerfeld. Heisenberg obteve o grau, mas foram-lhe negados a

distinção e o louvor que em circunstâncias normais teria tido. Foi instado a estudar óptica, um assunto aborrecido para alguém com os seus interesses. Mas esta história tem continuação.

Um ano mais tarde, em 1925, Heisenberg inventou a mecânica das matrizes, primeiro passo em direção à nova teoria quântica dos átomos. Mais tarde, trabalhando com Max Born e Pascual Jordan, em Göttingen, inventou uma versão completa da nova teoria quântica, uma nova dinâmica que podia ser aplicada nos cálculos das propriedades dos átomos, tal como a mecânica de Newton tinha sido utilizada para calcular as órbitas dos planetas. Embora a mecânica quântica, como mais tarde foi chamada, estivesse em excelente acordo com as experiências, os seus criadores tinham dificuldade em interpretá-la como sendo uma imagem da realidade. A imagem simples da realidade material que se obtém da velha mecânica de Newton (os planetas orbitando em torno do Sol ou o movimento das bolas de bilhar) não tinha analogia na mecânica quântica. As convenções visuais da nossa experiência ordinária não são aplicáveis ao micromundo dos átomos; temos de tentar compreender esse mundo de outra maneira.

Heisenberg e Niels Bohr esforçaram-se por descobrir um novo quadro conceptual adaptado ao mundo quântico que estivesse de acordo com a nova mecânica quântica. Através de tentativas para resolver estes problemas interpretativos, Heisenberg descobriu o «princípio da incerteza», o qual revelou uma propriedade profunda da mecânica quântica que não está presente na mecânica newtoniana.

De acordo com o princípio da incerteza, certos pares de variáveis físicas, como a posição e o momento (massa x velocidade) de uma partícula, não podem ser medidos simultaneamente com precisão arbitrária. Por exemplo, se repetirmos a medição da posição e do momento de uma partícula quântica isolada — um eletrão, por exemplo —, descobriremos que a medição flutua em torno de valores médios. Estas flutuações são então a medida da nossa incerteza ao determinarmos a posição ou o momento. O princípio da incerteza afirma que o produto destas incertezas nas medições não pode ser reduzido a zero. Se o eletrão obedecesse às leis da mecânica newtoniana, então as incertezas poderiam ser reduzidas a zero e a posição e o momento do eletrão poderiam ser determinados com precisão. Porém, ao contrário da mecânica newtoniana, a mecânica quântica permite-nos conhecer apenas uma distribuição de probabilidade destas medições — é inerentemente estatística. A maneira como Heisenberg ilustrou este notável princípio da incerteza foi considerado o poder de resolução de um microscópio — o mesmo problema que o tinha atrapalhado no exame oral.

Imaginemo-nos a observar uma partícula minúscula com um microscópio. A luz atinge a partícula e é dispersada para dentro do sistema óptico do microscópio. Para um sistema óptico específico o poder de resolução do microscópio — a mais pequena distância que pode distinguir — está limitado pelo comprimento de onda da luz utilizada. É óbvio que não se pode ver a partícula e determinar a respetiva posição a uma distância que seja inferior ao seu comprimento de onda; a luz com comprimento de onda maior encurva-se à volta da partícula e não sofre dispersão significativa.

Logo, para determinarmos a posição da partícula com grande precisão temos de utilizar luz com um comprimento de onda extremamente pequeno, pelo menos menor do que as dimensões da partícula.

Todavia, como Heisenberg constatou, a luz pode também ser imaginada como um feixe de partículas — quanta de luz, chamados «fotões» — e o momento de um fotão é inversamente proporcional ao seu comprimento de onda. Assim, quanto mais pequeno for o comprimento de onda da luz, tanto maior será o momento dos fotões correspondentes. Se um fotão de pequeno comprimento de onda e grande momento atinge a partícula sob o microscópio, comunica-lhe parte do seu momento, o que faz com que a partícula se mova, criando, assim, uma incerteza quanto ao conhecimento do seu momento. Quanto mais pequeno fizermos o comprimento de onda da luz, melhor conheceremos a posição da partícula, mas também menos certos ficaremos acerca do seu momento final. Inversamente, se sacrificarmos o nosso conhecimento da posição da partícula e utilizarmos luz com maior comprimento de onda, então poderemos determinar o seu momento com maior certeza. No entanto, se a mecânica quântica estiver correta, não poderemos determinar com precisão absoluta ambas as grandezas — a posição da partícula e o seu momento.

O «microscópio de Heisenberg», como foi chamado mais tarde este dispositivo, ilustra a base física do princípio da incerteza. Heisenberg teria, provavelmente, descoberto o princípio de incerteza mesmo que nunca tivesse estudado óptica. Contudo, sem o embaraço do exame oral, é pouco provável que lhe tivesse

ocorrido uma ilustração tão simples e fisicamente intuitiva das suas ideias matemáticas. A invenção do microscópio de Heisenberg mostra o poder criativo do génio, que transforma as derrotas em vitórias de outro tipo.

O microscópio de Heisenberg utiliza uma propriedade bastante geral do mundo quântico: para «vermos» o mundo quântico atómico temos de bombardear com outras partículas quânticas os objetos que queremos observar. Não é, pois, surpreendente que tenhamos necessidade de pequenas sondas para explorar o microcosmos das partículas quânticas, e as sondas mais pequenas de todas são as próprias partículas quânticas. Os físicos exploram o micromundo observando as colisões das partículas quânticas. Quanto maior é a energia e o momento das partículas que colidem, tanto menor é o comprimento de onda e mais pequenas as distâncias que podem ser resolvidas. Por esta razão, os físicos que tentam explorar distâncias cada vez mais pequenas necessitam de máquinas que acelerem partículas quânticas até energias cada vez mais elevadas, fazendo-as depois colidir com outras partículas-alvo. Foram construídos alguns protótipos de aceleradores de partículas antes da Segunda Guerra Mundial, mas só depois é que a física das altas energias entrou na sua idade heroica. Enormes aceleradores de alta energia foram construídos nos Estados Unidos, na Europa Ocidental e na União Soviética com o fim expresso de explorar o micromundo das partículas quânticas. Deste modo, os físicos das nações que pouco antes tinham estado em guerra juntaram forças para um ataque comum ao micromundo, um mundo que ninguém tinha visto antes.

Estas máquinas abriram uma janela sobre o mundo para lá do núcleo atómico — o minúsculo núcleo de um átomo, com apenas um décimo milésimo das dimensões deste. O núcleo é composto por dois tipos de partículas, o protão, que possui uma carga elétrica unitária, e o neutrão, semelhante ao protão em muitos aspetos, mas sem carga elétrica. Os protões e os neutrões têm interações muito fortes que os unem estreitamente para formar o núcleo. Os físicos estavam ansiosos por estudar esta força tão forte, pois pensavam que encerrava a pista para a estrutura última da matéria. Mas ninguém podia prever o rico e complexo mundo de partículas a que esta forte força nuclear deu origem, nem tão-pouco quanto tempo levaria a descobrir uma teoria verdadeiramente fundamental que explicasse essa força. Aguardavam-nos décadas de frustração. No entanto, foi na forja da frustração e da ignorância que forjaram a confiança na teoria correta quando ela, finalmente, chegou.

No começo destas explorações, nos últimos anos da década de 40, os físicos descobriram mais algumas partículas com interação forte, além dos protões e neutrões, às quais chamaram mesões π. Na década de 50, ao construírem aceleradores de mais alta energia, descobriram mais partículas com interações fortes, como os hiperões, os mesões *K*, os mesões *p*, partículas estranhas — um «jardim zoológico» de partículas, provavelmente em número infinito. A todas estas partículas com interações fortes foi dado o nome coletivo de «hadrões», que significa «fortes, pesados, espessos». A maior parte delas são altamente instáveis e decaem rapidamente em outros hadrões mais estáveis. Que é que a Natureza poderia estar a dizer-nos? Esta proliferação de espécies diferentes de partículas subatómicas parecia uma anedota. Os

físicos têm a convicção íntima de que a Natureza deveria tornar-se mais simples aos níveis mais fundamentais, e não mais complicada.

A crença na simplicidade da Natureza está hoje justificada. Os hadrões, neles incluídos os familiares protão e neutrão, revelaram-se, não unidades fundamentais e irredutíveis da matéria, mas sim corpos compostos por unidades ainda mais pequenas — que parecem ser irredutíveis —, chamadas «quarks». Este modelo, quark, da estrutura dos hadrões, proposto em 1963 por Murray Gell-Mann, e independentemente por George Zweig, foi maravilhosamente confirmado numa série de experiências realizadas no Stanford Linear Accelerator Center em 1968. Estas experiências detetaram quarks pontuais residindo dentro dos protões e dos neutrões «como passas num pudim».

Os físicos veem hoje os hadrões como manifestações da dinâmica de alguns quarks orbitando em torno uns dos outros, agrupados numa minúscula região do espaço — uma simplificação imensa, se a compararmos com o número e variedade infinitos das partículas. Em muitos aspetos, esta simplificação era semelhante à conseguida pelos químicos do século XIX quando compreenderam que era possível formar milhares de compostos moleculares com apenas cerca de oito dúzias de elementos atómicos.

No fim da década de 70, após importantes descobertas teóricas e experimentais, estava intacta uma nova imagem do micromundo subatómico. As unidades fundamentais da matéria eram os quarks, os leptões e os gluões. Em princípio, qualquer componente material do universo poderia ser explicada com base nas interações destas

partículas quânticas, o que constituiu um marco importante na tentativa de compreender a Natureza e forneceu a ferramenta conceptual necessária para compreender o *big bang*.

O modelo matemático que descreve estas partículas e as respetivas interações é chamado o «modelo padrão». Descrevê-lo-ei em pormenor no próximo capítulo. Mas antes é importante encontrar um processo de imaginar o micromundo das partículas quânticas. De que espécie de «material» são elas feitas? Como podemos pensar acerca do mundo quântico ao nível das distâncias subnucleares? Para tratar estas questões os físicos inventaram uma linguagem altamente matemática, chamada «teoria quântica relativista dos campos», a qual fornece o quadro conceptual para o estudo das interações das partículas quânticas, da mesma maneira que a física newtoniana fornece o quadro conceptual para o estudo do movimento dos planetas.

Os físicos teóricos inventaram a teoria quântica relativista dos campos nos anos 20, que se desenvolveu a partir da tentativa de articular a nova teoria quântica com a teoria da relatividade restrita de Einstein. Tal tarefa provou ser muito mais difícil do que alguém poderia imaginar. Como notou Steven Weinberg:

> A mecânica quântica sem a relatividade permitir-nos-ia conceber um grande número de sistemas físicos possíveis [...] Contudo, quando se junta a mecânica quântica à relatividade, descobre-se que é praticamente impossível conceber qualquer sistema físico. A Natureza consegue, de alguma maneira, ser simultaneamente relativista e quântica, mas estes dois requisitos restringem-na tanto que só lhe

deixam uma escolha limitada — oxalá uma escolha muito limitada!

Como estas observações sublinham, quer o princípio da relatividade, quer os princípios da teoria quântica, eram requisitos muito restritivos e não era de maneira nenhuma claro que pudessem ser articulados com sucesso numa descrição matemática do mundo. Mas foram articulados, com consequências profundas.

Um dos primeiros passos foi dado em 1926 pelos físicos teóricos Max Born, Werner Heisenberg e Pascual Jordan, que mostraram como aplicar os novos conceitos quânticos ao campo eletromagnético, um campo que já obedece às exigências da teoria da relatividade restrita de Einstein, e demonstraram que a ideia inicial de Einstein do fotão como partícula de luz podia ser descrita matematicamente.

Os passos importantes seguintes foram dados, em 1928, por Jordan e Eugene Wigner e, em 1929-1930, por Heisenberg e Wolfgang Pauli, que demonstraram que cada campo — o campo eletromagnético, o campo do eletrão, etc. — tinha uma partícula associada. As partículas eram manifestações de um campo «quantizado». É esta a ideia fundamental da moderna teoria dos campos, que afastou para sempre a velha ideia de que as partículas e os campos eram entidades separadas. Os campos são as entidades fundamentais, mas as partículas são as suas manifestações no universo.

Os físicos teóricos esforçaram-se durante décadas por aprofundar a compreensão destas novas teorias de campo, uma aventura intelectual que ainda hoje continua, trazidas pelas descobertas dos colegas experimentalistas — descobertas que careciam de explicação —, bem como pelo próprio anseio de descobrirem uma linguagem coerente que descrevesse o mundo quântico. Qual é a imagem das partículas quânticas que emerge destas novas ideias?

É difícil não imaginar as partículas quânticas como objetos vulgares, que por acaso são muito pequenos. Ao refletirem sobre as partículas, os físicos têm sempre tendência para pensarem deste modo, por ser tão fácil. Todavia, as partículas fundamentais não são feitas de um determinado «material», do mesmo modo que uma cadeira é feita de madeira, parafusos e cola. Qualquer imagem simples desaparece completamente logo que se abordam as questões de pormenor. É então que o misterioso mundo da realidade quântica se manifesta.

A primeira maneira que os físicos têm de pensar estas partículas é em termos das propriedades intrínsecas que permitem classificá-las — a massa, o *spin*, a carga elétrica, etc. A segunda é em termos de interação com as outras partículas. Uma vez conhecidas as propriedades intrínsecas de uma partícula quântica e todas as suas interações, então conhecer-se-á tudo o que é possível conhecer sobre a partícula. Mas como descrevem os físicos aquilo que conhecem?

As propriedades observadas das partículas quânticas podem ser descritas com precisão em linguagem matemática, e dentro desta linguagem a noção de simetria tem vindo a desempenhar um papel cada vez mais importante. Porquê a simetria? Julga-se que as partículas quânticas fundamentais, como os eletrões e os fotões, não têm estrutura — não são feitas de outras partes mais simples —, mas, no entanto, possuem certas simetrias, do mesmo modo que um cristal as tem. Mais ainda: o eletrão, seja ele o que for, é muito pequeno, talvez seja uma partícula pontual. Para descrever qualquer coisa sem partes e muito pequena os conceitos de simetria tornam-se extremamente úteis. Por exemplo, imaginemos uma esfera suspensa no espaço. Aparentará ser a mesma se nos movermos em torno dela — tem a propriedade de ser esfericamente simétrica. Se for comprimida até dimensões muito pequenas, mesmo até um único ponto, esta propriedade da simetria esférica mantém-se. Se, em vez de uma esfera, imaginarmos um elipsoide, que só tem simetria relativamente a um eixo, então, à medida que for comprimido até ter dimensão nula, a sua simetria conservar-se-á também. Aprendemos destes exemplos que mesmo corpos muito pequenos e sem estrutura podem, no entanto, possuir simetrias específicas. O conceito de simetria pode ser aplicado com grande efeito mesmo a um objeto composto e complicado, como um átomo, composto por eletrões e um núcleo, ou um núcleo atómico, constituído por protões e neutrões. As interações entre os constituintes do átomo e do núcleo possuem simetrias específicas que ajudam a determinar a estrutura composta, tal como as simetrias dos ladrilhos cerâmicos determinam os padrões em que podem ser dispostos.

Como veremos mais adiante, as partículas quânticas elementares são definidas em termos de como se transformam sob as «operações de simetria» matemáticas, por exemplo, como se modifica uma partícula quântica se a rodarmos em torno de um eixo no espaço — uma operação de simetria. O papel da simetria na descrição das propriedades das partículas quânticas é decisivo na física moderna. Como observou o físico teórico C. N. Yang:

> A Natureza parece extrair vantagens das representações matemáticas simples das leis de simetria. Quando refletimos sobre a elegância e a maravilhosa perfeição do raciocínio matemático envolvido e a comparamos com as consequências físicas, complexas e de grande alcance, sentimos sempre um profundo respeito pelo poder das leis da simetria.

Para compreendermos melhor a relação entre as simetrias matemáticas abstratas e a maneira como são representadas pelas partículas elementares reais, recordemos a aplicação das ideias de simetria, no século XIX, às várias espécies de cristais que se formam na Natureza, como o sal, os diamantes ou os rubis. Um cristal pode ser imaginado como sendo uma rede espacial, uma verdadeira estrutura periódica dos átomos no espaço.

Esqueçamos de momento os cristais e imaginemos uma rede matemática de pontos unidos por linhas, como a malha de um ecrã de rede, uma grelha quadrada, enchendo todo o espaço, ou então imaginemos uma grelha feita de estruturas de diamante ou de triângulos, desde que periódica. Os matemáticos determinaram e classificaram todas as estruturas reticulares periódicas possíveis num espaço tridimensional através do uso de ideias de simetria

abstratas. Aqui trata-se da simetria ou invariância que poderíamos observar realizando uma deslocação no espaço, como seja movendo-nos ao longo da aresta de um cubo numa rede cúbica infinita e descobrindo que a rede não se alterou. Tais simetrias podem, assim, ser encaradas como sendo operações matemáticas abstratas num espaço a três dimensões.

Se considerarmos agora novamente os cristais reais que encontramos na Natureza, todos os cristais reais possíveis são representações destas simetrias matemáticas, pois são, eles também, estruturas periódicas no espaço. Do mesmo modo, as simetrias matemáticas abstratas das leis naturais são representadas pelas partículas quânticas reais observadas na Natureza — eletrões, protões e neutrões. Falando sem grande precisão, podemos dizer que as partículas quânticas são como cristais microscópios, podendo ser descritas completamente em termos das suas propriedades de simetria. A simetria é a chave que abre ao entendimento humano a porta do mundo microscópico.

A teoria quântica relativista dos campos — a linguagem que descreve as simetrias das partículas quânticas — é uma disciplina matemática complexa, mas as ideias base podem ser assimiladas em linguagem simples. Comecemos por examinar o significado das palavras «relativista», «quântica» e «campos», que qualificam a palavra «teoria», que é uma imagem da realidade. A parte restante deste capítulo trata destes conceitos matemáticos, que, embora ajudem a compreender a origem do universo, não são essenciais para o leitor comum, podendo ser omitidos numa primeira leitura.

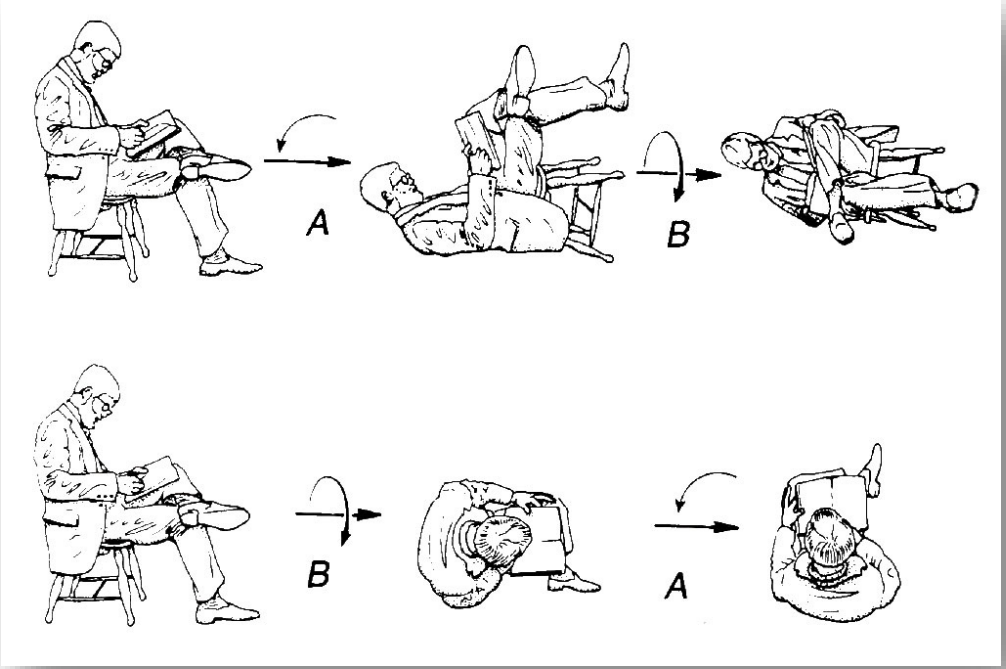
1. Relativista

«Relativista» refere-se à teoria da relatividade restrita de Einstein, de 1905. Einstein encontrou as leis corretas da transformação do espaço e do tempo que governam as medições de distância e de tempo levadas a cabo por observadores movendo-se um relativamente ao outro. Se a velocidade relativa de dois observadores fosse pequena, comparada com a velocidade da luz, então as transformações do espaço-tempo de Einstein corresponderiam às da mecânica newtoniana. Contudo, para grandes velocidades relativas, comparáveis à velocidade da luz, as transformações do espaço-tempo de Einstein eram diferentes das newtonianas e implicavam uma imagem radicalmente diferente do espaço e do tempo. Posteriormente, as transformações de Einstein foram interpretadas matematicamente como correspondendo a rotações no espaço-tempo quadridimensional de Hermann Minkowski. Qualquer teoria das partículas fundamentais deveria obedecer às regras de transformação do espaço-tempo de Einstein, o que era um requisito profundo e extremamente restritivo da descrição matemática das partículas. Estas regras de transformação do espaço-tempo de um observador para outro estão associadas com aquilo a que os matemáticos chamam um «grupo de simetria». Mas que é um grupo de simetria?

As simetrias estão à nossa volta — a simetria bilateral aproximada dos nossos corpos, a simetria esférica de uma bola, a simetria cilíndrica de uma lata de comida. Uma simetria tem a ver com a maneira como alguns objetos permanecem imutáveis ou invariantes mesmo que lhes apliquemos uma transformação. Por

exemplo, se rodarmos uma esfera perfeita em torno de qualquer eixo que passe pelo seu centro, ou um cilindro em torno do seu eixo longitudinal, permanecem imutáveis — uma manifestação da sua simetria específica. Tais operações são chamadas «operações de simetria». Matemáticos puros, como Hermann Weyl, que também aprenderam a nova teoria quântica, fizeram desde cedo notar que as simetrias poderiam ter uma profunda influência na elucidação dos problemas quânticos.

No século XIX os matemáticos tinham já inventado a descrição matemática de todas essas possíveis operações de simetria no âmbito de uma nova disciplina chamada «teoria dos grupos». Uma ideia base da teoria dos grupos consiste em descrever simbolicamente as operações de simetria, como as rotações, utilizando a álgebra. Por exemplo, suponhamos que representamos por R_1 a rotação de um objeto em torno de um determinado eixo designado por 1, segundo um dado ângulo. Sejam R_2 e R_3 outras rotações de outros ângulos em torno de eixos diferentes, designados por 2 e 3. Então, se escrevermos algebricamente o produto R_2xR_1, isso significa: primeiro realize-se a rotação R_1 e faça-se seguir a esta operação a rotação R_2. A operação conjunta R_2xR_1 é também uma rotação. Repare-se que, para as rotações, R_2xR_1 não é igual a R_1xR_2, mas não assumimos que assim fosse.



As operações de rotação não verificam a propriedade comutativa algébrica AXB=BXA. Neste exemplo A corresponde a uma rotação de 90º, no sentido contrário ao dos ponteiros de um relógio, em torno de um eixo perpendicular ao plano da página e B corresponde a uma rotação de 90º em torno de um eixo horizontal. Rodando o leitor segundo a operação B e depois segundo a operação A, vemos que o resultado final não é o mesmo que se obtém quando as operações são feitas por ordem inversa.

Suponhamos em seguida que realizamos a rotação R_3; assim, a rotação resultante é agora $R_3x(R_2xR_1)$, o que significa a rotação R_2xR_1 seguida de R_3. Suponhamos agora que começamos de novo e realizamos R_1 e a seguir a rotação conjunta representada por R_3xR_2, de tal maneira que o resultado final seja $(R_3xR_2)xR_1 = R_{3x}(R_2xR_1)$; concluímos, assim, que a «propriedade associativa» é válida para estas operações de rotação. Esta regra é um dos axiomas da teoria de grupos.

Mais ainda: verificámos que existe uma rotação mais simples do objeto, que consiste em deixá-lo inalterado — a operação identidade representada por I, que corresponde a não se executar

qualquer rotação. É evidente que $IxR_1=R_1xI=R_1$. A existência da operação identidade I é o segundo axioma da teoria de grupos.

Finalmente, assumimos que existe uma operação inversa que pode desfazer qualquer rotação e que consiste em rodar o objeto ao contrário. A operação inversa da rotação R_1 é representada por R_1^{-1} e tem a propriedade $R_1xR_1^{-1}=I=R_1^{-1}xR_1$.

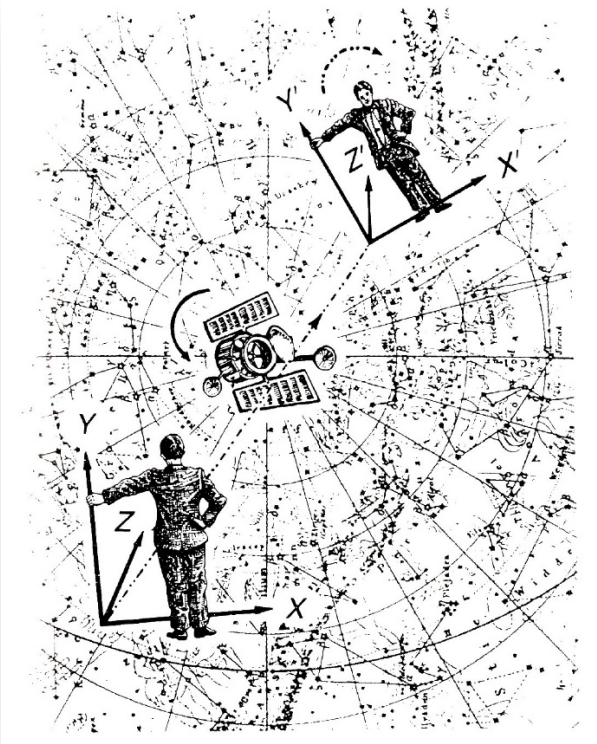
Destes três axiomas enganadoramente simples — a propriedade associativa, a existência da identidade e do elemento inverso — emerge a maravilhosa estrutura da teoria matemática dos grupos, muito como as belezas da geometria plana emergem dos axiomas de Euclides. Embora tenhamos ilustrado os axiomas algébricos da teoria de grupos com as rotações de um objeto no espaço a três dimensões, os axiomas são muito mais gerais e aplicam-se a muitas espécies de transformações de simetria em espaços multidimensionais — a troca de objetos, as reflexões espaciais, etc. Podem aplicar-se métodos algébricos formidáveis à noção de simetria, e todas essas simetrias possíveis foram classificadas e estudadas pelos matemáticos. Mas que têm tais ideias matemáticas abstratas a ver com a física?

Imaginemos dois físicos em diferentes locais no espaço, ambos observando um mesmo objeto situado num terceiro local. Os dois fazem medições sobre este objeto, relativamente às posições de cada um, e decidem depois transmitir os resultados um ao outro. Uma vez que cada um fez as medições relativamente ao seu sistema de coordenadas de medição, para comunicarem entre si necessitam de transformar, ou traduzir, as medições feitas num sistema nas

medições feitas no outro. A transformação de coordenadas mais geral para dois físicos em repouso relativo (como assumimos aqui) é uma translação (um deslocamento em linha reta no espaço) e uma rotação em torno de um eixo. É fácil compreender que quaisquer destas translações e rotações, quando descritas algebricamente, obedecem aos axiomas da teoria de grupos.

Como vemos, a teoria de grupos e a simetria entram em campo no momento em que nos interrogamos como é que várias medições feitas em sistemas de coordenadas diferentes se transformam umas nas outras — as leis gerais das transformações do espaço e do tempo.

É importante compreender que os conceitos de simetria se aplicam às leis gerais da física, e não a acontecimentos ou configurações específicos. No exemplo apresentado é importante que quaisquer dois sistemas de coordenadas (e não apenas alguns) possam ser transformados um no outro por uma translação e uma rotação. Mais ainda: se quaisquer dois físicos, usando sistemas de coordenadas diferentes, deduzirem as mesmas leis da física, podemos concluir que as leis da física são invariantes para a translação e para a rotação — aplicam-se indiferentemente do local onde cada um está situado ou da orientação de cada um no espaço. As simetrias das leis da física exprimem, assim, invariâncias.



Dois físicos encontram-se em repouso relativo no espaço, ligados aos respetivos sistemas de coordenadas, representados por três setas perpendiculares entre si. Se quiserem comunicar um ao outro os resultados das medições efetuadas relativamente aos seus sistemas de coordenadas, têm de saber como é que os dois sistemas estão relacionados. A transformação de coordenadas mais geral que transforma um sistema no outro é uma translação, que sobrepõe as origens dos dois sistemas de coordenadas, seguida de uma rotação.

Até aqui considerámos translações e rotações no espaço tridimensional normal. Contudo, refletindo um pouco, compreendemos que estas mesmas ideias devem ser generalizáveis e aplicar-se ao espaço-tempo quadridimensional de Minkowski, que é relevante para as leis de Einstein da transformação do espaço-tempo entre observadores que se movem. Os físicos aprenderam que o conteúdo mais profundo da teoria da relatividade restrita de Einstein consiste em que as leis da física são invariantes apenas para as operações de simetria correspondentes às rotações e translações no espaço-tempo a quatro dimensões. Se impusermos

esta exigência de simetria (que é o mesmo que admitir a validade da relatividade restrita), descobriremos qualquer coisa de notável.

O físico de Princeton Eugene Wigner esteve entre os primeiros que exploraram este aspeto de «grupo de simetria» das transformações de Einstein e o aplicaram às partículas quânticas. Em 1939 escreveu um artigo em que mostrou como estas considerações puramente matemáticas da teoria dos grupos implicam que as partículas quânticas podem ser classificadas — um progresso assinalável e profundo. A realização de Wigner é, de certa forma, semelhante à da anterior geração de cientistas que classificaram todos os cristais possíveis por meio dos grupos de simetria, os chamados «grupos cristalográficos» de redes espaciais periódicas. Enquanto os cristais podem ser representados por redes espaciais, objetos como as partículas quânticas (ou, aliás, qualquer objeto, uma vez que existe num espaço-tempo a quatro dimensões) devem ser representações das simetrias do espaço-tempo correspondentes incorporadas nas transformações de Einstein. Wigner mostrou que isto conduz a uma classificação das partículas quânticas.

Wigner mostrou primeiro que todas as partículas quânticas podem ser classificadas segundo a massa em repouso. Se a partícula estiver em movimento e a massa em repouso não for zero, podemos então imaginar-nos a correr atrás da partícula até que, relativamente ao nosso próprio movimento, esta fica em repouso, podendo então medir com rigor a sua massa em repouso. Por outro lado, se a massa em repouso da partícula for exatamente zero — como a massa do fotão, a partícula de luz —, mover-se-á sempre à

velocidade da luz e nunca poderemos alcançá-la. Assim, todas as partículas podem ser classificadas segundo a sua massa em repouso, consoante esta seja ou não nula.

O trabalho de Wigner também admite a existência de «taquiões», partículas hipotéticas que se movem *sempre* mais rapidamente do que a luz. Os taquiões nunca foram observados, ninguém conseguiu ainda formular uma teoria matemática consistente das suas interações. Gerald Feinberg, o físico que os batizou, comentou um dia comigo que o único lugar onde os taquiões podiam ser encontrados era no dicionário de Webster.

O segundo princípio importante da classificação de Wigner é que cada partícula quântica deve ter um determinado *spin*. Podemos imaginar as partículas como pequenos piões em rotação. Em unidades especiais esta rotação (*spin*) podia ter apenas os valores 0, $^1/_2$, 1, $^3/_2$, 2, $^5/_2$, 3 ... — um número inteiro ou semi-inteiro; estava «quantizada». A descoberta de uma partícula com *spin* $^1/_6$ implicaria uma violação do princípio da relatividade restrita e seria uma séria derrota para as leis da física.

As partículas com *spin* inteiro, 0, 1, 2 ..., são designadas por «bosões», ao passo que as partículas com valores semi-inteiros do *spin*, $^1/_2$, $^3/_2$, $^5/_2$ ..., são denominadas «fermiões», uma distinção da maior importância, dado que as interações das partículas de cada conjunto com as restantes partículas diferem acentuadamente. Por exemplo, o número total de fermiões que entram numa reação tem de ser igual ao número total dos que saem — o número de fermiões

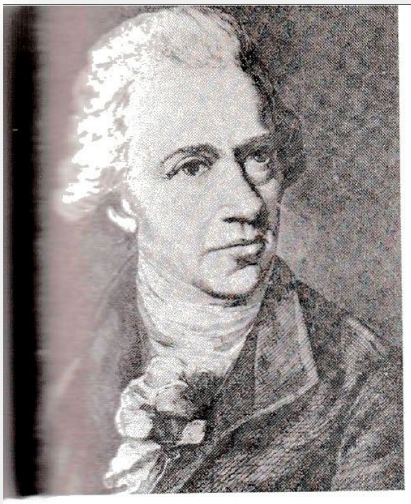
conserva-se. Mas os bosões não obedecem a nenhuma lei de conservação comparável.

Do ponto de vista da teoria quântica, o significado do sistema de classificação de Wigner, de 1939, reside no facto de as várias propriedades que usou para classificar as partículas — a massa, o *spin*, etc. — não estarem sujeitas ao princípio de incerteza de Heisenberg. Pode medir-se simultaneamente a massa e o *spin* de uma partícula com inteira precisão. Daí que tais propriedades (mas não outras) tenham valores não ambíguos para cada partícula; podem ser pensadas como os atributos das partículas quânticas.

Wigner baseou o trabalho na ideia de que as transformações de Einstein eram um grupo de simetria do espaço-tempo de Minkowski — uma das primeiras aplicações frutuosas dos princípios da simetria à moderna física das partículas. Foi uma ideia particularmente útil quando aplicada a sistemas de muitas partículas, como, por exemplo, o núcleo atómico, composto por vários protões e neutrões. A importância da ideia de Wigner residia no facto de a imposição do grupo de simetria à descrição matemática do mundo implicar automaticamente não só que os princípios da relatividade restrita fossem obedecidos, mas também que as partículas desse mundo pudessem ser classificadas de maneira simples. De uma única imposição decorria toda uma estrutura rica em consequências.

Ouvi um dia a seguinte anedota acerca de como Wigner aplicou pela primeira vez a teoria de grupos aos problemas da física atómica. Wigner e John von Neumann, o eminente matemático,

tinham frequentado o liceu juntos na sua cidade natal, Budapeste. Foram companheiros de quarto e amigos íntimos durante os estudos posteriores em Berlim. Wigner defrontava-se, no âmbito da mecânica quântica, com o problema de tratar muitas partículas com *spin* e apresentou-o a von Neumann, um génio matemático, na opinião de muita gente, o qual, mais tarde, deu várias contribuições fundamentais à matemática e inventou os conceitos fundamentais da programação de computadores. No Pentágono (era muitas vezes consultado acerca de problemas de defesa) dizia-se que ele valia, pelo menos, uma divisão americana. Von Neumann compreendeu imediatamente o problema de Wigner, fixou os olhos na parede e começou a resmungar consigo mesmo — estava a pensar. Passado um bom bocado, levantou, finalmente, o olhar para Wigner e perguntou: «Já alguma vez ouviste falar do lema de Schur?» O lema de Schur, um dos resultados fundamentais da teoria de grupos, era a pista para resolver o problema de Wigner em mecânica quântica. Foi, evidentemente, von Neumann quem orientou Wigner para a matemática da teoria de grupos.



William Herschel, que começou como músico, tornou-se o grande astrónomo do século XVIII, cujas observações abriram caminho para a visão dinâmica e evolucionária do universo que hoje possuímos. Descobridor de Úrano, mostrou também que as estrelas binárias distantes se movem de acordo com as leis de Newton, que as estrelas não estão simetricamente dispostas à volta do Sol e que este se move. Iniciou a elaboração de um importante catálogo de nebulosas. O filho, John Herschel, foi o continuador do seu trabalho *(AIP/Niels Bohr Library, E. Scott Barr Collection)*

George Ellery Hale, o astrónomo americano que construiu os primeiros grandes telescópios de reflexão capazes de revelar que algumas nebulosas eram, na realidade, gigantescos sistemas estelares — as galáxias. A fotografia mostra-o num momento de lazer em Mount Wilson, no lugar onde se encontra hoje o telescópio *Hooker* de 2,5 m *(Hale Observatories)*





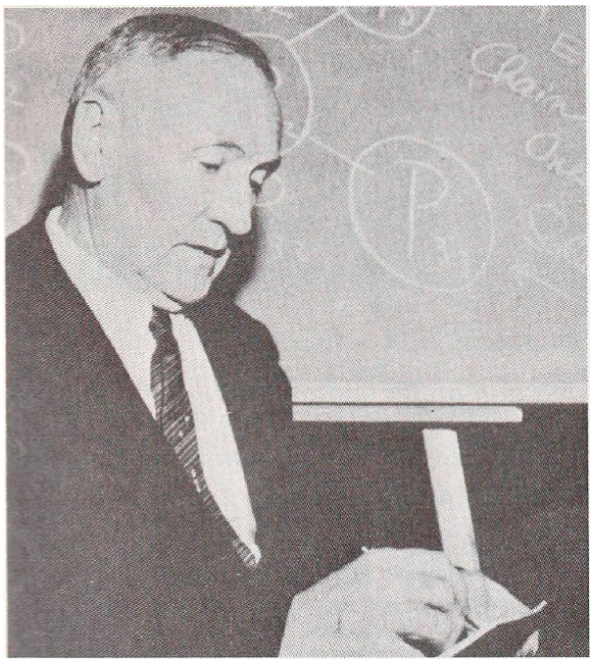
«A gargalhada do cavalo» — Harlow Shapley num momento de boa disposição. Através de medições cuidadosas da distribuição dos aglomerados globulares, estabeleceu a forma da nossa galáxia, a Via Láctea, que mostrou ser um disco achatado, e localizou o seu centro. Defendeu, no entanto, a opinião incorrecta de que todas as outras nebulosas eram parte da nossa galáxia, e não outros «universos-ilhas» distantes, semelhantes ao nosso. Foi, evidentemente, Shapley quem deu ao seu amigo, o poeta Robert Frost, a ideia para o poema *O Fogo e o Gelo (John Hubley)*



Edwin Hubble dedicou-se à observação das galáxias, os mais distantes objetos conhecidos pelos astrónomos da altura. Através das suas observações e das de muitos outros astrónomos, os cientistas ficaram convencidos de que as galáxias eram sistemas de estrelas exteriores e de que as anteriores especulações teóricas de que o universo estava a expandir-se eram corretas. Sob a sua liderança, a cosmologia observacional tornou-se uma ciência. Antes de se tornar astrónomo, Hubble foi soldado, treinador de basquetebol e estudante de Direito (Hale Observatories, Courtesy AIP/Niels Bohr Library).



Subrahmanyan Chandrasekhar aplicou em 1930 os conceitos das novas teorias quântica e da relatividade à explicação da matéria superdensa das estrelas anãs brancas. Na altura em que elaborou a teoria só se conheciam três destas estrelas peculiares. Hoje são conhecidas centenas. As anãs brancas são um dos três possíveis produtos finais da evolução estelar, sendo os outros dois as estrelas de neutrões e os buracos negros (AIP/Meggers Gallery of Nobel Laureates)

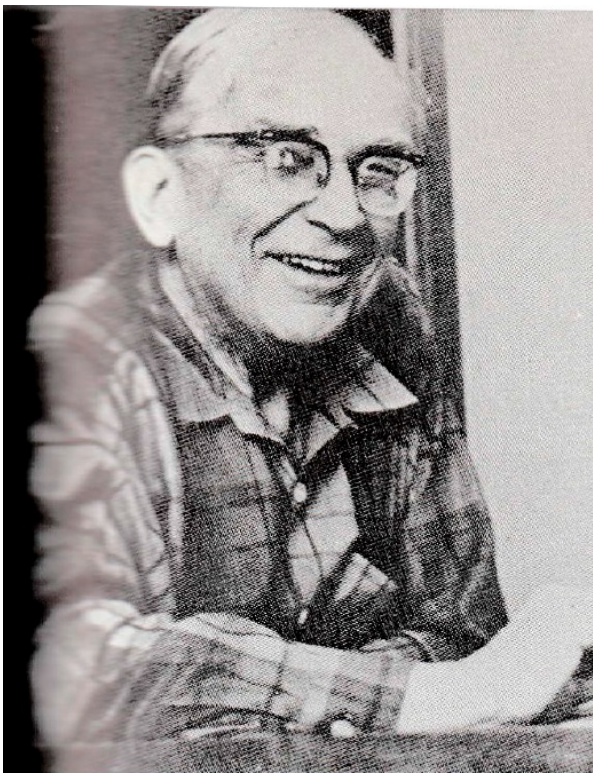


Fritz Zwicky teve uma longa e ativa carreira na astronomia. Juntamente com Walter Baade, conjeturou a existência de estrelas de neutrões em 1933. Através dos seus estudos sobre o movimento das galáxias nos aglomerados, concluiu que alguma forma de matéria escura tinha de estar presente nos aglomerados — o primeiro indício do enigma da «massa que falta» do universo.

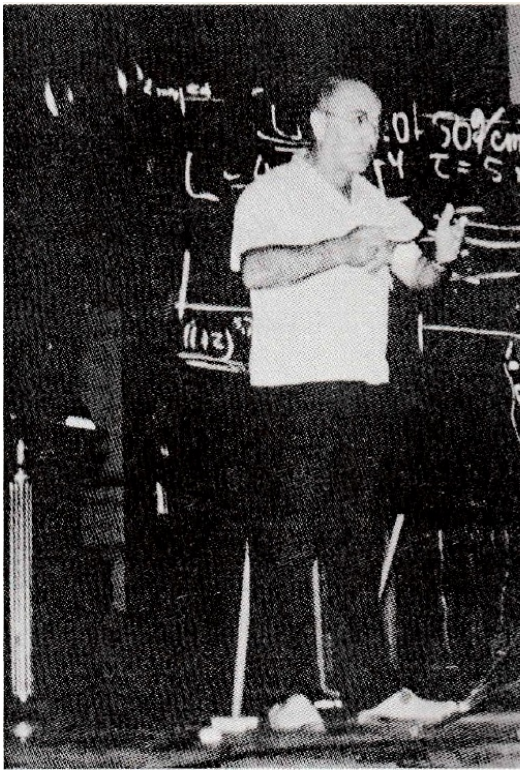


Walter Baade, que colaborou com Zwicky na teoria das estrelas de neutrões, realizou importante trabalho de observação em astronomia. Durante a Segunda Guerra Mundial descobriu que as estrelas se agrupam em duas populações principais – estrelas vermelhas velhas e estrelas azuis jovens. Posteriormente, mostrou que a incapacidade de distinguir as duas populações de estrelas conduzirá os astrónomos

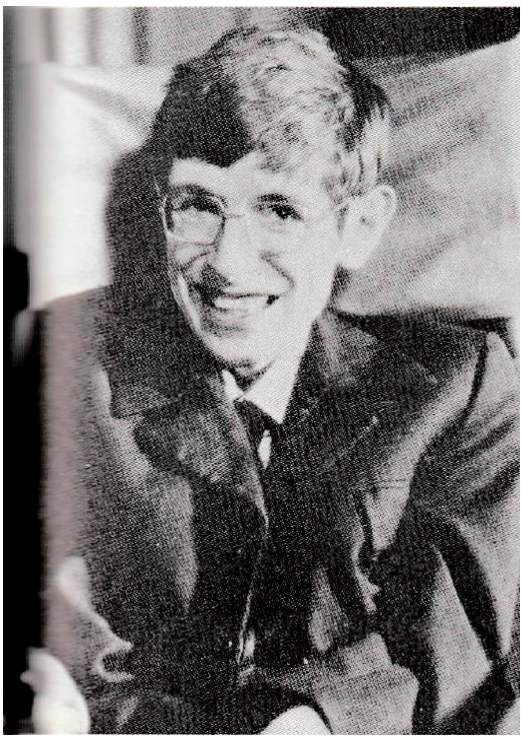
precedentes a subestimar tanto as dimensões como a idade do universo. Esta reavaliação deu uma idade superior para o universo, que hoje excede a idade das estrelas mais velhas, e resolveu uma das principais objeções à teoria do *big bang* da origem do universo. (Dorothy Davis Locanthy)



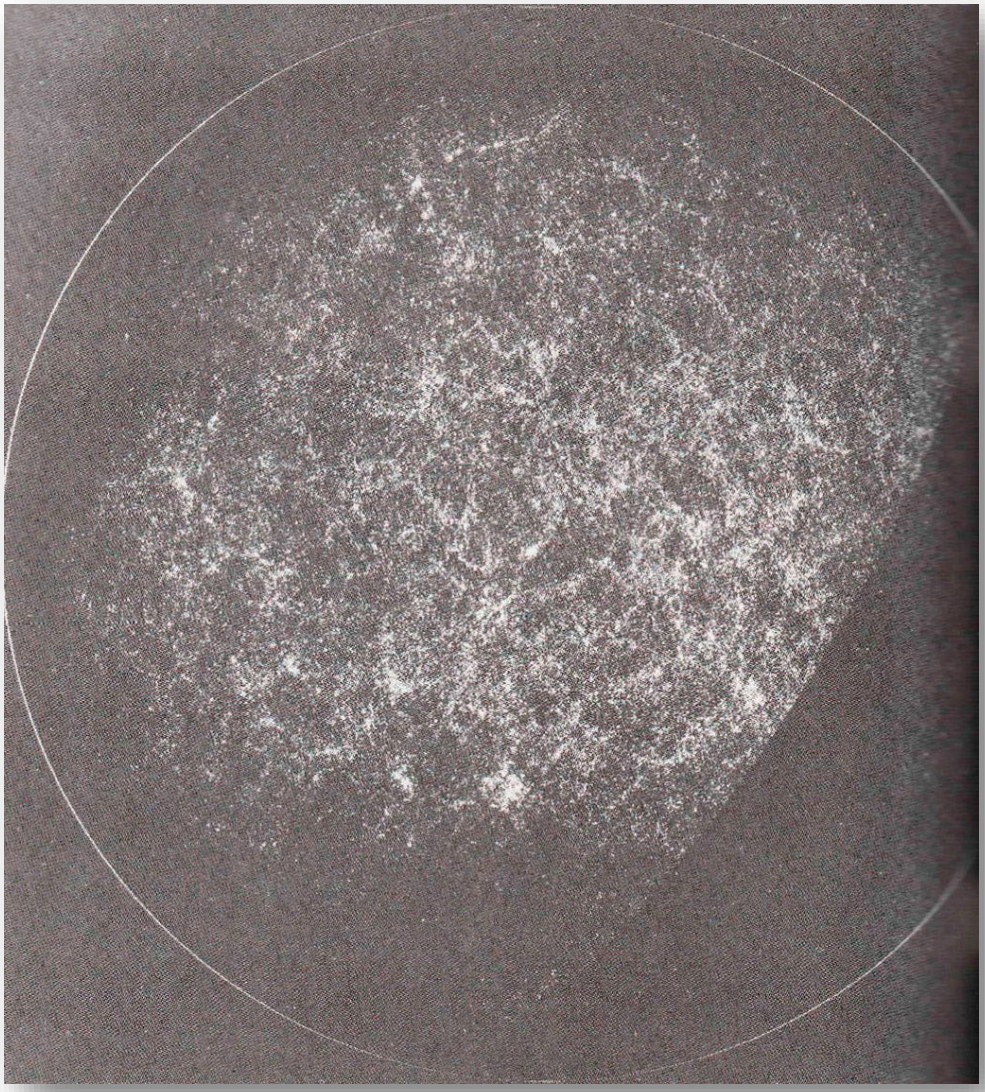
George Gamow aprendeu a teoria quântica no instituto Niels Bohr, em Copenhaga, no final da década de 20 e veio a explicar o "efeito túnel" −a maneira como as partículas nucleares podiam penetrar através das barreiras nucleares. A compreensão deste processo era crucial para compreender a fornalha nuclear no núcleo das estrelas. Na parte final da década de 40 Gamow e os seus dois estudantes Ralph Alpher e Robert Herman elaboraram a versão moderna da teoria do *big bang*. Gamow adorava anedotas e partidas. A sua melhor anedota era a de que a teoria do *big bang* acabou por mostrar-se correta (AIP/Niels Bohr Library).



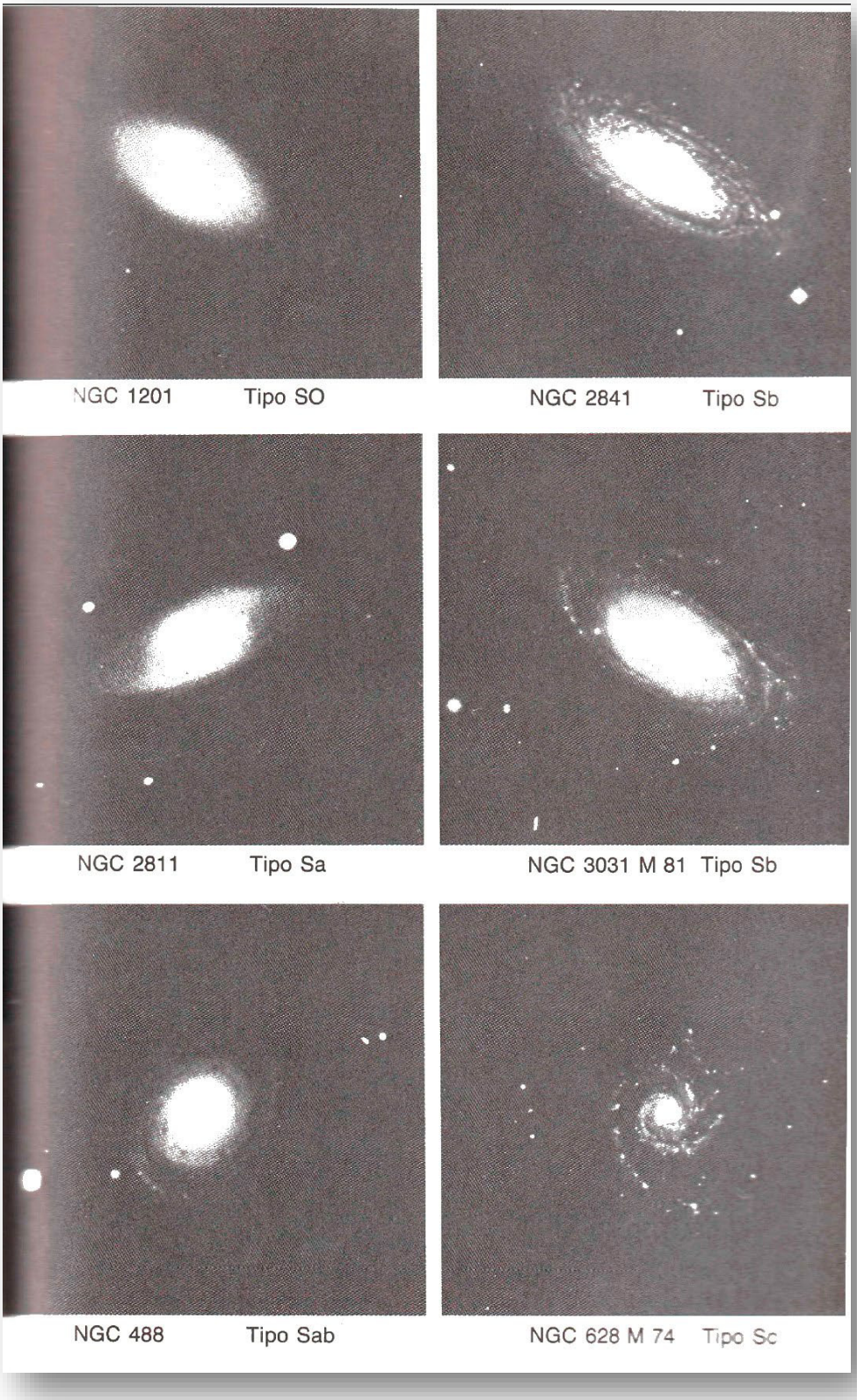
Yacob B. Zel'dovich e os seus colaboradores na União Soviética foram pioneiros do estudo teórico do universo primitivo. Foram os primeiros a salientar que as altas energias e temperaturas do *big bang* facultavam um novo campo de ensaios para testar a teoria quântica do campo das altas energias. Nesta fotografia vê-se Zel'dovich numa conferência na Checoslováquia (Leo Goldberg).



As investigações matemáticas de Stephen Hawking tiveram um profundo impacto na nossa compreensão atual do universo. Com Roger Penrose, outro físico matemático, Hawking demonstrou o "teorema da singularidade", que implica que o nosso universo deve evoluir a partir de ou para um estado muito denso. Foi o pioneiro da moderna teoria dos buracos negros e mostrou que estes não são absolutamente negros, podendo irradiar partículas quânticas – a radiação de Hawking. Atualmente, dedica-se à origem das galáxias e ao enigma da origem do universo a partir do nada (Flanklyn Institute).



Um milhão de galáxias. Nesta ilustração, processada a duas dimensões, estão representadas as posições de um milhão de galáxias no céu, sendo cada galáxia representada por um ponto branco. A linha que atravessa o campo é o limite do horizonte do telescópio Lick, onde os dados foram recolhidos. As regiões brancas são aglomerados e superaglomerados de galáxias, enquanto as regiões escuras são vazio. Alguns astrónomos afirmam que há provas de uma estrutura filamentar na distribuição. Recentemente, os astrónomos fizeram mapas tridimensionais da distribuição das galáxias. Fonte original desta ilustração: M. Seldner, B. L. Siebers, E. J. Groth e P. J. E. Peebles, *Astronomical Journal*, 82, 249, 1977



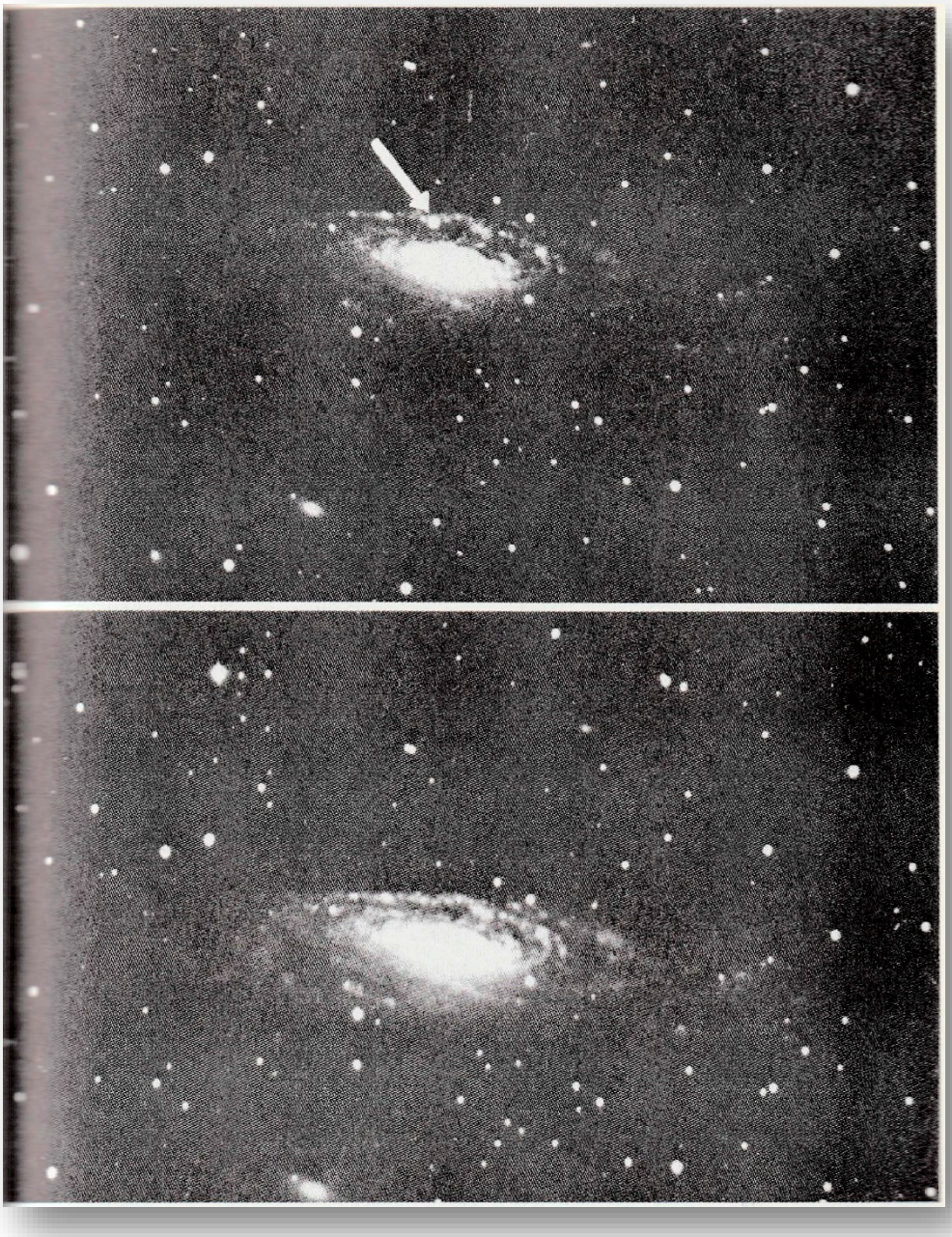
Vários tipos de galáxias normais. A do tipo SO é a galáxia lenticular, enquanto as outras são diferentes espirais, classificadas de acordo com o maior ou menor estreitamento dos respetivos braços (*Hale Observatories*).



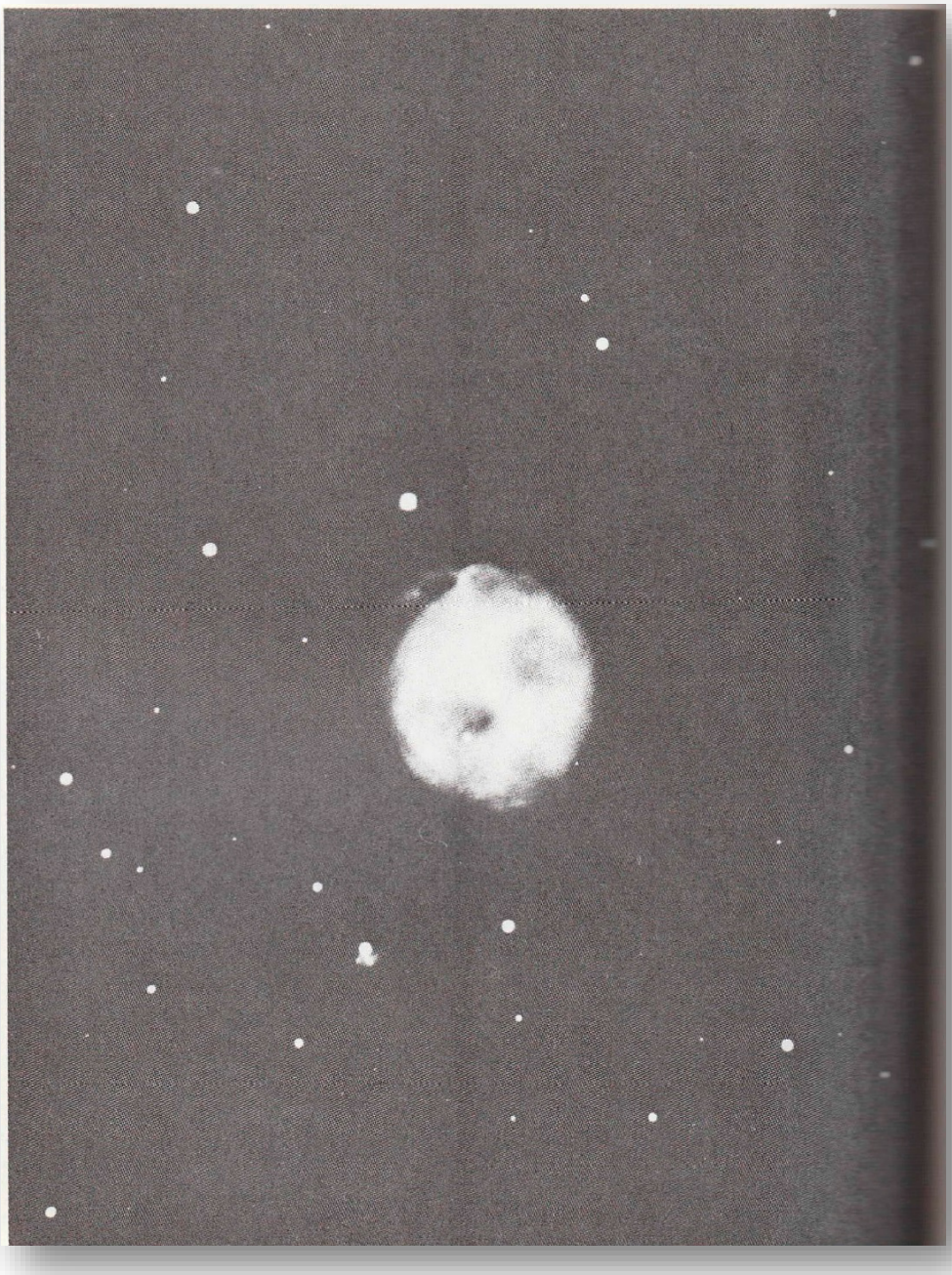
Uma parte do aglomerado de galáxias na Cabeleira de Berenice. O superaglomerado da Cabeleira é um dos mais densos, contendo 1300 galáxias principais (*Hale Observatories*).



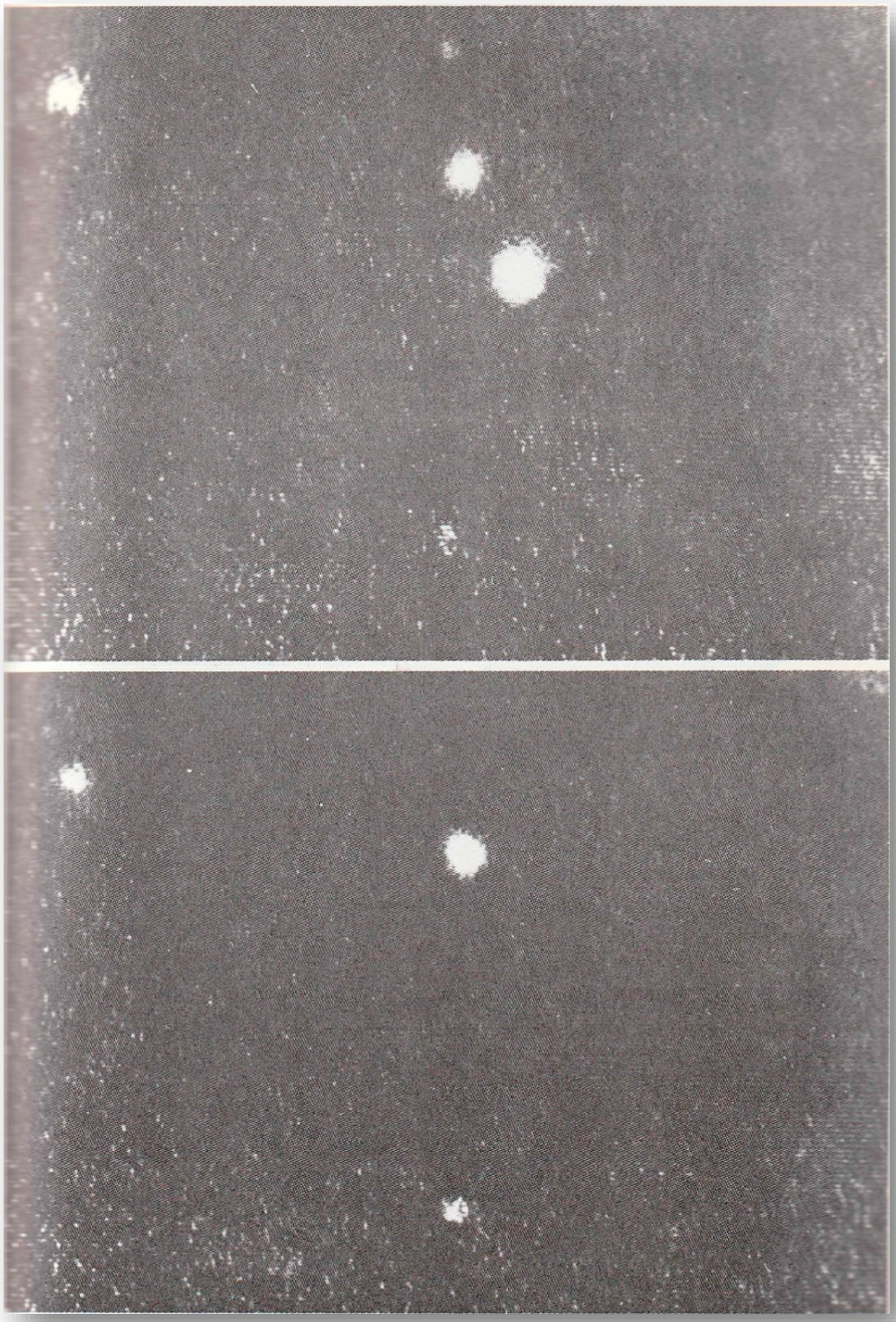
O aglomerado globular M 92. Centenas destes aglomerados de estrelas, contendo entre 50.000 e 1 milhão de estrelas, rodeiam o disco da nossa galáxia, concentrando-se perto do centro. Contêm estrelas velhas e constituem, provavelmente, uma pista para o processo de formação da própria galáxia (*Lick Observatory*).



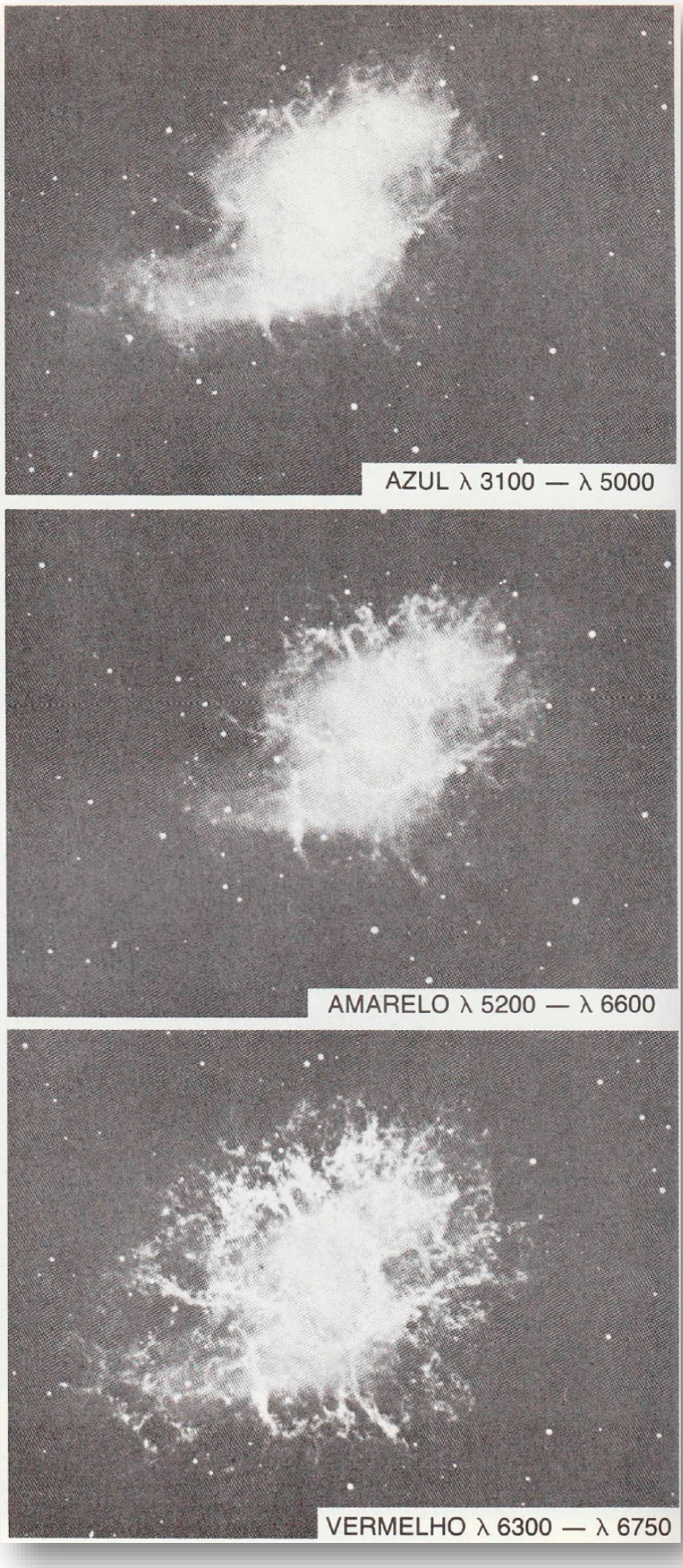
Duas exposições da galáxia NGC 7331 antes e durante a máxima intensidade da supernova de 1959. Centenas de supernovas como esta — cada uma delas é a explosão de uma estrela para formar uma estrela de neutrões — têm sido observadas noutras galáxias, mas nenhuma, nos tempos modernos, na nossa galáxia (*Lick Observatory*).

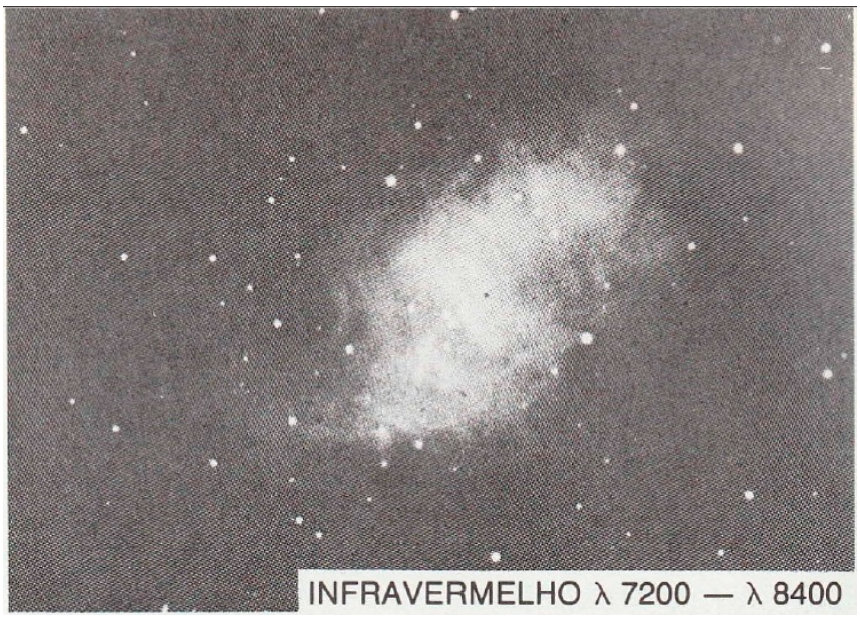


Uma nebulosa planetária na Ursa Maior, M 97. Uma nuvem de gás expelida para o espaço por uma estrela moribunda que outrora foi uma gigante vermelha e que agora se transformou numa anã branca (*Hale Observatories*).



Agora vemo-la, agora não a vemos. Fotografias de televisão do pulsar da nebulosa do Caranguejo (estrela de Baade) tiradas com um mínimo e um máximo de luz. Esta pulsação da estrela de neutrões na região do visível foi descoberta por Don Taylor, John Cocke e Michael Disney (*Lick Observatory*).


AZUL λ 3100 — λ 5000
AMARELO λ 5200 — λ 6600
VERMELHO λ 6300 — λ 6750



Nas quatro imagens, em cima, a nebulosa do Caranguejo, no Touro, M 1. Quatro fotografias em luz azul, amarela, vermelha e infravermelha. A nebulosa do Caranguejo é a supernova remanescente do pulsar (estrela de Baade) situado próximo do seu centro (*Hale Observatories*).



A galáxia elíptica gigante M 87 no centro do aglomerado de galáxias da Virgem. Esta longa exposição revela os aglomerados globulares na parte exterior. No centro desta galáxia pode estar situado um gigantesco buraco negro (*Lick Observatory*).



A nebulosa da Cabeça de Cavalo (IC 434) e a nebulosidade brilhante da NGC 2024 em luz vermelha. Estas imensas nuvens de gás e de poeiras, situadas nos braços da nossa galáxia, são os locais de nascimento das estrelas (*Lick Observatory*).

2. Campos

Voltemo-nos agora para o conceito de «campo», que foi desenvolvido no século XIX, muito antes da invenção da mecânica quântica ou da teoria da relatividade restrita. Os campos mais familiares são entidades físicas, como o campo elétrico ou o campo magnético, cuja presença se pode fazer sentir na vida de todos os dias. São invisíveis e, no entanto, influenciam a matéria; um campo magnético, por exemplo, atrai o ferro. Os físicos acreditam hoje que

todas as partículas quânticas — eletrões ou quarks — são manifestações de diferentes espécies de campos. Mas que são os campos?

Imaginemos um grande volume de ar, como, por exemplo, a massa de ar sobre um continente. A cada ponto do volume de ar podemos atribuir um único número, correspondente à temperatura do ar nesse ponto, a qual ilustra aquilo a que os físicos chamam um «campo escalar» — uma função numérica que exprime uma grandeza (a temperatura do ar), que varia de ponto para ponto, no espaço. Podemos supor também que este campo de temperatura é uma função do tempo: a temperatura muda continuamente de hora a hora.

São também possíveis outras espécies de campos. Por exemplo, suponhamos que o ar se move, como, normalmente, acontece. Então em cada ponto do ar podemos especificar um vetor, um objeto matemático com amplitude, que exprime a velocidade do ar nesse ponto, e uma direção, a do movimento do ar nesse ponto. Podemos imaginar um vetor como uma flecha associada a cada ponto do espaço. A velocidade do ar na atmosfera é um exemplo de um «campo vetorial» — tem amplitude e direção, podendo também variar com o tempo.

Campos como os da temperatura e da velocidade do ar podem ser estáticos e não evoluir, ou evoluir lentamente, ou então evoluir de tal maneira que se propaga um campo de ondas no meio. Matematicamente, a evolução dos campos no espaço e no tempo é descrita por um conjunto de «equações de campo». Os campos

podem também interagir uns com os outros. Por exemplo, se a temperatura é baixa em certa região, então o ar mais quente começa a mover-se em direção a ela; o campo escalar da temperatura influencia, assim, o campo vetorial da velocidade, e vice-versa.

Os físicos do século passado conheciam os campos, como o escalar e o vetorial, que acabei de descrever, no caso do ar. Cada campo tem necessariamente um meio associado, sendo o da temperatura a média da temperatura do ar. Os campos de ondas propagam-se sempre num meio, da mesma maneira que as ondas sonoras se propagam no ar. Parecia impossível que pudessem existir campos sem suporte material.

James Clark Maxwell, o físico escocês oitocentista que primeiro escreveu as equações que descrevem os campos elétricos e magnéticos e mostrou ser a luz um campo de ondas eletromagnético, refletiu também sobre o que seriam os campos. Concebeu modelos mecânicos do campo eletromagnético — máquinas feitas com engrenagens e parafusos que imitavam as propriedades do campo. Maxwell tanto admitia que os campos elétrico e magnético careciam do meio «éter», que se acreditava perpassar todo o espaço, como não. Muitos físicos, que acreditavam na existência do éter, tentaram deduzir as suas propriedades a partir das propriedades da luz ao propagar-se neste estranho meio. No seu trabalho de 1905 sobre a teoria da relatividade restrita, Einstein mostrou, porém, que, caso as suas ideias estivessem corretas, qualquer tentativa para detetar o éter fracassaria — tratava-se de um conceito supérfluo. Os campos eletromagnéticos

não necessitam de meio para se propagarem e, neste sentido, eram entidades fundamentais e irredutíveis. Ao contrário dos campos de temperatura e velocidade do ar, que podiam ser reduzidos às propriedades dos átomos do ar em movimento, o campo eletromagnético não tinha peças «atómicas».

Hoje em dia, como consequência do trabalho de Einstein, a atitude dos físicos em relação aos campos fundamentais mudou completamente. Tais campos não podem ser explicados em termos de qualquer outra entidade, como o éter. Os campos fundamentais (existem muitos, além do campo eletromagnético) são, sim, as entidades primárias com base nas quais tentamos explicar todas as outras coisas. Como disse Steven Weinberg, «*a realidade essencial é um conjunto de campos* [...] tudo o mais pode ser derivado, como consequência da dinâmica quântica desses campos».

É tão desprovido de sentido perguntar de que são feitos os campos como o é indagar de que «material» são feitas as partículas quânticas. Pensa-se atualmente que os campos são irredutíveis — não têm partes, são os entes mais simples. Os campos, como o eletromagnético e outros, que encontraremos, são entidades físicas definidas simplesmente em termos das equações que descrevem a sua evolução e classificadas segundo o modo como se transformam sob diferentes operações de simetria e como interagem com outros campos. Uma vez especificadas estas propriedades, o campo fica perfeitamente definido.

Que espécies de campos existem? Se impusermos que os campos obedeçam à teoria da relatividade restrita de Einstein,

então podemos utilizar o sistema de classificação de Wigner. Como seguidamente irei expor, cada campo corresponde a uma partícula quântica distinta, com massa e *spin* específicos — a base da sua classificação. Alguns, correspondentes a partículas quânticas sem massa, incluindo os campos eletromagnético e gravitacional, são de longo alcance, estendem-se por grandes distâncias, de modo que podemos facilmente detetar a sua presença. Outros, descrevendo as interações de partículas quânticas com massa, têm alcance bastante curto, estendem-se apenas por distâncias microscópicas atómicas, ou subnucleares.

Examinando a maneira como os campos se transformam se os rodarmos, podemos atribuir-lhes um *spin*. Não é surpreendente que o *spin* assim atribuído corresponda ao verdadeiro *spin* da partícula quântica associada ao campo. O campo eletromagnético tem *spin* 1, igual ao dos fotões; o campo de Dirac tem *spin* $^1/_2$, o mesmo que o eletrão; outras espécies têm *spin* 0, ou $^3/_2$, ou 2. O trabalho de Wigner conduz à classificação de todas as espécies de campos em função da massa e do *spin*.

Os campos têm também outras propriedades que ajudam a classificá-los, entre as quais se encontram as diferentes espécies de cargas, como a carga elétrica. Da mesma maneira que o *spin* de um campo está relacionado com a sua simetria espácio-temporal, também as cargas dos campos estão relacionadas com simetrias adicionais, chamadas «simetrias internas». Como podemos imaginar estas simetrias «internas» de carga adicionais? Que são elas?

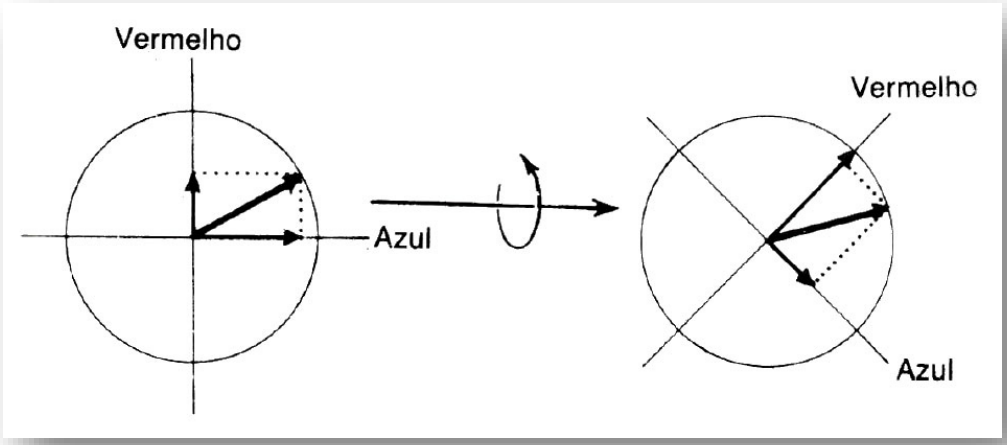
Até aqui temos discutido campos individuais com massa e *spin* específicos. Imaginemos, no entanto, vários campos, cada um deles com a mesma massa e o mesmo *spin*. Neste caso, os físicos continuam a falar de um campo individual, mas com várias componentes «internas». A ideia base de uma simetria interna é que a sua operação transforma as diferentes componentes do campo umas nas outras, de tal maneira que a situação física permanece inalterável.

Com o fim de ilustrar o que acabo de dizer, imaginemos dois campos da mesma espécie permeando todo o espaço e chamemos «vermelho» a um e «azul» ao outro. Usar cores como rótulos não é relevante; poderíamos também numerá-los como campos 1 e 2. Por analogia com o campo de temperatura no ar, suponhamos que no ponto x do espaço existe uma «temperatura vermelha» $T_V(x)$ e uma «temperatura azul» $T_A(x)$, que são as amplitudes dos dois campos no ponto x. Contudo, vamos supor que a energia dos dois campos depende somente da grandeza T(x), que é dada pela fórmula $T^2(x)=T^2_V(x)+T^2_A(x)$, isto é, o quadrado de T é a soma dos quadrados de T_V e de T_A.

Poderíamos continuar a imaginar que «vermelho» e «azul» são rótulos de eixos num espaço «interno» a duas dimensões (que nada tem a ver com o espaço físico real) e que as amplitudes dos campos vermelho e azul em todo o espaço real são medidas nos eixos «vermelho» e «azul» correspondentes no espaço interno. Uma rotação dos eixos neste espaço interno imaginário — uma operação de simetria interna — altera a amplitude relativa das componentes

vermelha e azul do campo, mas deixa a grandeza *T*(x) inalterada, por ser o raio de um círculo, que não varia quando os eixos rodam.

Suponhamos que realizamos tal rotação matemática, transformando as componentes vermelha e azul uma na outra. Tal rotação deixa *T*(x) e, portanto, a energia total inalteradas. A situação física descrita pelas equações de campo também permanece inalterada — as interações dos dois campos componentes são indiferentes à rotação dos campos vermelho e azul. Estamos então em presença de uma nova simetria — o mundo não é alterado por rotações neste espaço interno das componentes do campo. Que significa isto?



A amplitude dos campos «vermelho» e «azul» num dado ponto do espaço é indicada pelo comprimento das setas nos eixos vermelho e azul de um espaço «interno» imaginário. Quando os eixos deste espaço interno são rodados, as amplitudes dos campos vermelho e azul mudam. Contudo, a energia total do campo, que depende apenas do comprimento do raio, não muda, o que implica uma «simetria» interna das componentes do campo.

Os físicos sabem que invariâncias nas operações de simetria, como a rotação que acabámos de descrever, implicam a existência de grandezas conservadas, como a carga elétrica, que estão

associadas ao campo com várias componentes. A razão não é difícil de entender. Uma simetria implica que alguma coisa não varia — uma invariância do mundo, que acarreta a conservação de algo; no caso das simetrias internas, de diferentes cargas. Vemos que as simetrias dos campos com várias componentes implicam que os campos possuem cargas que são conservadas nas suas interações. Foi a matemática Emmy Noether que tornou matematicamente precisa esta relação da simetria com as leis da conservação, fornecendo, assim, uma das grandes razões para que os físicos teóricos procurem novas simetrias.

Emmy Noether, cujo trabalho neste campo foi tão importante, foi a primeira mulher recomendada para um lugar permanente na Universidade de Göttingen nas primeiras décadas deste século, apoiada por David Hilbert, um dos grandes matemáticos de todos os tempos. Durante uma reunião de professores debateu-se se era ou não adequada a indigitação de Noether para um corpo docente exclusivamente masculino; Hilbert manteve-se silencioso. Finalmente, sem terem encontrado qualquer solução, os professores dirigiram-se a Hilbert a fim de ouvirem a sua opinião; este respondeu, recordando aos colegas que eram o corpo docente de uma distinta universidade alemã, e não «um clube de natação» (onde naqueles tempos se praticava a segregação sexual). Noether obteve o lugar e o título, mas não um salário, tendo mais tarde sido obrigada a abandonar a Alemanha para fugir aos nazis.

Embora os campos com várias componentes e «simetrias internas» possam interagir e confundir-se, as suas cargas associadas nunca mudam. Daí que estas cargas — uma

consequência da simetria — sugiram mais um rótulo permanente com o qual os campos podem ser classificados. Por exemplo, se disséssemos a um físico que um campo tem uma massa de 0,51 milhões de eletrões-volts, *spin* $^1/_2$ e uma carga elétrica de —1, ele reconheceria este campo como o do eletrão.

O espaço interno pode ter muitas dimensões, correspondendo às muitas componentes de campo, em vez de apenas duas, como no caso dos campos «vermelho» e «azul». Em vez de uma simples rotação de eixo no plano a duas dimensões, a transformação pode ser muito mais complicada, mas a ideia base permanece a mesma: se as componentes de um campo com várias componentes podem ser transformadas umas nas outras sem alteração das interações de campo, então estamos em presença tanto de uma simetria como de uma lei de conservação de carga associada.

Vemos já quão importante é o papel da simetria na nossa compreensão dos campos que, na realidade, são definidos pela maneira como se transformam sob várias operações de simetria. Os campos não são substâncias etéreas que enchem o espaço e evoluem no tempo, são entidades irredutíveis que possuem massa, *spin* e carga — propriedades definidas pelas operações de simetria. Logo que estas propriedades tenham sido especificadas, o campo está completamente determinado.

O conceito de campo clássico é uma das ideias profundas da ciência moderna. Fornece uma linguagem simbólica e matemática para descrever o mundo físico real, uma linguagem que, quando

completamente dominada, não deixa lugar para uma posterior redução de significado. Se o conceito de campo tiver de ser transcendido, como poderá vir a acontecer no futuro, isso exigirá uma profunda alteração dos nossos conceitos de espaço, tempo e simetria. Atualmente, a teoria dos campos é a linguagem que os físicos utilizam para falar acerca da ordem material fundamental do cosmos.

3. Quântica

Até aqui descrevi o moderno conceito de campo puramente em termos de campos «clássicos» — os conceitos quânticos não desempenharam qualquer papel. Mas que têm tais campos a ver com as partículas quânticas — os quarks, os eletrões e outras partículas —, das quais o mundo é realmente feito? Os físicos descobriram a resposta a esta pergunta quando impuseram aos campos clássicos os princípios da teoria quântica: cada campo, quando «quantizado», isto é, tornado sujeito aos requisitos da teoria quântica, descreve uma partícula quântica associada. O *quantum* associado ao campo eletromagnético clássico de Maxwell era o fotão — uma partícula de luz: o *quantum* associado ao campo clássico de Dirac era o eletrão. Desta maneira, o desagradável dualismo de partículas e campos foi ultrapassado. A maneira como as partículas quânticas são classificadas — massa, *spin* e carga — é idêntica à dos campos associados.

A teoria quântica forneceu também uma interpretação para o campo clássico: a *intensidade* de um campo em determinado ponto no espaço era igual à *probabilidade* de encontrar a sua partícula

quântica associada nesse ponto. Os campos eram ondas de probabilidade para as suas partículas quânticas. Se o campo era intenso num determinado ponto, então era mais provável que a sua partícula quântica aí se encontrasse. Esta «interpretação estatística» da teoria quântica tem como consequência uma indeterminação essencial nas leis da física, dado que a *distribuição* dos acontecimentos quânticos é absolutamente determinada pelas equações da teoria quântica, ao passo que os acontecimentos *individuais* o não são. Por exemplo, a teoria não esclarece em que ponto de um ecrã irá bater um determinado fotão que passa através de um orifício; só a distribuição de muitos destes impactos pode ser determinada com precisão.

Com a aplicação feliz da teoria quântica à teoria dos campos ficaram resolvidos os principais mistérios com que os físicos se tinham confrontado durante as primeiras décadas deste século. Um poderoso instrumento matemático, um conjunto de conceitos profundos que abriram uma perspetiva imprevista da realidade, caíram nas mãos dos físicos. Emergiu uma nova ordem do mundo.

O mundo, visto segundo esta nova perspetiva, é uma vasta arena dos campos em interação, manifestando-se como partículas quânticas vagueando e interagindo umas com as outras. A experiência demonstrou que esta descrição matemática abstrata é capaz de descrever corretamente o mundo material microscópico, tal como é observado no laboratório. A teoria quântica relativista dos campos representa o culminar de décadas, se não séculos, de trabalho científico e tem demonstrado até agora notável capacidade

de sobrevivência. Os seus princípios fundamentais têm sido contestados, mas nunca derrubados.

As leis fundamentais da teoria quântica relativista dos campos já existiam na sua forma atual na década de 30. A partir de então, estas ideias têm sido acrescentadas, ampliadas e aplicadas ao mundo real das partículas quânticas. Mencionarei alguns dos mais destacados desenvolvimentos que guiarão o nosso pensamento quando nos ocuparmos da descrição da origem do universo.

4. Antipartículas

Um dos primeiros sucessos da teoria quântica relativista dos campos foi a previsão das antipartículas — novos quanta, que eram como que uma versão das partículas vistas ao espelho. As antipartículas têm a mesma massa e *spin* que as parceiras normais, mas as cargas têm sinais contrários. A antipartícula do eletrão chama-se positrão e tem carga oposta à do eletrão. Se juntarmos eletrões a positrões, aniquilam-se, libertando a energia imensa encerrada na respetiva massa, de acordo com a equivalência massa-energia de Einstein.

Como previram os físicos a existência de antipartículas? Recordemos que a «interpretação estatística» tinha como consequência que a intensidade de um campo determina a probabilidade de encontrar as suas partículas associadas. Assim, podemos pensar num campo, em determinado ponto do espaço, como descrevendo a criação ou aniquilação das suas partículas

quânticas, com uma determinada probabilidade. Desenvolvendo esta descrição matemática da criação ou aniquilação das partículas quânticas no contexto da teoria quântica relativista dos campos, descobrimos que não é possível criar uma partícula quântica sem criar também uma nova espécie de partícula — a sua antipartícula. A existência da antimatéria é-nos simplesmente imposta pela exigência de uma descrição matemática consistente do processo de criação e aniquilação que esteja de acordo com as teorias quânticas e da relatividade.

A necessidade da existência de antipartículas foi primeiro compreendida pelo físico teórico Paul Dirac, que também deu muitas outras contribuições de vulto à nova teoria quântica: encontrou a equação relativista que hoje tem o seu nome e que governa o campo do eletrão — uma conquista comparável à descoberta por Maxwell das equações do campo eletromagnético.

Ao resolver a equação, Dirac descobriu que, além de descrever o eletrão, a equação tinha soluções adicionais que descreviam uma outra partícula com uma carga elétrica de sinal contrário à do eletrão. Que poderia isto significar? Naquela altura a única partícula conhecida com esta propriedade era o protão. Dirac, não querendo fazer proliferar as partículas conhecidas, sugeriu que as soluções adicionais da sua equação descreviam o protão. Porém, após investigação cuidadosa, tornou-se claro que as partículas descritas pelas soluções adicionais tinham de ter precisamente a mesma massa que o eletrão, o que pôs de lado o protão, que tem uma massa de, pelo menos, 1800 vezes a do eletrão. As soluções adicionais tinham, portanto, de corresponder a uma partícula

completamente nova, com a mesma massa do eletrão, mas com carga de sinal contrário — um antieletrão! Esta hipótese recebeu uma maravilhosa confirmação experimental em 1932, quando Carl Anderson, físico do Cal Tech, detetou o antieletrão, hoje chamado positrão.

O advento das antipartículas mudou para sempre a maneira como os físicos pensavam a matéria. Acreditava-se que esta era permanente e imutável: as moléculas podiam ser alteradas, os átomos podiam decair por processos radioativos, mas os quanta fundamentais não mudavam. Com a descoberta da antimatéria por Paul Dirac esta imagem teve de ser substituída. Heisenberg pôs a questão desta forma:

> Creio que a descoberta, feita por Dirac, das partículas e antipartículas mudou toda a nossa perspetiva da física atómica [...] Penso que até essa altura todos os físicos tinham concebido as partículas elementares na linha do pensamento da filosofia de Demócrito, nomeadamente considerando-as unidades imutáveis que existem na Natureza, que são sempre o que são, nunca mudando, nunca podendo ser transformadas em nenhuma outra coisa. Não são sistemas dinâmicos, existem apenas. Depois da descoberta de Dirac tudo parecia diferente, dado que poderia perguntar-se por que é que um protão não pode ser, por vezes, um protão mais o par eletrão-positrão, e assim sucessivamente [...] Deste modo, o problema da divisão da matéria apareceu sob uma luz diferente.

A mutabilidade da matéria tornou-se a pedra angular da nova física das partículas. O facto de as partículas e antipartículas poderem ser criadas em conjunto a partir do vácuo, desde que se

forneça energia suficiente, é importante não somente para compreender como são as partículas criadas nos aceleradores de altas energias, mas também para descobrir os processos quânticos que tiveram lugar no *big bang* quente.

Paul Dirac era célebre pela sua reserva; raramente falava, mas, quando o fazia, ia sempre ao cerne das questões[1]. Ouvi a seguinte história acerca de Dirac, que, embora não confirmada, é certamente crível. Richard Feynman, um dos inventores da eletrodinâmica quântica que, como a maior parte das pessoas, gosta de uma boa conversa, estava sentado ao lado de Dirac ao jantar. Este não falava. Feynman, talvez em desespero, perguntou-lhe, finalmente: «Sentiu-se feliz quando escreveu aquela equação?», referindo-se, claro está, à eterna equação de Dirac. Passados alguns instantes, Dirac disse: «Sim.» E, após uma pausa ainda maior, indagou, quase incoerentemente: «O senhor também está a trabalhar numa equação?» Até mesmo Feynman, brilhantemente inventivo como era[2], não teve resposta.

5. Renormalização

Embora as ideias da teoria quântica relativista dos campos tivessem previsto com êxito a existência da antimatéria, os físicos teóricos das décadas de 30 e 40 encontraram inúmeras dificuldades matemáticas e problemas ao tentarem aplicá-las. Quando calculavam processos quânticos de interação, utilizando estas novas ideias, obtinham números infinitos, o que queria dizer que alguma coisa estava errada. A Natureza não tem grandezas físicas infinitas. O problema residia na ideia mesma de um campo de ondas

oscilando no espaço. Por mais pequeno que seja o volume do espaço que examinemos, estão sempre presentes alguns comprimentos de onda do campo muito pequenos, sendo a presença permanente destas ondas muito curtas em número infinito diretamente responsável pelos números infinitos que os físicos obtinham. Alguns chegaram a pensar que a teoria dos campos devia estar errada.

Outros, porém, continuaram a tentar resolver este problema até conseguirem domar estes números infinitos, recorrendo a um truque matemático, chamado «processo de renormalização». Demonstraram que os números infinitos apareciam somente nos cálculos de algumas grandezas, como a massa ou a carga elétrica das partículas quânticas em jogo, e que, se essas grandezas fossem redefinidas ou «renormalizadas» pela subtração de um número infinito, poder-se-iam então obter resultados finitos para todas as grandezas experimentalmente mensuráveis. A subtração destes números infinitos parecia um truque matemático, mas resultou.

No fim da década de 40 os físicos teóricos, entre os quais se destacam Freeman Dyson, Richard Feynman, Julian Schwinger e Sin-itiro Tomonaga, tinham já elaborado um exemplo viável de uma teoria quântica relativista de campos «renormalizada», que descrevia as interações de apenas duas partículas quânticas, o eletrão e o fotão; chamou-se-lhe «eletrodinâmica quântica», não somente por existirem dados experimentais surpreendentes relativos às interações dos fotões e dos eletrões, que careciam de explicação, mas também porque os fotões e os eletrões constituíam uma boa aproximação a um pequeno subsistema de todas as

partículas quânticas. Podiam, portanto, ignorar-se as suas interações com outras partículas quânticas, o que conduzia a uma simplificação considerável. Se o processo de renormalização valesse alguma coisa, devia funcionar neste caso.

Uma vez aplicado cuidadosamente o processo de renormalização, os resultados dos cálculos da eletrodinâmica quântica puderam ser comparados com experiências rigorosas. Para espanto de muitos, a teoria, apesar dos truques matemáticos abstratos, foi confirmada, em pormenor, pela experiência. Desde os tempos das previsões de Newton relativas aos movimentos planetários que a teoria e as observações não se harmonizavam tão completamente. Os próprios físicos ficaram admirados com o sucesso experimental da eletrodinâmica quântica.

Após o sucesso da eletrodinâmica quântica, os físicos propuseram-se aprofundar a compreensão do processo de renormalização, de maneira a estabelecê-lo menos como um truque matemático e mais como um aspeto profundo das interações entre as partículas quânticas. Um passo importante foi dado por Kenneth Wilson, da Universidade de Cornell, nos últimos anos da década de 60. O seu trabalho tinha como consequência que nas teorias renormalizáveis o valor da massa ou da carga de uma partícula quântica dependia da escala de distâncias à qual a examinássemos. Vista de longe, como geralmente acontece, uma partícula tem uma massa definida. Vista a distâncias microscópicas, como se faz num acelerador de altas energias, a mesma partícula pode ter uma massa efetiva maior ou menor do que o valor obtido para grandes distâncias. Isto parecia estranho. Como pode a massa de uma

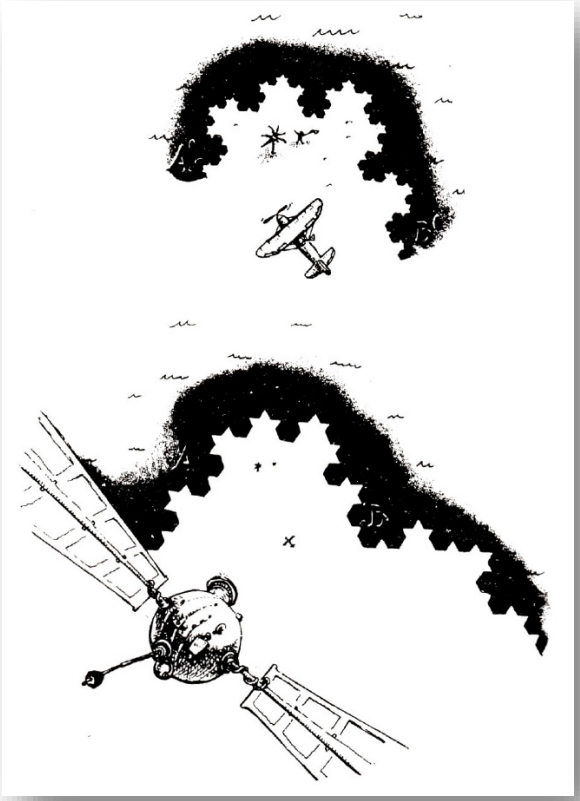
partícula depender da escala de distâncias à qual é observada? Normalmente imaginamos a massa como algo fixo e definido.

Imaginemos um segmento de reta com 15cm de comprimento desenhado numa folha de papel. Também ele parece fixo e definido. Visto de mais longe, parece mais curto. Reduzindo a distância a metade, o comprimento parece aumentar para o dobro. É claro que não estamos a ser enganados por este segmento de reta em crescimento — o segmento original continua a ter 15cm de comprimento. De facto, utilizando o conhecimento que temos da distância à folha de papel — a nossa escala de distâncias — e do ângulo subtendido pelo segmento, podemos calcular facilmente o seu comprimento.

Suponhamos agora que reduzimos a metade a nossa distância para o segmento de reta e, em vez de o crescimento se fazer segundo o fator 2, passa a fazer-se pelo fator 1 1/2 ou mesmo 2 1/2. Que cálculos fazemos então? Qual é o «verdadeiro comprimento» do segmento de linha reta?

É claro que os segmentos de reta não fazem isso, mas suponhamos que tiramos antes uma fotografia a uma linha de costa — uma linha muito torta — a partir de um satélite e medimos o seu comprimento entre dois pontos. Reduzamos agora a metade a distância a que nos encontramos e tiremos outra fotografia, medindo o comprimento entre os mesmos pontos. Podíamos pensar, em analogia com o caso do segmento de reta, que este comprimento duplicaria. Mas, facto assinalável, não o faz: aumenta mais do dobro. Se reduzirmos a escala de distâncias a metade,

encontraremos o mesmo desvio relativamente à duplicação esperada.



Um náufrago numa «ilha de Koch» visto em duas escalas de distância diferentes. A linha de costa da ilha de Koch apresenta comportamento fractal, ou de dimensão anómala — o comprimento aparente não varia proporcionalmente à distância de observação. Tal como numa linha de costa real a distância entre os mesmos dois pontos aumenta mais do que a proporção esperada à medida que nos aproximamos dela — torna-se efetivamente maior. As propriedades das partículas quânticas, como as respetivas massas e forças de acoplamento, dependem também das escalas de distância às quais são medidas. Os físicos teóricos conjeturam que muitas das interações se tornem iguais a escalas de distância muito pequenas, realizando, assim, a unificação das forças da Natureza.

Podemos descrever matematicamente este desvio no comportamento de escala por aquilo a que o matemático Benoit B. Mandelbrot chama «fractais» e os físicos chamam «dimensões anómalas». «Fractais», ou dimensões anómalas, são números que especificam com precisão, em qualquer exemplo dado, o desvio em

relação à lei de escala esperada. Mandelbrot descobriu vários exemplos deste estranho comportamento das escalas no mundo natural — muitas vezes é mais a regra do que a exceção. E as partículas quânticas, descritas por interações renormalizáveis, obedecem também a esta regra.

Ao examinarmos as partículas, a sua massa e força de acoplamento (que mede a sua interação com as outras partículas) variam, consoante as escalas de distâncias às quais são examinadas, tal como no caso da linha de costa. Os físicos Curtis Callen, de Princeton, e Kurt Symanzik, da Universidade de Hamburgo, Alemanha, deduziram em 1968 um conjunto de equações que descrevem este comportamento anómalo da dimensão nas teorias quânticas relativistas dos campos. As suas equações estão baseadas nas ideias de Wilson e em trabalhos anteriores dos físicos Murray Gell-Mann, Francis Low e A. Petermann. Estes desenvolvimentos matemáticos contribuíram grandemente para apoiar a convicção de que o processo de renormalização era mais do que um truque matemático — tinha conteúdo físico.

Nem todas as teorias quânticas relativistas dos campos são renormalizáveis — a matemática da renormalização funciona somente para alguns tipos de interações de partículas quânticas, de entre um número infinito de possibilidades. É notável que as interações renormalizáveis sejam precisamente aquelas que observamos. Estará a Natureza a tentar dizer-nos algo ao usar apenas interações renormalizáveis? Alguns físicos, admirados com este facto, pensam que a possibilidade de renormalização é uma

imposição fundamental da Natureza, como o princípio da relatividade restrita. Outros não estão tão seguros disto. Mas, pelo menos, uma coisa é clara: a Natureza, ao escolher interações renormalizáveis entre os *quanta*, tem sido simpática para com os físicos teóricos, que, em princípio, podem agora calcular as interações das partículas quânticas sem obterem como resposta um disparate.

6. Campos de Padrão[3]

Com o aparecimento da teoria quântica como linguagem da Natureza, a simetria e a teoria de grupos passaram a desempenhar um papel cada vez mais importante na física. No entanto, o uso mais profundo da simetria só foi descoberto em 1954 e a sua aplicabilidade à física só foi compreendida em 1968, com a descoberta da «teoria dos campos de padrão não abelianos[4]», inventada pelos físicos matemáticos C. N. Yang e Robert Mills.

A ideia base era generalizar a noção de simetria interna. Suponhamos que temos um campo com três componentes: às duas componentes a que anteriormente chamámos «vermelha» e «azul» juntamos uma terceira, chamada «amarela». Podemos imaginar que o vermelho, o azul e o amarelo correspondem a três eixos num «espaço interno» a três dimensões. A operação de simetria interna corresponderia a fazer uma rotação arbitrária neste espaço interno a três dimensões das componentes do campo. Se rodarmos matematicamente os eixos neste espaço interno, então as componentes vermelha, azul e amarela do campo rodam para o mesmo ângulo no espaço real. Se, nestas condições, a energia total

do campo permanecer constante, estaremos em presença de uma simetria. Neste caso, dizemos tratar-se de uma «simetria interna global», porque as diferentes componentes do campo foram rodadas para o mesmo ângulo em *todo* o espaço físico.

Imaginemos agora, como o fizeram Yang e Mills, que, em vez de rodarmos as componentes do campo para o mesmo ângulo em todo o espaço, deixamos o ângulo de rotação das componentes do campo variar de ponto para ponto no espaço físico, uma operação de «simetria interna local», pois difere localmente de ponto para ponto e não é a mesma em todo o espaço. Contudo, depois descobrimos que a energia total do campo variou, que a simetria original se perdeu.

Yang e Mills descobriram que a simetria perdida podia ser notavelmente restaurada se introduzíssemos no espaço real mais um campo com várias componentes, o chamado «campo de padrão não abeliano». Deixando também que as diferentes componentes deste campo rodem e se transformem umas nas outras de ponto para ponto no espaço real, restauramos a simetria perdida. O papel do campo de padrão consiste em que ele compensa a perda de simetria ocorrida quando transformamos a rotação interna global numa rotação local. Vemos, assim, que a imposição de uma simetria interna local — uma rotação das componentes do campo que difere de ponto para ponto no espaço físico — tem como consequência um novo campo — o campo de padrão. A existência de campos de padrão poderia, assim, ser deduzida unicamente a partir da exigência de simetria. Desta conclusão dramática, que concede ao conceito de simetria primazia sobre o conceito de campo, deriva a

maior parte da investigação contemporânea em teoria quântica relativista dos campos.

Uma maneira de visualizar o efeito do campo de padrão de Yang-Mills é imaginar um triângulo sobre uma grelha espacial (ver a figura). O triângulo simboliza o campo original com várias componentes, e a grelha, um sistema de coordenadas que pode representar as rotações no espaço interno. Uma rotação global do sistema de coordenadas da grelha não altera a forma física do triângulo — a situação física permanece inalterável. Contudo, se o ângulo de rotação da grade variar localmente de ponto para ponto, a forma do triângulo muda e a simetria perde-se. O efeito do campo de padrão é restaurar a simetria perdida, de tal modo que a forma do triângulo não se altera mesmo quando se distorce a grelha de coordenadas diferentemente em cada ponto. Inversamente, se impusermos que a forma do triângulo permaneça inalterada para distorções arbitrárias da grelha de coordenadas, somos obrigados a introduzir um campo de padrão de compensação para restaurar a simetria.

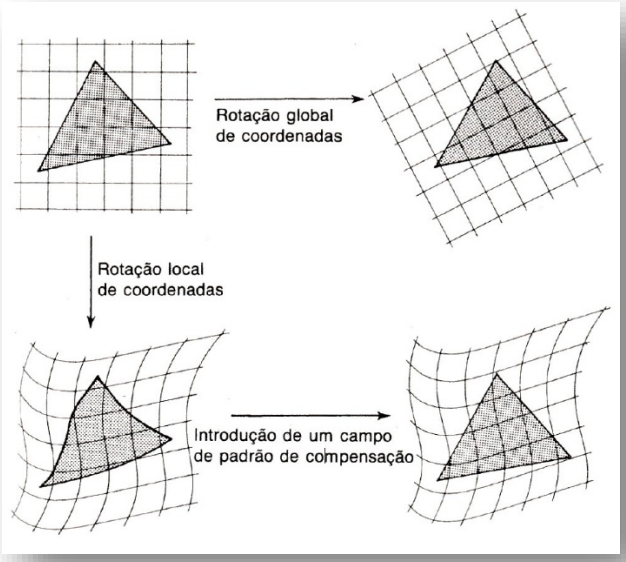
Quando Yang e Mills escreveram o artigo, em 1954, pouca atenção lhe foi dada. Os físicos admiraram o maravilhoso papel dos conceitos de simetria que incorporava, mas não viram como estas ideias podiam ser aplicadas aos problemas com que se debatiam — a elaboração de teorias realistas das interações nucleares forte e fraca.

Dois grandes obstáculos teóricos perfilavam-se no caminho da aplicação do conceito de campo de padrão à física das partículas

quânticas. O primeiro era o problema da renormalização — a teoria do campo de padrão não abeliano não se prestava ao processo de renormalização, que funcionara excelentemente no caso da eletrodinâmica quântica. Este problema só foi ultrapassado nos primeiros anos da década de 70, quando os físicos teóricos, utilizando alguns novos truques, provaram que a teoria do campo de padrão de Yang-Mills era também renormalizável.

O outro problema consistia em que em parte alguma da Natureza era visível a simetria do tipo Yang-Mills. Os físicos teóricos acreditavam que, se a simetria de Yang-Mills fosse exata, então o campo quântico correspondente, as partículas, não podiam ter massa. Nenhuma das partículas observadas experimentalmente parecia possuir as propriedades dos quanta sem massa de Yang-Mills.

Hoje conhecemos a explicação. As simetrias do campo de Yang-Mills não aparecem diretamente na Natureza, manifestando-se indiretamente de duas maneiras: podem ser simetrias exatas, mas completamente escondidas, ou então podem ser simetrias quebradas. Examinemo-las.



O triângulo simboliza um campo com várias componentes e a grelha um sistema de coordenadas que pode representar as rotações no espaço interno. Uma rotação global da grade deixa o triângulo inalterado. Contudo, uma rotação local que varie de ponto para ponto distorce o triângulo e então a situação física muda. Quando se introduz um campo de padrão de Yang-Mills, a simetria perdida é restaurada.

7. Simetrias de Padrão e Quebra de Simetria

Os físicos teóricos mostraram recentemente (por simulações da teoria de campos em computadores) que, se a simetria de Yang-Mills é exata, então fica completamente escondida — todas as componentes do campo que se transformam sob a operação de simetria (como as componentes vermelha, azul e amarela) têm as partículas quânticas associadas confinadas a uma minúscula região do espaço, nunca se manifestando como verdadeiras partículas. Permanecem, em vez disso, ligadas entre si e formam uma bola ou saco — uma partícula compacta. Como veremos no capítulo seguinte, tais objetos existem, correspondem aos hadrões observados, isto é, às partículas com interação forte, como o protão e o neutrão. De qualquer modo, a simetria exata de Yang-Mills

obriga ao confinamento dos *quanta* dos campos associados, sendo essa a razão pela qual não aparecem diretamente na Natureza.

A segunda possibilidade para o campo de Yang-Mills é que a simetria se quebre espontaneamente — as equações do campo possuem simetria, mas a solução das equações não a possui. Uma vez que são as soluções das equações que descrevem o mundo real das partículas quânticas, podemos concluir que no mundo real a simetria original é quebrada e é por isso que não a vemos. Mas como pode uma simetria quebrar-se?

O físico paquistanês Abdus Salam dá o exemplo seguinte: suponhamos que algumas pessoas são convidadas para jantar a uma mesa circular e que entre cada prato e o seguinte há um de salada. Os pratos de salada estão situados simetricamente entre os pratos rasos. A primeira pessoa a sentar-se, desconhecendo as regras do jantar, poderia, com igual facilidade, pegar no prato de salada à sua direita ou no prato de salada à esquerda, e, uma vez feita a escolha, a simetria original ficaria quebrada. As outras pessoas teriam de agir concertadamente; caso contrário, alguém ficaria sem salada. Não importa qual a escolha feita — direita ou esquerda —, qualquer escolha quebra a simetria original direita-esquerda. A solução para uma configuração simétrica quebra a simetria.

Outro exemplo de uma simetria quebrada espontaneamente e que está mais próxima da física real é o «ferromagneto de Heisenberg». Um íman é composto por uma imensidade de pequenos domínios magnéticos que, para o fim que nos propomos

atingir, podemos imaginar como sendo pequenas agulhas magnéticas — pequenos ímanes em forma de barra livres de rodar em torno de um eixo. Suponhamos que colocamos milhares dessas agulhas magnéticas sobre uma mesa, cada uma delas podendo mover-se livremente. Imaginemos também que a mesa está isolada do campo magnético da Terra, de tal maneira que o único campo magnético a que uma agulha pode reagir é o campo produzido pelas vizinhas na mesa.

Inicialmente, todas as agulhas apontam em direções aleatórias. O campo resultante produzido por todos os pequenos ímanes orientados ao acaso é, em média, nulo, dado que os campos se subtraem tantas vezes quantas se adicionam. Por não haver campo magnético resultante, mesmo que rodemos o plano da mesa, não encontraremos uma direção norte-sul preferencial. O sistema físico é, assim, rotacionalmente invariante, ou simétrico, no plano da mesa.

Suponhamos agora que conseguimos orientar algumas agulhas magnéticas numa região, de modo que todas apontem na mesma direção, produzindo o próprio campo magnético resultante. Podemos consegui-lo introduzindo um forte campo magnético externo nessa região e interrompendo-o em seguida. O campo magnético resultante de todas estas agulhas magnéticas orientadas cedo fará com que todas as outras as acompanhem e apontem na mesma direção. A simetria rotacional original ficou quebrada, dado que existe agora uma direção norte-sul preferencial, a direção do campo magnético resultante. Mais: esta nova configuração das pequenas agulhas — esta simetria de rotação quebrada — é

claramente a configuração estável. Se mudarmos manualmente a orientação de uma ou duas agulhas, estas voltarão à orientação original logo que libertas. O ferromagneto de Heisenberg ilustra as ideias fundamentais da quebra espontânea da simetria: embora a situação física original seja simétrica, é também instável; a situação de simetria quebrada é estável.

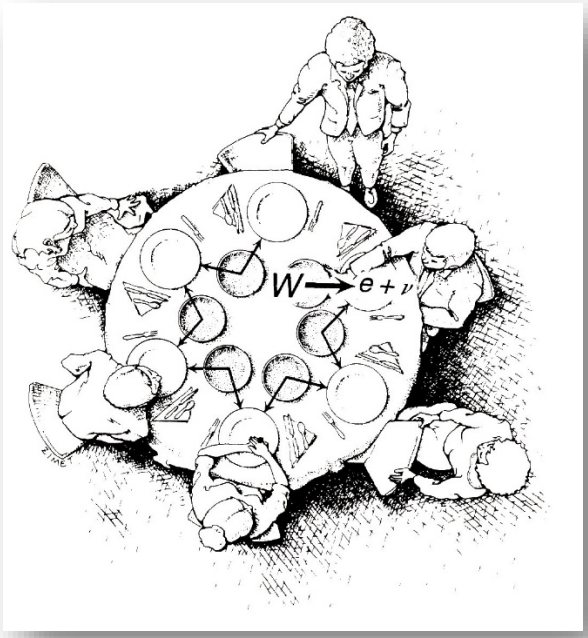


Ilustração de uma simetria quebrada espontaneamente — neste caso, a simetria direita-esquerda dos pratos da salada dispostos entre os pratos rasos. Se uma pessoa escolher um prato de salada, a simetria direita-esquerda é «espontaneamente quebrada».

A primeira sugestão de que as simetrias de padrão podiam quebrar-se espontaneamente proveio dos trabalhos de Peter Higgs, em 1965, físico britânico, e de Richard Brout e P. Englert, físicos da Universidade de Bruxelas. Higgs e eu estávamos ambos na Universidade da Carolina do Norte nessa altura. Lembro-me de o ouvir queixar-se do problema que estava a ter com outro físico, o *referee*[5]" do seu artigo, que pensava que o trabalho estava errado.

Felizmente, Higgs levou a melhor e o artigo, embora com demora, foi publicado. Lembro-me também de que nessa altura não podia imaginar que o trabalho tivesse algo a ver com a física real. Parecia-me um bonito trabalho de matemática, uma curiosidade. O próprio Higgs não estava certo da sua utilidade; naquela altura a maior parte dos físicos também não viam a sua aplicação à física real.

Sucedeu então que em 1967-1968 Steven Weinberg e Abdus Salam empregaram a ideia de Higgs num modelo de teoria de campo de padrão de Yang-Mills que pela primeira vez unificava duas interações distintas entre partículas quânticas — a força eletromagnética (que descreve as interações dos fotões com a matéria) e a força fraca (responsável pelo decaimento das partículas quânticas). O modelo eletrofraco de Weinberg e Salam incorporava as ideias de muitos outros físicos, entre os quais se destacavam Julian Schwinger, Sheldon Glashow e John Ward. Crê-se hoje que este modelo descreve o mundo real, embora tenha permanecido praticamente ignorado até se ter demonstrado em 1971 que as teorias do tipo Yang-Mills são renormalizáveis. Os físicos puderam então usar o modelo para efetuar cálculos pormenorizados das interações fraca e eletromagnética, no espírito do que tinham feito aquando do advento da eletrodinâmica quântica. Teve então início uma revolução em física teórica — a revolução dos campos de padrão —, que ainda hoje continua.

A ideia de Higgs consistia em acrescentar um novo campo ao campo de *padrão*, campo este hoje denominado «campo de Higgs»: tem *spin* zero e massa nula. A virtude do campo de Higgs reside em que os físicos podem utilizá-lo matematicamente para estudar o

processo da quebra de simetria com grande pormenor. Num certo sentido, o campo de Higgs é o «quebrador de simetria» — a primeira pessoa a pegar num prato de salada ou o campo magnético externo que força as agulhas magnéticas a alinhar numa direção comum. Introduzindo convenientemente o campo de Higgs, pode demonstrar-se matematicamente que a solução que conserva simetria das equações do campo é instável — a simetria «quer» quebrar-se, da mesma maneira que todas as agulhas magnéticas querem apontar na mesma direção. A solução instável é análoga a um lápis em pé — tem simetria cilíndrica em torno do bico, mas é instável. Um pequeno toque colocá-lo-á numa configuração assimétrica, mas estável. O campo de Higgs, como o lápis, escolhe a solução estável, mas de simetria quebrada.

A quebra de simetria do campo de Higgs afeta os campos de padrão de Yang-Mills, quebrando também a sua simetria perfeita: os campos de Yang-Mills têm todos massa zero no caso simétrico, mas, quando a simetria padrão se quebra, alguns destes campos de padrão, anteriormente sem massa, adquirem-na.

No caso do modelo eletrofraco, os *quanta*, com massa, dos campos de padrão correspondem às partículas *W* e *Z*, descobertas experimentalmente no CERN, laboratório europeu de altas energias, em 1983. Têm massas enormes, de mais de 90 vezes a massa do protão — uma consequência da quebra de simetria. Facto assinalável, as massas observadas das partículas *W* e *Z* estão de acordo com as previsões da teoria, incutindo grande confiança aos teóricos. Raras foram as ocasiões recentes em que os teóricos de partículas tiveram o prazer de ver ideias matemáticas abstratas

concretizadas na Natureza com tão maravilhosa perfeição. A ideia da simetria de padrão quebrada parece ter vindo para ficar.

Cada êxito em física cria novos problemas e enigmas a um nível mais profundo. O grande enigma é a gravidade. Já vimos como a teoria quântica relativista dos campos é o fruto do casamento da relatividade restrita com a teoria quântica. Mas para termos uma teoria que inclua a gravidade teremos de inventar uma que case a teoria da relatividade geral com a teoria quântica. Apesar de alguns dos melhores cérebros da física terem lutado com este problema durante décadas, ninguém conseguiu ainda resolvê-lo de maneira consistente e coerente. Fez-se alguma luz, mas a teoria quântica da gravidade escapa ainda ao nosso entendimento. Serão, evidentemente, necessários novos e profundos princípios para que os físicos possam incorporar a gravidade na teoria quântica.

A teoria quântica relativista dos campos é a linguagem matemática abstrata das partículas quânticas. Com o êxito do processo de renormalização e com a revolução da teoria dos campos de padrão, a maior parte dos enigmas com que se defrontavam os físicos de partículas na década de 60 foram resolvidos na década de 70. No capítulo seguinte irei descrever o «modelo padrão» — o consenso atual relativamente às teorias de campo que descrevem o mundo real dos quarks, leptões e gluões, os *quanta* dos quais todas as coisas são feitas. O sucesso destas teorias entusiasma os físicos, mesmo quando perguntam: «Que se segue?»

[1] Paul Dirac faleceu em 20 de Outubro de 1984. (*N. do T.*)

[2] Richard Feynman faleceu em 15 de Fevereiro de 1988. (*N. do T.*)

[3] Do inglês *gauge fields*. (N. do T.)

[4] As operações de simetria «abelianas», cujo nome deriva do matemático norueguês Niels Henrik Abel, obedecem à propriedade comutativa $R_1xR_2=R_2xR_1$, o mesmo não se verificando com as operações de simetria «não abelianas» mais gerais. $R_1xR_2 \neq R_2xR_1$.

[5] Especialista a que as revistas científicas recorrem para apreciação dos artigos que lhes são enviados para publicação. (*N. do T.*)

CAPÍTULO 10

O Modelo Padrão

> Além disso, na pesquisa de novas leis tem-se sempre a excitação psicológica de sentir que [...] ninguém pensou ainda na louca possibilidade que nesse momento contemplamos.
>
> *RICHARD P. FEYNMAN, Nobel Lecture, 1965*

Sucede, ocasionalmente que, após anos de experimentação e de luta intelectual, uma imagem nova e coerente do mundo físico emerge da confusão anterior. Foi este o caso da invenção na década de 20 da mecânica quântica, que, finalmente, clarificou o estranho mundo do átomo, que tinha intrigado os físicos durante décadas. Um exemplo mais recente foi a revolução dos campos de padrão, que culminou nos últimos anos da década de 70 na invenção das teorias quânticas relativistas dos campos das forças forte, fraca e eletromagnética. Estas teorias, procuradas durante décadas, trouxeram ordem ao mundo subnuclear. O resultado de tais revoluções é muitas vezes o estabelecimento de um novo consenso científico, uma visão coletiva da ordem do mundo. É importante para a ciência dispor desses consensos, que fornecem tanto um alvo definido para a crítica como um terreno firme para os apoiantes se poderem lançar em novos voos especulativos.

Hoje em dia o «modelo padrão» das interações das partículas nucleares representa um desses consensos. Tem sido um sucesso experimental — nenhuma experiência infirma o modelo padrão. Alguns físicos são da opinião de que o modelo não está ainda adequadamente testado, mas a maior parte pensa que está correto. O modelo é uma teoria quântica relativista dos campos, cujos *quanta* se chamam quarks, leptões e gluões — um conjunto de partículas fundamentais. Irei descrevê-lo pormenorizadamente neste capítulo, mas por agora exponho apenas a ideia fundamental.

Os físicos identificaram quatro interações fundamentais na Natureza: a interação nuclear forte, as interações fracas responsáveis pelo decaimento dos núcleos atómicos e das partículas quânticas, a força eletromagnética e a gravidade. O modelo padrão ocupa-se de três destas quatro forças: a forte, a fraca e a eletromagnética. (A gravidade, seguramente a mais fraca destas forças, traz consigo o problema, até agora insolúvel, de uma teoria quântica da gravidade e está explicitamente excluída.) No modelo padrão cada uma destas três forças tem como portadores um conjunto de partículas quânticas chamadas gluões, que são *quanta* de um campo de padrão de Yang-Mills. A força forte tem como portador um conjunto de oito «gluões coloridos», a força fraca, um conjunto de «gluões fracos», chamados *W* e *Z*, e a força eletromagnética, o fotão, a partícula de luz, também um gluão. Todos estes gluões interagem com um conjunto de partículas chamadas quarks e leptões. Os leptões distinguem-se pela propriedade que têm de interagir somente com os gluões fracos e com o fotão, e não com os gluões coloridos das interações fortes. Os quarks interagem com os três conjuntos de gluões, mas

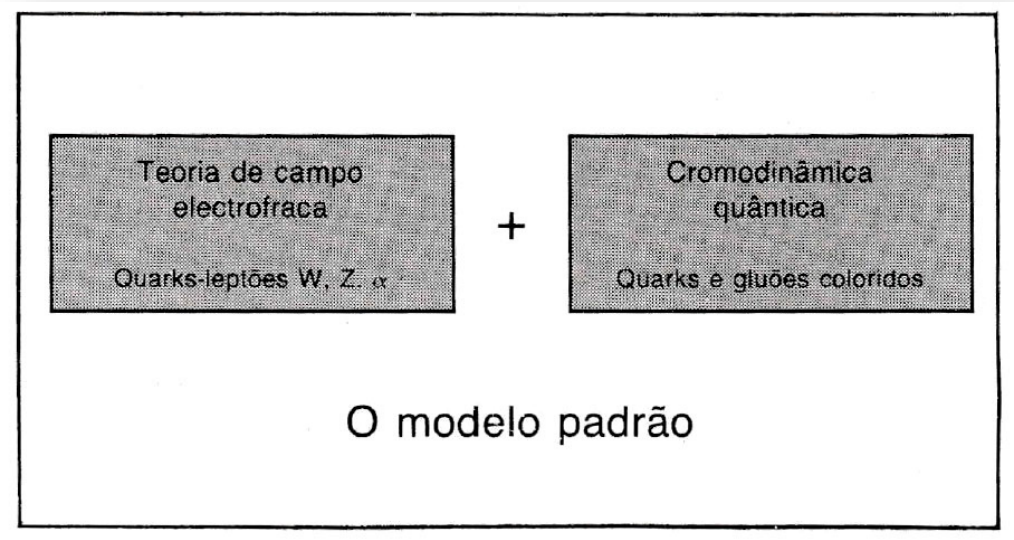
predominantemente com os gluões coloridos com interação forte. Os gluões, como o nome sugere[1], fazem com que os quarks e os leptões se «mantenham juntos». Sem gluões o universo «descolar-se-ia». Consistiria então num gás de quarks e leptões, sem interações, e seria pouco interessante.

O modelo padrão integra elegantemente duas teorias quânticas relativistas dos campos: por um lado, a cromodinâmica quântica, uma teoria de campo de quarks interagindo com os gluões coloridos fortes e, por outro lado, o modelo Weinberg-Salam das interações fraca e eletromagnética unificadas. Juntas, estas teorias de campos podem, em princípio, descrever tudo o que observamos no mundo material, exceto a gravidade.

A cromodinâmica quântica descreve, matematicamente, como os quarks se ligam entre si tão estreitamente que ficam permanentemente confinados em minúsculos «sacos». Estes objetos, semelhantes a sacos, com os quarks retidos lá dentro, são os hadrões, o grande zoo das partículas com interação forte, observadas nos laboratórios com aceleradores de altas energias. Entre estes hadrões estão os protões e os neutrões, cuja relativa estabilidade os distingue dos restantes membros da família e que se juntam para formar os núcleos atómicos. Num certo sentido, os núcleos são sistemas de quarks e gluões coloridos.

O modelo electrofraco unifica a anterior teoria dos fotões e eletrões, chamada «eletrodinâmica quântica», com uma teoria de Yang-Mills das reações exclusivamente fracas, que descreve o decaimento dos quarks e dos leptões. Foi o primeiro exemplo de

uma teoria de campo unificado, na qual duas interações distintas, neste caso as interações eletromagnética e fraca, se tornam não mais do que manifestações separadas de simetrias de campo subjacentes. Este modelo electrofraco serviu de inspiração às tentativas posteriores de unificação dos campos.



O leptão carregado estável é o eletrão, que pode combinar-se com os núcleos, constituídos por protões e neutrões, para formar os átomos. Os átomos podem formar estrelas, planetas, moléculas e vida. O modelo padrão é o primeiro passo da receita para cozinhar um universo.

No capítulo anterior descrevi a teoria quântica relativista dos campos, um quadro conceptual para pensar o mundo microscópico. A teoria dos campos fornece uma linguagem geral para estudar as partículas quânticas, a sintaxe ou as regras a que deve obedecer tal descrição. Mas é questão diferente descobrir as palavras

apropriadas — as partículas quânticas que aparecem na Natureza — que dão conteúdo real a essa linguagem. Nas suas explorações experimentais e teóricas, os físicos descobriram os quarks, os leptões e os gluões — unidades de matéria, aparentemente irredutíveis, a partir das quais tudo pode ser feito. Estas partículas quânticas encaixam como palavras no quadro conceptual da linguagem da teoria dos campos e obedecem às suas regras.

Antes de considerarmos cada uma destas partículas quânticas mais pormenorizadamente, vale a pena recordar que, apesar da linguagem abstrata que os físicos usam para as descrever, estas partículas existem realmente. Ninguém as conhece tão intimamente como os físicos experimentais, que diariamente as estudam no laboratório. A história seguinte ilustra a diferença entre o modo como os teóricos e os experimentadores veem as partículas quânticas: Arthur Eddington, astrofísico teórico, e Ernest Rutherford, físico experimental, conversavam a uma mesa de jantar na década de 20. Eddington observou que os átomos e os eletrões eram, provavelmente, meros conceitos. Rutherford levantou-se da mesa como se a mulher que amava tivesse sido insultada. Repreendendo Eddington, disse-lhe que os átomos não eram somente conceitos; encontrava-os todos os dias no laboratório e eram seus amigos.

1. Os Leptões

Os físicos recolheram provas da existência de um total de seis leptões — a palavra significa «ligeiro», «célere». O grupo dos leptões, composto pelo eletrão, muão, tau e respetivos neutrinos,

pode ser pensado como minúsculas partículas pontuais sem estrutura (pelo menos, nenhuma estrutura foi jamais vista). Na classificação de Wigner têm todos *spin* $^1/_2$. Três dos seis leptões têm massa e uma carga elétrica de -1: o eletrão, o muão e o tau. Os outros três não têm carga, são os três neutrinos (é uma questão em aberto saber se têm massa ou não). Os seis leptões podem ser divididos em três «famílias», com um par de leptões em cada família. Em cada família há um membro com carga elétrica, enquanto o outro é um neutrino eletricamente neutro.

A primeira família de leptões, a «família do eletrão», é composta pelo eletrão, simbolizado por *e*, e pelo seu neutrino associado, simbolizado por um ν_e. De acordo com as regras da teoria quântica relativista dos campos, tem de existir uma antipartícula para cada partícula. O antieletrão, ou positrão, é simbolizado por e, e o neutrino antieletrão, uma partícula distinta do neutrino eletrão, é simbolizado por $\bar{\nu}_e$.

O eletrão, a primeira partícula elementar, foi descoberta há muito, em 1897. É, tanto quanto se pode saber, absolutamente estável e não decai em outras partículas mais leves. A estabilidade absoluta do eletrão é garantida pela lei de conservação da carga elétrica — a carga elétrica total deve conservar-se nas interações entre partículas. O eletrão é a mais leve das partículas quânticas com carga e não pode decair em partículas mais leves porque não existe nenhuma partícula para transportar a sua carga elétrica. Como todas as leis da física, a lei de conservação da carga elétrica está sujeita ao teste da experiência, mas até agora ninguém observou a sua violação.

As leis de conservação, como a lei de conservação da carga elétrica, desempenham um papel importante na compreensão das interações entre partículas e, como veremos, da origem do universo. De acordo com o trabalho de Noether, as leis de conservação são uma consequência da simetria exata. A lei de conservação da carga elétrica, se for absoluta, é o resultado de uma simetria exata das equações da teoria dos campos. Encontraremos outras leis de conservação da carga ao longo da descrição das partículas, sendo todas elas consequência das simetrias internas do modelo padrão.

Os eletrões são, possivelmente, as mais familiares de todas as partículas elementares, devido ao seu uso nos instrumentos eletrónicos. Uma corrente elétrica é simplesmente o movimento de eletrões ou outras partículas carregadas. Descobriu-se que o enxame dos eletrões que rodeia o núcleo atómico é o responsável pelas propriedades químicas dos átomos. Por serem tão abundantes na Natureza e tão fáceis de extrair dos átomos, os eletrões foram amplamente estudados e as respetivas propriedades determinadas com precisão.

O neutrino eletrão, o parceiro da família do eletrão, é extremamente leve — é possível que tenha exatamente massa zero. Os experimentadores tentam tenazmente fixar limites para a sua massa, ou medi-la, caso não seja nula. Os neutrinos, dado que não têm carga elétrica, não interagem diretamente com o campo eletromagnético. Têm apenas interações muito fracas com a restante matéria e, consequentemente, atravessam-na: a terra e tudo o mais que se coloque no seu caminho. Utilizando detetores

extremamente sensíveis, os físicos experimentais conseguiram o êxito assinalável de identificarem eventos subatómicos induzidos pelos neutrinos; obtiveram até feixes de tão fugidias partículas.

A família seguinte de leptões, a «família do muão», compõe-se do muão, simbolizado por μ, e do neutrino muão, simbolizado por ν_μ,. O muão é uma partícula, tanto quanto se pode determinar, idêntica ao eletrão, exceto no facto de ter uma massa 207 vezes maior. Tem a mesma carga elétrica e o mesmo *spin* $^1/_2$ que o eletrão. O parceiro da família do muão, o neutrino muão, tem, provavelmente, massa nula. O muão e o seu neutrino, tal como sucede na família do eletrão, têm antipartículas parceiras. Os muões são, porém, instáveis e não vagueiam como os eletrões, porque, quando decaem até aos eletrões, neutrinos eletrões e neutrinos muões, a carga elétrica pode transitar para o eletrão.

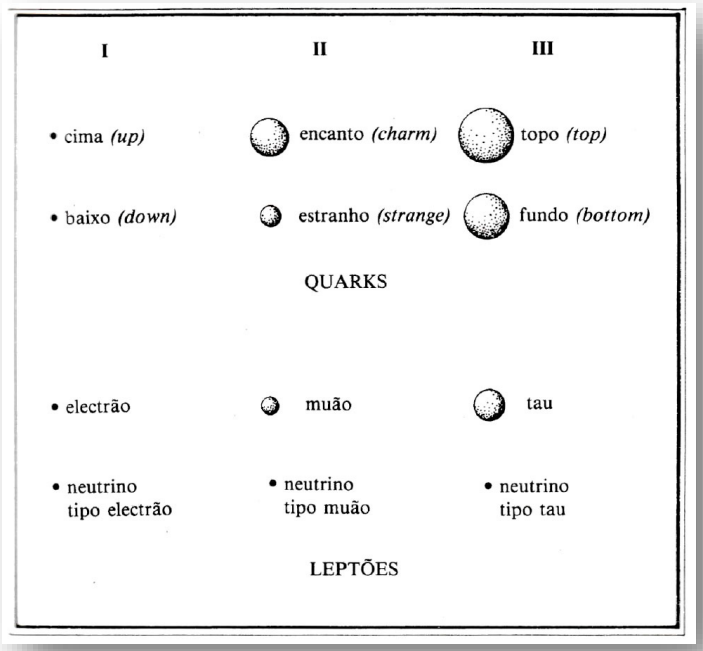
O decaimento do muão exibe ainda outras leis de conservação semelhantes à lei de conservação da carga elétrica — a lei de conservação dos números eletrónico e muónico. Suponhamos que atribuímos uma «carga de número muónico» de +1 ao muão e ao seu neutrino, -1 às antipartículas correspondentes e 0 a todas as outras partículas. De modo semelhante, podemos atribuir uma «carga de número eletrónico» de + 1 ao eletrão e seu neutrino, -1 às suas antipartículas e 0 a todas as outras partículas. Notamos então que o decaimento do muão conserva tanto o número muónico como o número eletrónico. A contabilidade da conservação da carga pode ser apresentada sob a forma da tabela seguinte:

	μ	$\longrightarrow$	e	$+$	v_e	$+$	v_μ
Carga elétrica	-1		-1		0		0
Número muónico	1		0		0		1
Número eletrónico	0		1		-1		0

É claro que este é apenas um processo de decaimento, mas, quando os físicos olham para muitas interações diferentes de muões e eletrões, verificam que todas as novas leis de conservação se mantêm válidas. Alguns pensam que, ao contrário do que sucede com a lei de conservação da carga, estas novas leis de conservação serão violadas, o que será observado um dia, embora a interação que as viola tenha poucas probabilidades de ocorrer. Estas leis de conservação correspondem a simetrias no modelo padrão, o que quer dizer que o modelo terá de ser modificado se as leis forem violadas.

Recentemente, em 1977, foi descoberta outra família de leptões, a família do tau. O membro da família eletricamente carregado chama-se tau, é simbolizado por τ e o neutrino associado por $V\tau$. O tau tem uma massa 3491 vezes maior do que a do eletrão. Se o muão for um eletrão «pesado», então o tau será um muão pesado. O tau, tal como o muão, é instável, apresentando vários modos de decaimento possíveis. Mas o «número tauónico» é, evidentemente, conservado em todos estes processos. Os números tauónico, muónico e eletrónico são todos cargas conservadas no modelo padrão. A soma destas cargas denomina-se «número leptónico»; uma vez que cada carga individual é conservada, o mesmo acontece à soma.

Podemos resumir a nossa classificação dos leptões, o nosso primeiro conjunto de blocos para construir o universo, dispondo-os num quadro. Com três famílias, indicadas por I, II e III, os leptões agrupam-se no esquema seguinte:



Os quarks e leptões detetados agrupam-se em três famílias. Os quarks e os leptões parecem ser partículas pontuais fundamentais. Os leptões, como o eletrão, podem ser detetados diretamente, enquanto os quarks existem apenas ligados em partículas chamadas «hadrões».

onde a seguir ao símbolo que identifica a partícula indicamos a massa em unidades de massa do eletrão:

Quadro dos leptões

Carga eléctrica	Família I		Família II		Família III	
−1	e	1	μ	207	τ	3491
0	ν_e	menos de 0,00012	ν_μ	menos de 1,1	ν_τ	menos de 500

2. Os Quarks

Os leptões podem ser criados e detetados diretamente no laboratório. Os físicos experimentais obtêm feixes de eletrões, de muões e até de neutrinos e disparam-nos contra outras partículas alvo. Se nos debruçarmos agora sobre os quarks, outro conjunto de partículas quânticas pontuais de *spin* $^1/_2$, descobriremos que, em muitos aspetos, se assemelham aos leptões. Também parecem ser partículas pontuais, sem estrutura. Mas existem diferenças significativas no modo como os quarks e os leptões aparecem na Natureza.

Primeiro, ao contrário do que sucede com os leptões, ninguém conseguiu obter feixes de quarks, nem sequer foi visto qualquer quark a vaguear. Existem, mas parecem estar sempre tão estreitamente ligados a outros que não podem separar-se. Tudo leva à conclusão de que os quarks estão permanentemente confinados em sistemas ligados, pequenos sacos que podem ser identificados como hadrões — partículas com interação forte que *podemos* observar no laboratório. Foi utilizando aceleradores de

altas energias como se fossem microscópios e olhando para dentro de hadrões, como o protão e o neutrão, que primeiro foi detetada a presença dos quarks.

Existem muitos hadrões — um número infinito —, mas, notavelmente, todos podiam ter sido formados a partir de apenas alguns quarks orbitando em torno uns dos outros, dentro do saco, numa infinidade de configurações orbitais diferentes. O modelo padrão afirma que existem seis quarks (cinco foram detetados experimentalmente), os quais, ao contrário dos leptões, têm cargas elétricas fracionárias, pouco vulgares, de $-{}^{1}/_{3}$ e ${}^{2}/_{3}$ da unidade de carga. Parecem estar também organizados em três famílias, consistindo cada uma delas num par de quarks.

Aos seis quarks foram dados os símbolos seguintes:

u = u = quark «cima» *(up)*

d = d = quark «baixo» *(down)*

c = c = quark «encanto» *(charm)*

s = s = quark «estranho» *(strange)*

b = b = quark «fundo» (algumas vezes chamado «beleza») (*bottom* ou *beauty*);

t = t = quark «topo» (algumas vezes chamado «verdade») (*top* ou *truth*)[2]

Os antiquarks correspondentes são simbolizados por $\bar{u}$, $\bar{d}$, $\bar{c}$, etc. Estão agrupados nas famílias I, II e III, como a seguir se indica:

Quadro dos quarks

Carga eléctrica	Família I		Família II		Família III	
$^2/_3$	*u*	2	*c*	1 500	*t*	?
$-^1/_3$	*d*	6	*s*	200	*b*	3 000

Todos os quarks possuem massa. Não podemos, claro está, medir diretamente a massa de uma coisa que está permanentemente confinada, mas os físicos teóricos podem atribuir massas aos quarks com base nas suas propriedades, mesmo estando eles ligados. O quark *up* é o mais leve, com uma massa

aproximadamente igual a duas vezes a massa do eletrão. O outro quark leve, o quark *down*, tem uma massa de cerca de seis vezes a do eletrão, enquanto os outros quatro quarks têm massas bastante mais elevadas. O valor aproximado da massa está indicado na tabela de quarks a seguir ao símbolo da partícula, exceto para o quark *top*, cuja massa é desconhecida, embora existam provas de que tenha sido detetado[3].

Estes quarks não existem como tal na Natureza, combinando-se, em vez disso, para formar os hadrões, que podem ser observados. As regras para formar os hadrões a partir dos quarks são bastante simples e constituem aquilo a que se chama o «modelo quark dos hadrões». Foi inventado em 1963 por Murray Gell-Mann e, independentemente, por George Zweig, com o fim de compreender as relações sistemáticas que se observavam entre os hadrões. Seguem-se as regras do modelo dos quarks.

Recordemos que todas as partículas quânticas têm *spin* inteiro ou semi-inteiro. Os hadrões com *spin* semi-inteiro são chamados «bariões»; os que têm *spin* inteiro são os «mesões». Se representarmos cada um dos quarks *u*, *d*, *s*, *c*... pelo símbolo genérico *q*, então os bariões são constituídos por três quarks:

$$qqq$$

enquanto os antibariões são constituídos por três antiquarks:

$$\bar{q}\bar{q}\bar{q}$$

Os elementos da outra grande subfamília dos hadrões, os mesões de *spin* inteiro, são constituídos por um quark e um antiquark, a saber:

$$\bar{q}q$$

Com estas regras elementares pode verificar-se que, dados os quarks com as cargas fracionárias de $^{2}/_{3}$ e $-^{1}/_{3}$ e os antiquarks de cargas $-^{2}/_{3}$ e $-^{1}/_{3}$, todos os bariões e mesões têm valores inteiros de carga de 0, ± 1, ± 2, exatamente as cargas observadas em laboratório. Por exemplo, o protão é constituído por três quarks:

protão ~ *uud*

enquanto o neutrão é composto por três quarks:

neutrão ~ *udd*

Uma vez que o quark *u* tem uma carga de $^{2}/_{3}$ e o quark *d* tem uma carga elétrica de $-^{1}/_{3}$, pode ver-se que a carga do protão é $^{2}/_{3}+^{2}/_{3}-^{1}/_{3} = +1$ e que a carga do neutrão é $^{2}/_{3}-^{1}/_{3}-^{1}/_{3}=0$, o que está certo.

Outros hadrões, as chamadas partículas «estranhas», podem também formar-se a partir dos quarks se substituirmos um quark *d* (down) por um quark *s* (strange). Por exemplo, o lambda, uma partícula «estranha» observada no laboratório, é formado a partir

de quarks, substituindo um dos quarks *d* do neutrão por um quark *s*:

lambda — *usd*

Identicamente, todos os hadrões observados podem ser formados a partir dos quarks.

Dentro dos hadrões, os quarks podem orbitar em torno uns dos outros, com uma infinita variedade de configurações discretas, cada uma das quais correspondendo a um hadrão diferente. Normalmente, só as configurações orbitais de baixa energia, como o protão, o neutrão ou o lambda, são observadas em laboratório. As configurações de energia mais elevada são muito instáveis e decaem rapidamente nas configurações de energia mais baixa.

Embora os quarks não possam ser detetados diretamente, podem ser usados para construir todos os hadrões com interação forte. Os leptões, por outro lado, podem ser detetados diretamente. Mas os quarks e os leptões, se ignorarmos as diferenças respeitantes à maneira como aparecem na Natureza, parecem-se uns com os outros — são ambos partículas pontuais com *spin* $^1/_2$ e parecem organizar-se em três famílias. Esta observação será o trampolim para saltos especulativos no sentido de formular uma teoria unificada dos quarks e dos leptões.

E que dizer das interações entre os quarks e os leptões? Estas interações devem ser importantes, porque unem os quarks entre si,

mas não os leptões. De acordo com o «modelo padrão», estas interações são mediadas pelo outro conjunto de partículas quânticas chamadas gluões, o derradeiro conjunto de quanta do modelo padrão.

3. Os Gluões

Os gluões são uma classe nova de partículas quânticas com *spin* igual a 1. Curiosamente, do ponto de vista da teoria quântica relativista dos campos, a sua existência deve-se a razões de simetria.

Recordemos que todas as partículas quânticas têm um campo associado, sendo o dos gluões o campo de padrão de Yang-Mills. No capítulo anterior descrevi como a existência de campos de Yang-Mills pode ser matematicamente deduzida se postularmos a existência de uma simetria «interna», não apenas globalmente, em todo o espaço-tempo, mas localmente, em cada ponto do espaço-tempo. A imposição desta simetria interna local acarreta a existência de um campo de padrão de Yang-Mills, sendo os *quanta* deste campo os gluões, os quais são, portanto, uma consequência da simetria.

Enquanto mediadores das interações entre os quarks e os leptões, os gluões podem ser imaginados como partículas quânticas que são trocadas entre duas outras partículas quânticas como uma bola trocada entre dois jogadores. Os gluões têm uma força de acoplamento característica dos quarks e dos leptões, que mede a

sua aderência (*stickiness*) — a maior ou menor força com que os gluões acoplam às outras partículas através da interação forte. Esta força de acoplamento dos gluões é proporcional às diferentes cargas — generalizações da ideia de carga elétrica — que os quarks e os leptões possuem. Por exemplo, o fotão, a partícula de luz, é um gluão e acopla à carga elétrica de outras partículas com uma intensidade proporcional a essa carga. Mas existem outros gluões que acoplam a outras espécies de carga.

Os físicos aprenderam que os gluões trocados entre partículas quânticas são responsáveis por todas as forças da Natureza. Cada uma das três forças descritas pelo modelo padrão — a forte, a eletromagnética e a fraca — tem um conjunto de gluões associado e uma teoria matemática de campo que descreve as interações destes. A força forte, que liga os quarks, é mediada por um conjunto de oito «gluões coloridos» e é descrita matematicamente pela teoria de campos chamada «cromodinâmica quântica». As forças eletromagnética e fraca são mediadas por gluões conhecidos por fotão e bosão fraco (simbolizados por *W* e *Z*) e descritas matematicamente pela teoria do campo electrofraco unificado. Consideremos agora mais pormenorizadamente estes gluões e as teorias de campo que descrevem as suas interações.

4. Gluões Coloridos e Cromodinâmica Quântica

Os quarks interagem predominantemente com um conjunto de oito gluões «coloridos». Mas que é a «cor»? Assume-se que cada quark aparece com uma de três cargas, chamada a «carga de cor». Os quarks não são, na realidade, coloridos, mas é útil imaginar que

cada quark possa ser vermelho, azul ou amarelo. Por exemplo, existe um quark *up* vermelho, um quark *up* azul e um quark *up* amarelo. O papel dos oito gluões coloridos é poderem trocar as cargas de cor dos quarks. Por exemplo, se um quark vermelho interage com um dos oito gluões coloridos, pode transformar-se num quark *azul*. Os oito gluões coloridos interagem e mudam as cargas de cor entre os quarks, como também interagem entre eles, trocando as respetivas cargas de cor.

A teoria de campo que descreve matematicamente estas interações entre os quarks coloridos e os oito gluões coloridos chama-se «cromodinâmica quântica», ou, abreviadamente, QCD (do inglês *quantum chromodynamics*). De acordo com a QCD, os gluões coloridos unem os quarks uns aos outros em pequenos sistemas ligados, que podem ser identificados, os hadrões. Estudos da QCD feitos com o auxílio de computadores indicam que a ligação entre quarks devida aos gluões coloridos é tão forte que nunca podem separar-se; é esta a razão pela qual os quarks existem num estado ligado permanentemente dentro dos hadrões.

Que acontecerá se tentarmos libertarmos um quark da sua prisão hadrónica? Ao tentarmos «puxar» um quark do interior de um hadrão, descobrimos que a força dos gluões coloridos aumenta relativamente à distância, o que significa que se torna cada vez mais difícil separá-los, sendo necessário fornecer cada vez mais energia para o conseguir. Esta acaba por ser tão grande que, de acordo com a equivalência massa-energia de Einstein, a energia nos gluões coloridos que ligam o quark se transforma num par quark-antiquark de grande massa, cada um dos quais se torna então parte

de um hadrão. Em vez de libertarmos um quark, acabámos por criar dois hadrões!

Além de fornecer uma explicação para o confinamento dos quarks, a cromodinâmica quântica explica também o modelo quark dos hadrões — dá automaticamente as regras que nos dizem quais as combinações de quarks que se ligam para formar os hadrões. A ideia básica contida na matemática da QCD é que os quarks e os gluões, embora tenham carga colorida, preferem formar combinações de «cor neutra» e não têm carga de cor. Uma maneira simples de pensar acerca disto é imaginar as três cores de um quark como sendo as três cores primárias — vermelho, azul e amarelo — e os três antiquarks como tendo as três cores complementares. Se misturarmos, em partes iguais, as três cores primárias (ou os seus três complementos), obteremos branco neutro. Assim, três quarks com cores diferentes juntos, ou três antiquarks juntos, formam uma combinação de «cor neutra». Mas três quarks juntos formam um barião e três antiquarks formam um antibarião. Esta é precisamente a regra do modelo dos quarks, obtida aqui pela imposição de que a cor esteja «confinada» — só podem existir combinações de cor neutra. Notemos a seguir que a cor e o seu complemento formam uma mistura cinzenta — mais uma combinação de cor neutra, o que corresponde a combinar um quark com um antiquark para fazer um mesão. Vemos que a imposição do confinamento da cor ou da neutralidade fornece as regras para formar os hadrões. Os quarks e os gluões podem resplandecer em «cores» radiantes, mas ocorrem somente nas combinações a preto e branco correspondentes aos hadrões observados.

A cromodinâmica quântica, enquanto teoria das forças que ligam os quarks, surgiu no princípio da década de 70, proposta independentemente por um grande número de físicos teóricos, que conheciam o modelo dos quarks e sabiam que a ideia dos quarks coloridos fornecia as regras necessárias para a construção dos hadrões. O que antes faltara para a invenção da QCD fora a prova de que a teoria do campo de Yang-Mills dos gluões coloridos era renormalizável. Logo que os físicos matemáticos provaram que as teorias de campo, como a QCD, eram, na realidade, renormalizáveis, o entusiasmo começou a instalar-se — a QCD tornara-se uma teoria viável. Quando os físicos teóricos começaram a explorar matematicamente as propriedades de renormalização da QCD, fizeram uma descoberta notável e completamente imprevista: a energias muito elevadas ou a distâncias correspondentemente pequenas, a força de acoplamento dos gluões coloridos torna-se cada vez mais fraca, até que a energias infinitas se anula — a interação entre os quarks e os gluões coloridos desaparece. Esta estranha propriedade, exclusiva das teorias de campo do tipo Yang-Mills, chama-se «liberdade assintótica» — a energias assintoticamente elevadas os quarks e os gluões coloridos comportam-se como se fossem livres e não interagissem. Esta descoberta matemática foi feita por Hugh David Politzer, da Universidade de Harvad e, independentemente, por David J. Gross e Frank Wilczek, da Universidade de Princeton, em 1973. Baseava-se, em parte, no trabalho anterior de Curtis Callen, Kurt Symanzik e Ken Wilson sobre a teoria da renormalização. Esta descoberta acerca da QCD harmonizava-se maravilhosamente com as experiências feitas anteriormente no Stanford Linear Accelerator Center[4] e noutros laboratórios de física de altas energias que

mostravam já que os quarks dentro do protão e do neutrão se comportavam realmente como se fossem partículas livres dentro das suas prisões em forma de saco. Estas experiências vieram dar apoio à ideia da liberdade assintótica e à convicção crescente de que a QCD era a teoria correta para descrever as interações que ligavam os quarks.

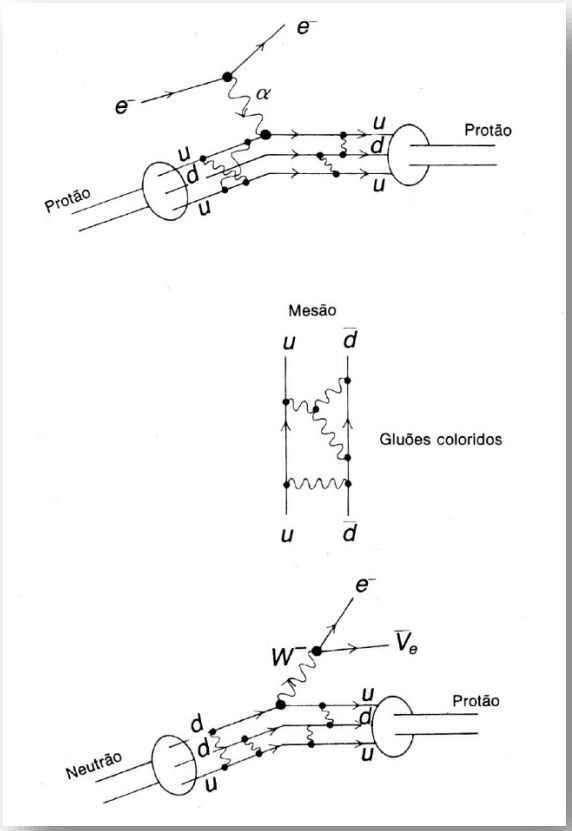


Diagrama das interações das partículas quânticas de acordo com o modelo padrão. Em conformidade com a cromodinâmica quântica, os hadrões de interação forte, como o protão e o neutrão, são combinações de quarks ligados pelos gluões coloridos, representados aqui pelas linhas onduladas que ligam os quarks. O primeiro diagrama mostra um eletrão (e^-) dispersado por um protão feito de quarks (*uud*) através da troca de um fotão (γ) — um *quantum* de luz, que é também um gluão. Foi através de processos deste género que a presença de quarks foi primeiro detetada nos laboratórios de física das altas energias. O segundo diagrama é uma representação de um mesão π e de outra partícula quântica de interação forte composta por um quark (u) e por um antiquark ($\bar{d}$), que aqui se veem trocando gluões coloridos. O último diagrama representa o decaimento de um neutrão num protão, num eletrão (e^-) e num neutrino antieletrão (ν_e). O decaimento ocorre porque um gluão fraco (*w*) pode transformar um quark down (*d*) num quark *up* (u). Os gluões coloridos não podem fazer isto.

A QCD apresenta um conjunto de simetrias internas que se traduzem na existência de leis de conservação da carga — leis que se manifestam nas interações fortes dos hadrões. A carga de cor é uma grandeza conservada, mas, uma vez que todos os hadrões são de cor neutra, não existe maneira de ver esta lei da conservação em ação no laboratório. Seria como postular uma lei de conservação da carga elétrica para um mundo composto apenas de partículas neutras — não veríamos nunca nenhuma carga elétrica para observar se era ou não conservada.

Mas existem outras leis de conservação da carga aplicáveis aos hadrões que podem ser testadas em laboratório. O número de quarks de cada espécie — *up*, *down*, *strange*, *charm*, etc. — é necessariamente conservado (os antiquarks são contados com sinal menos), o que significa que numa interação entre os hadrões em forma de saco o número de quarks *up* e *down* permanece o mesmo. Podem saltar de hadrão para hadrão no momento da colisão, quando os sacos se sobrepõem, mas o número total não muda. Uma vez que as antipartículas são contadas com sinal menos, a lei de conservação do número de quarks *up* permite também a criação de um par quark-antiquark *up* a partir de energia pura. A aplicação das diferentes leis de conservação dos números de quarks às interações fortes dos hadrões observados foi confirmada por milhares de experiências laboratoriais e ninguém as põe em dúvida.

As interações fracas, que ainda não descrevi, violam as diferentes leis de conservação do número de quarks. Por exemplo, um quark *charm* pode transformar-se num quark *down* através da interação fraca, o que viola a conservação do número de quarks

charm e do número de quarks *down*. Mas até a interação fraca respeita a lei de conservação do número total de quarks. O número de quarks menos o número de antiquarks é rigorosamente conservado no modelo padrão.

Esta lei, aparentemente absoluta, de conservação do número total de quarks tem como consequência uma severa lei de conservação que afeta as interações correspondentes entre hadrões — a lei de conservação do número de bariões. Os bariões são a família de hadrões de *spin* semi-inteiro, implicando esta lei que o número total deva ser conservado em todas as interações. O protão, o mais leve dos bariões, deve ser absolutamente estável, como consequência da lei de conservação do número de bariões. Não existe nenhuma partícula a que possa ser cedida a carga bariónica, da mesma maneira que o eletrão não dispõe de nenhuma partícula mais leve à qual possa ceder a sua carga elétrica. E isto é bom! Os protões constituem a maior parte da matéria visível do universo; se pudessem decair rapidamente, o universo decompor-se-ia.

Mais tarde, à luz de outras teorias de campo que vão para além da QCD e do modelo padrão, reexaminaremos a lei de conservação do número de bariões e a estabilidade do protão. Estas outras teorias de campo implicam que o protão *seja* instável, embora a sua taxa de decaimento seja tão lenta que o universo não é afetado por este facto.

Segundo os teóricos, a cromodinâmica quântica faculta um modelo matemático completo de todas as interações fortes dos

hadrões. A descoberta de uma teoria exata da força nuclear forte realiza um sonho de décadas. Infelizmente, embora a QCD seja uma teoria de campo matematicamente elegante e elaborada com base em belas simetrias, é muito difícil resolver as suas equações e extrair resultados precisos que possam ser comparados com os resultados das experiências. Mas, para os poucos pormenores das interações fortes que *podem* ser extraídos da teoria e comparados com a experiência, a concordância é rigorosa. Outras comparações entre a teoria e a experiência terão de esperar até que os teóricos resolvam as equações da QCD em poderosos computadores — uma área de investigação de grande atividade. A despeito das dificuldades atuais, a maior parte dos físicos teóricos esperam ter encontrado a teoria certa da forte força que liga os quarks. Por vezes, gosto de irritar os meus colegas experimentalistas dizendo-lhes que as suas difíceis experiências envolvendo a interação forte não são mais do que cálculos analógicos para resolver as equações da QCD.

5. Os Gluões Eletrofracos e o Modelo de Weinberg-Salam

O mundo dos quarks e dos gluões coloridos, cujas interações são tão fortes que nos permitem praticamente ignorar tudo o mais, constitui, por direito próprio, um mundo exclusivo de interações. Todavia, existem outras interações — as interações fraca e eletromagnética — que envolvem tanto quarks como leptões. As interações eletromagnética e fraca costumavam ser consideradas distintas. Hoje, porém, são vistas como uma interação electrofraca unificada, mediada por um conjunto de quatro gluões: o fotão,

simbolizado por *y*, que é o *quantum* do campo eletromagnético, e os três gluões fracos, que consistem em dois gluões carregados eletricamente, simbolizados por W^+ e W^-, e num neutro, Z^0, que são *quanta* de um campo de Yang-Mills. São estes os gluões electrofracos, responsáveis por novas forças entre os leptões e os quarks.

O fotão, o primeiro gluão a ser detetado diretamente como partícula quântica, acopla à carga elétrica das partículas. Diferentemente dos gluões coloridos, que produzem uma força entre os quarks cuja intensidade aumenta com a distância, os fotões produzem uma força entre as partículas cuja intensidade diminui com a distância — a força de Coulomb, descrita no século XVIII pelo físico francês C. A. de Coulomb. Por esta razão, as partículas eletricamente carregadas, como os eletrões, podem ser separadas do resto da matéria e produzir campos elétricos de longo alcance, que se estendem por distâncias macroscópicas. O mesmo se passa no caso dos campos magnéticos. Por exemplo, o campo de um íman em forma de barra pode estender-se até grandes distâncias, grandes, quando comparadas com as distâncias atómicas. Graças ao longo alcance dos campos eletromagnéticos, ao facto de a luz ser uma onda eletromagnética e ao importante papel do campo eletromagnético na ligação dos eletrões ao núcleo, a interação eletromagnética pode ser facilmente estudada em laboratório. Foi a primeira a ser domada pelos teóricos e hoje é a mais bem compreendida de todas as interações.

A primeira teoria quântica relativista dos campos moderna, chamada «eletrodinâmica quântica», foi inventada na década de 40

e resume tudo o que sabemos sobre a interação da luz com os eletrões. A eletrodinâmica quântica teve tal êxito na explicação das propriedades experimentalmente observadas da interação eletromagnética que se tornou o paradigma de todas as futuras teorias de campo.

Que são as interações fracas? Ao contrário da interação eletromagnética, a interação fraca tem um alcance muito curto e só se torna observável à escala subnuclear de distâncias através do decaimento das partículas quânticas. Um exemplo elementar e cuidadosamente estudado da interação fraca é o decaimento de um neutrão livre num protão, um eletrão e um neutrino antieletrão, decaimento esse que tem lugar, em média, a cada 1000 segundos. Quase todas as partículas quânticas acabam por decair em partículas mais leves.

Ao contrário da interação eletromagnética, as interações fracas eram muito complexas e enigmáticas, pelo que os físicos gastaram décadas para descobrir as respetivas propriedades. Parte da dificuldade reside na fraquíssima intensidade da força fraca; outra dificuldade reside no seu curtíssimo alcance. No entanto, se não fosse pela existência da interação fraca, que se sabe ser mediada pelos gluões fracos de grande massa, simbolizados por W^+, W^- e Z^0, os quarks mais pesados e os leptões seriam absolutamente estáveis e não decairiam em partículas mais leves. Muitas formas exóticas de matéria poderiam então existir indefinidamente e não apenas durante breves instantes em colisões de alta energia entre as partículas quânticas. Se a interação fraca, que elimina as partículas «estranhas» e «encantadoras» do mundo, permitindo-lhes decair,

fosse «desligada», o mundo seria, na verdade, muito estranho. Todas estas partículas exóticas se poderiam tornar blocos químicos para a construção de novas formas de matéria «estranha».

O papel dos gluões de interação fraca, W^+, W^- e Z^0, é o de transformarem quarks em outros quarks e leptões em outros leptões. Interagindo com um gluão *W*, um quark *charm* pode ser transformado num quark *strange* ou *down*. De igual modo, os gluões fracos interagem com os leptões. Através da interação com um gluão *W*, um *tau* pode ser transformado num *neutrino tau*. Regra geral, a existência destes gluões fracos significa que todos os diferentes quarks — *up*, down, *charm*, *strange*, etc. — podem ser transformados uns nos outros e leptões diferentes, dentro de uma família, podem transformar-se igualmente uns nos outros. Mas os quarks não podem transformar-se em leptões, e vice-versa, porque estas interações fracas respeitam várias leis de conservação de número.

A tendência da Natureza é no sentido de as partículas mais pesadas, uma vez que têm maior massa-energia, libertarem esta energia por decaimento em partículas mais leves, sendo a energia da partícula original transformada na energia do movimento das partículas mais leves. Segue-se daqui que os hadrões mais pesados, «estranhos» ou «encantadores» — hadrões que contêm um quark *strange* ou *charm* —, decairão em hadrões mais leves, que não contêm estes quarks de massa elevada. Em última análise, só o eletrão, o neutrino e o quark *up*, de menor massa, são estáveis. Dado que os quarks só podem ocorrer dentro dos bariões, tal significa que somente o barião mais leve — o protão — é estável. A

estabilidade destas diferentes partículas é garantida pelas leis de conservação das cargas e pelo facto de elas serem as partículas mais leves que transportam uma carga conservada. É importante que tais partículas estáveis existam. De contrário, não teríamos nada com que construir o universo visível.

A teoria quântica relativista dos campos, que descreve as interações fraca e eletromagnética unificadas, foi descoberta por Steven Weinberg e Abdus Salam em 1967-1968. Não somente descrevia corretamente as interações fraca e eletromagnética como eram então conhecidas, como previu ainda aspetos completamente novos destas interações, aspetos que, quando observados experimentalmente mais tarde, deram grande credibilidade à teoria. Talvez a descoberta mais emocionante tenha sido a que em 1983 foi realizada pelos experimentadores do CERN, quando detetaram os gluões *W* e *Z* com os valores de massa previstos — uma maravilhosa confirmação da teoria electrofraca.

Todos os gluões que descrevi anteriormente — os gluões coloridos da interação forte e os fotões — têm massa zero. Por outro lado, os gluões fracos *W* e *Z* têm massa elevada. Como pode existir tão grande diferença de massa entre os gluões *W* e *Z* e o fotão se as interações fraca e eletromagnética mediadas por estes mesmos gluões estão, de facto, unificadas? A resposta a esta pergunta reside na noção de «simetria quebrada».

A ideia de simetria quebrada está no âmago da síntese electrofraca de Weinberg e Salam. Esta ideia pode ser explicada de maneira bastante simples. De acordo com a teoria matemática dos

campos quânticos, que estes cientistas elaboraram, o fotão e os gluões fracos são todos *quanta*, associados com diferentes componentes dos campos de Yang-Mills, que possuem uma simetria de Yang-Mills que os transforma uns nos outros, de tal maneira que todos são manifestações de campos unificados subjacentes e estão, neste sentido, relacionados. Mas o modelo electrofraco apresenta, além de campos Yang-Mills, de leptões e de quarks, um novo ingrediente, o campo de Higgs. Este campo, com massa de *spin* zero, determina de forma crucial a maneira como aqueles campos se vão manifestar na Natureza. O campo de Higgs tem também várias componentes; se escrevermos as suas equações, descobriremos que possuem igualmente a simetria de Yang-Mills. Mas a solução estável destas equações para o campo de Higgs não tem simetria; em vez disso, apresenta uma simetria quebrada. Uma vez que o campo de Higgs interage com os outros campos, as equações destes, embora simétricas, também têm agora soluções nas quais a simetria é quebrada. A quebra de uma simetria afeta todos os campos.

O campo de Higgs é uma espécie de «quebrador de simetrias» que destrói a simetria de Yang-Mills original. Contudo, no caso das interações fracas, a simetria quebrada é justamente o que é preciso para diferenciar as interações eletromagnéticas das fracas. Os campos *W* e *Z* adquirem grandes massas devido a este «mecanismo de Higgs», que quebra as simetrias, enquanto o fotão, que conserva um resíduo da simetria original exata de Yang-Mills, permanece sem massa. Embora a teoria subjacente seja simétrica e unificada, a sua manifestação no mundo real não o é. Como comentava Weinberg, «mesmo que uma teoria postule um elevado grau de simetria, não

é necessário que [...] os estados das partículas possuam essa simetria [...] Nada na física me parece tão prometedor como a ideia de que é possível uma teoria ter um elevado grau de simetria oculto no mundo real.

O campo de Higgs parece desempenhar um papel crucial na quebra da simetria das equações, e as partículas quânticas associadas a este campo deverão um dia ser detetadas pelos experimentadores. Infelizmente, ao contrário do que sucede com as massas do *W* e do *Z*, a teoria não dá uma previsão precisa para a massa da partícula de Higgs. No entanto, produzir e detetar em laboratório uma partícula de Higgs seria uma poderosa confirmação destas ideias. Os experimentadores do laboratório de altas energias do CERN vão procurá-la, agora que o *W* e o *Z* foram detetados.

Eis, pois, como as coisas são: o modelo padrão consiste no conjunto formado pela cromodinâmica quântica para as interações fortes dos quarks e pela teoria electrofraca para as interações eletromagnética e fraca dos quarks e dos leptões. Parece concordante com todas as experiências; contudo, falta fazer mais testes. Em 1984 os experimentadores relataram alguns acontecimentos de alta energia observados no CERN que podem ser difíceis de explicar no quadro do modelo padrão. Mas o significado destes acontecimentos não é claro. A maior parte dos teóricos de partículas tomam o modelo padrão como uma verdade absoluta, que descreve como se constrói um universo, e qualquer modelo futuro deverá incluí-lo.

6. Para Além do Modelo Padrão

A maior parte dos físicos, porém, sente que o modelo padrão é, no fundo, insatisfatório e não pode ser a última palavra. Pensam que devemos ir para além dele, a fim de alcançarmos um conhecimento ainda mais profundo das partículas quânticas. A razão fundamental desta insatisfação com o modelo padrão, malgrado os seus êxitos, é que deixa em aberto um certo número de questões fundamentais. Enquanto estas questões se mantiverem sem resposta, os físicos sabem que não acabaram o trabalho.

O modelo padrão comporta 19 parâmetros — as massas dos quarks e dos leptões, as forças de acoplamento, etc. — de entrada, que devem ser determinados previamente por via experimental. Dados estes números, pode-se, em princípio, descrever matematicamente todas as interações fortes, fracas e eletromagnéticas que observamos em laboratório. Isto é uma grande conquista. No entanto, a maior parte dos físicos sente que a teoria final não deveria ter parâmetros de entrada, nem qualquer constante fundamental adimensional, e que todas as massas dos quarks e todas as forças de acoplamento deveriam ser previsíveis por uma tal teoria mestra.

Einstein interrogava-se se Deus teria tido alguma possibilidade de escolha ao criar o mundo da maneira como o fez e certa vez escreveu:

> No que respeita a tais [constantes adimensionais], gostaria de enunciar um teorema que, presentemente, apenas pode

> ser baseado na fé na simplicidade, isto é, na inteligibilidade da Natureza: não existem constantes arbitrárias desta espécie, quer dizer, a Natureza é constituída de tal maneira que é logicamente possível estabelecer leis tão fortemente determinadas que dentro delas só podem ocorrer constantes completamente determinadas racionalmente (e não, portanto, quaisquer constantes cujo valor numérico possa ser mudado sem destruir a teoria).

Se Einstein estiver certo e a teoria mestra, ainda por descobrir, não comportar parâmetros livres, então Deus não teve escolha — não podia ter «ajustado» os parâmetros, como as massas dos quarks, para fazer universos diferentes.

Uma outra razão pela qual o modelo padrão não satisfaz muitos físicos é por ser incompleto em dois aspetos. Primeiro, não inclui a gravidade nem os princípios da teoria da relatividade geral. Segundo, porque a unificação dos campos no modelo padrão é ainda incompleta. Embora os campos eletromagnético e fraco estejam unificados no modelo de Weinberg-Salam, o campo colorido forte não está unificado com nenhum destes.

Alguns físicos, encorajados pelo sucesso da síntese electrofraca, trataram de englobar a interação forte numa unificação mais completa dos campos e construíram modelos conhecidos por «teorias de grande unificação» (grand unified theories), ou, abreviadamente, GUT. De acordo com as GUTs, os gluões fraco, eletromagnético e colorido forte são *quanta* associados a um único campo de padrão de Yang-Mills com várias componentes. No mais simples de tais modelos, chamado «modelo SU(5)», o campo unificado único de Yang-Mills tem 24 componentes, das quais 12

correspondem aos três gluões fracos, ao fotão e a oito gluões coloridos do modelo padrão, correspondendo as outras 12 a gluões completamente novos. Este campo de Yang-Mills interage com os quarks, com os leptões e também com um novo conjunto de campos de Higgs que quebram a simetria única que diferencia as interações forte, fraca e eletromagnética.

De acordo com este modelo, os 12 novos gluões adquirem enorme massa quando a simetria é quebrada, uma massa tão grande que estas novas partículas nunca serão detetadas em nenhum acelerador de partículas, porque nenhum acelerador terá alguma vez energia suficiente para os criar. Estes 12 novos gluões têm uma importante propriedade que os gluões forte, fraco e eletromagnético do modelo padrão não têm: podem transformar quarks em leptões, e vice-versa. As suas interações podem violar as leis de conservação dos números de quarks e de leptões, o que significa que podem fazer com que um dos quarks dentro do protão se transforme num leptão. O protão, que antes era absolutamente estável no modelo padrão, torna-se agora instável no modelo GUT mais simples. O decaimento do protão parece ser uma consequência natural da ideia da grande unificação, embora seja possível elaborar modelos especiais das GUTs nos quais o decaimento do protão não ocorre.

A vida média do protão foi calculada em 10^{31} anos no modelo GUT mais simples, o SU(5) — na realidade, muito tempo: milhares de milhões de vezes a idade atual do universo. Mas já em 1982 os experimentadores que investigam o decaimento do protão com detetores gigantescos — piscinas cheias de água rodeadas por

tubos fotoelétricos capazes de detetar os produtos do decaimento do protão — mostraram que o seu tempo de vida excede 10^{31} anos. O modelo SU(5), a GUT mais simples, está, portanto, excluído. Os experimentadores continuam a fazer funcionar os detetores, na esperança de encontrarem um decaimento. Seria muito excitante para os físicos se o decaimento do protão viesse a ser observado — este facto viria em apoio da ideia das GUTs. Então, através do estudo dos pormenores do processo de decaimento do protão, os físicos poderiam obter indicações experimentais de qual a GUT mais promissora, caso existisse alguma.

Os físicos teóricos foram para além do modelo padrão de outras maneiras, até mesmo para além das GUTs. Um desses conjuntos de ideias dá pelo nome de «supersimetria» e «supergravidade». A supersimetria é uma nova espécie de simetria que transforma os fermiões — campos com *spin* semi-inteiro — em bosões — campos com *spin* inteiro —, e vice-versa. A supergravidade é uma versão local da supersimetria, da mesma maneira que a simetria de Yang-Mills é uma versão local da «simetria interna» vulgar, e incorpora a gravidade no programa de unificação dos campos. Alguns teóricos brilhantes inventaram modelos de supergravidade que unificam todas as interações, os quais não são, porém, muito realistas.

As GUTs e a supersimetria são «ideias loucas» no limite da investigação atual em física teórica. Podem estar certas ou erradas; não existem provas experimentais a favor ou contra elas. No entanto, encerram a esperança de uma nova e mais coerente descrição da Natureza e atraíram muitos jovens e brilhantes físicos

para trabalharem nesse campo. Irei descrever esses esforços com mais pormenor num dos próximos capítulos.

É claro que muitas destas ideias loucas nunca serão testadas nos aceleradores convencionais de altas energias. As novas teorias descrevem regiões do espaço e do tempo tão pequenas e energias tão elevadas que nenhum acelerador construído na Terra alguma vez poderá sondá-las. A única vez em que semelhantes energias estiveram em jogo foi antes do primeiro milionésimo de bilionésimo de segundo do *big bang*. Cada vez mais os físicos teóricos se viram para a dinâmica do universo primordial e aí procuram pistas que possam contradizer ou apoiar estas ideias loucas.

Antes de descrever estas ideias loucas, vejamos quais as consequências do mais bem testado modelo padrão no que toca ao universo primitivo. Mas antes temos de aprender um pouco de termodinâmica e de cosmologia.

---

[1] Do inglês glue, cola. (*N. do T.*)

[2] Conservamos as designações inglesas, por estarem consagradas na literatura. (*N. do T.*)

[3] A descoberta do quark *top* não foi até agora confirmada. (*N. do T.*)

[4] Centro do Acelerador Linear de Stanford. *(N. do T.*)

CAPÍTULO 11

Termodinâmica e Cosmologia

A segunda lei da termodinâmica prevê um único fim para o universo, uma «morte térmica», na qual a temperatura será tão baixa que tornará a vida impossível.

SIR JAMES JEANS

Nos últimos dois capítulos descrevi a teoria quântica relativista dos campos — a linguagem das partículas quânticas — e o conteúdo do mundo quântico microscópico — o «modelo padrão» dos quarks, leptões e gluões. De posse desta informação, estamos quase prontos para recuar no tempo a fim de estudarmos o *big bang*. Contudo, antes de o fazermos, é preciso introduzir mais alguns conceitos: a termodinâmica e o papel da geometria cósmica global do espaço e do tempo. Juntaremos então estas três ideias — a física das partículas quânticas, a cosmologia e a termodinâmica — e elaboraremos um modelo matemático do universo primitivo.

O nosso modelo do universo primitivo será um gás de partículas quânticas enchendo uniformemente o universo inteiro. O que a cosmologia traz de novo a este modelo é a possibilidade de o espaço se contrair ou expandir ao longo do tempo — uma particularidade que influencia o gás contido no espaço. Já descrevi as partículas quânticas e as suas interações e as cosmologias FRW para a

estrutura global do espaço e do tempo. Discutirei agora algumas particularidades da termodinâmica dos gases necessárias para completar o nosso modelo.

Esqueçamos por agora o universo e a cosmologia e consideremos um gás encerrado num recipiente, com temperatura, pressão e volume definidos, propriedades macroscópicas que caracterizam o estado do gás. As leis físicas da termodinâmica que relacionam entre si propriedades macroscópicas de um gás eram já compreendidas pelos físicos do século XIX. No entanto, foi só quando os físicos adotaram um ponto de vista mais profundo que o significado destas leis termodinâmicas foi, finalmente, reconhecido.

Para tal, lembremo-nos de que os gases não são o meio contínuo que superficialmente aparentam ser, consistindo, na verdade, em enormes quantidades de partículas vagueando e colidindo umas com as outras ou com a parede do recipiente. Os físicos derivaram matematicamente as leis da termodinâmica acima referidas, assumindo que cada partícula obedece às leis do movimento da mecânica de Newton e fazendo a média sobre o movimento de todas as partículas. Este novo avanço, chamado «mecânica estatística», deu origem a uma nova e profunda visão da Natureza e das propriedades coletivas da matéria. Por exemplo, de acordo com a mecânica estatística, a temperatura de um gás é proporcional à energia média do movimento de todas as partículas (quanto mais rapidamente as partículas se moverem, tanto mais elevada será a temperatura) e a pressão é proporcional ao seu movimento médio. Desta maneira, as variáveis macroscópicas que descrevem um gás

podem ser entendidas como traduzindo as propriedades coletivas de todas as partículas do gás.

Os gases possuem outras propriedades macroscópicas, além da temperatura e da pressão. Entre elas encontra-se a entropia — uma medida estatística da desarrumação das partículas que vagueiam. Para ilustrar a entropia, suponhamos que o recipiente de gás contém dois gases diferentes: chamemos-lhes A e B. Podemos imaginar que, numa configuração inicial, todas as partículas A estão agrupadas numa metade do contentor e que todas as partículas B estão agrupadas na outra metade, separadas por uma barreira. Retiremos essa barreira. As partículas A e B começam a misturar-se e em breve o contentor está cheio de uma mistura uniforme. Como podemos descrever o que aconteceu em termos de entropia?

A entropia é a medida do grau de desordem de um sistema físico. Mas como se mede a desordem? A resposta fundamental vem-nos da teoria das probabilidades, o estudo matemático do acaso. As configurações improváveis de todas as partículas de gás são consideradas «ordenadas» e é-lhes atribuída uma entropia baixa; enquanto configurações prováveis, são as mais «desordenadas» e têm entropia elevada. Por exemplo, se nos imaginarmos a dar as várias mãos de cartas num jogo de póquer, o mais provável é que a maior parte das mãos sejam conjuntos desordenados de cartas. Estas configurações «desarrumadas» têm alta entropia. As poucas mãos desejáveis têm uma baixa probabilidade de ocorrer — um conjunto de configurações de cartas com baixa entropia.

Apliquemos agora este tipo de raciocínio ao nosso recipiente com gás. Quando a barreira é removida, as partículas A e B estão ainda separadas, uma configuração improvável, correspondente a um estado relativamente ordenado. As partículas começam a misturar-se, dado que uma configuração mista das partículas A e B é o estado mais provável. Constatamos, assim, que a entropia, que é uma medida do grau de desordem, aumenta quando os gases se misturam. Este aumento da entropia — a transição de um sistema físico fechado de uma configuração de relativa baixa probabilidade para outro com probabilidade mais elevada — constitui a segunda lei da termodinâmica e é uma das pedras angulares da mecânica estatística.

Quando um gás, como o que estamos a considerar, atinge um estado de entropia máxima, isto é, quando as partículas estão completamente misturadas e desarrumadas ao máximo, diz-se estar em «estado de equilíbrio». Não há nada que possamos fazer para aumentar a desarrumação; logo, o gás está em equilíbrio, adquirindo a estabilidade da desordem completa. Falando com rigor, devemos referir-nos a esta situação com um estado de «equilíbrio térmico», querendo significar que a temperatura é uniformemente a mesma em qualquer ponto da massa de gás.

Os gases em estado de equilíbrio térmico têm muitas propriedades importantes, que podem ser rigorosamente provadas com a matemática da mecânica estatística. Por exemplo, poder-se-ia pensar que as propriedades de um tal gás dependem dos pormenores das interações entre todas as diferentes partículas do gás, como se chocam, como embatem contra as paredes do

recipiente. No entanto, e notavelmente, de acordo com a mecânica estatística, o conhecimento destes pormenores não tem qualquer importância. Tudo o que é necessário saber para determinar o estado físico do gás é o facto de as diferentes partículas interagirem e colidirem de maneira a poderem transferir energia de umas para as outras.

O que será então importante para especificar o estado físico do gás? O que é espantoso num gás em equilíbrio térmico é que, logo que conhecemos a sua temperatura e as densidades das grandezas conservadas nas interações entre as partículas, o estado fica definido. No gás que temos vindo a considerar o número de partículas A e o número de partículas B são grandezas conservadas. Para determinar as respetivas densidades basta dividirmos o número total de partículas pelo volume que ocupam. Logo que conheçamos a densidade destas partículas e a sua temperatura, ficamos também a conhecer o estado do gás.

Estas mesmas considerações aplicam-se aos gases de partículas quânticas. As grandezas necessárias para definir o estado do gás são a temperatura, o número de espécies de partículas quânticas no gás e a densidade de partículas conservadas nas interações — a carga elétrica, o número leptónico, o número bariónico. Eis por que as respetivas leis de conservação irão provar ser muito úteis quando aplicarmos a mecânica estatística ao universo primitivo.

As partículas quânticas, contudo, obedecem às leis da mecânica quântica, não às leis newtonianas, o que modifica algumas das equações da mecânica estatística. Os físicos levaram a cabo todas

estas modificações de maneira a que a mecânica estatística possa ser aplicada com precisão aos gases de partículas quânticas. Mas essas alterações não afetarão as propriedades qualitativas dos gases que atrás descrevi.

A entropia de um gás de partículas em equilíbrio é fácil de calcular; de acordo com a mecânica estatística, é proporcional ao número total de partículas. Quanto mais partículas o gás contiver, mais desordenado se pode tornar e maior será a entropia. Se um gás se compuser de partículas A e B, então podemos considerar entropias separadas para as partículas A e B, dado que os números de partículas A e B podem diferir. Fala-se então correntemente de «entropia específica», que é a razão entre a entropia total e a das partículas A ou B.

Até aqui descrevi um gás em equilíbrio com volume fixo e a temperatura fixa. Que acontecerá se agora fizermos aumentar o volume? Imaginemos um êmbolo saindo do contentor, de tal maneira que o volume deste aumenta. Suponhamos ainda que variamos o volume a um ritmo lento, comparado com o tempo de colisão médio entre as partículas. Isto implica que o gás permaneça sempre em equilíbrio térmico, dado que as partículas têm tempo suficiente para transferir a sua energia de umas para as outras durante a expansão. Esta expansão lenta chama-se «expansão adiabática» e durante ela pode mostrar-se que a entropia do gás permanece constante.

Em resumo, as ideias fundamentais acerca dos gases são as seguintes: em primeiro lugar, numa situação de equilíbrio térmico

um gás é descrito pela temperatura e pela densidade das várias grandezas conservadas; em segundo lugar, numa expansão ou contração adiabáticas a entropia total, proporcional ao número total de partículas, permanece constante. Estas propriedades elementares dos gases em equilíbrio térmico podem ser aplicadas à descrição de um gás que enchesse o universo inteiro. Mas o universo não é um simples recipiente com gás — não tem arestas nem limites bem definidos. Além do mais, a natureza da sua expansão em volume é diferente da do êmbolo movendo-se para fora do recipiente, o que exige modificações, que devemos ter em conta.

Imaginemos que o universo inteiro está uniformemente preenchido por um gás de partículas e que o espaço-tempo do universo é o de um dos modelos homogéneos e isotrópicos FRW: espaços que se contraem ou expandem consoante nos deslocamos no tempo para a frente ou para trás. Podemos aplicar as regras da termodinâmica e da mecânica estatística a este gás que enche o universo, desde que tomemos nota de uma importante diferença entre o universo e um contentor de gás. Ao contrário do recipiente com gás, o universo não tem limite: ou é infinito, ou se fecha sobre si próprio. O universo expande-se porque o próprio espaço estica — o deslocamento de Hubble —, e não porque o seu limite se mova, como sucedia com o êmbolo no recipiente. Se traçássemos um triângulo colossal com feixes de luz laser e o fizéssemos flutuar no espaço, então, à medida que o universo envelhecesse, esse triângulo expandir-se-ia no espaço. Da mesma maneira, o gás que enche o universo está sujeito ao mesmo deslocamento de Hubble.

Suponhamos que retiramos as paredes do nosso recipiente com gás. Então a pressão do gás sobre as paredes baixaria até zero e o gás espalhar-se-ia no espaço circundante. O gás de fotões que enche o universo também tem uma pressão, mas não existem paredes para o conter. Que é que produz a pressão? Poderíamos ser tentados a pensar que esta pressão é causada pela expansão do universo. Mas isso é incorreto. A expansão do universo é a expansão do próprio espaço, e não a expansão de alguma coisa *dentro* do espaço do universo. O gás de fotões *acompanha* a expansão geral do espaço — o seu movimento não é análogo à expansão do gás dentro de um recipiente. Os fotões podem produzir uma pressão simplesmente porque são partículas com energia movendo-se à velocidade da luz, voando em todas as direções e atingindo tudo no seu caminho. Este bombardeamento por fotões não produz uma pressão de radiação.

Uma vez compreendida a sua correta aplicação ao universo como um todo, a termodinâmica torna-se um poderoso instrumento conceptual e de cálculo. Aplicando este ponto de vista termodinâmico ao universo como hoje se apresenta, os físicos interpretam todos os corpos celestes como se fossem partículas de um gás que o enchesse. Este gás tem duas componentes fundamentais. A primeira é a matéria: as galáxias, as estrelas e toda a matéria escura invisível — essencialmente, um «gás» de objetos com muita massa e que não se deslocam muito. Este gás de matéria é «frio» — tem temperatura zero —, porque a temperatura é a medida da energia média do movimento aleatório. A segunda é a radiação — o gás de fotões das micro-ondas de fundo detetado por Penzias e Wilson.

Qual destas duas componentes, matéria ou radiação, contribui predominantemente para a densidade de massa do universo? A resposta a esta questão é importante, dado que, segundo a teoria da gravitação de Einstein, é a densidade da massa do universo que controla a sua taxa de expansão: quanto maior é a densidade de massa, tanto mais lenta é a expansão. Se medirmos a contribuição da matéria para a densidade de massa do universo atual e a compararmos com a densidade de massa-energia do gás de fotões, descobrimos que a densidade de massa é maior, pelo menos, por um fator de 1000 — o universo é «dominado pela matéria», e não «pela radiação». Concluímos, assim, que a dinâmica gravitacional de todo o universo — a expansão — é hoje controlada pelo seu conteúdo de matéria, e não pelo de radiação.

Embora a matéria claramente domine a densidade de energia do universo atual, no passado distante, durante o período do *big bang*, a radiação era a componente dominante, foi ela que controlou a dinâmica da expansão. Como é que sabemos isto? À medida que recuamos no tempo, o universo contrai-se, aquecendo o gás de partículas nele contido e fazendo subir a temperatura deste. A densidade de energia da matéria aumenta, mas a densidade da energia de radiação aumenta mais rapidamente e acaba por ultrapassar a densidade da energia da matéria. Não é difícil explicar porquê.

Um fotão, parte do gás de radiação de fundo, é caracterizado por um comprimento de onda inversamente proporcional à sua energia. Os fotões «quentes» são azuis e têm comprimentos de onda curtos; os «frios» são vermelhos, com comprimentos de onda longos. Se

considerarmos um gás de fotões com muitos comprimentos de onda diferentes, então a temperatura do gás é a energia média dos fotões no gás. Daqui poderemos concluir que o comprimento de onda médio de um fotão no gás é inversamente proporcional à temperatura do gás. Se imaginarmos o universo em contração, então todos os fotões do universo estarão com deslocamento para o azul: os comprimentos de onda diminuem; a energia média de radiação, *E*, aumenta; portanto, a temperatura, *T*, aumenta proporcionalmente: *E*,~ *T*.

Se a seguir considerarmos um qualquer volume do espaço, *V*, ocupado pelo gás de fotões, também este se contrairá com o espaço. Uma vez que um volume é o cubo de um comprimento e como todos os comprimentos, como o comprimento de onda dos fotões, se contraem com o inverso da temperatura, concluímos que qualquer volume no espaço diminui com o inverso do cubo da temperatura: $V^{\sim -3}$. A densidade de energia do gás de fotões é a energia dos fotões dividida pelo volume do espaço que ocupam. Dado que a energia média dos fotões, *E*, é proporcional à temperatura e esta deve ser dividida pelo volume, *V*, concluímos que a densidade de energia, E,/V do gás de fotões é proporcional à quarta potência da sua temperatura: $E,/V \sim T^4$— uma relação conhecida por lei de Stefan-Boltzmann, dos nomes dos seus descobridores. O que isto tudo quer dizer é que, se soubermos a temperatura de um gás de fotões, também saberemos a sua densidade de energia. Uma vez que sabemos que, para o universo atual, a temperatura deste gás é de cerca de 3 K, podemos, assim, calcular a densidade de energia da radiação e compará-la com a densidade da energia da matéria.

Como atrás afirmei, a densidade atual da matéria pode conter contribuições das formas visível e invisível da matéria. Só a densidade da matéria visível, estimada a partir das observações astronómicas, é aproximadamente igual a 1000 vezes a densidade de energia da radiação, calculada pela lei de Stefan-Boltzmann. A inclusão de possível matéria invisível apenas aumentará a densidade da matéria. Este é, pois, o modo pelo qual somos levados a concluir que o universo de hoje é dominado pela matéria. E no passado?

Para compararmos as densidades de energia da matéria e da energia da radiação no passado temos de saber como é que a densidade da matéria depende da temperatura. A densidade da energia da matéria, uma vez que a matéria pode ser vista como fria e sem movimento, é precisamente a massa-energia E_m da matéria — uma grandeza fixa independente da temperatura — dividida pelo volume que ocupa, V. Donde se segue que a densidade de energia da matéria no universo, E_m/V, é proporcional ao cubo da temperatura dos fotões: $E_m/V \sim T^3$. Se recuarmos no tempo, o universo contrai-se e a temperatura do gás de fotões aumenta. A densidade de energia do fotão, que é proporcional à quarta potência da temperatura, acaba por ultrapassar a densidade de energia da matéria, que é proporcional apenas à terceira potência. A transição para um universo dominado pela radiação ocorreu quando o universo tinha apenas um milésimo do tamanho atual e quando a temperatura era de cerca de 3000 K, em vez dos cerca de 3 K atuais. Essa temperatura está acima do ponto de fusão da maioria dos metais — um universo quente, na verdade.

Embora a matéria domine hoje a radiação, a história é bem diferente se compararmos a entropia da matéria com a entropia da radiação. Para um gás em equilíbrio a entropia total é proporcional ao número total de partículas. Comparemos então a entropia da matéria (essencialmente o número total de partículas nucleares que constituem as galáxias) com a entropia dos fotões (que é proporcional ao número total de fotões). A densidade do número de partículas nucleares — protões e neutrões — no universo atual é de cerca de uma partícula nuclear por metro cúbico (este número está afetado de considerável incerteza — poderá ser dez vezes maior — , mas essa incerteza não afeta significativamente a nossa descrição). O número de fotões por metro cúbico é de cerca de 400 milhões, um número determinado a partir da temperatura de 3 K do universo atual. Daí que a razão entre a entropia dos fotões e a entropia da matéria nuclear (que é independente do volume), a que se chama «entropia específica», tenha o valor de 400 milhões (com um fator de incerteza de cerca de 10). Vemos então que a entropia do universo atual está quase toda no gás dos fotões da radiação, e não na matéria.

O valor da entropia específica é extremamente importante porque determina a natureza do universo. Se a entropia específica fosse centenas de vezes maior do que é, poderíamos mostrar que o universo primitivo teria sido demasiado quente para que as galáxias se tivessem formado, pelo que hoje não existiriam estrelas. Por outro lado, se a entropia específica fosse muito inferior ao valor atual, quase todo o hidrogénio teria sido transformado em hélio no *big bang*. As estrelas poderiam então existir, mas estrelas feitas só de hélio não seriam muito luminosas. Concluímos então que, se a

entropia específica tivesse um valor muito diferente do que tem atualmente, o universo seria extremamente diferente e, provavelmente, hostil à evolução da vida.

O universo é um sistema fechado; portanto, a sua entropia, que verificamos existir principalmente no gás de fotões, aumenta com o tempo, de acordo com a segunda lei da termodinâmica. As galáxias formam-se e as estrelas ardem, lançando, assim, fotões para o espaço, que vão adicionar-se ao gás de fotões preexistente. Este processo aumenta a entropia total do universo. Mas o facto notável é que o aumento da entropia total do universo, através destes processos, integrados ao longo da vida inteira de todas as galáxias e estrelas, é somente um décimo milionésimo da entropia já existente nos fotões de fundo — uma fração minúscula. Para todos os fins e propósitos, toda a entropia do universo atual está no gás de fotões e tem permanecido efetivamente constante desde o *big bang*. A entropia é essencialmente uma grandeza conservada no nosso universo.

Não há muito tempo, os cientistas falavam da «morte térmica» do universo. Na década de 30 o físico James Jeans, refletindo a opinião da maioria dos seus colegas, fez notar o seguinte:

> Independentemente de todas as considerações astronómicas, o princípio físico geral conhecido por segunda lei da termodinâmica prevê um único fim para o universo — uma «morte térmica», na qual a energia total do universo estará uniformemente distribuída e toda a substância do universo estará à mesma temperatura. Esta temperatura será tão baixa que tornará a vida impossível.

Os físicos, como Jeans, que conceberam o universo sujeito à segunda lei da termodinâmica não estavam enganados acerca da morte térmica. Mas nada sabiam na década de 30 acerca da existência do gás de fotões a 3 K. Sabemos agora que a «morte térmica» do universo *aconteceu há muito*, com o *big bang*, que criou o gás de fotões. Quase toda a entropia do universo está nesse gás de fotões. Todas as estrelas, ao arderem, apenas podem contribuir com uma minúscula fração para a entropia total, que já se encontra presente.

Embora a «morte térmica» do universo não seja o problema que outrora foi, o nosso conhecimento atual do universo cria outros problemas. Uma vez que a entropia total numa expansão adiabática é conservada, a entropia do universo tem sido sempre enorme. Donde veio ela? E por que é que a entropia específica das partículas nucleares é tão pequena, comparada com a da radiação de fotões? Se esta entropia específica tivesse um valor muito diferente, então o universo seria hoje também muito diferente. Embora coloque aqui pela primeira vez estas questões, não lhes responderei agora. Voltaremos a elas quando tratarmos de um estado *pré-big bang* do universo, chamado «universo inflacionário». Muitos físicos pensam que este estado *pré-big bang* contém as respostas a muitos destes formidáveis enigmas que não são resolvidos pelo cenário do *big bang*.

Através do exame destes problemas da termodinâmica e da cosmologia, começamos a descortinar um tema emergente no nosso estudo do cosmos: a estreita relação existente entre os mais pequenos corpos que conhecemos, as partículas quânticas que

enchem o universo, e a dinâmica do universo como um todo, o maior objeto conhecido. Este tema encontra o seu pleno significado quando agora formos examinar o problema da origem do universo.

CAPÍTULO 12

O Big Bang

Não posso negar o sentimento de irrealidade ao escrever sobre os primeiros três minutos como se soubéssemos realmente do que estamos a falar.

STEVEN WEINBERG

Imaginemos um supercomputador que tenha programadas todas as leis da física que hoje conhecemos. O programa contém o modelo padrão dos quarks, leptões e gluões, juntamente com alguns números de entrada (*input*) obtidos experimentalmente, como as massas dos quarks e leptões e as forças de interação dos gluões. Utilizando esta informação, o supercomputador pode calcular as propriedades das partículas hadrónicas, determinar como se afastam umas das outras e depois construir um modelo dos núcleos e dos átomos.

O supercomputador também tem programadas as equações de Einstein, deduzindo, assim, que existem três espaços homogéneos e isotrópicos que poderiam descrever globalmente o universo — as cosmologias FRW. Contudo, teríamos de lhe dizer qual destas três cosmologias FRW se aplica ao nosso universo. Para sermos precisos dir-lhe-emos que o parâmetro cósmico $\Omega = {}^{1}/_{10}$ corresponde à cosmologia FRW aberta. O nosso supercomputador informa-nos que

os cálculos do universo primitivo não serão sensivelmente alterados se a densidade total da matéria for um pouco superior ou um pouco inferior a este valor.

Finalmente, o supercomputador contém na programação as leis da mecânica estatística e da termodinâmica. Dizemos ao computador que nos primeiros tempos da sua história o universo pode ser tratado como um gás homogéneo de partículas quânticas governado pelas leis da mecânica estatística. Trata-se de uma simplificação imensa, que resulta numa enorme poupança de tempo. O computador determina que esse gás está aproximadamente em equilíbrio e, portanto, para tudo calcular precisa apenas de saber a temperatura do gás, a entropia específica das várias partículas quânticas, as leis de conservação para as interações entre as partículas quânticas e as massas destas. Damos-lhe esta informação de entrada e fica pronto a trabalhar.

Resumidamente, o nosso supercomputador simula o universo quase da mesma maneira que os computadores utilizados pelos astrofísicos simulam a evolução das estrelas. Tal como o telescópio de Galileu, numa Era mais recuada, os computadores, graças à sua capacidade de manipularem informação complexa, abrem-nos uma nova janela sobre a realidade. Mostram-nos uma imagem de um mundo que, de outro modo, nunca veríamos. Apesar de o supercomputador ser útil ao facultar-nos um modelo quantitativo do universo, é também vantajoso ter presente uma imagem mental simples que nos ajude a interpretar a sua saída (*output*), especialmente para o período do *big bang*, que não deve ser visualizado como uma explosão com origem num determinado

ponto do espaço e que se expande para fora. Uma maneira melhor de visualizar o *big bang* é imaginar que o espaço do universo é fechado e consiste unicamente na superfície bidimensional de uma esfera. Na superfície dessa esfera encontra-se o gás homogéneo de partículas quânticas, a uma temperatura definida, interagindo de acordo com as leis da mecânica estatística. A expansão ou contração do universo é visualizada como a expansão ou contração da esfera. À medida que a esfera se contrai com o tempo, o gás, na sua superfície, torna-se mais quente e, se ela se expandir, torna-se mais frio. É claro que, se supusermos que o universo é aberto, então, em vez da superfície fechada da esfera, temos de imaginar uma superfície infinita. O ponto principal, contudo, é que o *big bang* é espacialmente homogéneo e isotrópico — acontece em toda a parte ao mesmo tempo, em todo o universo. Não há um «lado de fora» do universo onde possamos sentar-nos confortavelmente e observar a sua evolução.

Deixamos agora correr o supercomputador, que calcula as propriedades do universo à medida que este evolui no tempo. Podemos examinar a saída, apresentada em ecrãs e gráficos, a nosso bel-prazer, e ver em pormenor o que está a passar-se. Uma coisa que notamos imediatamente é que a saída mais interessante é, para os tempos primordiais, medida em minutos, segundos e microssegundos. Isto passa-se porque nesses tempos iniciais a temperatura do gás de fotões se elevou o suficiente para que o gás interagisse com a matéria. Do ponto de vista da física microscópica das partículas quânticas, não aconteceu muita coisa durante os milhares de milhões de anos que medeiam entre esses instantes iniciais e os dias de hoje. As estruturas macroscópicas mais

importantes — galáxias, estrelas, planetas e a vida — brotaram posteriormente desse gás primordial.

Examinando o universo primordial, descobrimos que o parâmetro fundamental que governa os processos físicos é a temperatura do gás das partículas quânticas em interação que preenche todo o espaço do universo. A temperatura, por ser proporcional à energia média das partículas que colidem, estabelece quais as novas partículas quânticas que podem ser criadas a partir da energia das colisões. Para que seja possível criar partículas com uma certa massa a partir de energia pura é necessário um nível mínimo de energia. Estes limiares de energia observam-se nas experiências com aceleradores de altas energias, nas quais é necessário um mínimo de energia para produzir novas partículas. Estas energias específicas, ou limiares de temperatura, podem ser calculadas a partir da massa-energia conhecida das partículas quânticas observadas em laboratório. Uma vez que a temperatura do universo aumenta à medida que recuamos no tempo, a existência destes limiares de temperatura para a criação de partículas implica que o universo primitivo possa ser visualizado como uma série de estádios ou Eras, cada um deles separado do anterior por um desses limiares.

Por exemplo, consideremos o limiar que ocorre no início da «Era dos leptões», quando o universo tem apenas cerca de um segundo de idade e a temperatura é de cerca de 10^{10} K[1]. Abaixo desta temperatura o universo consiste, na sua maior parte, num gás radiante de fotões. Contudo, à medida que o universo ultrapassa esta temperatura, algo de novo acontece. Os fotões em colisão

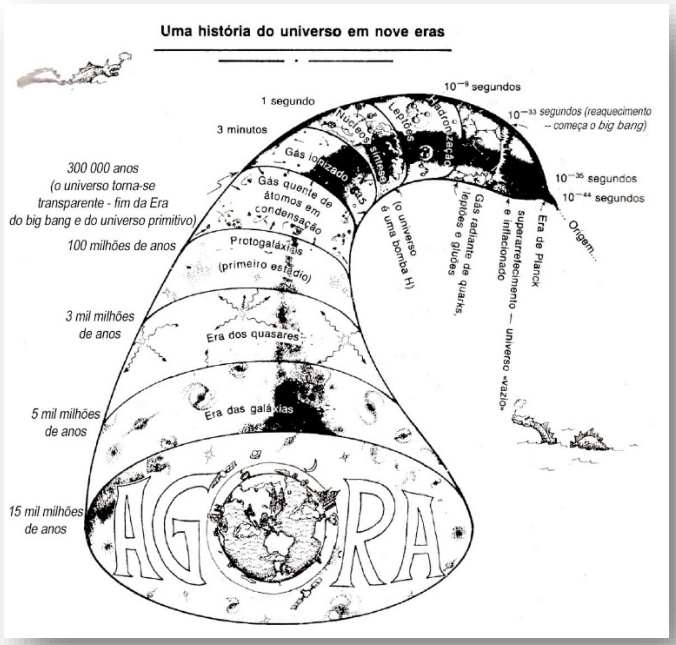
tornam-se tão energéticos que pares de fotões colidem e se convertem em pares eletrão-positrão (antieletrão) com massa. Sabemos com precisão qual a temperatura à qual este processo se inicia, porque sabemos que a energia mínima de um fotão para que esta transformação se dê (a energia é proporcional à temperatura do fotão) é precisamente igual à massa do eletrão, uma grandeza conhecida, multiplicada pelo quadrado da velocidade da luz — a equação de Einstein $E=mc^2$ é aqui utilizada. É claro que os eletrões e os positrões aniquilar-se-ão em fotões mal acabem de ser criados, mas deambulam o tempo suficiente para influenciarem a dinâmica do gás.

Esta imagem da produção de pares partícula-antipartícula será um tema fulcral da história do *big bang* à medida que a temperatura vai aumentando para lá do limitar da Era dos leptões, que se situa a 10^{10} K. A temperaturas ainda mais elevadas, pares muão-antimuão são produzidos pelos fotões. Ao aquecer, o universo enche-se de toda a espécie de partículas quânticas e respetivas antipartículas — uma cena de vasta carnificina e criação. Há várias características importantes desta imagem que devemos ter em mente.

Primeiro, as partículas quânticas produzidas, ainda que possuindo cada uma delas uma massa de repouso característica, podem ser tratadas como se não tivessem massa — como os fotões —, uma vez que a temperatura do universo é significativamente maior do que essa massa-energia de repouso. A razão por que se pode fazer esta aproximação útil é que as partículas se movem tão rapidamente a alta temperatura que quase toda a sua energia está

na energia cinética de movimento, e não na massa-energia de repouso. Efetivamente, as partículas materiais tornam-se semelhantes à radiação — sem massa e movendo-se à velocidade da luz.

Segundo, as novas partículas, uma vez criadas, partilham a energia total do *big bang* disponível com os fotões. Por exemplo, logo que o limiar de temperatura para a produção de eletrões e positrões é atravessado, o universo passa a ser constituído por números aproximadamente iguais de fotões, eletrões e positrões, tendo cada um destes grupos sensivelmente a mesma energia. Esta equipartição dos números de diferentes partículas e da respetiva energia é uma consequência de o universo estar em equilíbrio enquanto se expande — a taxa de colisão das partículas é maior do que a taxa de expansão do universo. Então a energia disponível pode ser dividida equitativamente por cada classe de partículas que participam nas interações. Se, por exemplo, imaginarmos que num dado momento os fotões ultrapassam em número os eletrões e os positrões, então mais eletrões e positrões serão criados até que o equilíbrio seja atingido. Steven Weinberg compara o estabelecimento deste equilíbrio com a lei da oferta e da procura na economia clássica — num mundo de concorrência perfeita a oferta e a procura atingem sempre o equilíbrio, a situação estável.



A história do universo pode ser vista como uma sucessão de Eras. Pouco depois da sua origem (presumivelmente, a partir do nada) o universo passa por um período inflacionário — as suas dimensões expandem-se enormemente e arrefece muitíssimo. Segue-se uma fase de reaquecimento, a criação do gás de partículas quânticas e o princípio do *big bang* — uma série de Eras durante as quais se formam os hadrões e os núcleos atómicos. O *big bang* dura até o universo ter cerca de 300 000 anos. As recombinações de eletrões com os núcleos formam então os primeiros átomos e o universo torna-se transparente. Posteriormente, formam-se as estrelas e as galáxias.

Utilizando as leis da mecânica estatística (devidamente modificadas, de forma a ter em conta as diferentes estatísticas quânticas das partículas de *spin* inteiro e semi-inteiro), pode determinar-se com precisão o número de partículas por unidade de volume para cada uma das diferentes partículas quânticas em equilíbrio em qualquer momento durante o *big bang*. O poder da mecânica estatística está em podermos determinar tais números unicamente a partir do facto de as partículas estarem em equilíbrio — os pormenores das complicadas interações não precisam de ser conhecidos.

A importância das leis de conservação exata, como a conservação da carga, conservação do número leptónico e conservação do número bariónico, a que anteriormente me referi, torna-se também evidente. Imaginemos todas as partículas quânticas interagindo a uma temperatura muito alta, que desce à medida que o universo se expande. Quando a temperatura baixa, atravessamos um limiar de produção de uma determinada espécie de partículas, a qual deve então deixar de existir. Por exemplo, quase todos os eletrões e os positrões se aniquilam em fotões quando a temperatura cai abaixo do seu limiar de produção. Todavia, dado que hoje existe um pequeno excesso de eletrões, estes devem ter sobrevivido ao massacre final eletrão-positrão, ocorrido quando a temperatura caiu abaixo do limiar eletrão-positrão, o que significa que devem ter existido sempre, devido à rigorosa lei de conservação da carga. Os eletrões têm carga elétrica unitária negativa, enquanto os positrões têm carga elétrica unitária positiva. Só se existirem mais eletrões à partida é que existe um excesso de carga negativa. Então a conservação da carga elétrica garante que alguns eletrões devem sobreviver e, de facto, conseguem-no. Este excesso de eletrões combinar-se-á muito mais tarde com os núcleos para formar os primeiros átomos.

Vimos que, devido à existência de limiares de energia, o *big bang* se organizou ordeiramente numa série de Eras, cada uma delas separada da anterior por um tal limiar. O que acontece durante essas Eras depende de forma crucial do intervalo de temperatura que caracteriza a Era e do modelo específico das partículas quânticas que programámos no nosso supercomputador. O modelo que primeiro examinaremos é o «modelo padrão». Ponhamos a

trabalhar o nosso supercomputador no princípio do tempo e deixemo-lo correr para a frente.

1. A Singularidade

De acordo com o modelo padrão, no instante zero (por definição) o universo tinha densidade de matéria infinita, curvatura infinita e temperatura infinita — um estado conhecido por «singularidade». Uma singularidade parece algo de monstruoso e mesmo misterioso e os físicos têm-se perguntado se tal singularidade pode ser evitada.

Visando uma resposta a esta interrogação, os físicos matemáticos Roger Penrose, Stephen Hawking e George Ellis mostraram que, em condições muito gerais (por exemplo, que o universo, considerado como um gás de partículas, sempre teve densidade de massa e pressão positivas), todas as soluções para as equações de Einstein acabam por desenvolver uma singularidade — um estado em que o universo entrou em colapso até um ponto matemático —, resultado este conhecido como o «teorema da singularidade». Embora o trabalho não tenha provado que estas condições extremas fossem realmente prevalecentes no princípio do tempo, o modelo padrão satisfaz certamente as condições do «teorema da singularidade», o que significa que, se se adotarem as equações de Einstein juntamente com algumas condições gerais relativas à matéria do universo, então a singularidade é inevitável.

O aparecimento de uma tal singularidade é uma boa razão para rejeitar o modelo padrão da origem do universo. Porém, isto não significa que não seja um bom modelo para as interações das partículas muito depois da origem, uma vez que a densidade da matéria toma um valor elevado, mas finito.

Estas singularidades aparecem de forma não ambígua na matemática. Mas será que ocorrem realmente na Natureza? Até os físicos clássicos encontram tais singularidades nas suas descrições matemáticas da Natureza (por exemplo, uma partícula pontual eletricamente carregada tem uma densidade infinita de energia do campo elétrico no ponto onde se encontra). Mas, com base na experiência passada, podemos afirmar que tais singularidades nas descrições matemáticas das entidades físicas refletem simplesmente um conhecimento físico incompleto. O aparecimento das singularidades matemáticas na descrição da Natureza constitui um verdadeiro desafio, que leva os físicos a inventarem melhores descrições matemáticas baseadas em leis físicas mais profundas, de forma a evitar a singularidade. A singularidade na origem do universo, consequência de alguns modelos, deve ser vista como um desafio, e não como um véu de ignorância para trás do qual não podemos olhar.

De acordo com o modelo padrão, depois da singularidade inicial a densidade da matéria e a temperatura do universo eram enormes, mas finitas. Ambas continuaram a declinar rapidamente à medida que o universo se expandia. O modelo implica que o gás radiante das partículas quânticas em interação se compunha de quarks, leptões e gluões, todos com enorme energia, que lhes permitia

converterem-se livremente uns nos outros de uma maneira que respeitava as leis de conservação. Os gluões coloridos convertiam-se em pares quark-antiquark, que quase imediatamente se aniquilavam em gluões. Os gluões fracos convertiam-se em pares leptão-antileptão, etc. — uma vasta cena de criação e destruição de todos os quanta do modelo padrão.

Trata-se, na realidade, de um universo simples, sem estrutura, um gás totalmente caótico e muito uniforme. Graças a esta simplicidade, pode ser facilmente tratado matematicamente. Nada de grande interesse acontece no modelo padrão do universo até a temperatura descer para cerca de 10^{15} K — uma temperatura ainda muito alta, muito para além da temperatura no interior de uma estrela. Mas 10^{15} K corresponde a uma massa-energia igual à dos bosões fracos, *W* e *Z*, a maior escala de massas no modelo padrão e o primeiro de vários limiares de energia que cruzaremos. A esta temperatura o universo tinha apenas a idade de um décimo de nanossegundo (um décimo milésimo de milionésimo de segundo!).

2. O Limiar da Quebra da Simetria Electrofraca: 10^{15} K

A temperaturas acima dos 10^{15} K os gluões eletromagnéticos e fracos interagem simetricamente. Logo que a temperatura cai abaixo de cerca de 19^{15} K, a simetria quebra-se e a distinção entre estas duas interações torna-se manifesta — os bosões fracos, *W* e *Z*, deixam de estar em equilíbrio com as outras partículas na sopa quântica, porque têm uma massa demasiado elevada para serem criados, enquanto os fotões permanecem, devido ao facto de não terem massa e serem facilmente criados.

Segundo o modelo de Weinberg-Salam, a diferenciação das interações eletromagnética e fraca é, em parte, consequência de uma quebra espontânea de simetria. Como exemplo de tal simetria quebrada, descrevi o alinhamento de todos os pequenos domínios magnéticos num íman, produzindo um campo magnético — o ferromagneto de Heisenberg. Mas, se aquecermos um íman normal, os seus domínios magnéticos tornam-se agitados e desorientados, começando a alinhar-se em direções aleatórias. A uma determinada temperatura crítica todo o íman perde completamente todos os vestígios de magnetismo, porque os domínios não mais se alinham numa direção preferencial: a simetria rotacional original para a qual não há direção preferencial foi restaurada. Este exemplo revela uma importante propriedade das simetrias espontaneamente quebradas: a uma dada temperatura crítica são restauradas.

A simetria espontaneamente quebrada da teoria de Weinberg-Salam não é exceção e, tal como a do íman, é restaurada a uma certa temperatura crítica, como foi primeiro salientado pelos físicos soviéticos D. A. Kirzhnits e Andrei Linde. Mas, ao contrário da do ferromagneto, esta temperatura é tão alta (10^{15} K) que só poderia ter sido atingida antes do primeiro nanossegundo do *big bang*. Acima dessa temperatura crítica a distinção entre as interações eletromagnética e fraca não é importante. Os gluões fracos, *W* e *Z*, tornam-se efetivamente *quanta* sem massa, como os fotões, os gluões coloridos e outros *quanta*. A transição para o estado simétrico à temperatura crítica é bastante suave. Como no caso do íman, à medida que a temperatura aumenta, cada vez se vê menos a simetria quebrada, até que esta, à temperatura crítica, desaparece completamente e a simetria original é restaurada.

Vemos aqui pela primeira vez uma notável propriedade da moderna teoria da origem do universo: quanto mais recuamos no tempo, mais quente o universo se torna, e as simetrias quebradas são restauradas. O universo e todas as suas interações entre partículas tornam-se cada vez mais simétricos à medida que nos aproximamos do *big bang*. Esta propriedade milita a favor da esperança de que o universo se torne cada vez mais simples, mais simétrico e mais tratável na sua história inicial, uma esperança a que os físicos se apegam na construção dos modelos.

Inversamente, progredindo no tempo, vemos que a temperatura cai e que as simetrias perfeitas se quebram. As diferenças físicas entre as várias interações — forte, fraca e eletromagnética — tornam-se agora evidentes.

O universo atual, com a sua temperatura relativamente baixa, é o remanescente gelado do *big bang*. Como um cristal de gelo que congelou a partir de vapor de água uniforme, apresenta muitas estruturas — as galáxias, as estrelas e a própria vida. Do ponto de vista moderno, porém, até os protões e os neutrões — a própria substância da matéria — são fósseis congelados do *big bang*. Também eles foram criados à medida que a temperatura descia. Este acontecimento chama-se «hadronização».

3. Hadronização: 10^{14} K

Depois de a simetria electrofraca ter sido quebrada, o universo consistia num gás composto por números aproximadamente iguais

de leptões, de quarks, das respetivas antipartículas, de gluões coloridos e de fotões continuamente a serem criados e destruídos. Efetivamente, os quarks eram livres para deambular e interagir com outras partículas — um breve período de liberdade condicional.

Recordemos que a cromodinâmica quântica (QCD), a teoria quântica relativista dos campos que descreve as interações dos quarks e dos gluões coloridos, tem a propriedade chamada «liberdade assintótica». A altas energias, a força de acoplamento dos gluões coloridos torna-se mais fraca — a aderência dos gluões diminui. Altas energias correspondem a altas temperaturas e a temperaturas superiores a 10^{14} K a força de acoplamento diminui tanto que a interação forte se torna fraca. A estas temperaturas elevadas os hadrões tornam-se, literalmente, «descolados» e os quarks são libertados.

Contudo, logo que a temperatura desceu abaixo dos 10^{14} K e o universo continuou a expandir-se, formaram-se em torno dos quarks as prisões gluónicas — os pequenos sacos a que chamamos «hadrões» —, aprisionando-os para todo o sempre. Esta transformação de um gás de partículas quânticas de quarks livres e gluões coloridos num de quarks aprisionados, ou hadrões, é o que se chama «hadronização» e assinala o início da Era dos hadrões.

4. A Era dos Hadrões: 10^{14} K a 10^{12} K

A temperaturas inferiores a 10^{14} K os quarks estão encerrados nos hadrões — o universo tem apenas um centésimo de

microssegundo de idade. Todos os hadrões são agora parte da sopa quântica; os quarks livres e os gluões coloridos desapareceram completamente. É como se um filme a cores (os quarks e gluões coloridos) de repente se tornasse preto e branco (os hadrões).

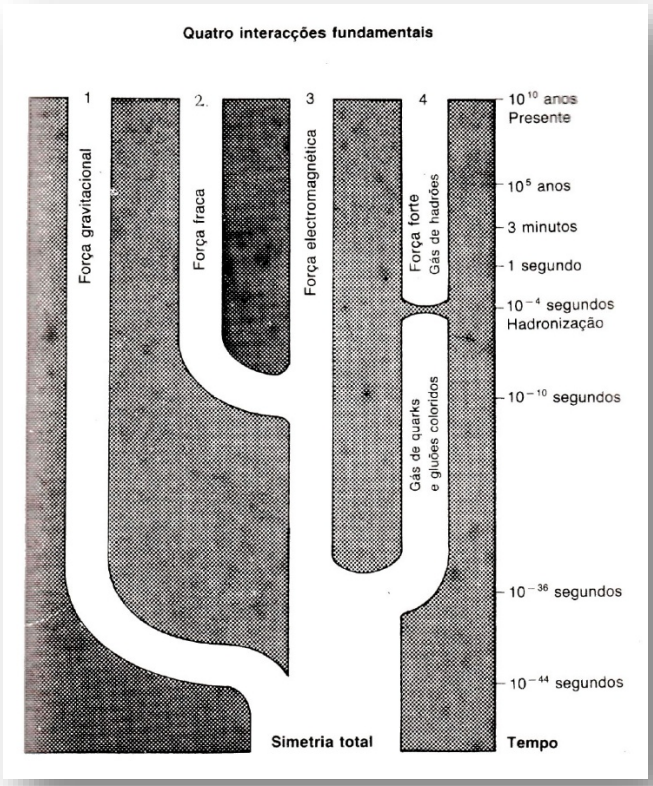
Os hadrões são as partículas quânticas associadas à força forte responsável pela coesão do núcleo atómico. Os primeiros hadrões que os físicos descobriram foram os nucleões — o protão e o neutrão. Depois vieram os piões, partículas com *spin* zero, de vida muito mais curta do que o neutrão e tendo apenas um sétimo da massa deste. Em breve muitos outros hadrões semelhantes foram também descobertos. Hoje todos estes hadrões experimentalmente detetados são considerados sistemas de quarks permanentemente ligados.

Às altas temperaturas da Era dos hadrões os fotões e outras partículas do gás de partículas quânticas são suficientemente energéticos para produzir pares hadrão-anti-hadrão. Estas partículas partilham a energia total disponível com todas as outras partículas. Logo, mesmo que novas partículas entrem em cena, o número *total* de partículas, que é proporcional à entropia, permanece constante. Dado que as diferentes partículas, incluindo todos os hadrões, partilham a energia disponível a uma dada temperatura, isto significa que os números de partículas das diferentes espécies são aproximadamente iguais.

Por exemplo, quando a temperatura foi suficiente para criar piões, no início da Era dos hadrões, existiam, aproximadamente, tantos piões como fotões, eletrões, positrões, muões, etc.

Concluímos então que, quando a temperatura foi suficiente para criar nucleões (cerca de 10^{13} K), o número de nucleões e antinucleões era, aproximadamente, o mesmo que o número de cada uma das outras partículas; em particular, o número de nucleões era, aproximadamente, igual ao número de fotões. Isto é realmente notável, se nos lembrarmos de que os fotões excedem atualmente os nucleões na proporção de cerca de 400 milhões de fotões para 1 nucleão. Portanto, no fim da Era dos hadrões, quando o universo tinha cerca de um décimo milésimo de segundo de vida, todos aqueles nucleões «em excesso» se aniquilaram com antinucleões, deixando apenas uma minúscula fração de protões e neutrões sobreviventes, que duraram até aos nossos dias. Por que razão sobreviveram?

O modelo padrão incorpora uma rigorosa lei de conservação do número de bariões, o que implica que em cada interação entre partículas o número de bariões menos o número de antibariões é conservado. Uma vez que os nucleões, o protão e o neutrão, são os bariões mais leves, todos os outros bariões podem desintegrar-se neles, transmitindo-lhes a respetiva carga de número bariónico. O neutrão pode ainda decair num protão e passar-lhe o seu número bariónico. Mas este processo de decaimento do neutrão demora cerca de um milhar de segundos, o que é muito, comparado com a duração total da Era dos hadrões. Eis por que o neutrão é efetivamente estável durante este período.



A evolução do universo como uma sucessão de simetrias quebradas. Se as ideias da moderna teoria do campo estiverem certas e as simetrias quebradas forem restauradas a alta temperatura, então nos primeiríssimos tempos as quatro forças conhecidas da Natureza, hoje vistas como distintas, estavam unificadas.

A lei de conservação exata do número bariónico impõe, assim, que, se a Era dos hadrões termina com um minúsculo excesso de bariões relativamente aos antibariões, esse excesso deverá ter existido desde sempre, muito antes da Era dos hadrões. Este pequeníssimo excesso está hoje espelhado na grande entropia específica do universo — o excesso do número de fotões sobre o número de nucleões. A existência dos nucleões, a matéria visível das estrelas e das galáxias, parece ser um acidente, uma remanescência feliz de uma Era mais recuada do universo.

No seu esforço de compreenderem o universo, os físicos tentaram explicar o pequeníssimo excesso de matéria nuclear

relativamente à antimatéria. Começaram por supor que o universo teve início num estado simétrico, no qual o número bariónico era efetivamente zero. Todavia, se o modelo padrão estivesse correto e o número bariónico fosse conservado, então o número de bariões seria hoje também zero — um desastre, porque então não haveria matéria visível no universo. Estes físicos apelam, portanto para as GUTs, que vão para além do modelo padrão e das quais resulta a não conservação do número bariónico. Uma consequência das GUTs é que o protão pode decair e, se pode decair, também pode ser criado. Outra consequência foi primeiro notada pelo físico soviético Andrei Sakharov, antes mesmo de as GUTs terem sido inventadas: o minúsculo excesso de bariões pode realmente ser criado a partir de um universo com o número bariónico nulo, desde que o gás das partículas quânticas em interação tenha propriedades especiais. Tais ideias levam-nos bem para além do modelo padrão que estou a descrever. Retomá-las-emos num próximo capítulo.

À medida que a temperatura diminui, durante a Era dos hadrões, muitos limiares energéticos, correspondentes às massas de vários hadrões, são franqueados. Alguns hadrões deixam de estar em equilíbrio com as outras partículas e, a menos que uma lei de conservação o proíba, aniquilam-se na sua totalidade. Na altura em que a temperatura tinha caído até cerca de 10^{12} K a maioria dos hadrões mais pesados tinham sido já aniquilados (com exceção dos protões e dos neutrões), chegando-se, assim, ao limiar da Era dos leptões.

5. A Era dos Leptões: 10^{12} K a 10^{10} K

No início da Era dos leptões o universo tem um décimo milésimo de segundo de idade, a temperatura é de 1 bilião de Kelvin (10^{12} K) e cada centímetro cúbico da sopa quântica cósmica pesa cerca de 1000 t. O universo consiste numa mistura de quantidades aproximadamente iguais de fotões, eletrões, neutrinos tipo eletrão, muões, neutrinos tipo muão, algumas outras partículas, como os piões — fósseis hadrónicos leves da Era anterior —, e respetivas antipartículas, mais uma «contaminação» relativamente pequena de números iguais de protões e de neutrões que já não estão em equilíbrio com as outras partículas. Contudo, os poucos protões e neutrões continuam a interagir com todas as outras partículas. Por exemplo, um protão que interage com um eletrão converte-se num neutrão e num neutrino tipo eletrão. Os neutrões, ao interagirem com positrões, convertem-se em protões e neutrinos tipo antieletrão. Embora o número total de protões e neutrões seja muito pequeno — um por cada cerca de 100 milhões das outras partículas —, os protões e os neutrões convertem-se rapidamente uns nos outros devido às suas interações com os leptões.

À medida que a temperatura desce do valor que tinha no início da Era dos leptões, franqueamos o limiar de produção dos muões. Todos os muões e antimuões se aniquilam agora em eletrões, positrões e neutrinos tipo muão e eletrão. Qualquer excesso de carga dos muões pode ser passada para os eletrões (o eletrão, recordemos, é a partícula carregada mais leve e não tem nenhuma outra partícula mais leve a quem possa passar a carga). Por esta razão, nenhum muão sobrevive à chacina dos muões. Mas os

neutrinos tipo muão, uma vez que transportam um número leptónico muónico — uma grandeza conservada —, devem continuar a existir, mesmo que já não interajam. Mais ainda: a sua densidade numérica é, aproximadamente, igual à dos fotões, pois era este o valor imediatamente antes de as interações cessarem. Grandes quantidades de neutrinos tipo muão estão agora livres para vaguearem pelo universo, quase sem interagirem, como o fazem hoje os fotões.

De igual modo, a uma temperatura ainda mais baixa os neutrinos tipo eletrão deixam de estar em equilíbrio com os eletrões e os positrões. Juntam-se então aos neutrinos tipo muão e aos neutrinos tipo tau (que deixaram de estar em equilíbrio ainda mais cedo do que os neutrinos tipo muão) para vaguearem pelo universo, constituindo um gás sem interações. À medida que o universo se expande, o comprimento de onda destes neutrinos é desviado para o vermelho e, tal como no caso dos fotões, a sua temperatura desce. Hoje os neutrinos são um gás de radiação de fundo, como o gás dos fotões, enchendo o universo a uma temperatura ligeiramente mais baixa do que a dos fotões, isto é, cerca de 2 K. A razão pela qual esta temperatura é ligeiramente mais baixa é que os neutrinos deixaram de interagir antes que os eletrões e os positrões se aniquilassem em fotões. Este processo de aniquilação, ao dar-se, aquece o gás de fotões cerca de 30% acima da temperatura do de neutrinos.

Se esta imagem está correta, por que não conseguem os cientistas detetar o gás de neutrinos? Infelizmente, apesar de haver hoje centenas de neutrinos fósseis por centímetro cúbico, estes

quase não interagem. As suas interações são tão fracas que seria necessário melhorar a tecnologia atual de deteção de neutrinos por um fator, não de 10, mas de milhões, para os encontrar. Mas os físicos estão a estudar este problema, e talvez um dia se encontre um processo inteligente de detetar estes neutrinos fósseis. É claro que, se estas ideias estiverem certas, o grosso da presente entropia do universo reside não somente no gás de fotões, mas também no gás de neutrinos.

Estes neutrinos livres são normalmente imaginados como não tendo qualquer massa. A possibilidade de poderem possuir uma pequena massa é curiosa. Se fosse este o caso, então o gás fóssil dos neutrinos poderia ser a matéria escura que poderia fechar o universo. Os físicos experimentais tentaram medir a massa do neutrino tipo eletrão e concluíram que não podia ser suficientemente grande para constituir a matéria escura. Contudo, os limites experimentais das massas dos neutrinos tipo muão e tau são muito menos restritivos, podendo estes neutrinos desempenhar essa tarefa.

Se houver neutrinos com massa, sendo esta cerca de $^{1}/_{5000}$ da massa do eletrão (não pode ser maior, porque então a densidade do gás de neutrinos excederia o limite observado da densidade de massa média do universo), pode demonstrar-se que a sua atração gravitacional mútua tenderá a formar aglomerados gigantescos, com massa, aproximadamente, igual à dos superaglomerados de galáxias atuais. Desponta, assim, uma imagem da formação dos superaglomerados: os neutrinos com massa, libertados durante a Era dos leptões, começam a formar aglomerados por ação

gravitacional, desenvolvendo-se em corpos com as dimensões de superaglomerados de galáxias. Mais tarde, depois de o *big bang* ter terminado, hidrogénio e hélio gasosos precipitam-se gravitacionalmente no interior destes corpos neutrínicos gigantescos, criando objetos com formas de «panquecas», com o tamanho de superaglomerados de galáxias, que se fragmentam então em objetos com as dimensões de galáxias, através de interações complicadas envolvendo o hidrogénio gasoso. Isto acontece entre 1 e 10 milhões de anos depois do *big bang*, a Era da formação das galáxias. Discuti algumas das consequências deste modelo das panquecas para a formação dos superaglomerados e das galáxias num capítulo anterior.

Quer esta imagem das panquecas esteja ou não correta, vemos aqui um ótimo exemplo de intercâmbio especulativo entre a física das partículas e a astronomia. As maiores de todas as estruturas, os superaglomerados de galáxias, podem muito bem estar a dizer-nos algo sobre as estruturas mais pequenas, os neutrinos, e sobre as propriedades do universo com menos de um segundo de idade. A física quântica fundamental passa a ter como campo de ensaios a totalidade do universo.

Além de libertar os neutrinos para todo o sempre, a Era dos leptões estabelece também uma razão neutrão-protão de 2 neutrões para cada 10 protões, razão esta que é importante para o estabelecimento da quantidade final de hélio produzido durante a Era seguinte — a Era dos fotões. No início da Era dos leptões existem números iguais de protões e neutrões, uma vez que se convertem livremente uns nos outros. Mas um neutrão tem

ligeiramente mais massa do que um protão — cerca de 0,14% — e pode decair num protão, num eletrão e num neutrino tipo antieletrão. Por alturas do fim da Era dos leptões a temperatura baixou suficientemente para que a pequena diferença de massa entre o protão e o neutrão causasse uma importante diferença nas suas abundâncias relativas. Devido a esta pequena diferença de massa, tornou-se mais provável que um neutrão se convertesse num protão do que o contrário. Cálculos pormenorizados concluíram que existiam somente 2 neutrões para cada 10 protões no fim da Era dos leptões, quando a temperatura baixou para 10 mil milhões de Kelvin (10^{10} K).

Alguns físicos teóricos que fizeram estes cálculos pormenorizados sublinham que esta razão neutrão-protão depende crucialmente do número de espécies diferentes de neutrinos. E, uma vez que a quantidade de hélio produzido depende diretamente desta razão, dependerá também do número de neutrinos diferentes. De acordo com os cálculos, se existirem mais de quatro neutrinos, então ter-se-á produzido demasiado hélio — alguns poucos por cento acima da quantidade observada. Ora, agora existem somente três neutrinos diferentes, como no modelo padrão — tipo eletrão, tipo muão e tipo tau —, e, assim, os cálculos, baseados no modelo padrão, reproduzem a abundância observada de hélio. Outros pensam que as incertezas nos cálculos e a incerteza na abundância estimada de hélio primordial observada atualmente no universo sugerem que não existem fundamentos rigorosos para impor limites ao número de espécies dos neutrinos. No entanto, estas estimativas sublinham novamente a relação íntima entre as

propriedades do universo observado (a abundância de hélio) e a física quântica fundamental (o número de espécies de neutrinos).

6. A Era dos Fotões: 10^{10} K a 10^{3} K

No fim da Era dos leptões, todos os leptões pesados, muões e taus, desapareceram, enquanto hordas de neutrinos inundaram o universo, sem, contudo, interagirem com o que quer que fosse. Os únicos hadrões que restaram foram apenas uma minúscula contaminação de protões e neutrões, com 10 protões para cada 2 neutrões. Os fotões, eletrões e antieletrões estão ainda em equilíbrio, criando-se e destruindo-se uns aos outros. Quando a temperatura desce abaixo do limiar de produção de pares eletrão-positrão, a maioria destes pares aniquila-se em fotões (reaquecendo ligeiramente o universo). Este limiar de temperatura assinala o início da Era dos fotões. Os positrões são eliminados da sopa por aniquilação e o pequeno número de eletrões negativamente carregados que resta é igual ao número de protões positivamente carregados (assumindo que a carga elétrica total do universo, uma grandeza conservada, era inicialmente zero). Uma vez que existe apenas cerca de 1 protão para 400 milhões de fotões, segue-se que existe só 1 eletrão para 400 milhões de fotões. O universo é agora radiação, dominado pelos fotões, que continuam a interagir, e pelos neutrinos, que não interagem.

No primeiro segundo (que marca o princípio da Era dos fotões, que durou cerca de 300 000 anos) a temperatura dos fotões era de 10 mil milhões de Kelvin e a densidade da radiação cerca de 100 kg por centímetro cúbico — um fluido de luz muito espesso e viscoso.

O universo inteiro está prestes a tornar-se um gigantesco reator termonuclear. Quase todo o hélio que hoje observamos foi produzido pela fusão do hidrogénio (protões) nos cerca de 100 segundos que se seguiram.

As estrelas, que se formaram bastante depois do *big bang*, também queimam hidrogénio para produzirem hélio, mas fazem-no a um ritmo muito mais lento. Desde há 10 mil milhões de anos, quando as primeiras estrelas começaram a existir, só 2% ou 3% do hidrogénio do universo foram convertidos em hélio nas fornalhas estelares. No entanto, 25% de toda a matéria visível do universo são compostos por hélio produzido em poucos minutos durante o *big bang*. A fusão do hidrogénio, produzindo hélio, liberta energia como uma bomba de hidrogénio. Mas até a contribuição dessa imensa energia para a densidade total de energia já existente no gás de fotões foi minúscula e não aqueceu significativamente o universo.

No início da Era dos fotões existe uma pequena contaminação de protões, neutrões e eletrões no espesso fluido de luz pura — o que resta da sopa de partículas quânticas. O número de protões é igual ao número de eletrões, de forma que a carga elétrica total é zero, continuando a existir 2 neutrões por cada 10 protões.

Durante os primeiros segundos os protões e os neutrões bombardeiam-se uns aos outros e aos fotões. Ao colidirem, os protões e os neutrões podem formar um núcleo de deutério, composto por um único protão e um único neutrão ligados. O deutério tem um núcleo muito frouxamente ligado — o protão e o

neutrão são facilmente libertados quando atingidos pelos sempre presentes fotões. Durante os primeiros segundos os núcleos de deutério vão sendo desfeitos à medida que se formam.

Em contraste com o deutério, o núcleo de hélio, composto por 2 protões e 2 neutrões, está fortemente unido — é necessária uma quantidade enorme de energia para desfazer um núcleo de hélio. Uma vez feito, fica para durar bastante tempo. Podemos produzir facilmente um núcleo de hélio a partir da colisão de dois deuterões que fiquem «colados». O problema de produzir hélio, fundindo deutério, durante os primeiros segundos consiste no facto de o deutério ser muito menos estável — destrói-se ao mesmo ritmo a que se forma —, não havendo, portanto, muito deutério presente. É este o «estrangulamento do deutério» na produção do hélio.

Cerca de 100 segundos depois a temperatura baixa para mil milhões de Kelvin. De súbito, os fotões já não são suficientemente energéticos para quebrarem os deuterões quando estes se formam. Mas um neutrão, se estiver livre e não ligado a nenhum núcleo, como aqui sucede na maioria dos casos, decai num protão, num eletrão e num neutrino antieletrão em cerca de 1000 segundos — um período de tempo apenas 10 vezes maior do que a idade do universo neste momento. Deste modo, alguns neutrões (inicialmente 2 por cada 10 protões) tiveram uma oportunidade de decair em protões na altura em que o universo tinha 100 segundos de idade. Em resultado disto, de cada 16 partículas nucleares 14 são agora protões e 2 neutrões. Os 2 neutrões podem formar dois núcleos de deutério, juntando-se a 2 protões. Uma vez que a temperatura do universo baixou, o estrangulamento do deutério

desaparece, tornando-se este suficientemente estável para colidir e formar hélio. Quase todo o deutério se funde rapidamente em hélio e na altura em que o universo tem 200 segundos de idade o processo de fusão está completo. Das 16 partículas nucleares originais, 4 delas, 2 neutrões e 2 protões, ficaram ligadas, constituindo hélio, enquanto as restantes 12 são protões. Vemos, assim, que 4 das 16 partículas, ou seja, 25% da matéria nuclear do universo, são hélio e a maior parte do que sobra é hidrogénio, precisamente o que se observa hoje — um resultado quantitativo dos primeiros minutos do universo e uma poderosa confirmação da ideia do *big bang*.

Vemos também que a quantidade de hélio produzida depende principalmente da razão inicial protão-neutrão no início da Era dos fotões e também do ritmo a que a temperatura desce. Esta quantidade não é particularmente sensível à razão entre o número de fotões e o número de partículas nucleares — a entropia específica do universo.

Contudo, uma pequena porção de deutério, somente cerca de um centésimo de 1% de todo o hidrogénio (uma proporção observada hoje em dia), escapou sem se fundir em hélio. Curiosamente, esta pequena quantidade vestigial de deutério depende bastante sensivelmente da entropia específica — a razão entre os números de fotões e de partículas nucleares. Quando o número de partículas nucleares é relativamente alto (entropia específica baixa), existem mais deuterões para colidir uns com os outros e poucos sobrevivem ao holocausto termonuclear. Todavia, se o número de partículas nucleares é relativamente baixo (entropia

específica elevada), então muito poucos deuterões estão presentes para colidir e produzir hélio e muitos mais sobrevivem. Passando despercebido, mais deutério escapa à fornalha.

O facto de a abundância relativa de deutério observada ser elevada (um centésimo de 1% é elevado) implica que a entropia específica seja elevada, cerca de 400 milhões de fotões por partícula nuclear. Algum do deutério produzido no *big bang* pode ser destruído pela queda em estrelas durante a evolução posterior do universo. Assim, é até possível que no *big bang* tenha sido produzido mais deutério do que aquele que hoje observamos.

Estas incertezas levam a que a maioria dos físicos e dos astrofísicos pensem que a abundância de deutério observada hoje é um limite inferior da quantidade produzida durante a Era dos fotões. Contudo, se assim for, então devemos concluir que o valor da entropia específica implica um valor atual de densidade da matéria nuclear visível correspondente a ter um parâmetro cósmico $\Omega={}^{1}/_{10}$ — insuficiente para fechar o universo. Se ambicionarmos um valor maior para Ω, então terá de haver matéria escura, possivelmente neutrinos com massa ou outras partículas exóticas fora do quadro do modelo padrão. Temos aqui um outro exemplo de como a física do micromundo — a produção de deutério na Era dos fotões — tem consequências cósmicas: o valor do parâmetro cósmico Ω.

Depois dos primeiros minutos o imenso reator termonuclear que é o universo pára. A síntese nuclear fica então completa; a

temperatura continua a baixar à medida que o universo se expande. O universo consiste agora num gás de fotões, de eletrões, de protões e de núcleos de elementos leves, como o hélio e o deutério. Não acontecem muitas coisas neste estado de plasma (semelhante ao interior de uma estrela) até se passarem cerca de 300.000 anos e a temperatura baixar até uns meros 3000 K. Acontece então algo de espetacular: o universo torna-se transparente. Este evento chama-se «recombinação».

7. A Recombinação e o Fim do Big Bang

Os primeiros 300.000 anos do universo foram um mundo ardente de escuridão. O universo era opaco à transmissão da luz, semelhantemente ao interior do Sol, que também é opaco — não podemos ver através dele. Se alguns eletrões se combinaram com protões ou núcleos de hélio para formar hidrogénio atómico ou hélio, estes foram imediatamente destruídos pelos energéticos fotões. Daqui que os fotões não tenham conseguido viajar até muito longe antes de interagirem. Por esta razão, os telescópios ópticos nunca verão a luz de eventos anteriores aos primeiros 300.000 anos, aproximadamente, pela mesma razão por que não podemos ver o interior do Sol.

Todavia, logo que a temperatura cai abaixo dos 3000 K, os eletrões combinam-se com os núcleos para formar verdadeiros átomos (é o evento «recombinação»), porque os fotões não são já suficientemente energéticos para os separar. Agora os fotões deixam efetivamente de interagir e são livres de voar em todas as direções à velocidade da luz. Subitamente, o universo torna-se

transparente, banhado por uma brilhante luz amarela, que é a cor correspondente à matéria aquecida até 3000 K. Por convenção, este acontecimento assinala o fim do *big bang* e da expansão não estruturada do universo; em breve algumas estruturas — as protogaláxias — começarão a emergir.

Quase ao mesmo tempo ocorre outro acontecimento importante: a densidade de energia da matéria sob a forma de hidrogénio e hélio atómicos ultrapassa a densidade de energia dos fotões. O universo é dominado pela matéria, deixando de o ser pela radiação, característica que retém até aos dias de hoje, dado que a densidade de matéria é hoje mil vezes maior do que a densidade de radiação.

Depois da recombinação a temperatura do universo continua a baixar e a sua cor muda do amarelo para o laranja, do laranja para o vermelho, deste para o vermelho escuro e depois para a escuridão das profundezas do espaço. O nosso computador diz-nos que aos 10 milhões de anos de idade do universo a densidade de massa era um milhão de vezes a que é hoje, isto é, cerca de um átomo de hidrogénio por centímetro cúbico, o que quer dizer que nessa altura a densidade da matéria de todo o universo era equivalente à densidade da matéria das galáxias de hoje. Isto implica que galáxias semelhantes às de hoje não poderiam ter existido quando o universo tinha apenas 10 milhões de anos de idade — teriam estado empilhadas umas sobre as outras nessa altura. Galáxias ou protogaláxias formaram-se, provavelmente, algures entre os 100 milhões de anos e o milhar de milhões de anos, quando os átomos de hidrogénio e de hélio se precipitaram na massa informe da matéria preexistente, escura e invisível.

À medida que o nosso computador avança no tempo a partir do primeiro milhão de anos — quando a matéria no universo era um gás uniforme de hidrogénio e hélio — até aos primeiros mil milhões de anos, podemos ver a formação das galáxias. Enormes massas de hidrogénio e hélio gasosos ter-se-iam formado a partir do gás uniforme, com as dimensões, quer dos superaglomerados (de acordo com o «modelo panqueca»), quer somente das galáxias individuais (de acordo com outros modelos que apresentei num capítulo anterior). Poderiam ter-se formado as estrelas do tipo população II — as estrelas feitas de puro hidrogénio e hélio sem elementos pesados. As estrelas de grande massa arderiam rapidamente, entrando em colapso para dar origem a buracos negros ou estrelas de neutrões; isto geraria ondas de choque no restante gás, comprimindo-o e criando, assim, as condições para a formação de mais estrelas novas. Gigantescos buracos negros poderiam formar-se nos núcleos das galáxias, consumindo estrelas e emitindo enormes quantidades de luz — os primeiros quasares. O universo está agora a caminho de desenvolver estruturas cada vez mais complexas — galáxias, estrelas, planetas e, por fim, a vida: os habitantes do jardim de Herschel. A temperatura do gás de fotões continua a baixar até atingir o valor atual de 2,7 K. Embora esta temperatura não seja senão uma insignificante lembrança da primitiva glória, a sua deteção foi a descoberta que deu credibilidade a todo o modelo do *big bang* que acabo de descrever.

8. Para Além da Síntese Electrofraca

Considerando todo o período de vida do universo, verificamos que o período mais bem compreendido pelos cosmólogos — o

verdadeiro *big bang* — estende-se por volta do primeiro nanossegundo — a quebra da simetria electrofraca — até aos primeiros 300.000 anos, a altura da recombinação. Quer antes, quer depois desse período do *big bang*, as coisas não são bem compreendidas. Por exemplo, o período de formação das galáxias é difícil de compreender, dada a sua genuína complexidade. Só o desenvolvimento futuro de novos telescópios fornecerá as informações de que os cientistas precisam para penetrar nesta Era complicada. De igual modo, as temperaturas e energias anteriores à quebra da simetria electrofraca são tão elevadas que ainda não foram reproduzidas em nenhum laboratório de física das altas energias. O que aconteceu nesse período inicial é um jogo de suposições para os teóricos dos campos.

Suponhamos que recuamos no tempo até ao primeiro nanossegundo e, usando o nosso supercomputador, deixamos o tempo recuar de maneira que a temperatura vá subindo. Que acontece?

Segundo o modelo padrão, não acontece muita coisa. O gás radiante de quarks, leptões e gluões continua a contrair-se e a sua temperatura a aumentar. Uma vez que a densidade e a pressão deste gás verificam as condições do teorema da singularidade de Penrose-Hawking, a singularidade na própria origem do universo acaba por ser atingida e o nosso computador vomita números infinitos — exatamente um disparate. Para desenvolvermos uma imagem do universo anterior ao primeiro nanossegundo precisamos de ir além do modelo padrão das partículas quânticas e utilizar um

novo modelo. Como podemos construir um novo modelo? A que critérios deverá obedecer?

O modelo padrão dos quarks, leptões e gluões tem a virtude de ter sido parcialmente testado nos laboratórios de altas energias. Se quisermos ir para além dele, até aos regimes de mais altas energias encontrados antes do primeiro nanossegundo, temos então de abandonar o terreno seguro, testado e examinado nos nossos laboratórios, e aventurar-nos no desconhecido, tendo por guia a nossa imaginação. Mas não somente a imaginação. Podemos também abordar o problema de uma maneira lógica. Devem ter-se dado acontecimentos importantes antes do primeiro nanossegundo, acontecimentos estes que estabeleceram as condições adequadas que permitiram ao universo evoluir até àquilo que hoje vemos. Se não formos cuidadosos, os voos da nossa imaginação rapidamente cairão por terra.

Inicialmente, poder-se-ia pensar que construir um novo modelo que inclua o modelo padrão, mas que vá também além dele, seria tarefa fácil. Mas não é. A dificuldade está em que, se não formos muito cuidadosos, o novo modelo irá prever um estado do universo atual em completa contradição com os factos.

O presente estado do nosso universo depende criticamente de certas grandezas físicas que tomam valores num intervalo estreito. Já atrás mencionei uma dessas grandezas, a entropia específica de 400 milhões de fotões para cada partícula nuclear. Se este valor fosse muito diferente do atual, então o universo, como o observamos, não existiria. No modelo padrão do universo primitivo

o valor da entropia específica é uma entrada — corresponde à carga de número bariónico inicial do universo. Outros modelos capazes de ir além do modelo padrão permitiriam determinar a entropia específica mas, infelizmente, fornecem um valor errado, conduzindo-nos, assim, a um universo que não existe. Os ambiciosos construtores de modelos devem ser também cuidadosos.

Outros exemplos de tais grandezas físicas críticas são os valores das massas dos quarks. Por exemplo, o quark *down* tem uma massa maior do que o quark *up*, razão por que o neutrão, que contém mais quarks *down* do que o protão, é mais pesado do que ele. Isto implica que um neutrão livre possa decair num protão, libertando, assim, energia. Mas, se em vez disso, o quark *up* fosse mais pesado do que o quark *down*, então o neutrão (e não o protão) seria o nucleão estável. Mas então o átomo de hidrogénio poderia não existir, porque o seu núcleo é constituído por um único protão e esse protão poderia agora decair num neutrão. Cerca de 75% do universo visível são hidrogénio; concluiríamos então que não existiria se os valores das massas dos quarks fossem apenas ligeiramente diferentes.

Existem vários exemplos de tais grandezas físicas, que não podem situar-se fora de uma estreita gama de valores sem que o universo deixe de ser como é; as estrelas, as galáxias e a vida poderiam não existir. Do ponto de vista do modelo padrão, admite-se simplesmente que estas grandezas têm os valores que têm. São entradas do nosso supercomputador e, logicamente, poderiam ter outros valores. Contudo, os físicos querem compreender os valores

observados destas constantes específicas com base numa teoria física mestra, e não tê-las apenas como entradas. Tal teoria mestra, se existir, vai claramente para além do modelo padrão, uma vez que deve fixar logicamente os valores exatos destas constantes. Uma teoria mestra realizaria o sonho de Einstein de que «não há constantes *arbitrárias*».

Para realizar este sonho alguns físicos teóricos ambiciosos estão a explorar «ideias loucas», ideias que não gozam, atualmente, de nenhum apoio experimental, mas que, por outro lado, não são inconsistentes com a experiência. Estas ideias loucas, na fronteira da investigação atual, podem dizer-nos algo acerca do que se passou antes do primeiro nanossegundo de vida do universo e revelar talvez o próprio ato da criação. O universo inteiro é o seu campo de ensaios. Pondo de lado precauções conservadoras, examinemos agora estas ideias loucas.

---

[1] Aqui, e no que se segue, todas as temperaturas e tempos serão valores aproximados.

PARTE TRÊS

Ideias Loucas

É um sentimento maravilhoso reconhecer as características unificadoras de um complexo de fenómenos que se apresentam como desligados da experiência direta dos sentidos.

ALBERT EINSTEIN, 1901

CAPÍTULO 13

Teorias do Campo Unificado

É um sentimento maravilhoso reconhecer as características unificadoras de um complexo de fenómenos que se apresentam como desligados da experiência direta dos sentidos.

ALBERT EINSTEIN, 1901

O meu primeiro contato com a teoria do campo unificado foi aos 13 anos. Li um artigo num jornal, em 1953, que anunciava, com grandes parangonas, que Einstein tinha, finalmente, descoberto uma teoria de campo unificado, a culminar o trabalho da sua vida. A história, intitulada «Einstein apresenta uma nova teoria para unificar a lei do cosmos», continha até as equações do campo unificado, que eram completamente ininteligíveis para mim. No entanto, o artigo transmitiu-me a ideia de que, para lá da complexidade do mundo e da pluralidade das sensações, existia uma ordem unificadora e que Einstein, de certa maneira, a tinha compreendido. Isto era simplesmente maravilhoso — a essência do universo, o mistério da existência, estavam agora incorporados em algumas equações. Fiquei entusiasmado. Anos mais tarde soube que a teoria do campo unificado de Einstein poucos resultados tinha dado e que a maioria dos físicos pensam que está errada. O próprio Einstein ficou embaraçado com a atenção que o artigo suscitou.

Mas a minha exuberância juvenil gerou qualquer coisa que não posso deixar perder-se, mesmo hoje — a ideia de que uma lei física única e simples regula a totalidade da existência material. Tal lei física explicaria a origem do universo, o seu conteúdo e o seu destino. Todas as outras leis naturais poderiam ser deduzidas logicamente desta lei única. Se tal lei fosse descoberta, seria o triunfo final da física: a descrição lógica dos fundamentos da existência estaria então completa.

Ninguém, nem sequer os físicos, tem a mais leve prova de que tal lei mestra exista. É fácil imaginar uma multidão de problemas. Talvez a própria ideia de lei física falhe num ponto qualquer. Por exemplo, a descrição matemática da Natureza, que até aqui não desiludiu os físicos, talvez seja inadequada à tarefa de expressar tal lei. Outra possibilidade é que a lei mestra exista, mas o espírito humano seja incapaz de a descobrir. Mesmo uma superinteligência artificial com capacidades excedendo as do espírito humano seria limitada pela própria lei mestra. Portanto, não poderia descobrir a lei mestra — uma espécie de *catch-22*[1] na compreensão do universo.

As leis físicas podem ser comparadas às regras dos desportos. Porém, diferentemente destas regras, que são elaboradas por seres humanos, as leis físicas parecem ser inerentes à ordem do universo, que não foi inventada por ninguém. Muitas vezes as regras dos desportos são mudadas pelos jogadores, que podem, por exemplo, conceder vantagem (*handicap*) se houver disparidade de destreza. Neste caso existe uma regra não escrita que governa a mudança das regras — a regra de os jogadores quererem tornar o jogo mais

estimulante e interessante, nivelando os adversários. De modo análogo, podemos imaginar que a lei física muda, mas então existirá uma nova lei que governa essa mudança. É concebível que, ao descobrirem novas leis, que são generalizações lógicas das leis anteriores, os físicos possam concluir que o processo nunca termina. Em vez de uma lei universal absoluta no âmago da existência, poderão encontrar uma regressão infinita das leis, ou, pior ainda, a confusão total, ou a ausência de leis — um universo fora da lei.

Não há, por isso, a garantia de que uma lei física simples esteja à nossa espera. No entanto, e malgrado esta possibilidade, a noção de uma lei simples que descreva toda a existência atrai-nos como o Santo Graal. E, tal como a demanda do Santo Graal, a busca pode ser mais interessante do que o objeto procurado. Como procuram os físicos as leis da Natureza?

1. Uma Invenção Livre do Espírito

Não há muito tempo, muitas das pessoas que refletiam sobre estes assuntos pensavam que os físicos deduziam, logicamente, as leis da Natureza diretamente das experiências e das observações. As leis fundamentais estariam intimamente relacionadas com a experiência. Hoje em dia este método foi abandonado e os físicos não deduzem diretamente as leis das experiências. Em vez disso, tentam intuir as leis fundamentais através de raciocínios matemáticos. Ninguém exprimiu melhor este abandono do empirismo estrito do que Einstein na sua palestra «Herbert Spencer» em 1933. Notou ele:

> É minha convicção que a construção matemática pura nos permite descobrir os conceitos e as leis que os ligam, o que nos dá a chave para a compreensão da Natureza [...] Num certo sentido, portanto, considero ser verdade que o pensamento puro pode compreender a realidade, como os antigos sonharam.

Einstein foi profundamente influenciado pela sua própria invenção da teoria da relatividade geral. Criou uma construção matemática pura, aquilo a que chamaríamos um modelo, uma «invenção livre» do espírito, para descrever o mundo físico. A partir deste modelo deduziu logicamente algumas consequências quantitativas para a experiência e para a observação — um pequeno desvio na órbita do planeta Mercúrio, o encurvamento da luz em torno do limbo do Sol e o facto de que os relógios deveriam andar mais devagar num campo gravitacional. Se as observações não confirmarem estas consequências do modelo, este falha; logo, é um modelo testável. Mas o modelo, em si, é livremente criado, e não deduzido da experiência. Einstein comentou em seguida:

> Se a base da física teórica não pode ser a inferência experimental, tendo de ser antes uma invenção livre, temos algum direito de esperar encontrar o caminho correto? Além disso, será que esta abordagem correta existe realmente ou apenas na nossa imaginação? A isto respondo, com toda a confiança que, na minha opinião, o caminho correto existe. Mais ainda, está ao nosso alcance descobri-lo.

Encontrar esse «caminho correto» é a ambição dos construtores contemporâneos dos modelos das teorias de campo. Esse caminho parece estar a conduzi-los ao começo do universo; se é ou não um

caminho errado, o tempo o dirá. Recentemente, os físicos teóricos têm-se esforçado ao máximo, levando os modelos teóricos para além das energias efetivamente investigadas em laboratório, para alcançar as muito mais elevadas energias reinantes no primeiro nanossegundo de vida do universo.

A maior parte destes modelos, invenções livres dos espíritos, são criados por cientistas jovens, com visão concentradora, energia livre ilimitada e notável capacidade de sublimar os impulsos mais primitivos na ambição intelectual de conhecer. No seu jogo de cartas conceptual com a Natureza os físicos ganharam já algumas vazas e querem agora «fugir sem pagar»[2], recuar mesmo até ao princípio do tempo. É difícil dizer se estão a fazer *bluff* ou se dispõem realmente de todas as cartas de que precisam. Poderá ser necessária uma profunda revisão do nosso conceito de realidade material para que uma explicação da origem do universo seja possível. Mas é já claro que as teorias quânticas relativistas dos campos e as suas intrincadas simetrias estão a trazer surpresas conceptuais, uma inesperada riqueza de poder explanatório que traz entusiasmados os físicos, cujo tema de trabalho tem sido a unificação dos campos quânticos e das respetivas interações através da aplicação dos princípios da simetria.

Em princípio, parece fútil tentar reduzir a diversidade das forças da Natureza a uma única força subjacente simples. No entanto, os físicos fizeram consideráveis progressos relativamente a este problema, malgrado os obstáculos. Um exemplo anterior de semelhante unificação de forças e campos foi a unificação matemática, por Maxwell, dos campos elétrico e magnético num

único campo eletromagnético. Anteriormente ao trabalho de Maxwell os campos elétrico e magnético eram vistos como inter-relacionados, mas distintos. Depois de Maxwell os físicos compreenderam que esta inter-relação era mais profunda — os campos elétrico e magnético transformam-se, na realidade, um no outro à medida que variam no tempo. Se os campos elétrico e magnético oscilavam no tempo, podiam propagar-se como uma onda eletromagnética no espaço, uma onda que podia ser identificada com a luz. A unificação de Maxwell dos campos elétrico e magnético conduziu, assim, diretamente à notável descoberta de que a luz é uma onda eletromagnética, respondendo de uma nova maneira à velha pergunta «Que é a luz?».

A unificação dos campos foi apenas um aspeto da descoberta de Maxwell. Um outro foi a «redução de parâmetros». Nos seus primeiros estudos de eletricidade e magnetismo os físicos experimentais determinaram duas constantes físicas: as «suscetibilidades elétrica e magnética» do espaço vazio. Estas duas constantes aparecem nas equações das ondas eletromagnéticas; Maxwell podia, portanto, calcular a velocidade da onda — a velocidade da luz — em função dessas constantes. Três diferentes parâmetros experimentais, que anteriormente se pensava serem independentes — as suscetibilidades elétrica e magnética e a velocidade da luz —, estavam agora relacionados de uma maneira fixa e determinada. Em vez de três parâmetros independentes, havia agora só dois. Esta redução de parâmetros é outro dos objetivos do programa de unificação dos campos. O objetivo é descobrir uma teoria mestra sem parâmetros arbitrários, de maneira que todas as constantes físicas possam ser calculadas.

Os físicos estão longe de alcançar este objetivo último, mas o programa da unificação dos campos progride continuamente. Das quatro forças observadas — a gravitacional, a eletromagnética, a fraca e a forte —, os físicos inventaram teorias que unificam três, ficando a gravidade de fora. Trabalha-se presentemente no sentido de elaborar teorias de unificação dos campos que incluam também a gravidade, embora estas não sejam até agora muito realistas.

2. Teorias de Grande Unificação (GUTs)

A moderna demanda das teorias do campo unificado principiou com o trabalho de Einstein nas décadas de 20 e 30. Partindo da teoria da relatividade geral para descrever a gravidade e da teoria de Maxwell para descrever o eletromagnetismo, Einstein procurou uma teoria unificadora mais abrangente que integrasse ambas as forças. Na altura em que fez este trabalho as forças forte e fraca estavam ainda a começar a ser compreendidas, forças essas que os físicos de hoje consideram tão fundamentais como a gravidade e o eletromagnetismo. Einstein via a teoria dos campos unificados como algo que iria emergir como consequência da combinação da mecânica quântica com a relatividade geral.

Embora os esforços de Einstein para construir uma teoria do campo unificado tivessem falhado, ele inspirou outros físicos, mostrando-lhes a possibilidade de as diversas forças da Natureza poderem ser manifestações de um único campo unificado. Quando nas décadas seguintes os físicos exploraram as forças fraca, forte e eletromagnética, a ideia da unificação esteve sempre no seu inconsciente, acenando como uma terra prometida.

Nas décadas de 50 e 60 alguns físicos, entre os quais Julian Schwinger, Murray Gell-Mann, Sheldon Glashow, Abdus Salam, John Ward e Steven Weinberg, começaram a fazer notar que a unificação das interações eletromagnética e fraca era mais plausível do que a das interações eletromagnética e gravitacional. E isto porque as interações eletromagnética e fraca tinham uma importante propriedade em comum: pensava-se que ambas eram mediadas por gluões de *spin* 2. Estas intuições deram frutos quando o modelo electrofraco foi inventado em 1967-1968 e o programa de unificação de campos entrou num caminho novo e profundo. Pensava-se que a real estrada para a unificação consistia agora em fundir os campos sob a égide de uma simetria espontaneamente quebrada. A noção de simetria quebrada explicava como era possível a forças fundamentalmente unificadas e simétricas manifestarem-se tão diferentemente na Natureza. O modelo electrofraco, quando tomado juntamente (embora sem unificação) com a cromodinâmica quântica, a teoria da força forte, constitui o que se tornou conhecido como o «modelo padrão» dos quarks, leptões e gluões interagentes.

Enquanto na década de 70 se acumulavam provas experimentais a favor do modelo electrofraco unificado, alguns físicos, crendo assegurado o futuro sucesso experimental do modelo, procuravam já generalizar matematicamente o esquema da unificação, de modo a englobar a força forte aglutinadora dos quarks e mediada pelos gluões coloridos numa síntese das forças fraca, eletromagnética e forte. Uma vez que os gluões coloridos, tal como os gluões eletromagnético e fraco, eram todos *quanta* dos campos de padrão de Yang-Mills, é natural imaginar-se que todos os gluões do modelo

padrão mais não sejam do que componentes de um único campo unificado. É esta a moderna ideia base da unificação de campos.

As teorias de campo que unificam realmente as três forças — as forças eletromagnética, fraca e forte — são chamadas «teorias de grande unificação», ou GUTs (Grand Unification Theories). Propositadamente, as GUTs não realizam a unificação total dos campos, uma vez que não englobam a gravidade, de longe a mais fraca das quatro forças conhecidas.

Os físicos teóricos que hoje exploram as GUTs acreditam que elas têm profundas consequências sobre a natureza do universo primitivo, mesmo antes de este ter um nanossegundo de idade. As GUTs são igualmente responsáveis por novas propriedades das partículas quânticas, como o decaimento do protão (os protões são estáveis no modelo padrão), bem como pela existência de toda uma nova classe de partículas quânticas, os monopólos magnéticos.

Ao contrário do progenitor, o modelo electrofraco, as GUTs carecem de confirmação experimental direta, por mais atraentes que sejam; são exemplos de ideias loucas, ideias novas que são coerentes com os princípios físicos gerais e com as experiências, mas que não têm ainda provas diretas que as apoiem.

De todas as ideias loucas que irei descrever, a das GUTs é certamente a mais amadurecida. Os físicos teóricos elaboraram várias GUTs engenhosas, modelos matemáticos específicos para calcular as propriedades do decaimento do protão, e os seus colegas

experimentadores tentam arduamente observar este processo de decaimento, o que até agora ainda não aconteceu. Se a ideia das GUTs estiver errada, como muito bem pode acontecer, será grande o desapontamento, e os físicos virar-se-ão para um novo caminho.

A propriedade comum da força colorida forte e da força electrofraca é que ambas são mediadas por gluões — os *quanta* dos campos de Yang-Mills, campos que são uma consequência da simetria. O problema de unificar estas interações é, assim, o de encontrar uma simetria única (e não múltipla) que, por quebra espontânea, resulte em subssimetrias mais pequenas correspondentes às interações forte e electrofraca.

Uma maneira de visualizar este processo de unificação e de quebra de simetria consiste em representar a simetria da força forte colorida por uma circunferência, que é simétrica para rotações em torno do centro. De maneira análoga, a simetria do modelo electrofraco é representada por uma outra circunferência. Podemos imaginar que os raios destas duas circunferências distintas sejam inversamente proporcionais às intensidades das interações correspondentes. As duas circunferências não têm diretamente nada a ver uma com a outra; a simetria da interação forte e a simetria da interação electrofraca são completamente independentes uma da outra no modelo padrão.

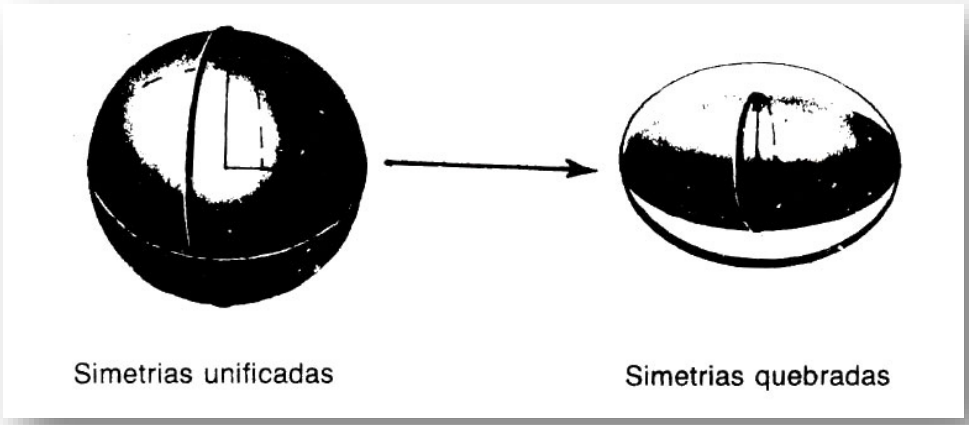
Podemos agora compreender a ideia da unificação dos campos se imaginarmos que estas duas circunferências, aparentemente independentes, correspondem, na realidade, a diferentes círculos máximos — equadores — de uma mesma esfera. Uma

circunferência é simétrica para rotações em torno de um ponto, enquanto uma esfera, simétrica em relação a qualquer eixo, tem simetria mais elevada. As circunferências anteriores são agora vistas como subssimetrias da simetria esférica, que as incorpora e unifica a ambas. Dado que ambas têm o mesmo raio, as interações forte e electrofraca têm a mesma intensidade nesta imagem — estão unificadas.

A quebra espontânea de simetria pode também ser compreendida no contexto desta imagem. A simetria esférica perfeita, conquanto seja uma solução da equação de campo, não é estável: a esfera transforma-se numa figura elipsoidal, que é, ela sim, a solução estável. A esfera transformada pode ser caracterizada por duas circunferências, de raios diferentes, correspondentes às diferentes intensidades das várias interações — uma manifestação da simetria quebrada. Mas a estrutura inicial subjacente é a esfera perfeita.

As subssimetrias descritas pelo modelo padrão são mais complicadas do que simetrias circulares simples e a simetria unificadora das GUTs é mais complicada do que a de uma esfera, mas a ideia base é semelhante. Vistas desta maneira, as interações forte, fraca e eletromagnética devem a sua progenitura à quebra espontânea de uma simetria GUT mestra. Há várias maneiras de mergulhar as simetrias do modelo padrão numa única simetria GUT maior; descobrir a maneira e qual a simetria maior escolhida pela Natureza, se é que alguma existe, é o jogo a que se entregam os modernos construtores de modelos.

Inicialmente, unificar os gluões fortes coloridos (que medeiam as interações entre os quarks) com os gluões electrofracos (que medeiam as interações tanto entre quarks como entre leptões) afigura-se enigmático. A unificação de todos estes gluões sob a égide de uma grande simetria parece exigir que unifiquemos igualmente os quarks e os leptões. Para que a ideia das GUTs funcione, todos os quarks e leptões deveriam ser igualmente encarados como componentes de campos de lepto-quarks de grande unificação. Do ponto de vista da grande simetria quebrada, os quarks e os leptões seriam indistinguíveis, transformar-se-iam uns nos outros.



Uma esfera perfeita tem uma simetria rotacional completa e os raios das circunferências podem ser imaginados como representando as intensidades das interações, neste caso iguais. Se a simetria for quebrada, a esfera transforma-se e os raios maior e menor podem agora ser imaginados como representando as diferentes forças das interações.

Os físicos Jogesh C. Pati, da Universidade de Maryland, e Abdus Salam foram os primeiros a especular acerca de uma tal unificação de quarks e leptões. Em 1973 sugeriram que os leptões deviam ser vistos como uma «quarta cor», a ser adicionada às três cores dos quarks. Quando a simetria de quatro cores entre os quarks e os

leptões se quebrasse espontaneamente na simetria exata de três cores da interação forte, quarks e leptões passariam a ser distinguíveis.

Este modelo tinha ainda a notável consequência de o protão poder decair em partículas mais leves. A razão desta nova propriedade do decaimento do protão podia ser diretamente relacionada com o facto de nos seus modelos os quarks e os leptões estarem unificados enquanto componentes de um único campo lepto-quark; logo, os quarks, em particular os que se encontram dentro do protão, podem converter-se em leptões. Outra maneira de enunciar esta propriedade consiste em dizer que a lei de conservação da carga bariónica que impunha a estabilidade do protão no modelo padrão era violada nestes novos modelos. Esta foi a primeira indicação de que a maioria das GUTs acarretavam o decaimento do protão.

Infelizmente, Pati e Salam não previram qualquer valor preciso para a taxa de decaimento do protão, a qual era um parâmetro arbitrário do seu modelo, nem realizaram uma verdadeira unificação das interações forte e eletrofraca, no sentido de as três interações serem manifestações de uma única simetria de padrão quebrada. Mas apenas um ano mais tarde, em 1974, Howard Georgi e Sheldon Glashow, da Universidade de Harvard, inventaram o primeiro modelo que implementava a verdadeira unificação — o «modelo SU (5) mínimo» —, que, pela sua economia, se tornou exemplo para as futuras GUTs e é, como observou Georgi, «muito bonito».

O modelo Georgi-Glashow era a mais simples de todas as GUTs. Os 8 gluões coloridos e os 4 gluões eletrofracos foram incorporados como as 12 componentes de um único campo de Yang-Mills de 24 componentes: alcançou-se, assim, a grande unificação. Esta grande simetria única foi então espontaneamente quebrada pelos campos de Higgs destruidores da simetria. O resultado foi que 8 dos 24 gluões originais puderam ser identificados com os gluões coloridos e 4 com os gluões electrofracos, W^+, W^-, Z e y. 12 novos gluões, batizados de gluões X, adquiriram massas elevadas, tão elevadas que nenhum acelerador poderia criá-las. No entanto, os novos gluões X poderiam converter os quarks em leptões, e vice-versa, através da sua interação com os quarks e leptões normais. Os gluões X de massa elevada desestabilizariam, portanto, o protão. Uma vez mais a grande unificação implicava o decaimento do protão.

Georgi e Glashow apontaram ainda uma consequência importante da sua GUT. O modelo padrão, que não unifica as três interações, comporta dezanove parâmetros arbitrários. Um destes parâmetros, o «ângulo θ_w, da interação fraca», deixou de ser arbitrário no modelo SU(5), sendo o seu valor calculado igual a 37,7º, uma consequência da grande simetria unificadora exata. Esta previsão foi um passo no sentido da realização do sonho de Einstein: «não existem constantes *arbitrárias*».

Infelizmente, tal previsão não concordava com o valor que os experimentadores tinham obtido para este ângulo. Muitas pessoas pensaram que o modelo SU(5), conquanto realizasse maravilhosamente a ideia das GUTs, não era uma descrição correta

da Natureza. Porém, a ideia da grande unificação começava a enraizar-se firmemente nos espíritos dos teóricos, os quais trabalhavam com afinco na exploração das GUTs.

3. O Deserto

Em breve Howard Georgi, Helen Quinn e Steven Weinberg aprofundaram a ideia das GUTs, examinando quais as suas consequências sobre o mundo microscópico a distâncias extremamente pequenas. As GUTs eram teorias de campo «renormalizáveis», o que significava que os teóricos podiam calcular matematicamente as propriedades das partículas quânticas a energias muito elevadas, ou, correspondentemente, a distâncias muito pequenas. Uma vez feito isto, abriu-se aos teóricos uma nova e notável perspetiva do microcosmos.

De acordo com a ideia de renormalização, as propriedades físicas das partículas quânticas, como sejam as suas massas ou forças de interação, podem variar, consoante a escala de distâncias à qual são medidas ou observadas. (Recordemos a analogia com o comportamento «fractal» do comprimento da linha de costa feita num capítulo anterior.) O que os teóricos descobriram foi o seguinte: se a Natureza é, de facto, descrita por uma GUT, então as interações quânticas, a despeito das diferentes intensidades a grandes distâncias, revelariam, quando vistas numa escala de distâncias muito pequena, a simetria GUT exata subjacente, o que significava que a energias elevadíssimas, correspondentes a pequenas distâncias, as interações forte, fraca e eletromagnética se fundiam numa única interação unificada.

Se, por exemplo, definirmos adequadamente as constantes de acoplamento para as interações de gluões fortes, fracos e eletromagnéticos, simbolizadas por g_s, g_w e g_e, respetivamente, estas, embora diferentes a baixas energias (traduzindo a quebra de simetria), tornam-se todas iguais a energias muito elevadas. Este ponto importante — as simetrias subjacentes de uma teoria de campo espontaneamente quebrada seriam restauradas a distâncias muito pequenas — foi primeiro posto em relevo por Kenneth Wilson no seu trabalho pioneiro sobre a teoria da renormalização. Era agora a vez de as suas ideias desempenharem um papel importante na interpretação física das GUTs.

Estas ideias de renormalização e de escalas de distâncias foram de imediato aplicadas ao modelo SU(5), de Georgi e Glashow, com resultados notáveis. Lembremos que este modelo previa, infelizmente, um valor incorreto para o ângulo θ_w, da interação fraca. Todavia, à luz das novas ideias, tornou-se claro que esta previsão numérica, feita com base na simetria exata, se aplicava somente a escalas de distâncias pequeníssimas, às quais a simetria GUT se torna exata, e não às escalas de distância muito maiores observadas em laboratório e para as quais o ângulo fora medido. O problema consistia então em calcular o ângulo θ_w às escalas de distâncias observadas no laboratório.

No modelo SU(5) a escala de distâncias à qual as intensidades das três interações se tornam iguais — a chamada «escala das GUTs» — podia ser estimada por meio da teoria da renormalização, sendo o valor encontrado 10^{-29} cm. Esta distância é incrivelmente pequena, se constatarmos que um protão tem cerca de 10^{-14} cm de

diâmetro e é já muito pequeno (embora seja observável nos aceleradores modernos). Utilizando a teoria da renormalização, os teóricos podiam extrapolar o valor previsto para o ângulo θ_w da interação fraca à escala das GUTs de 10^{-29} cm para a escala de distâncias de 10^{-14} cm observada nos laboratórios. O valor do ângulo calculado para as escalas de distâncias laboratoriais foi de 27,2º. Embora traduzisse uma melhoria, este valor ainda não concordava completamente com o valor experimental. Contudo, à medida que eram feitas mais experiências, o valor experimental do ângulo foi-se alterando até ser de 27,7º — o valor obtido pelos teóricos, dentro da incerteza experimental. A cotação do modelo SU(5) subiu; os teóricos ficaram convencidos de que tinham dado com qualquer coisa. O modelo tornou-se objeto de intensa investigação matemática.

Uma curiosa imagem do mundo microcósmico quântico das pequenas distâncias emergiu destas investigações. O modelo padrão usual diz-nos o que acontece até uma escala de distâncias de 10^{-16} cm —a «escala electrofraca» —, as dimensões apropriadas para as partículas de massa mais elevada que desempenham um papel neste modelo, os gluões fracos, *W* e *Z*. Quando o modelo padrão é incorporado na GUT SU(5), as suas interações distintas são unificadas. Utilizando a GUT SU(5) como guia, os teóricos descobriram que nada de importante acontece quando se vai da escala electrofraca de 10^{-16} cm até à escala das GUTs de 10^{-29} cm, correspondente ao tamanho dos gluões *X*. Não aparece nenhuma partícula nova e as forças de acoplamento das três interações começam a convergir suavemente para o mesmo valor à medida que a simetria GUT se instaura.

Sheldon Glashow alcunhou apropriadamente esta vasta região microscópica de «o deserto», porquanto não encerra qualquer física nova. À medida que nos aproximamos da escala das GUTs, 10^{-29} cm, os gluões *X* fundem-se com os outros gluões do modelo padrão; todas as interações gluónicas se tornam perfeitamente simétricas e unificadas. Finalmente, à incrível escala de distância de 10^{-33}cm os efeitos da gravidade quântica (que não foram explicitamente incluídos nas GUTs) devem tornar-se importantes. Chamou-se-lhe «escala de Planck», em alusão a Max Planck, que foi o primeiro a constatar ser esta uma distância importante associada a efeitos gravitacionais. Visto que a escala de Planck é cerca de 10.000 vezes menor do que a incrivelmente pequena escala das GUTs, a maior parte dos físicos não tem problemas em ignorar os efeitos da gravidade quântica nas especulações acerca das GUTs.

De acordo com estas ideias, o micromundo possui uma hierarquia de escalas de distância, marcos na estrada que leva a distâncias sempre menores. As energias relativamente baixas dos atuais aceleradores de partículas permitem aos físicos ir até às escalas de distância do modelo padrão (cerca de 10^{-16} cm). Para energias mais elevadas os físicos calcularam as outras escalas de distância que apresentámos atrás, pelo que temos uma hierarquia de distâncias:

A escala electrofraca $\simeq 10^{-16}$ cm;

A escala das GUTs $\simeq 10^{-29}$ cm;

A escala de Planck $\simeq 10^{-33}$ cm;

Entre estas distâncias microscópicas não acontece grande coisa, ao contrário do que se passa a distâncias muito maiores, às quais existem complexas interações entre partículas que produzem hadrões, núcleos e átomos.

Calculando as razões entre as três escalas de distância (para o que é irrelevante o facto de termos escolhido exprimi-las em centímetros), obtemos os grandes números puros 10^{13} e 10^{17}. Estes números carecem de explicação. Que fazem eles na Natureza? Do ponto de vista das GUTs, ou de qualquer outra teoria, estes números, que representam a hierarquia das escalas de distância na Natureza, não têm hoje qualquer explicação. São «constantes *arbitrárias*». Os físicos querem explicar estes números e resolver o «problema da hierarquia», mas até agora, malgrado algumas sugestões intrigantes, a solução tem-lhes escapado. Até as GUTs têm «constantes *arbitrárias*» e não podem ser as derradeiras teorias do campo unificado.

Apesar de as GUTs deixarem insondáveis mistérios por desvendar, representam um longo caminho em direção à unificação das várias partículas quânticas. O grande número de gluões, quarks e leptões, por exemplo, incomoda muita gente. Parte do atrativo da ideia das GUTs é que esta proliferação de partículas quânticas é, na verdade, superficial e que todos os gluões, bem como os quarks e os leptões, podem ser vistos simplesmente como componentes de alguns campos unificadores fundamentais. Sob a operação de simetria das GUTs estas componentes do campo transformam-se umas nas outras. A razão pela qual as partículas quânticas parecem ter diferentes propriedades na Natureza é que a simetria de

unificação foi quebrada. Os diferentes gluões, quarks e leptões são análogos às facetas de um diamante lapidado, que parecem diferentes consoante a maneira como pegamos no diamante, mas que, de facto, são todas manifestações do mesmo objeto subjacente.

4. O Decaimento do Protão e a Assimetria Matéria-Antimatéria do Universo

Além de facultar uma imagem conceptual definida das distâncias microscópicas, o modelo SU(5) tinha como consequência que o protão fosse instável e tivesse de decair. Os teóricos calcularam o tempo de vida média do protão em função de outras grandezas conhecidas e estimaram que o protão levaria 10^{30} anos a desintegrar-se num positrão e num pião neutro, o que corresponde a 100 milhares de biliões de vezes a idade do universo — muito tempo, na verdade. No entanto, esta previsão permitia testar diretamente o modelo SU(5).

Na altura em que os teóricos fizeram esta previsão os experimentadores sabiam que o período de vida do protão era superior a 10^{29} anos — a previsão de 10^{30} anos do modelo SU(5) concordava com as observações feitas. Os experimentadores enfrentavam, assim, um desafio, lançado pelos teóricos e pelas novas GUTs, no sentido de melhorarem as suas observações do possível decaimento do protão e testarem a previsão.

A pesquisa do decaimento dos protões não é uma experiência da física de alta energia para ser feita num dos aceleradores gigantes. Em vez disso, o que é necessário é uma observação cuidadosa de um grande volume de matéria — quanto maior, melhor — para ver se algum protão dessa matéria se desintegra. A busca do decaimento do protão exemplifica uma espécie diferente de experiências feitas para verificar as leis fundamentais da física, experiências que se afastam das de alta energia e se aproximam das de muito baixa energia, mas envolvendo um grande volume de matéria. Ao mesmo tempo que os físicos experimentais instalavam sistemas de deteção sensíveis em grandes blocos de matéria para procurar possíveis decaimentos do protão, os teóricos especulavam sobre o significado do decaimento do protão no maior volume que existe — a totalidade do universo. Descobriram então algo de notável.

Recordemos que na minha descrição do *big bang* uma das entradas do nosso supercomputador era a entropia específica a razão entre o número de fotões e o número de bariões (protões e neutrões), que é de cerca de 400 milhões para 1. Este número diminuto de bariões corresponde a toda a matéria visível do universo atual. O enigma levantado por este reduzido número de bariões é o facto de não ser zero. Por que razão sobreviveu no *big bang* um tão pequeno, mas não nulo, número de protões e neutrões? Este enigma torna-se ainda mais complicado pelo facto de a lei de conservação do número de bariões implicar que este se manteve constante desde o princípio do universo, o que parece uma condição inicial bastante arbitrária.

Uma condição inicial bem mais atraente é a simetria perfeita entre bariões e antibariões, isto é, zero bariões no instante da criação. Porém, se o número bariónico se conservasse, então o número total de bariões igualaria o número de antibariões e viveríamos num mundo com simetria matéria-antimatéria. Contudo, o universo observado não tem tal simetria matéria-antimatéria, compõe-se principalmente de matéria; logo, começar com tal simetria é já um problema.

As GUTs resolvem este problema. Na maior parte das GUTs o número bariónico não é conservado e existe a possibilidade entusiasmante de o universo poder ter começado num estado de simetria matéria-antimatéria e ter posteriormente criado a própria assimetria matéria-antimatéria. A possibilidade do decaimento do protão nas GUTs — o primeiro exemplo de um processo que viola a conservação do número bariónico — implica também que os protões podem ser criados e fornece a chave para responder à questão de saber por que é que o universo visível existe.

Em 1968, antes mesmo de as GUTs terem sido inventadas, o cientista soviético Andrei Sakharov verificou que, se o número bariónico não se conservasse, tal explicaria, em parte, como a assimetria matéria-antimatéria que hoje vemos poderia ter nascido de um estado de simetria perfeita. Sakharov constatou ainda que a não conservação do número bariónico, embora seja uma condição necessária para a criação da matéria, não era uma condição suficiente: havia que satisfazer outras condições.

A primeira destas condições é que o universo tem de fabricar mais matéria do que antimatéria. Para tal a matéria e a antimatéria, que são a imagem uma da outra num espelho, têm de ser diferenciadas por alguma interação que nos diga de que lado do espelho se encontra o universo atual. Os experimentadores detetaram realmente tais interações (chamam-se interações que violam a simetria para a inversão no tempo); portanto, esta condição *é* satisfeita.

Outra condição seria que o universo se encontrasse num estado de não equilíbrio durante um estádio muito inicial do seu desenvolvimento, quando os processos que violam a conservação do número bariónico foram mais eficazes. Isto significa que em dada altura da sua história primitiva o universo deve ter sofrido uma «transição de fase», uma mudança do seu estado básico, que aconteceu tão rapidamente que o ritmo de colisões entre as partículas quânticas no gás primordial não conseguiu acompanhá-la. Se tal transição de fase ocorreu, qualquer assimetria matéria-antimatéria gerada durante a transição conseguiu sobreviver, pois, uma vez acabada a transição, os processos que violam a conservação do número bariónico tornaram-se menos eficazes e a conservação do número de bariões foi efetivamente restaurada. Os protões e neutrões «em excesso», gerados a partir do nada, ficam agora encerrados no universo.

Em resumo, a assimetria matéria-antimatéria poderia ser gerada a partir de um estado simétrico, desde que (1) o número de bariões não fosse conservado, (2) existissem interações que violassem a simetria para a inversão no tempo e (3) o universo tivesse estado

em certa altura num estado de não equilíbrio de expansão extrema. O modelo padrão não satisfaz a primeira e a terceira condições, razão por que a origem da matéria visível permanece um enigma neste modelo. Contudo, as GUTs, que vão para além do modelo padrão, podem violar a conservação do número de bariões. Isto conduziu ao reavivar do interesse em explicar a assimetria matéria-antimatéria observada. Em 1978 muitos físicos teóricos, tendo tomado consciência de que este velho enigma podia agora ser explicado no contexto das GUTs, trabalhavam já arduamente no cálculo da assimetria para ver se esta concordava com as observações.

Alguns anos atrás uma *T-shirt* com o slogan «A cosmologia aceita as GUTs»[3] gozava de modesta popularidade entre os cosmólogos. O slogan significava que as GUTs podiam resolver o problema da origem da assimetria matéria-antimatéria e explicar, assim, a génese do universo visível. No entanto, outro significado do slogan era que muitas GUTs (o modelo SU(5) é uma exceção) implicavam que as massas dos neutrinos não eram estritamente nulas e estes podiam, portanto, constituir a matéria escura do cosmos. As GUTs poderiam fornecer a resposta à origem não só da matéria visível, mas também da invisível.

Utilizando o modelo SU(5), os físicos teóricos estimaram o número de partículas nucleares (a entropia específica) criadas no universo primitivo, obtendo um número que era 10 a 100 vezes demasiado pequeno. Contudo, esta estimativa dependia de pormenores do universo primitivo quando este tinha apenas 10^{-35} segundos de idade e que não eram bem compreendidos. Assim, o

facto de a estimativa errar por um fator tão grande foi visto, não como um falhanço, mas sim como um sinal de que as novas ideias estavam mesmo a resultar. Posteriormente outros modelos de GUTs deram os números certos. Da mesma forma que os físicos no fim da década de 60 calcularam a génese no *big bang* de elementos como o hélio, o deutério e o lítio a partir de protões e neutrões, os físicos dos últimos anos da década de 70 calcularam a génese de bariões, como os protões e os neutrões, a partir de quarks e leptões. Cada vez mais físicos confiam nas GUTs, especialmente no modelo SU(5).

Entretanto, os experimentadores melhoraram o limite medido para a vida média do protão, que sabiam exceder 10^{29} anos — muito tempo. Como podem os experimentadores esperar limitar o período de vida de qualquer coisa que excede já o tempo de vida do universo em 10 milhões de biliões de anos?

Maurice Goldhaber, físico do Laboratório Nacional de Brookhaven, que muito tem refletido sobre o possível decaimento do protão, notou uma vez que «sabemos: pelos nossos ossos»[4] que o período de vida média do protão excede 10^{16} anos. O corpo humano contém cerca de 10^{28} protões; se a vida média do protão fosse menor que 10^{16} anos, isso corresponderia a 30.000 decaimentos por segundo e o nosso próprio corpo seria um perigoso objeto radioativo. Como até esta grosseira estimativa sugere, poderemos obter um limite mais rigoroso para a vida média do protão se usarmos um volume contendo mais protões do que os nossos corpos e um sistema de deteção melhor do que o nosso estado de saúde.

Existem, felizmente, muitos protões à nossa volta. De acordo com a teoria quântica, os protões, caso decaiam, devem fazê-lo aleatoriamente, o que implica que, se os físicos observarem um grande número de protões, a sua vida média, medida em anos, é dada pelo número total de protões em observação dividido pelo número de decaimentos do protão realmente observados durante um ano. Daqui que as principais limitações na medição da vida média do protão sejam o número total de protões que pode ser observado e a eficiência de deteção na observação do decaimento dos protões, no caso de tal ocorrer.

Para minimizar os eventos de fundo devidos aos raios cósmicos (que podem ser confundidos com o decaimento do protão) e, portanto, aumentar a eficiência da deteção os físicos instalaram as experiências para observar o decaimento do protão a grandes profundidades. As experiências estão a ser realizadas na mina Soudan, no Minesota, no túnel do monte Branco, entre a Itália e a França, no campo aurífero de Kolar, a sul da Índia, na cordilheira do Cáucaso, na União Soviética, e na mina Silver King, no Utah. Uma dessas experiências, resultante da colaboração entre a Universidade da Califórnia, em Irvine, a Universidade de Michigan e o Laboratório Nacional de Brookhaven, utiliza 8000 t de água — uma imensa piscina — a 600 m de profundidade, na mina de sal de Moston, a leste de Cleveland, Ohio. Dentro deste volume de água estão colocados detetores fotomultiplicadores muito sensíveis capazes de detetar os reveladores produtos do decaimento de um único protão (um pião neutro π^o e um positrão e^+) de entre todos os presentes na água. Em 1982 os experimentadores comunicaram que nenhum desses eventos tinha sido observado e que a vida

média do protão no modo de decaimento em $\pi^o + e^+$ devia, portanto, exceder 10^{32} anos, o que implicava que o modelo GUT SU(5) mais simples, que estimava esse período de vida em cerca de 10^{30-31} anos, estava errado. Foi uma desilusão para muita gente, especialmente para os teóricos. Todavia, apesar desse desapontamento, continua-se a trabalhar nas GUTs.

Os teóricos têm muitas outras GUTs capazes de acomodar os novos limites experimentais impostos à vida média do protão. Alguns deste modelos implicam que os produtos dominantes do decaimento do protão não são $\pi^o + e^+$, mas sim partículas completamente diferentes, mais difíceis de ver. Contudo, uma vez que as experiências prosseguem, os limites irão melhorando e até alguns destes modelos mais elaborados poderão acabar por ter de ser postos de lado.

Se o decaimento do protão vier a ser observado, será grande a excitação. Estudando o decaimento, os físicos poderão obter informações adicionais pormenorizadas, que seriam muito úteis para determinar qual o modelo GUT que descreve a Natureza, se é que algum o faz. Por estranho que pareça, as piscinas de água no fundo de uma mina poderão fornecer pistas acerca da origem do universo.

5. Para Além das GUTs Simples

Muitos físicos duvidam de que a Natureza tenha escolhido uma GUT simples. Sentem-se pouco à vontade com um modelo que

descreve corretamente a Natureza à escala de distância de 10^{-16} cm (a escala dos gluões fracos, *W* e *Z*) e extrapola esse conhecimento para distâncias de 10^{-29} cm (a escala das GUTs). Esta extrapolação é de 13 ordens de grandeza, aproximadamente a relação que existe entre a espessura de um dedo comparada com a distância entre a Terra e o Sol. Se a ideia de uma GUT simples estiver certa, então a Natureza apresentará um microdeserto cobrindo uma região imensa. Muitos físicos veem isto como uma falta de imaginação da Natureza.

Os cientistas observaram que, desde a escala macroscópica dos superaglomerados de galáxias até à escala microscópica dos gluões *W* e *Z*, a Natureza revela novas estruturas físicas à medida que as escalas de distância mudam. No livro de Philip Morrison, *Powers of Ten*[5]*,* estão ilustradas a riqueza e a variedade das produções da Natureza todas as vezes que a escala de distância muda por um fator de 10. Existirá alguma razão (além da falta de imaginação ou de dados) para esta riqueza cessar subitamente quando descemos a distâncias inferiores ao tamanho dos gluões *W* e *Z*, só voltando a ter de novo interesse à escala das GUTs, 13 ordens de grandeza mais pequena? Desconhece-se a resposta a esta pergunta. A única maneira de a descobrir é continuar a realizar experiências com aceleradores a energias ainda mais elevadas.

Os físicos planeiam construir alguns grandes aceleradores. Os europeus do CERN, perto de Genebra, começaram a construção do LEP, uma máquina em que eletrões e positrões são acelerados até uma energia muito elevada em dois feixes circulares deslocando-se em sentidos opostos (os eletrões movendo-se num sentido e os

positrões em sentido inverso) e colidindo depois. Pelo estudo dos fragmentos resultantes da colisão os físicos esperam encontrar novas partículas, como as partículas de Higgs, ou novos quarks e leptões pesados.

Nos Estados Unidos os físicos estão também a planear um acelerador de supercolisão para bater todos os outros. A sua alcunha é «desertrão», não somente por poder explorar o deserto das GUTs, mas também por a máquina ser tão grande que terá de ser construída num deserto americano.

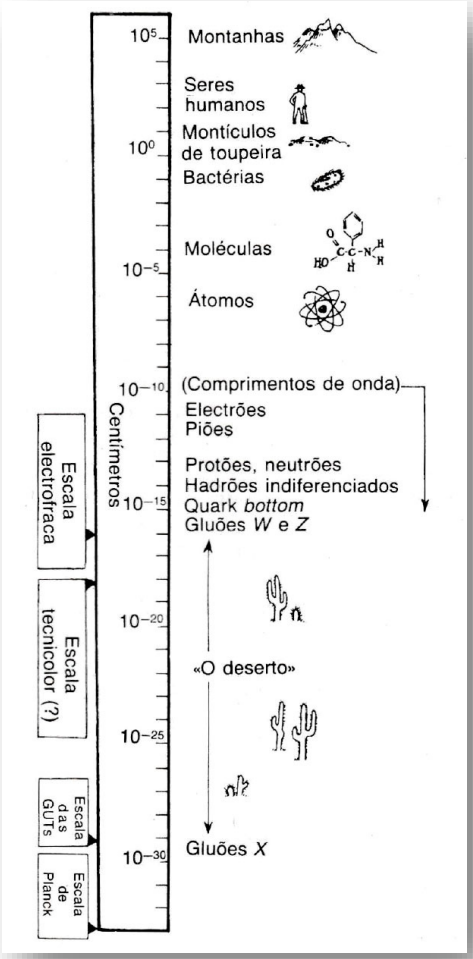
Alguns anos atrás, durante uma visita ao laboratório científico de Los Alamos, no Novo México, parei para ver um colega que estava à procura de locais para instalar esta máquina monstruosa. Na parede do gabinete estavam afixados mapas dos desertos americanos, com circunferências desenhadas à mão (a forma da máquina), indicando possíveis locais. O acelerador poderá vir a ter 48 km de diâmetro, pelo que será necessário encontrar uma vasta região desértica plana, de forma a minimizar os custos de escavação. Depois de decidirem qual o projeto que melhor serve o objetivo de explorar os reinos desconhecidos das partículas quânticas, os físicos irão apresentar ao governo uma proposta para o construir. As energias que esperam vir a alcançar com esta máquina serão comparáveis às do gás de partículas quânticas quando o universo tinha apenas algumas frações de um nanossegundo de idade.

Mesmo enquanto os aceleradores estão em fase de projeto os físicos teóricos, libertos desses problemas terra a terra, ocupam-se

da construção de modelos do que poderá vir a ser descoberto no «desertrão» de colisões — uma empresa imaginosa conhecida por «povoamento do deserto». Alguns modelos de GUTs, como o SU(5), implicam que somente existe um vasto micromundo de deserto para além da distância de 10^{-16} cm e preveem que nada de novo será encontrado pela máquina de supercolisão. Mas uma peculiaridade de todas essas GUTs simples, como o modelo SU(5), é que unificam apenas três das forças conhecidas na Natureza — a forte, a fraca e a eletromagnética. Talvez existam forças completamente novas, distintas destas três e da gravidade, que se manifestem apenas a distâncias muito pequenas. Tais forças não serão reveladas até que sejam construídos aceleradores com energia capaz de as sondar, energia que a máquina de supercolisão permitirá alcançar.

Uma tal força nova, que até agora só existe na imaginação, foi introduzida, independentemente, por Steven Weinberg, da Universidade do Texas, e Leonard Susskind, da Universidade de Stanford, e recebeu o nome de força «tecnicolor» ou «hipercor» Estes teóricos suspeitam de que ela poderá revelar-se a escalas de distância da ordem dos 10^{-17} ou 10^{-18} cm. A força tecnicolor assemelha-se à força colorida normal que liga os quarks uns aos outros. É, porém, mediada por um novo conjunto de gluões tecnicolores, que interagem com um novo conjunto de quarks, os «tecniquarks», que se ligam entre si para formar «tecni-hadrões», semelhantes aos hadrões vulgares, como os protões, neutrões e piões, com a diferença de as suas massas serem muito maiores. Se estas ideias acerca da tecnicolor estiverem certas, então os novos tecni-hadrões serão criados nos novos e poderosos aceleradores,

da mesma maneira que os vulgares hadrões foram criados nos velhos aceleradores.



As escalas de distâncias do universo em potências de 10, desde a escala de Planck até às dimensões do universo. É possível que haja um «deserto» no micromundo. Uma das principais descobertas da física moderna é perceber como o micromundo influencia o macromundo.

Os físicos não introduziram a força tecnicolor só para imaginarem novas partículas hipotéticas que poderiam povoar o deserto. Esperavam também aprofundar a compreensão dos parâmetros não explicados do modelo padrão, tais como os valores das massas dos quarks vulgares. Talvez as novas forças tecnicolores pudessem explicar as massas observadas dos quarks. Infelizmente, a ideia das

forças tecnicolores não tem tido muito êxito na elucidação destes problemas não resolvidos e levanta novos problemas se tentarmos combiná-la com as interações fracas usuais. A despeito destas dificuldades, a tecnicolor é uma ideia louca que poderá revelar-se importante quando as novas máquinas estiverem a funcionar. Ninguém pode excluir a existência de novas forças de muito curto alcance, e as partículas quânticas associadas a estas forças poderiam fazer o deserto florir.

As GUTs e as forças tecnicolores, que podem ser-lhes adicionadas, representam apenas uma (conquanto a principal) tentativa de unificar as forças da Natureza. Outra sugestão para unificar as várias partículas quânticas é a de que elas não são entidades elementares, mas sim compostas. No passado, sempre que os físicos encontraram o que tomaram por um objeto «elementar», como o átomo ou o protão, descobriram posteriormente que o objeto era, de facto, composto por corpos ainda mais pequenos. Por que não serão os quarks, os leptões e os gluões suscetíveis de decomposição?

Talvez sejam compostos. No entanto, os físicos que exploraram esta sugestão não encontraram ainda maneira de a aplicar. No passado a hipótese de uma partícula ser composta (na ausência de quaisquer provas experimentais diretas desse facto) explicava normalmente alguma propriedade de outro modo misteriosa dessa partícula. Por exemplo, a hipótese de os átomos serem feitos de eletrões e de núcleos ajudou a explicar o espectro da luz que emitiam. De modo análogo, a hipótese de os hadrões serem

compostos por quarks traduzia-se em relações corretas entre as massas observadas dos hadrões.

No seu trabalho de unificação Pati e Salam exploraram a ideia dos «preões», corpos mais pequenos a partir dos quais os quarks e os leptões poderiam ser construídos. O físico israelita Haim Harrari desenvolveu um modelo *rishon* de quarks e leptões, no qual todos os quarks e leptões observados podem ser vistos como sendo construídos a partir de dois *rishons*. Mas, apesar destes e de outros esforços, a ideia de que os quarks e os leptões são compostos não foi útil para explicar as respetivas massas ou quaisquer outras propriedades. Isto é desencorajador. Talvez os físicos ainda não tenham encontrado o modelo composto correto ou não estejam a aplicar devidamente as ideias da estrutura composta. A noção de quarks e leptões compostos é ainda outra ideia louca a colocar na prateleira até que alguém descubra como fazê-la funcionar.

Embora não exista nenhuma informação experimental que as apoie diretamente, as GUTs estão atualmente na primeira linha da investigação em física teórica. Os físicos sentem-se encorajados pelo facto de as GUTs lhes permitirem unificar três forças da Natureza, pelo menos em teoria. Duas décadas atrás tal unificação teria sido impensável. Hoje a unificação de todas as forças da Natureza faz parte do programa principal da física moderna.

A observação experimental do decaimento do protão será uma importante confirmação das GUTs. Contudo, mesmo que os experimentadores não vejam o decaimento dos protões, isto significa apenas que a vida média do protão pode ser mais longa do

que se consegue medir. A ausência de decaimento do protão imporia então uma nova restrição aos modelos dos teóricos, mas não seria o fim da ideia das GUTs.

Muitos teóricos trabalham atualmente numa generalização mais ampla das GUTs, de forma a incluírem uma nova espécie de simetria — a supersimetria —, que envolve a transformação dos campos de *spin* semi-inteiro em campos de *spin* inteiro, e vice-versa. Estes modelos têm o nome de «SUSY GUTs» (*supersymmetric grand unified theories*)[6]. O banco de ensaios destas SUSY GUTs é o próprio universo primitivo — a dinâmica do universo inflacionário, uma fase postulada do universo anterior ao *big bang* padrão. Outras teorias ainda mais completas do que as GUTs são as teorias da supergravidade, generalizações da teoria da relatividade geral de Einstein, que abarcam a quarta força, a gravitação, e são o tema de um capítulo posterior.

A ideia das GUTs nasceu do sucesso da revolução da teoria dos campos de padrão e do modelo padrão dos quarks, leptões e gluões, que se seguiu. Mas as GUTs não foram o único produto desta revolução científica. Os físicos que investigaram matematicamente as teorias dos campos de padrão descobriram toda uma nova classe de objetos que poderiam habitar o micromundo quântico. O mais apaixonante deles é o monopólo magnético, uma partícula que possui uma única carga magnética e que é diferente de qualquer outra coisa jamais vista. Os monopólos magnéticos são previstos por muitas GUTs, e, se estas teorias estiverem certas, poderemos encontrá-los um dia. A curiosa história dos monopólos magnéticos é apresentada em seguida.

[1] Expressão que, em gíria, traduz uma situação em que a vítima não pode triunfar (do romance homónimo de J. Heller, 1961). (*N. do T.*)

[2] No original, *shoot the moon*, expressão de gíria que significa fugir (de um prédio arrendado) durante a noite para evitar o pagamento da renda. (*N. do T.*)

[3] No original, *Cosmology takes GUTs*, expressão cujo segundo sentido («A cosmologia exige coragem») se perde na tradução. (*N. do T.*)

[4] No original, *we know in our bones*, que também significa «temos a certeza». (*N. do T.*)

[5] Literalmente, «potências de 10». (*N. do T.*)

[6] Teorias de grande unificação supersimétrica. (*N. do T.*)

CAPÍTULO 14

Monopólos Magnéticos

Do ponto de vista teórico, pensar-se-ia que os monopólos deveriam existir, devido à beleza da matemática. Muitas tentativas para os encontrar foram já feitas, mas nenhuma teve êxito. Devemos então concluir que a beleza da matemática, por si só, não é uma razão adequada para a Natureza ter feito uso de uma teoria. Temos ainda muito que aprender na busca dos princípios fundamentais da Natureza.

P. A. M. DIRAC, 1981

Quando Albert Einstein tinha 4 anos de idade, o pai deu-lhe uma bússola — uma pequena barra magnética livre para rodar —, que se comportava de uma maneira maravilhosa, que «não era o tipo de fenómeno que pudesse encontrar lugar no mundo inconsciente dos conceitos». A agulha magnética respondia ao campo magnético da Terra, invisível, mas detetável. Esta experiência estimulou-lhe o interesse pela física e Einstein deu mais tarde um importante contributo para o esclarecimento da natureza do magnetismo e da eletricidade. Como outros físicos, porém, estava intrigado com uma invulgar assimetria entre o magnetismo e a eletricidade — não existem cargas magnéticas comparáveis às cargas elétricas. O nosso mundo está cheio de partículas carregadas eletricamente, como os eletrões ou os protões, mas ninguém jamais detetou uma carga magnética isolada. O hipotético objeto que a possuiria chama-se «monopólo magnético».

Para termos uma ideia do que seria um monopólo magnético, caso exista, teremos de imaginar um íman em forma de barra, como a agulha de uma bússola, com polos norte e sul nas extremidades. O campo magnético de um íman em forma de barra pode ser visualizado como linhas de força saindo do polo norte e reentrando no polo sul, sendo depois canalizadas através do íman novamente para o polo norte. Se espalharmos limalha de ferro na vizinhança do íman, poderemos ver essas linhas de força na Natureza. Tal configuração de campo é um exemplo de um «campo dipolar» — tem dois polos, o norte e o sul, as linhas de campo nunca terminam, fecham-se à volta do magneto. Se o íman for cortado em dois, o resultado não é a separação do polo norte do polo sul, mas sim dois ímanes em forma de barra. Encontrar um polo norte ou sul isolados — um objeto com linhas de campo magnético só saindo ou só entrando — corresponderia a descobrir um monopólo magnético. Por razões que não são muito claras, a Natureza, ou não criou monopólos magnéticos, ou criou muito poucos.

Pelo contrário, os monopólos elétricos — partículas que transportam carga elétrica — são abundantes. Cada grão de matéria contém um número incrível de eletrões e protões, que são verdadeiros monopólos elétricos. Podem visualizar-se as linhas de força do campo elétrico emergindo de ou convergindo para uma partícula eletricamente carregada, começando ou acabando nela. Mais ainda: a experiência confirmou a lei de conservação da carga elétrica — a carga total de monopólos elétricos de um sistema fechado não pode ser criada nem destruída. Mas nada de semelhante aos monopólos elétricos existe no mundo do magnetismo, apesar de ser fácil imaginar um monopólo magnético.

James Clark Maxwell, o físico escocês que unificou matematicamente os campos elétrico e magnético em 1864, incluiu nas suas equações fundamentais do eletromagnetismo a existência de cargas elétricas, mas não a possibilidade de existência de cargas magnéticas. Teria sido fácil para ele fazê-lo; esteticamente, tal inclusão tornaria as equações maravilhosamente simétricas relativamente à eletricidade e ao magnetismo. No entanto, como aconteceu com outros físicos daquele tempo, não detetou qualquer prova da existência de cargas magnéticas na Natureza e resolveu bani-las das equações. A assimetria natural da eletricidade e do magnetismo é uma peculiaridade que desde então tem vindo a chocar os físicos.

Estes continuaram a aprofundar os seus conhecimentos do campo eletromagnético de Maxwell. Sabiam que as equações de Maxwell podiam ser simplificadas se os campos elétrico e magnético fossem derivados matematicamente a partir de um campo mais fundamental — um campo de padrão[1]. O campo de padrão eletromagnético é o primeiro e mais simples exemplo do conceito geral de campo de padrão descoberto muito mais tarde por Yang e Mills. Curiosamente, ao exigirem que as equações de Maxwell fossem expressas em termos do campo de padrão simples, os físicos constataram que a ausência de carga magnética tinha agora uma explicação matemática. Inversamente, podiam provar que a ausência de carga magnética implicava matematicamente a existência de um campo de padrão. O campo de padrão introduzia, assim, uma assimetria entre os campos elétrico e magnético.

Todavia, a introdução do campo de padrão como estrutura subjacente do eletromagnetismo parecia na altura uma inovação matemática, um truque conceptual, e não física verdadeira. Obteve-se da ideia de campo de padrão — nenhuma carga magnética — precisamente o que se tinha lá posto — nenhuma carga magnética. Na década de 20 o matemático Hermann Weyl mostrou que a incorporação de campos elétricos e magnéticos na nova teoria quântica requeria, na realidade, uma descrição em termos de um campo de padrão. Foi então que se começou a ver que o campo de padrão eletromagnético era fisicamente importante e matematicamente de muito interesse. A mecânica quântica parecia «feita» para os campos de padrão, os quais, o que era notável, tinham como consequência a ausência de monopólos magnéticos. Esta visão teórica concordava tão inteiramente com a experiência que o conceito de campo eletromagnético de padrão ganhou firmeza. Nessa altura apareceu Paul Dirac.

Em 1931 Dirac examinou as implicações físicas da «bela matemática» do campo de padrão eletromagnético no quadro da teoria quântica, tendo notado:

> Quando fiz este trabalho, esperava encontrar uma explicação para a constante de estrutura fina [a constante relacionada com a unidade fundamental de carga elétrica]. Mas isto falhou. A matemática conduzia inexoravelmente ao monopólo.

Contrariando a visão teórica dominante, Dirac descobriu que a existência de um campo de padrão eletromagnético, combinada com a teoria quântica, implicava que os monopólos magnéticos

pudessem, de facto, existir, desde que a unidade fundamental da carga magnética tivesse um valor específico. O valor da carga magnética que Dirac encontrou era tão grande que, se tal monopólo magnético existisse na Natureza, seria facilmente detetável através dos efeitos dos seus intensíssimos campos magnéticos.

Uma maneira de visualizar o resultado obtido por Dirac é imaginar um íman em forma de barra fino, com comprimento quilométrico e um campo magnético emergindo de cada extremidade. Aqui o campo magnético faz lembrar o de um monopólo magnético, porque o íman é muito fino e as extremidades estão muito afastadas. Mas isto não é um verdadeiro monopólo, pois as linhas do campo magnético não terminam realmente na ponta do estreito íman, são canalizadas através dele e emergem na outra extremidade.

Imaginemos a seguir que uma das extremidades deste estreito íman se estende até ao infinito e que a espessura da barra magnética é matematicamente reduzida a zero. O íman parece agora uma linha, ou corda matemática, com um campo magnético radial emergindo da ponta — um verdadeiro monopólo magnético pontual. Mas é então esta corda infinitamente fina (chamada «corda de Dirac») que canaliza o fluxo do campo magnético até ao infinito? Notavelmente, Dirac provou que, desde que a carga magnética do monopólo, com o valor g, obedeça à equação

$$ge=n/2 \qquad n=0,\pm1,\pm2\ldots$$

onde ***e*** é a unidade fundamental da carga elétrica (uma grandeza conhecida experimentalmente), a presença de tal corda pode nunca ser detetada fisicamente. Segundo ele, a corda torna-se então um simples artifício matemático sem realidade física, tal como as linhas coordenadas nos mapas são artifícios matemáticos para a descrição da superfície da Terra e não têm significado físico. Matematicamente, a corda de Dirac, com um monopólo magnético na ponta, seria uma linha no espaço ao longo da qual o campo de padrão eletromagnético não estaria definido.

Contudo, surpreendentemente, esta falta de definição não teria consequências mensuráveis, desde que a carga do monopólo magnético satisfizesse a condição de Dirac. Uma outra consequência do monopólo de Dirac era que a carga magnética, tal como a elétrica, seria rigorosamente conservada.

Um colega meu ouviu uma vez uma palestra de Dirac sobre o monopólo magnético. Para ilustrar a corda magnética Dirac mostrou uma corda real e, fixando-a por uma das extremidades — representando a localização do monopólo magnético —, moveu o resto da corda à volta, afirmando que a sua orientação não tinha consequências físicas. Depois da palestra o meu colega, em privado, perguntou a Dirac se este tinha guardado a corda com o fim expresso da demonstração nessa palestra, ao que Dirac, enigmaticamente, respondeu: «Não, tinha essa corda no bolso muito antes de começar a pensar acerca dos monopólos magnéticos.»

Depois do importante trabalho de Dirac os físicos teóricos aceitaram a possível existência de monopólos magnéticos, raciocinando que, se nenhuma lei da física proíbe a sua existência, então talvez existam. Os físicos experimentais foram então encorajados a procurá-los.

Talvez os monopólos magnéticos estejam incrustados na matéria vulgar e possam ser extraídos por meio de campos magnéticos fortes. Outros físicos tentaram produzi-los em aceleradores de partículas, mas nenhum foi encontrado. A busca dos monopólos estendeu-se aos raios cósmicos — chuvas de partículas quânticas e núcleos atómicos caindo sobre a terra. Em determinada ocasião uma trajetória numa emulsão fotográfica exposta na alta atmosfera pelos cientistas da Universidade da Califórnia, em Berkeley, parecia poder ser a de um monopólo magnético. Mas hoje existe o convencimento de que essa trajetória particular foi a de um núcleo pesado sulcando a emulsão, trajetória esta que poderia facilmente ser confundida com a que seria de esperar para um monopólo.

Uma técnica moderna e eficaz de pesquisa de monopólos magnéticos consiste em usar um anel supercondutor — um condutor elétrico fechado, sem resistência à corrente elétrica. Uma corrente elétrica estabelecida num tal anel persiste para sempre. Se esse anel supercondutor tivesse corrente elétrica nula e um monopólo magnético o atravessasse, então começaria a circular nele uma corrente elétrica que seria facilmente detetada. Mais ainda: a intensidade dessa corrente induzida estaria relacionada de forma precisa com o valor da carga do monopólo magnético.

Blas Cabrera, da Universidade de Stanford, montou um desses anéis supercondutores. Em 14 de Fevereiro de 1982, pouco antes das 14 horas, a corrente no anel saltou de zero para precisamente o valor esperado se um monopólo tivesse atravessado o anel. Os físicos ficaram novamente bastante excitados.

O fenómeno jamais se repetiu. Talvez a corrente tenha sido causada por um monopólo e estes sejam tão raros que o evento tenha sido um feliz acaso. Contudo, a física está baseada em resultados reprodutíveis. Até que a experiência possa ser repetida, a maior parte dos físicos considerarão o acontecimento uma casualidade. A busca experimental dos monopólos magnéticos continua, «mas sempre sem êxito».

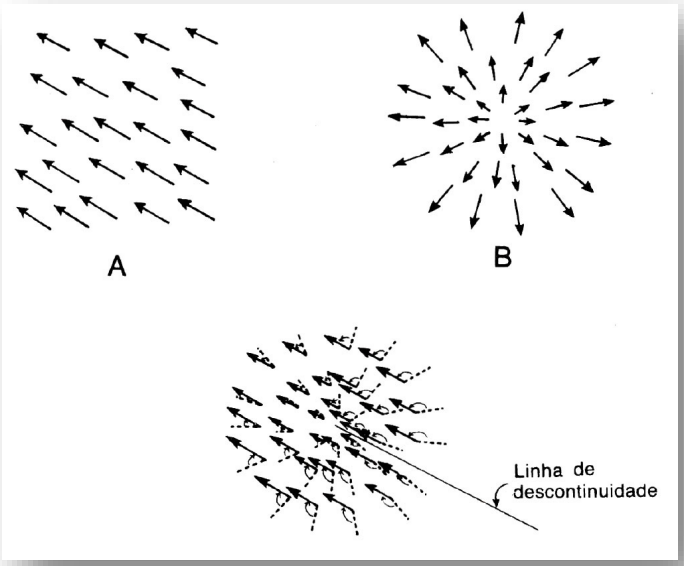
Depois da revolução dos campos de padrão do início da década de 70, os físicos alcançaram um entendimento inteiramente novo dos monopólos magnéticos que ultrapassou em muito o trabalho original de Dirac. Os inovadores foram Gerard't Hooft, jovem físico holandês que tinha já dado uma importante colaboração à teoria dos campos de padrão, e A.M. Polyakov, um brilhante jovem físico soviético. Em 1974 demonstraram matematicamente que algumas das novas teorias dos campos de padrão de Yang-Mills possuíam soluções matemáticas para as configurações de campo que correspondiam aos monopólos magnéticos se a sua simetria de campo de padrão fosse espontaneamente quebrada, como era muitas vezes o caso. Mais ainda: os físicos podiam calcular pormenorizadamente as propriedades destes monopólos — massa, *spin*, etc. Isto era diferente do que se passava com o monopólo de

Dirac, para o qual tais propriedades tinham sido deixadas por especificar.

Anteriormente a estas descobertas muitos físicos teóricos pensavam que as soluções das teorias de campos só descreviam as partículas correspondentes aos campos que entravam nas equações. Aqui estava, porém, uma classe inteiramente nova de soluções não previstas para as teorias de campo, identificadas como «solitões topológicos», que correspondiam aos monopólos magnéticos. Com o fim de descrever estas novas possibilidades da teoria dos campos de padrão, vou dizer algumas palavras acerca do conceito matemático de «topologia», bem como sobre o que se entende por «solitão».

A topologia, que é um ramo das matemáticas puras altamente desenvolvido, trata das propriedades imutáveis dos objetos matemáticos, como as figuras geométricas, que não dependem das transformações contínuas que o objeto sofra. Imaginemos que a superfície de uma esfera perfeita é deformada até se obter um elipsoide, ou um charuto comprido, ou qualquer outra forma. Imaginemos seguidamente a superfície de um *donut* e deformemo-la arbitrariamente no nosso espírito. Seja qual for a maneira como deformemos o *donut*, não há processo de o ir transformando continuamente até obter uma esfera sem rasgar a superfície de maneira a que dois pontos vizinhos se tornem separados. Não podemos ver-nos livres do buraco do *donut*. Este exemplo simples ilustra o facto de as superfícies da esfera e do *donut* serem topologicamente distintas — não existe processo de ir transformando uma forma na outra de forma contínua.

Outro exemplo de duas configurações topologicamente distintas consiste em imaginarmos um campo vetorial — grandes quantidades de pequenas setas — num plano bidimensional. Cada ponto no plano está associado a um vetor — um valor numérico e uma direção —, chamando-se o conjunto de todos estes vetores «campo vetorial». Consideremos as duas diferentes configurações de campos vetoriais mostradas na figura, com os campos vetoriais representados por pequenas setas, todas elas com o mesmo tamanho e cada uma ligada a um ponto. Podem estas duas configurações de campo ser continuamente transformadas uma na outra por rotação de cada seta em torno do ponto de origem? Se a transformação for contínua (semelhantemente a não rasgar a superfície da esfera do *donut* no exemplo anteriormente apresentado), então isto implica que a rotação que imprimimos a cada seta no plano pode diferir apenas infinitesimamente da rotação que imprimimos às setas infinitesimalmente vizinhas. Se tentarmos alinhar as setas na configuração B de tal maneira que façam lembrar a configuração A, com todas as setas a apontar no mesmo sentido, concluiremos que, por mais que tentemos, existe sempre uma linha de descontinuidade no plano (mostra-se um exemplo). De um lado da linha as setas são rodadas no sentido dos ponteiros de um relógio e do outro no sentido contrário. Uma vez que tal linha representa um salto descontínuo no sentido da rotação de vetores adjacentes, podemos concluir que as configurações de campo A e B não podem ser transformadas uma na outra de forma contínua; as configurações de campo são topologicamente distintas, como a esfera e o *donut*. Estes exemplos de configurações topologicamente distintas chegam para a nossa lição sobre topologia. A seguir perguntaremos: que é um solitão?



Configuração de campos de setas tipo «vácuo», A, e configuração tipo «ouriço-cacheiro», B. Estas configurações são topologicamente distintas. Se tentarmos converter a configuração B na configuração A por rotação dos campos de setas, concluiremos que tal não pode ser feito apenas com pequenas diferenças de rotação entre setas vizinhas; há sempre uma linha a separar as setas entre as quais há uma grande diferença de rotação. No caso do monopólo, uma tal linha pode ser identificada com uma corda de Dirac.

O primeiro exemplo de um solitão é uma onda solitária na água, uma porção de água que se mova de sua livre e estranha vontade. Tal onda solitária foi vivazmente descrita por J. Scott-Russell, em 1844, no seu *Relatório acerca das Ondas*:

> Estava eu a observar o movimento de um barco que era rapidamente rebocado ao longo de um estreito canal por uma parelha de cavalos quando o barco, de repente, parou, mas não a massa de água, no canal, posta por ele em movimento, que se acumulou à volta da proa em estado de violenta agitação, deixando-o rapidamente para trás e rolando para a frente com grande velocidade, tomando a forma de uma grande e solitária elevação de água, redonda e lisa, que continuou o seu curso ao longo do canal, aparentemente sem mudança de forma ou diminuição de velocidade. Segui-a a cavalo, ultrapassando-a quando rolava ainda a uma velocidade de 13—14 km/h, mantendo a

> forma original, com cerca de 9 m de comprimento e 30-45 cm de altura. A altura diminuiu gradualmente e depois de uma perseguição de 1,5-3 km perdi-a de vista nas voltas do canal. Esta foi, no mês de Agosto de 1834, a minha primeira oportunidade de observar aquele singular e belo fenómeno.

Em 1895 dois matemáticos holandeses, Diederick Johannes Korteweg e o seu aluno Gustav de Vries, desenvolveram uma equação matemática com o fim de explicar a onda de água perseguida por Scott-Russell. Este trabalho deu início ao estudo matemático dos solitões, como eram chamadas as soluções das equações, e muitos outros exemplos foram amplamente explorados pelos matemáticos nas décadas seguintes. Embora os físicos soubessem da existência deste trabalho matemático e das suas aplicações às ondas na água, ou às ondas nos plasmas elétricos das partículas carregadas, não suportavam a ideia de que as soluções do tipo solitão espreitavam nas equações de algumas das suas teorias do campo de Yang-Mills favoritas.

Estas novas soluções para as equações da teoria dos campos são mais bem descritas como «porções» de energia, da mesma maneira que o primeiro solitão era uma porção de água. Estas porções localizadas parecem-se com as outras partículas quânticas, mas, diferentemente delas — quarks, leptões e gluões —, não correspondem a *quanta* de um campo fundamental. Em vez disso, são construídas a partir de uma configuração curiosamente distorcida dos campos fundamentais. Estas porções torcidas da energia do campo, os solitões, são os monopólos magnéticos descobertos por Gerard't Hooft e Polyakov.

Se um solitão não passa de uma porção da energia do campo, por que razão não se separa em pedaços e se não dissipa nos seus campos constituintes? Que é que mantém coeso o solitão de Gerard't Hooft-Polyakov? Para respondermos a esta questão temos de invocar a topologia. Alguns dos campos no solitão são exatamente como o campo vetorial que descrevi no plano bidimensional, com a diferença de estarem num espaço real a três dimensões, dirigidos radialmente para fora a partir de um ponto. Polyakov chamou a tal configuração de campo um «ouriço-cacheiro», por os vetores apontarem para fora como os espinhos das costas do ouriço, quando em posição defensiva. A outra configuração de campo, na qual todos os vetores estão alinhados na mesma direção, corresponde ao «vácuo» — representa ausência de campos. O «ouriço-cacheiro» e o «vácuo» são configurações topologicamente distintas e para converter uma na outra é necessário que «rasguemos» os campos. Isto requer energia infinita e, como esta não está disponível, um solitão ouriço-cacheiro não pode dissipar-se no vácuo. Aprendemos que são as propriedades topológicas da configuração de campo do solitão que garantem a sua estabilidade. É uma porção que vive para sempre.

A única maneira de destruir ou criar um solitão topológico é aproximando-o de um antissolitão topológico. No antissolitão os vetores do ouriço-cacheiro apontam radialmente para dentro; se sobrepusermos esta configuração de campo ao solitão que aponta para fora, as duas configurações cancelam-se e podem ser continuamente transformadas no vácuo. Os solitões topológicos podem ser imaginados como tendo uma «carga topológica» que é absolutamente conservada — uma carga de mais uma unidade para

o solitão e de menos uma unidade para o antissolitão. A conservação desta carga topológica garante a estabilidade do solitão.

As soluções para as teorias do campo padrão descobertas por Gerard't Hooft e Polyakov tinham todas estas maravilhosas propriedades topológicas. Curiosamente, tudo o que restava dos campos longe da «porção» era um campo magnético radial. A porção, um solitão topológico, era um monopólo magnético! A carga topológica do solitão era precisamente a carga magnética do monopólo. Com o trabalho de Gerard't Hooft e Polyakov, o monopólo, relegado durante décadas ao estatuto de curiosidade, voltou à física, vingado, como veremos em breve. Depois deste trabalho teórico sobre o monopólo, outros solitões topológicos, chamados «instantões», «vórtices» e *kinks*[2], vieram desempenhar um papel cada vez mais importante na compreensão da teoria dos campos.

O monopólo magnético de Gerard't Hooft-Polyakov está relacionado com o monopólo original de Dirac. Vimos que era impossível tentar alinhar todos os vetores do ouriço-cacheiro no mesmo sentido através de uma transformação contínua. Existia uma linha, que terminava no ponto central, ao longo da qual a transformação não era contínua. Esta linha singular é precisamente a corda de Dirac, que era necessária ao tratamento que ele fazia do monopólo magnético. Ela surge logo que tentamos transformar uma configuração ouriço-cacheiro numa configuração vácuo. Dado que é necessária uma energia infinita para nos libertarmos da corda de

Dirac, devemos concluir que o monopólo tem de ser absolutamente estável. Não se pode deformar um *donut* até se obter uma esfera.

Algumas das novas teorias de campos de padrão exploradas pelos teóricos apresentam estes novos solitões topológicos, correspondendo aos monopólos magnéticos, enquanto outras não os possuem. Aplicando as lições de topologia da matemática pura, foi fácil determinar quais as teorias que tinham ou não soluções com o monopólo. O modelo padrão, que a maior parte dos físicos aceita como uma descrição adequada da Natureza para distâncias superiores a 10^{-16} cm — a escala das dimensões dos gluões *W* e *Z* —, não tinha soluções com monopólo. Porém, no momento em que os físicos foram para além do modelo padrão, na tentativa de construírem modelos que unificassem a força forte colorida com a força electrofraca, como nas GUTs, descobriram que a topologia das configurações de campo permitia um número enorme de monopólos, os quais, porque ocorrem nas GUTs, são chamados «monopólos GUTs».

É fácil estimar a massa de um monopólo GUT examinando as propriedades das soluções solitão topológico correspondentes ao monopólo. A massa é cerca de uma centena de vezes maior do que a escala de massas GUT (10^{-15} vezes a massa do protão), cerca de um micrograma. Isto é uma massa imensa, muito maior ainda do que a massa dos mais pequenos objetos macroscópicos, como as bactérias. Tais monopólos GUTs, caso existam, são objetos exóticos e maravilhosos, nunca anteriormente vistos.

Devido à sua imensa massa, é necessária uma energia imensa para produzir monopólos GUTs. Nenhum acelerador de partículas, existente ou planeado, poderia, possivelmente, fazê-lo; até mesmo a energia libertada numa supernova é insuficiente para criar monopólos magnéticos. Contudo, houve uma altura em que podiam ter sido produzidos, porque a energia estava disponível: o começo do *big bang*. Os teóricos, começando por Yakob B. Zel'dovich e M. Yu. Khlopov, na União Soviética, e John Preskill, na Universidade de Harvard, calcularam o número de monopólos GUTs produzidos no *big bang* e encontraram um número enorme. De acordo com estes cálculos, a maior parte da massa do universo devia encontrar-se hoje sob a forma de monopólos magnéticos — um resultado absurdo. Aqui, nas suas GUTs favoritas, os teóricos encontraram um resultado absurdo.

Foi então esta a vingança do monopólo magnético. Os monopólos, outrora considerados meras curiosidades, impunham-se agora aos físicos que pensavam acerca do universo primitivo. De acordo com os cálculos, deveriam existir bastantes, mesmo demasiados. Onde estão eles?

As pesquisas terrestres acabaram em nada. Mas por que razão restringir a busca dos monopólos magnéticos à Terra? Se os monopólos magnéticos existem, devem estar a voar no espaço interestelar e intergalático, ao lado de outros detritos deixados pelo *big bang*. Os físicos de raios cósmicos procuraram os monopólos nas cascatas de partículas que chovem na atmosfera superior todos os dias e não encontraram nenhum. Mas, ao observarem o interior da galáxia para além dos fluxos de raios cósmicos de partículas que

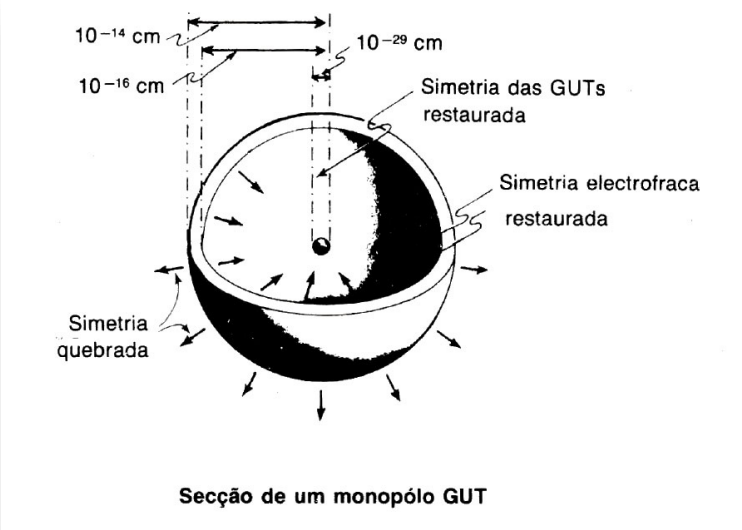
caem na Terra, obtêm uma restrição mais severa do número de monopólos magnéticos no universo.

A nossa galáxia tem um campo magnético que se enrola nos seus braços espiralados e cuja intensidade foi medida pela sua influência sobre as partículas e a luz. Os astrónomos asseguram-nos que este campo existe há muito tempo, certamente há 100 milhões de anos, ou mesmo há todo o tempo de vida da galáxia. A existência do campo magnético galático restringe severamente à existência de monopólos magnéticos, porque, se os monopólos existissem em número elevado, teriam há muito consumido o campo galático, que não existiria hoje.

Os monopólos magnéticos são acelerados por um campo magnético, como as partículas eletricamente carregadas o são por campos elétricos. À medida que os monopólos magnéticos são acelerados, ganham energia, que adquiriram do campo magnético. Como consequência desta perda de energia, a intensidade do campo magnético reduz-se e este cedo desaparece. Vemos, assim, que a presença de monopólos magnéticos em grandes quantidades no espaço exterior é contrária à existência a longo prazo do campo magnético galático. Daqui que a existência hoje de tal campo imponha um limite severo à densidade dos monopólos, um limite conhecido por «limite de Parker», o físico que primeiro o estabeleceu. Se o limite de Parker for válido, então podemos provar que nenhuma experiência feita na Terra, como as feitas com anéis supercondutores, alguma vez «apanhará» monopólos magnéticos: são, simplesmente, demasiado raros. O limite de Parker, que envolve a intensidade observada do campo magnético galático, é

ainda um outro exemplo de como as observações astronómicas constrangem os modelos loucos dos físicos teóricos.

Os monopólos magnéticos GUTs são corpos com extensão, podem ser visualizados como minúsculas esferas de energia do campo. Imaginemos que nos aproximamos de um monopólo. De longe, tudo o que se pode ver é o respetivo campo magnético, mas, à medida que nos aproximamos da sua «superfície», começamos a detetar pares partícula-antipartícula a serem criados e aniquilados pelo intenso campo magnético. Uma vez dentro do monopólo, encontramos toda a espécie de *quanta* interagentes — os leptões, os quarks e os gluões coloridos. Todavia, à medida que se vai penetrando, acontece uma coisa curiosa: as simetrias de padrão quebradas vão sendo restauradas. Por exemplo, penetrando até uma distância de 10^{-16} cm do centro do monopólo — a escala de distância da unificação eletrofraca —, verificamos que a simetria eletrofraca foi restaurada. No próprio núcleo do monopólo, que começa a 10^{-29} cm do centro — a escala das GUTs —, toda a simetria está completamente restaurada. Aqui, no minúsculo núcleo do monopólo, encontramos um estado de matéria com simetria completa, fazendo lembrar o estado do universo primitivo. Os surpreendentes monopólos GUTs transportam dentro de si a história térmica completa do universo. São «cebolas» surpreendentes, em que cada camada corresponde a uma Era do universo primitivo. Esta história remonta ao instante da criação dos monopólos, o momento do universo primitivo em que a temperatura era tão alta que a simetria GUT era exata.



Um monopólo GUT revela a história térmica completa do universo. À medida que nos aproximamos do centro do monopólo, as simetrias quebradas são restauradas. Fora do monopólo é o mundo da simetria quebrada. Lá dentro a simetria eletrofraca é restaurada e no próprio núcleo, com cerca de 10^{-29} cm de diâmetro, reina a simetria completa dos GUTs.

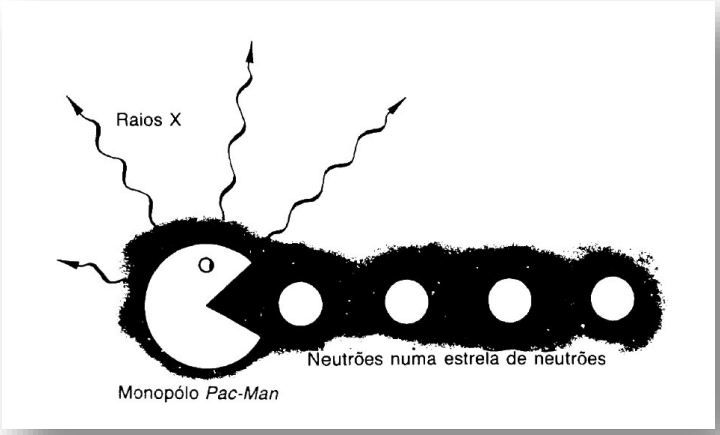
A simetria exata no núcleo do monopólo GUT tem implicações para o decaimento do protão. Existe uma probabilidade definida, mas pequena, de um protão atingir o núcleo de um monopólo numa colisão frontal. Se tal acontecer, as interações GUTs simétricas que violam a conservação do número de protões podem fazer com que o protão decaia rapidamente em partículas mais leves. O monopólo fica incólume, uma vez que a sua carga magnética é absolutamente conservada. Os monopólos catalisam, assim, o decaimento do protão, um processo conhecido como o efeito Rubikov-Callen, dos seus descobridores.

Se um monopólo GUT tivesse de atravessar uma piscina com 8000 t de água equipada para observar o decaimento do protão, deixaria um rasto de protões em decaimento na sua esteira. Tal

evento espetacular assinalaria não somente protões em decaimento, mas também a existência de monopólos.

Se os monopólos catalisam o decaimento dos protões, então alguns físicos estão convencidos de que devem ser ainda mais raros do que o indicado pelo limite de Parker. O raciocínio desenvolve-se como se segue. As estrelas de neutrões são como núcleos gigantes feitos de neutrões, os companheiros neutros dos protões. As estrelas de neutrões têm fortes campos magnéticos. Os monopólos, voando em torno da galáxia, são capturados pelo campo magnético da estrela de neutrões e precipitam-se nela. Dentro da estrela de neutrões encontram muitos neutrões, cujo rápido decaimento em partículas mais leves catalisam, tal como acontece com os protões. O monopólo atua como um catalisador e come neutrões, da mesma maneira que o *Pac-Man* come pontos vídeo. Este processo liberta grandes quantidades de raios X da estrela de neutrões. Daí que a luminosidade de raios X das estrelas de neutrões, tal como é observada pelos astrónomos de raios X, forneça um limite para o número de monopólos que andam à solta na galáxia e que são capturados pelas estrelas de neutrões. Os fluxos de raios X medidos no velho pulsar PSR 1929 + 10, utilizando o satélite Einstein de raios X, limitam o número de monopólos magnéticos nele presentes a menos de 1 bilião (10^{12}). Este limite traz consigo um melhoramento do anterior limite de Parker de um fator de 10 biliões e, se tal limite é razoável, então os monopólos são tão raros que não há qualquer esperança de alguma vez vir a encontrar um aqui na Terra. Esta linha de raciocínio, que articula as propriedades dos monopólos GUTs, o decaimento do protão, as estrelas de neutrões e os raios X, exemplifica novamente os poderosos

constrangimentos que as observações astronómicas colocam às teorias de campo do micromundo.



O *Pac-Man* come pontos vídeo da mesma maneira que o monopólo GUT come neutrões dentro de uma estrela de neutrões. Este processo gera raios X. A partir dos limites medidos para os raios X emitidos pelas estrelas de neutrões, os físicos podem estabelecer limites para a abundância de tais monopólos.

Os físicos extraíram várias conclusões das experiências, observações e teorias dos monopólos magnéticos. É possível que existam grandes quantidades de monopólos, mas em sítios restritos e escondidos dentro da Terra, das estrelas ou dos buracos negros no centro das galáxias. Se assim for, os anteriores limites da presença dos monopólos, dado que se referem à presença média de monopólos, não existem. Outra possibilidade é a de que as ideias do *big bang* e das GUTs estejam, pura e simplesmente, erradas. Finalmente, e esta é a opção que a maioria dos físicos teóricos correntemente apoia, qualquer coisa fez desaparecer os monopólos GUTs logo após terem sido criados no universo primitivo.

Se, depois de os monopólos terem sido criados, se tiver dado uma imensa expansão do universo imediatamente anterior à Era do

big bang, então os monopólos terão sido «apagados». Se tal tiver acontecido, a sua raridade atual será coerente com todos os limites astronómicos. Este mecanismo, que contribuiu para a ausência de monopólos, apoia o modelo do «universo inflacionário», que descreverei num dos capítulos seguintes.

Contudo, mesmo que os monopólos não existam, a tentativa para os compreender já ensinou muito aos físicos. Ao tentarem dominar matematicamente os monopólos GUTs, os físicos teóricos foram arrastados para a proibitiva, mas maravilhosa, matemática da topologia. Estão a ser agora reveladas ricas e complexas propriedades topológicas das teorias de campo e é pouco crível que as futuras teorias do micromundo não vão incorporar, de certo modo, estas características.

Estas explorações revitalizaram a simbiose da matemática pura e da física. Os matemáticos e os físicos teóricos dialogam, os físicos estão ansiosos por aprender os segredos dos matemáticos e os matemáticos intrigados pela perspetiva de as suas abstrações contribuírem para a nossa compreensão do mundo real. É um diálogo com uma longa tradição, que remonta aos tempos recuados em que se constatou pela primeira vez que as razões entre números simples estavam relacionadas com a música harmónica das cordas. Hoje a música do microcosmos é guiada mais por princípios abstratos do que por números simples. Não obstante, a música abstrata que hoje ouvimos entusiasma-nos, como as primeiras harmonias simples entusiasmaram os antigos que primeiro ouviram as cordas cantar.

[1] Do inglês: *gauge field*. (*N. do T.*)

[2] Singularidades que surgem como soluções de equações de ondas. (*N. do T.*)

CAPÍTULO 15

A Unificação da Gravidade

A força gravitacional é a mais antiga força conhecida do homem e a menos bem compreendida.

PETER VAN NIEUWENHUIZEN, 1981

Newton descobriu a lei da gravitação universal, que relaciona a queda dos corpos na Terra com o movimento da Lua e dos planetas. Mas a sua reivindicação de prioridade na descoberta desta grande lei foi contestada. O seu colega da Royal Society, Robert Hooke, sentiu que tinha precedido Newton, no que tinha alguma razão, e Newton sabia-o. Newton, que não era tímido, argumentou contra a reivindicação de Hooke. Os conteúdos das suas cartas só superficialmente parecem a correspondência educada entre dois cavalheiros ingleses; lendo-as nas entrelinhas, descobre-se muito insulto e azedume. Foi numa dessas cartas para Hooke que Newton proferiu a famosa afirmação: «Se vi mais longe do que qualquer homem, foi porque subi aos ombros de gigantes.» Esta declaração é ainda mais pertinente quando constatamos que Hooke era baixo. Newton pretendia não só diminuir a estatura científica de Hooke, como com a mesma frase representar a sua própria linhagem intelectual como remontando aos antigos.

A lei da gravitação de Newton manteve-se como o modelo de lei física clássica durante dois séculos. Então, em 6 de Novembro de 1919, perante uma reunião conjunta de membros da Royal Society e da Royal Astronomical Society, foram relatadas as observações do encurvamento da luz à volta do Sol, confirmando a teoria da gravitação de Einstein. A lei de Newton fora substituída.

De acordo com a nova teoria da relatividade de Einstein, aquilo que experimentamos como sendo a força da gravidade pode ser matematicamente representado pela curvatura de um espaço-tempo quadridimensional: a gravidade é geometria. Einstein deduziu da teoria geométrica da gravidade um desvio da trajetória da luz junto do Sol duplo do previsto pela aplicação da teoria de Newton. Foi este maior desvio que foi, de facto, observado e relatado em 1919. Desde então a teoria da relatividade geral tem sido repetidamente sujeita a severos testes experimentais, utilizando-se, entre outros métodos modernos, a medição de distâncias com radar aos planetas distantes e aos satélites artificiais. Sobreviveu, sucessivamente, a todos os testes. A maior parte dos físicos estão tão impressionados com o enorme sucesso quantitativo da relatividade geral no mundo macroscópico que ficariam surpreendidos se a teoria falhasse dramaticamente num futuro próximo.

No entanto, e a despeito do êxito experimental da teoria da gravidade de Einstein no sistema solar e da imagem cósmica abrangente que ela implica, do ponto de vista da física quântica, a teoria é extremamente intrigante. De acordo com a teoria quântica dos campos, cada campo tem uma partícula quântica associada, e

o *quantum* do campo gravitacional chama-se o «gravitão». A partícula gravitão, embora exista em teoria, interage tão fraquissimamente com a restante matéria (muito mais fracamente do que os neutrinos) que mesmo grandes progressos em relação à tecnologia dos nossos dias não seriam suficientes para produzir instrumentos que a detetassem.

Contudo, os físicos teóricos podem calcular matematicamente as interações do gravitão com a restante matéria, encontrando então números infinitos. Curiosamente, tais números, que desapontaram anteriormente os teóricos, quando calculavam as interações dos fotões com a matéria, podiam ser domados, «renormalizados», no que parece ser um procedimento matematicamente consistente. Mas os números infinitos encontrados nas interações do gravitão não podem ser «renormalizados»; a situação é bastante pior. A interação gravitacional quântica com a matéria não é uma teoria renormalizável, o que quer dizer que os físicos não a compreendem.

O problema dos números infinitos na teoria quântica é já suficientemente angustiante; existem, porém, problemas conceptuais ainda mais profundos. Um deles é que a própria definição de partícula quântica, que pode ser formulada com precisão na teoria da relatividade restrita de Einstein, falha, ou pelo menos deve ser profundamente modificada para ser coerente com a teoria da relatividade geral. Da combinação da teoria quântica com a relatividade restrita nasceram novos conceitos do micromundo, de alcance profundo. Ninguém sabe como devem estes conceitos ser modificados quando se tentar combinar a teoria

quântica com a teoria da relatividade geral. Ninguém sabe até hoje se tal combinação será possível.

Um exemplo simples ilustra a dificuldade da combinação dos conceitos quânticos com a relatividade geral. O ponto de partida da relatividade geral é o «princípio da equivalência», pelo qual um campo gravitacional local não é distinguível de um movimento acelerado. Se nos encontrarmos num foguetão em movimento uniformemente acelerado no espaço exterior, seremos pressionados contra o chão do foguete, tal e qual como se existisse um verdadeiro campo gravitacional no foguetão (como sucederia se o foguetão estivesse em repouso na superfície de um planeta).

Einstein reconheceu neste princípio de equivalência que a presença de um campo gravitacional local releva do problema de o observador estar ou não acelerado, quer dizer, depende do sistema de coordenadas que se escolha para descrever o movimento. Por exemplo, se escolhermos para sistema de coordenadas o foguetão em aceleração, então existe um campo «gravitacional», mas num sistema de coordenadas não acelerado não existe nenhum. As leis matemáticas fundamentais da física deverão, porém, tomar a mesma forma para todos os observadores, independentemente do facto de um observador estar acelerado, imóvel ou animado de qualquer movimento em relação a outro observador. De outro modo, as leis fundamentais dependeriam da escolha arbitrária feita por um observador de um sistema de coordenadas de medida e esta espécie de arbitrariedade não pode refletir-se nas leis fundamentais. Este princípio da «invariância geral de coordenadas» está incorporado na relatividade geral. Neste aspeto, ultrapassa a

primitiva teoria da relatividade restrita, que exigia somente que as leis matemáticas da física tivessem a mesma forma para observadores com movimento relativo uniforme — um movimento especial, a velocidade constante.

De acordo com a teoria quântica relativista dos campos, um campo gravitacional constante cria um banho radiante de partículas quânticas, como sejam fotões, a uma determinada temperatura. Seria como se estivéssemos dentro de um forno (felizmente, para a força da gravidade medida na Terra esta temperatura é muito baixa). Mas o princípio da equivalência implica que um campo gravitacional seja o mesmo que uma aceleração; portanto, um observador em movimento acelerado vê um banho de partículas quânticas criado pelo campo gravitacional, enquanto outro observador em repouso não vê nada. Assim, é a própria noção de criação e destruição das partículas quânticas que sofre alteração. Não fica claro o que restará do conceito de «partícula quântica» na relatividade geral; no entanto, esse conceito é fulcral no pensamento dos físicos acerca do micromundo.

Os físicos teóricos, entre os quais Einstein, que refletiram profundamente sobre estes problemas, estão convencidos de que a combinação da teoria quântica com a teoria da relatividade geral impõe uma modificação substancial das nossas ideias fundamentais da física. Einstein pensava que, se a teoria quântica e a teoria da relatividade geral pudessem ser combinadas, isso resultaria numa teoria de campo unificado de todas as forças. Até ao presente, e apesar dos enormes esforços dos melhores cérebros da física, ninguém conseguiu realizar esta fusão de uma forma matemática e

fisicamente consistente. Esta demora não é de estranhar se nos lembrarmos de que os físicos levaram décadas de luta intelectual para combinar a teoria quântica com a teoria da relatividade restrita, que é mais simples — uma fusão que resultou nas notáveis teorias quânticas relativistas dos campos e numa nova visão do microcosmos. É de crer que muitas mais décadas passarão antes que os físicos consigam uma fusão consistente da teoria quântica e da relatividade geral. Descobrir uma teoria quântica da gravidade é o grande enigma a desvendar pelos modernos físicos teóricos.

Mesmo na ausência de uma teoria quântica da gravitação viável, os físicos não conseguem resistir a especular acerca da natureza do espaço-tempo à escala de distância de 10^{-33} cm — a escala de Planck, à qual os efeitos da gravidade quântica se tornam importantes —, nem que seja para se convencerem de que os seus conceitos mais queridos desaparecem a estas distâncias. Alguns mantêm que, à escala de Planck, o espaço-tempo desenvolve uma estrutura semelhante à da espuma. A escalas de distância maiores, como as que observamos todos os dias, o espaço-tempo parece plano e regular, como a superfície do oceano visto de uma certa altitude, mas de perto, à escala de Planck, é agitado e espumoso, como o oceano em tempestade. Se os físicos quiserem descrever o micromundo à escala de distância de Planck, então pouco restará do conceito de espaço-tempo contínuo, no qual a descrição da Natureza se tem baseado até agora. Talvez seja necessário invocar novos conceitos que transcendam o espaço e o tempo.

No entanto, e apesar das atuais teorias quânticas relativistas dos campos falharem à escala de Planck, os físicos desconhecem que

algo possa impedi-los de descrever o micromundo para qualquer escala de distância maior do que a de Planck. Por esta razão, pode ser razoável ignorar a gravidade quando se pensa na unificação de todas as outras forças. Já existem teorias matemáticas consistentes — GUTs — que unificam as forças eletromagnética, fraca e forte a escalas de distância superiores à de Planck, embora tenham ainda de ser testadas.

Muitos físicos estão convencidos de que, apesar de teorias como as GUTs terem feito luz sobre a dinâmica do universo primitivo, só uma teoria totalmente unificada — que inclua a gravidade — permitirá fornecer uma explicação da origem do universo. Com efeito, se imaginarmos que recuamos no tempo até ao universo primitivo, a temperatura e a energia de interação das partículas quânticas podem aumentar sem limites, de tal maneira que a escala de distância de Planck possa vir a ser explorada. O problema da gravidade quântica parece inevitável se pretendermos compreender a origem mesma do universo.

Embora os físicos estejam longe de atingir o objetivo de inventar uma teoria de campo unificado completa que incorpore a gravidade, muitos sentem que se aproximaram desse objetivo na última década. O ponto de partida tem sido sempre a teoria da relatividade geral de Einstein e os seus conceitos associados, dado que esta tem tido enormes êxitos quando aplicada à física gravitacional macroscópica. O problema reside em modificá-la de maneira a não perder as previsões de tanto êxito da gravidade em grande escala, resolvendo ao mesmo tempo os enigmas da gravidade quântica das curtas distâncias e unificando-a com as outras forças da Natureza.

Duas respostas sugeridas para este problema têm atraído recentemente a atenção de muitos: a teoria da «supergravidade» e a teoria de «Kaluza-Klein» do espaço-tempo com mais de quatro dimensões. Se estas ideias irão conduzir a becos sem saída ou se serão passos em direção a uma teoria mestra do universo, só o tempo o dirá. São, no entanto, ideias loucas, que podem até ser suficientemente loucas para estarem certas.

1. Supersimetria e Supergravidade

Toda a partícula quântica pode ser imaginada como um pequeno pião rodando, tomando esta rotação (*spin*) apenas os valores discretos de 0, $^1/_2$, 1, $^3/_2$, 2, etc., medidos em certas unidades. O spin 0 significa que o pião não roda; o *spin* $^1/_2$ significa uma dada quantidade de rotação; o *spin* 1 significa duas vezes essa quantidade, etc. O pião, que é um hadrão de interação forte, tem *spin* 0; o protão, o neutrão, os quarks e os leptões têm todos spin $^1/_2$, enquanto o fotão e os gluões fracos *W* e *Z* têm *spin* 1 e o gravitão *spin* 2. A rotação foi uma das pedras angulares do sistema de classificação de Wigner para as partículas quânticas.

Os físicos dividem as partículas quânticas com *spin* em duas classes, os bosões, que são partículas com *spin* inteiro — 0, 1, 2, etc. —, e os fermiões, que são partículas com *spin* semi-inteiro — $^1/_2$, $^3/_2$, etc. A razão desta divisão é que os bosões e os fermiões comportam-se muito diferentemente segundo as leis da mecânica quântica. Bosões idênticos, por exemplo, podem ocupar a mesma posição no espaço, enquanto fermiões idênticos não podem. Os

bosões idênticos são «sociáveis» e preferem condensar-se em grupos, ao passo que os fermiões idênticos são «insociáveis», excluindo-se uns aos outros. Os eletrões apresentam este «princípio de exclusão» dos fermiões: ao orbitarem em torno do núcleo, não podem ocupar um mesmo estado de energia. O exclusivismo dos eletrões resulta numa repulsão mútua e explica por que é que os átomos não entram em colapso se forem comprimidos uns contra os outros.

Tanto os bosões como os fermiões desempenham papéis importantes na descrição da Natureza ao nível mais fundamental. O modelo padrão compreende estas duas espécies de partículas — os quarks e os leptões são fermiões de *spin* $^1/_2$, enquanto os gluões e as partículas de Higgs são bosões de *spin* 1 e *spin* 0, respetivamente. O modelo padrão compreende também simetrias importantes, como a simetria de padrão de Yang-Mills, que relaciona os diferentes gluões. Mas uma característica crucial de todas essas simetrias, que veem os vários campos quânticos como componentes de um único campo subjacente, é que os diferentes campos quânticos que se transformam uns nos outros por operações de simetria têm de ter o mesmo *spin*. Todos os campos de Yang-Mills, por exemplo, têm *spin* 1, de modo que essas operações de simetria não misturam campos de *spins* diferentes.

Recentemente, para surpresa de muita gente, os físicos teóricos descobriram outra simetria matemática, que transforma, de facto, uns nos outros, os campos com *spins* diferentes, à qual deram o nome grandioso de «supersimetria». Sob uma transformação de supersimetria, os campos de bosões de *spin* 0 podem ser

transformados num campo de fermiões de spin $^{1}/_{2}$, e vice-versa. Do mesmo modo que os diferentes gluões podem ser vistos como diferentes componentes de um único campo de Yang-Mills, também, de acordo com a supersimetria, os bosões e os fermiões de *spins* diferentes podem ser vistos como componentes diferentes de um único «supercampo». Sob uma operação de supersimetria, os diferentes componentes do supercampo, campos de diferentes *spins*, transformam-se uns nos outros. Pela primeira vez a imaginação matemática encarou a possibilidade de todas as partículas quânticas, e não somente as de idêntico *spin*, serem componentes de um único supercampo mestre. É esta perspetiva de unificar as partículas e os campos de *spins* diferentes, esta possibilidade de englobar todos os campos possíveis, que excita os físicos à medida que exploram as complexidades matemáticas da supersimetria.

A supersimetria foi descoberta, independentemente, por vários grupos de físicos. Foi discutida por Y. A. Golfond e E. P. Likhtman, do Instituto de Física Lebedev, de Moscovo, e mais tarde por D. V. Volkov e V. P. Akulov, do Instituto de Física Técnica de Kharkov. Pierre M. Ramond e John Schwarz, do Cal Tech, e André Neveu, da École Normale Supérieure, descreveram igualmente uma simetria bosão-fermião. Mas a supersimetria só atraiu a atenção de muitos físicos em 1973, quando Julius Wess e Bruno Zumino inventaram uma teoria quântica relativista dos campos simples, renormalizável e supersimétrica.

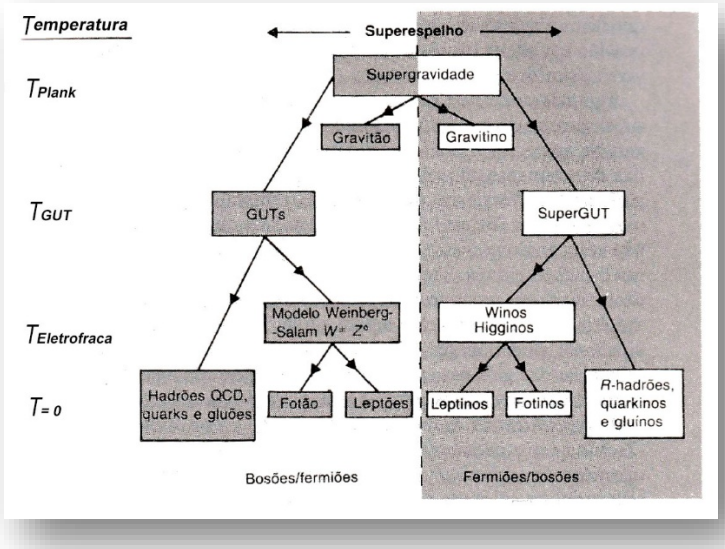
A partir dessa data, muitos teóricos laboriosos construíram várias outras teorias de campos supersimétricas. A teoria de campo

supersimétrica de Wess-Zumino não pretendia ser um modelo matemático realista das partículas quânticas existentes. Em vez disso, fornecia uma espécie de laboratório conceptual no qual os físicos matemáticos podiam explorar as implicações da supersimetria antes de entrarem na construção de modelos mais complicados, que esperavam poderem vir a ser experimentalmente relevantes. O ente matemático fundamental deste modelo era um único supercampo com componentes correspondentes aos campos de *spin* 0 e *spin* $^1/_2$. A supersimetria exata implicava que as partículas quânticas com *spin* 0 e *spin* $^1/_2$ deveriam ter igual massa (isto é igualmente verdade nos modelos supersimétricos mais complicados). Mas tais partículas de massa igual e *spins* diferentes nunca foram observadas na Natureza, razão pela qual o modelo simples não é experimentalmente relevante. Para conduzir a uma descrição das partículas observadas na Natureza a supersimetria deve, portanto, ser quebrada. Então as massas de partículas de diferentes *spins* relacionadas pela supersimetria não têm de ser iguais. Todavia, embora os físicos tenham construído modelos matemáticos em que a supersimetria é quebrada, não conseguiram relacionar nenhumas das partículas já conhecidas e com *spins* diferentes através da supersimetria.

Os físicos que estudaram teorias supersimétricas estão convencidos de que os quarks, os leptões e os gluões vulgares *não* estão relacionados uns com os outros por uma operação de supersimetria, não obstante terem *spins* diferentes. Se a supersimetria quebrada se manifestar na Natureza, então os «superparceiros» dos quarks, dos leptões e dos gluões — os *quanta* que estão realmente relacionados com eles pela supersimetria —

devem ser partículas quânticas completamente novas, nenhuma delas detetada até agora e na sua maior parte de grande massa. Os físicos teóricos divertem-se a prever quais serão as suas propriedades.

Se a Natureza for governada pela supersimetria, então o micromundo será regido por uma espécie de «superespelho». De um lado do superespelho estão as partículas vulgares, como os leptões, os quarks e os gluões; do outro lado, cada uma destas partículas tem um superparceiro imagens — novas partículas, alcunhadas de «leptinos», «quarkinos» e «gluínos». Enquanto os leptões e os quarks são fermiões de *spin* $^1/_2$, os seus superparceiros imagens são bosões. O superparceiro do fotão de spin 1 é o «fotino» de *spin* $^1/_2$, etc. A tarefa de dar um nome a estas partículas imaginadas tem qualquer coisa em comum com a tarefa de dar um nome às bestas imaginadas como mantecoras[1] e unicórnios pelos zoólogos medievais.



O superespelho. Se a supersimetria existir, então as partículas quânticas vulgares — quarks, leptões e gluões — têm superparceiros, quarkinos, leptinos e gluínos, que poderiam ter massa muito elevada. Talvez possam vir a ser detetados quando forem construídos novos aceleradores. Este diagrama ilustra a unificação das forças à medida que a temperatura do universo aumenta. À temperatura de Planck todas as forças estão unificadas sob a égide de uma teoria da supergravidade. A temperaturas mais baixas a simetria é quebrada, dando origem às diferenças entre partículas vulgares e as suas superparceiras.

Quando aparecerem novos superanéis de colisões, como o desertrão, talvez a energia seja suficiente para criar as novas Superpartículas e estas «bestas» imaginadas se tornem reais. Infelizmente, ninguém pode calcular com confiança a massa dos superparceiros, sendo, portanto, desconhecido o limiar de energia para produzir estas novas partículas exóticas. É concebível que a massa da maior parte dos superparceiros seja tão elevada que nem os superanéis de colisões possam criá-los; talvez a única ocasião em que se dispôs de energia suficiente tenha sido o universo primitivo.

Algumas superpartículas, como o fotino e o gravitino, o superparceiro do gravitão, poderão ter massas relativamente baixas. Se assim for, então a razão pela qual não foram detetadas diretamente não se deve a terem grande massa, mas sim ao facto de interagirem de maneira extremamente fraca com a matéria vulgar. Estas partículas podem desempenhar um papel na evolução do universo. Alguns físicos especulam que os fotinos ou os gravitinos, libertados das suas interações com o resto da matéria nos primeiros estádios do *big bang*, podiam ser a matéria escura do universo atual e que, por aglomeração gravitacional, poderiam ter desempenhado um papel importante na formação das galáxias. A estrutura do próprio universo torna-se o campo de ensaios para as ideias loucas da supersimetria.

A supersimetria manifestada no modelo protótipo Wess-Zumino é a supersimetria «global», querendo isto dizer que a operação de supersimetria é a mesma em todo o espaço. Nos seus artigos e palestras Zumino sublinhou que, tal como as antigas simetrias internas globais podiam ser generalizadas a simetrias locais do tipo Yang-Mills, também a supersimetria global deveria ter uma generalização de campo de padrão local. Mais ainda, era evidente, a partir das propriedades matemáticas da supersimetria, que tal generalização faria com que o campo gravitacional surgisse como o campo de padrão associado à supersimetria local. A primeira teoria matemática desta versão local da supersimetria, chamada «supergravidade», foi descoberta em 1976 por Sergio Ferraro, dos Laboratórios Frascati, perto de Roma, e por Daniel Z. Freedman e Peter van Nieuwenhuizen, da Universidade Estadual de Nova Iorque, em Stony Brook. Pouco depois Bruno Zumino, do CERN, e Stanley Deser, da Universidade Brandeis, obtiveram uma derivação alternativa mais simples da teoria da supergravidade.

A supergravidade revelou-se uma imaginosa extensão da teoria da gravidade de Einstein, fazendo desta uma teoria supersimétrica e, notavelmente, a relatividade geral foi, de facto, generalizada. Pode dizer-se que a teoria original de Einstein consistia na descrição do gravitão, o hipotético *quantum* de gravidade, como um bosão de *spin* 2. Na extensão supergravitacional da teoria de Einstein, o gravitão adquire um superparceiro, o gravitino, um fermião de spin $^{3}/_{2}$, e estas duas partículas transformam-se uma na outra sob transformações supersimétricas locais.

Ao efetuarem os cálculos quânticos, usando a teoria da supergravidade, os teóricos descobriram, para surpresa sua, que os números infinitos que infestavam a primitiva teoria da gravidade, que apenas continha o gravitão, eram agora cancelados por números infinitos simétricos produzidos pelo gravitino. Tais cancelamentos não eram fortuitos, mas uma consequência mais profunda da presença da supersimetria. Embora não se saiba ainda se a teoria da supergravidade é completamente renormalizável, este «abrandamento dos infinitos» parece ser um passo em direção a uma teoria viável da gravidade quântica.

A teoria simples da supergravidade inclui somente o gravitão e o gravitino, o que dificilmente corresponde ao mundo real, com as suas muitas partículas. Existirão outras teorias matemáticas da supergravidade? Na busca de teorias matemáticas das partículas quânticas, os físicos procuram sempre princípios gerais poderosos que possam limitar a área de pesquisa. Mostrou-se que o princípio da supersimetria local é tão restritivo que somente oito possíveis teorias da supergravidade existem. (No reino das ideias, isto é análogo a mostrar que só existem cinco sólidos tridimensionais regulares.) Estas oito teorias da supergravidade são rotuladas por um inteiro *N*=1,2, ...8, onde a supergravidade *N*=1 é a mais simples, compreendendo apenas os campos do gravitão de *spin* 2 e do gravitino de *spin* $^3/_2$. As teorias de supergravidade de *N* mais elevado requerem campos com *spin* 0, $^1/_2$ e 1. A supergravidade *N*=8, a mais complexa, compreende um total de 163 campos, todos mutuamente relacionados pela supersimetria. Somos tentados a identificar alguns destes 163 campos com os conhecidos quarks, leptões e gluões do modelo padrão, mas, infelizmente, esta

identificação simples falha. A maior parte dos que trabalham na supergravidade sentem que falta ainda uma ideia crucial; sem ela as teorias simplesmente não descrevem o mundo real.

Apesar de ainda não estar resolvido o problema de tornar a teoria da supergravidade realista, a compreensão teórica avançou. Cada uma destas teorias traz consigo a gravidade, de modo que a supergravidade é, potencialmente, uma teoria de campo completamente unificado. Todos os campos, incluindo os de *spin* 0, $^1/_2$ e $^3/_2$, são agora uma consequência de uma supersimetria local, enquanto anteriormente só os campos de Yang-Mills com *spin* 1 podiam ser deduzidos como uma consequência da simetria.

Os físicos que trabalharam muito tempo na supergravidade testemunham o seu poder conceptual e a complexidade matemática subjacente, características que partilha com o progenitor, a teoria da relatividade geral. Talvez postulando a existência de uma única supersimetria mestra se possa abarcar todo o universo. É ainda muito cedo para desistir deste sonho. Talvez seja uma teoria supersimétrica o «Santo Graal» da investigação em física.

2. A Quinta Dimensão e Para Lá Dela

Uma característica do nosso mundo físico é tão óbvia que passa despercebida à maior parte das pessoas — o facto de o espaço ser tridimensional. Na teoria da relatividade restrita de Einstein o espaço e o tempo tornam-se intimamente ligados, tão profundamente que Hermann Minkowski foi capaz de mostrar que

o tempo nesta teoria é como se fosse uma quarta dimensão (embora não seja uma dimensão espacial). Ninguém tem a mais leve ideia da razão pela qual o mundo em que vivemos tem um tempo e três dimensões de espaço, e não, por exemplo, onze dimensões. É claro que o mundo seria extremamente diferente se lhe alterássemos a dimensionalidade. Talvez dimensionalidades mais elevadas sejam letais para a vida e devamos estar agradecidos pelo nosso modesto lote de quatro.

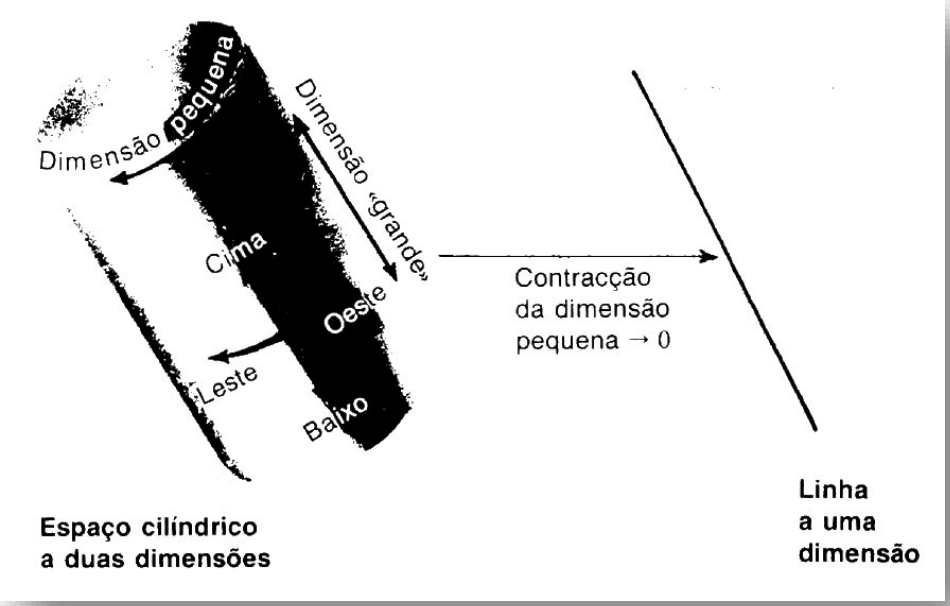
O facto, confirmado pela experiência, de o mundo em que vivemos ter três mais uma dimensão espaço-tempo é algo que a forma atual das leis da física traduz com simplicidade. Alguns físicos não estão satisfeitos e são da opinião de que a dimensionalidade do nosso mundo deveria ser logicamente dedutível de uma teoria mestra do universo, e não apenas um postulado de partida. Atualmente, não podem ainda calcular o número observado das dimensões do espaço-tempo a partir de primeiros princípios. Mas, invulgarmente, estão a desenvolver um quadro conceptual no qual tais cálculos possam um dia ter sentido. Este quadro conceptual, conhecido por «teoria de Kaluza-Klein», nasceu de outra generalização da relatividade geral a quatro dimensões, de Einstein, desta vez para espaços de dimensões superiores. Antes de descrever a singular teoria de Kaluza-Klein farei uma pequena digressão para descrever o que significam dimensões «grandes» e «pequenas».

As três dimensões espaciais que observamos são dimensões «grandes» — podemos mover-nos dentro delas. Se existirem dimensões adicionais, não poderão ser como as «três grandes»; se

o fossem, poderíamos mover-nos também dentro delas, o que, evidentemente, entra em conflito com a experiência. As dimensões adicionais que os físicos contemplam são dimensões «pequenas», tão pequenas que não podem ser vistas e, portanto, não influenciam diretamente a nossa perspetiva tridimensional do mundo. Que são então dimensões pequenas?

Para visualizar as dimensões pequenas, imaginemos um mundo que tenha apenas uma dimensão grande. O espaço desse mundo a uma dimensão é representado por uma linha de comprimento infinito. Imaginemos a seguir que esta linha está situada sobre a superfície de um cilindro, de tal maneira que o espaço completo é agora representado pela superfície bidimensional do cilindro. A segunda dimensão adicional corresponde a andar em torno do cilindro. Voltaremos então ao ponto de partida — a dimensão adicional é uma circunferência, e não uma curva.

Os espaços que se enrolam sobre si próprios, como o espaço a uma dimensão de uma circunferência ou a superfície bidimensional de uma esfera, são chamados pelos matemáticos «espaços compactos». Um cilindro pode ser imaginado como um espaço a duas dimensões, uma das quais é compacta (a circunferência) e a outra é não compacta (a linha reta). Podemos imaginar o raio da circunferência tão pequeno que se reduza a zero; então voltamos ao espaço a uma dimensão — a linha de comprimento infinito. É claro que, fazendo a circunferência muito pequena, podemos aproximar-nos do espaço a uma dimensão da linha, tanto quanto quisermos. A circunferência é a dimensão pequena adicional e a linha é a dimensão grande observada.



A superfície de um cilindro é um espaço a duas dimensões. A dimensão «grande» é a linha e a dimensão «pequena» é a circunferência. Se a dimensão pequena for comprimida até ter raio nulo, resta apenas uma linha, um espaço a uma dimensão. Ideia semelhante se pode aplicar ao nosso mundo, que poderá ter mais de quatro dimensões. As outras dimensões superiores podem ser dimensões pequenas, que não observamos diretamente.

A possibilidade de existirem dimensões pequenas adicionais para lá das «quatro grandes» do espaço-tempo — dimensões tão pequenas que não estão em conflito com a experiência — foi descoberta no contexto da relatividade geral de Einstein por Theodore Kaluza, matemático e linguista, que examinou as equações de Einstein generalizadas a um espaço-tempo de cinco dimensões, no qual a quinta dimensão adicional era compacta — não mais do que uma pequena circunferência. Imaginou que em cada ponto do espaço-tempo ordinário a quatro dimensões existiria uma pequena circunferência, tal como sobre cada ponto ao longo da linha sobre o cilindro considerámos existir uma pequena circunferência.

Tal como podemos mover-nos de ponto para ponto no espaço vulgar, também podemos imaginar uma partícula movendo-se à volta de uma pequena circunferência na quinta dimensão. É claro que não se afasta muito (e de maneira alguma nas dimensões grandes) porque a circunferência é muito pequena e tudo o que faz é andar continuamente à volta. No entanto, que significa a possibilidade deste grau de liberdade adicional? Kaluza mostrou que esta liberdade de movimentos adicional associada a uma simetria circular em cada ponto do espaço-tempo podia ser interpretada como a simetria simples de padrão do campo eletromagnético. Esta interpretação não é surpreendente, de um ponto de vista moderno, se pensarmos que a simetria (como a da pequena circunferência) implica automaticamente a existência de um campo de padrão (como o campo eletromagnético). A teoria das cinco dimensões de Kaluza não só descreve a curvatura do grande espaço-tempo a quatro dimensões em termos das equações gravitacionais usuais de Einstein, como unifica também fisicamente a gravidade e o campo de padrão eletromagnético de Maxwell, utilizando a estranha ideia de uma quinta dimensão circular. O trabalho de Kaluza impressionou Einstein, que lhe escreveu: «A ideia de conseguir [uma teoria unificada] por intermédio de um mundo cilíndrico a cinco dimensões nunca me ocorreu [...] A minha primeira impressão é que a sua ideia me agrada muitíssimo.» O próprio Einstein começou a trabalhar nessa ideia.

Kaluza só pôde demonstrar a unificação da gravidade e do eletromagnetismo por intermédio desta quinta dimensão compacta fazendo algumas suposições restritivas ao resolver as equações de Einstein. Em 1926 Oskar Klein fez progredir significativamente esta

teoria, demonstrando que estas suposições restritivas eram completamente desnecessárias. Mais ainda: Klein calculou o raio da pequena circunferência na quinta dimensão em função de grandezas conhecidas, a escala de distância de Planck e a carga eletrónica, e descobriu que este raio era cerca de 10^{-30} cm — um raio extremamente pequeno, garantindo que a quinta dimensão está em segurança, fora de vista. Porém, apesar das suas pequenas dimensões, a liberdade que os campos têm de se mover em torno dessa minúscula circunferência está sempre presente em cada ponto do espaço vulgar e é tudo quanto é preciso para garantir a existência do campo eletromagnético.

Depois da década de 30 a ideia de Kaluza-Klein caiu em desfavor e permaneceu adormecida durante anos. Mas, recentemente, ao procurarem-se todos os caminhos possíveis para a unificação da gravidade com as outras forças, essa ideia voltou de novo a ser proeminente. Hoje em dia, ao contrário do que sucedia na década de 20, os físicos querem fazer mais do que a unificação da gravidade com o eletromagnetismo, querem unificar também a gravidade com as interações fraca e forte, o que carece de ainda mais dimensões, além da quinta.

Os físicos teóricos generalizaram a teoria original a cinco dimensões a um número arbitrário de dimensões superiores. Todas as dimensões superiores são compactas: encurvam-se num minúsculo espaço multidimensional que existe em cada ponto do espaço vulgar e é, portanto, inobservável. Mas a liberdade de movimentos dentro destes minúsculos espaços compactos com simetrias mais gerais do que a simetria circular simples corresponde

exatamente à liberdade de realizar transformações de padrão de Yang-Mills. Facto assinalável, as simetrias de padrão locais são precisamente as simetrias do espaço composto de dimensão superior. Devido a este facto matemático, todas as teorias de padrão dos campos de Yang-Mills podem ser interpretadas estritamente de forma geométrica em termos de tais espaços compactos de dimensão superior.

Infelizmente, a teoria de Kaluza-Klein é extremamente restritiva, tão restritiva que ninguém foi ainda capaz de encontrar uma teoria de Kaluza-Klein realista que permita obter o modelo padrão. Conquanto recebam com agrado tais princípios restritivos como limitação das suas escolhas na busca da teoria correta, os físicos ficam frustrados pelo facto de tais restrições extremas terem até agora conduzido a teorias que não conseguem descrever o mundo quântico observado. Apesar de tudo, a pesquisa continua.

Uma notável descoberta foi feita em 1978 por Eugene Cremmer e Bernard Julia, dois físicos matemáticos franceses, quando combinaram a ideia de Kaluza-Klein com a teoria da supergravidade. Recordemos que existem oito teorias da supergravidade distintas, sendo a supergravidade $N=1$ a mais simples, com os campos do gravitão e do gravitino apenas, e a mais complicada a supergravidade $N=8$, com 163 campos diferentes. Cremmer e Julia verificaram que, se examinarmos a supergravidade $N=1$ num espaço a onze dimensões (em vez de quatro), supusermos que sete destas onze dimensões são compactas, à maneira de Kaluza-Klein, e que as restantes quatro são as dimensões grandes do espaço-tempo, então a teoria resultante a

quatro dimensões é a supergravidade *N*=8. Uma teoria simples da supergravidade *N*=1 a onze dimensões transforma-se na complicada teoria da supergravidade *N*=8 a quatro dimensões. Este resultado encoraja os investigadores, que têm esperança de que as complexas teorias de campo de que precisam para descrever o mundo real a quatro dimensões sejam o fruto de teorias muito mais simples de dimensionalidade superior. Alguns físicos creem que tudo o que precisam de fazer é descobrir a aplicação apropriada da ideia de Kaluza-Klein para que a teoria mestra do universo surja.

Apesar do apelo estético das ideias fundamentais, existem obstáculos matemáticos de grande monta que têm de ser ultrapassados para que a ideia da unificação multidimensional funcione. Ninguém sabe a razão profunda pela qual algumas dimensões são compactas e pequenas e outras — as quatro que vemos — grandes. Até mesmo na teoria de Kaluza-Klein se supõe simplesmente que essas quatro dimensões são grandes, enquanto as outras são compactas — uma suposição que os físicos esperam um dia poder eliminar. É concebível que a ideia de simetria quebrada — neste caso, a simetria quebrada de um espaço de dimensão elevada — desempenhe um papel na eliminação desta suposição. Talvez o mundo real, com as suas quatro grandes dimensões, corresponda à solução de simetria quebrada, mas estável, das equações que descrevem as simetrias de uma geometria multidimensional. Tais pistas, embora intrigantes, ainda não resolveram o enorme enigma da dimensionalidade que se observa no espaço-tempo.

Existem outros problemas. No seu trabalho pioneiro Klein calculou o raio da quinta dimensão em função do comprimento de Planck e também da carga eletrónica, que é uma medida da intensidade da interação eletromagnética. Se soubéssemos o valor do raio da quinta dimensão, então, invertendo os cálculos, poderíamos calcular a carga eletrónica. Recentemente, os físicos estimaram os raios das outras dimensões e utilizaram-nos para calcular as cargas que medem a intensidade das outras forças. Mas estas cargas calculadas são excessivamente grandes para que tenham algo a ver com as intensidades observadas das forças, o que sugere que estas teorias multidimensionais não são muito realistas.

Outro enigma ainda é que os campos de *spin* $^1/_2$ que observamos na Natureza são difíceis, se não mesmo impossíveis, de acomodar num espaço compacto multidimensional. Faltam alguns ingredientes cruciais para que a ideia de Kaluza-Klein possa ser posta em marcha, se é que isso é possível.

Estes enigmas estimulam, atualmente, a imaginação dos físicos teóricos. A ideia de que as várias simetrias de padrão que desempenham papel crucial na compreensão das forças da Natureza são simplesmente uma manifestação da simetria de um espaço de dimensionalidade superior exerce tal fascínio que o trabalho sobre esta ideia maravilhosa continuará até que, finalmente, se mostre que ela é inconciliável com a experiência ou até que apareça uma ideia muito melhor. A esperança de conseguir uma unificação geométrica da gravidade com as outras forças da Natureza através de uma generalização profunda da teoria da

relatividade geral de Einstein às dimensões adicionais não será facilmente abandonada. Talvez um dia físicos persistentes possam, inclusivamente, resolver o mistério de o nosso mundo ter três dimensões espaciais e uma dimensão temporal. Na arena de tais ideias loucas à solta até mesmo a solução desse enigma profundo não parece coisa muito estranha de esperar.

Vejamos agora as consequências que estas ideias loucas, a estarem certas, têm sobre a origem do nosso universo. Entremos no mundo antes de ele ter um nanossegundo de idade!

---

[1] Animal lendário com cabeça humana, corpo de leão, espinhos de porco-espinho e cauda de escorpião. (*N. do T.*)

CAPÍTULO 16

Antes do *Big Bang* – O Universo Inflacionário

> Ouvi dizer muitas vezes que não existem ofertas grátis. Parece hoje possível que o próprio universo seja uma oferta grátis.
>
> *ALAN GUTH, 1982*

O modelo do *big bang* do universo primitivo cobre um período que se estende desde o primeiro nanossegundo até aos primeiros 300.000 anos e prevê corretamente as abundâncias relativas dos elementos leves e a temperatura da radiação de fundo de micro-ondas. Este sucesso motivaram um exame mais aprofundado das hipóteses subjacentes ao modelo do *big bang*, tendo os físicos sido confrontados com alguns enigmas acerca da própria natureza do universo como um todo.

Se continuarmos a utilizar o modelo do *big bang* como uma descrição do universo antes do primeiro nanossegundo, então o supercomputador no qual imaginámos estar programado o modelo terá um *output* extremamente simples: um gás homogéneo, com todos os quarks, leptões e gluões, que se torna cada vez mais quente à medida que o universo se contrai até à singularidade. Mesmo que ignoremos o problema da bizarra singularidade na origem dos tempos, outros problemas nascem do modelo do *big bang* se pretendemos explicar algumas das mais dramáticas

propriedades do universo observado — a sua isotropia em grande escala (parecer, aproximadamente, o mesmo em todas as direções) e o facto de ser espacialmente plano. Introduzimos estas propriedades observadas no modelo padrão como suposições de partida quando presumimos uma geometria FRW isotrópica para o espaço-tempo com um valor de Ω próximo de 1 — equivalente a considerar o espaço quase plano. Vale a pena recapitular as provas observacionais que apoiam estas hipóteses, as quais, como veremos, são muito intrigantes.

À escala das maiores distâncias, o universo é, singularmente, homogéneo e isotrópico, como o revela a distribuição uniforme observada das galáxias e quasares no céu (que não estão agrupados em algumas regiões apenas). A melhor prova da isotropia do universo é a radiação de fundo de micro-ondas, que é isotrópica em mais de uma parte em 10.000. Uma vez que esta radiação é um fóssil do *big bang*, podemos concluir que o universo primitivo era também extremamente isotrópico durante o *big bang*. Se quisermos explicar esta notável isotropia que vem do início do universo até aos tempos atuais, acabaremos por ter de enfrentar um enigma fácil de ilustrar.

Suponhamos que o universo começou 12 mil milhões de anos atrás. Olhando para o céu à noite, poderíamos localizar um grupo de quasares à distância de 7 mil milhões de anos-luz. Se olhássemos na direção oposta, poderíamos ver outro grupo de quasares, semelhante ao primeiro, que estivesse também a 7 mil milhões de anos-luz. Refletindo sobre estes factos, ficamos em presença de um enigma. Os dois grupos de quasares estão à

distância de 14 mil milhões de anos-luz um do outro e, uma vez que o universo tem só 12 mil milhões de anos luz, a luz proveniente de um desses grupos de quasares não pode ter alcançado o outro. Dado que a luz se propaga à máxima velocidade observada no universo, concluiríamos que os dois aglomerados de quasares são «causalmente desligados» — jamais poderiam ter comunicado entre si ou ter-se influenciado mutuamente em qualquer instante passado. Se isto for verdade, por que razão parecem tão semelhantes no aspeto e estádio de evolução, como se tivessem tido uma origem comum e sido sujeitos a processos físicos idênticos?

Para apreciar quão estranha esta situação é, imaginemo-nos a assistir à realização de um filme — uma grande paisagem na qual está prestes a ser filmada uma cena de batalha. Todos os atores e figurantes receberam instruções precisas para coordenarem as ações de maneira a que o resultado tenha a aparência de uma batalha real. Isto requer planeamento cuidadoso e comunicação entre os diferentes grupos e o realizador. Imaginemos o que aconteceria se tal planeamento não tivesse sido feito ou fosse impossível de fazer. Iríamos presenciar o caos, com pessoas e grupos aparecendo em momentos e lugares errados. Ficaríamos bastante surpreendidos se o resultado fosse uma cena de cinema perfeita.

Deveríamos ficar igualmente surpreendidos que grupos diferentes de galáxias ou quasares da «paisagem» do universo que nunca tiveram a oportunidade de comunicar uns com os outros pareçam tão semelhantes. A isotropia do universo é, de facto,

espantosa. Por exemplo, a temperatura da radiação de fundo de micro-ondas é, aproximadamente, a mesma em todas as direções do céu, a despeito de esta radiação ter origem em milhares e milhares de milhões de regiões do universo que estavam desligadas causalmente aquando do *big bang*. Como receberam todas estas regiões distintas «instruções precisas» para terem hoje a mesma temperatura se nunca comunicaram entre si? Semelhantemente, que é que «disse» às galáxias e aos quasares para se distribuírem tão uniformemente? A isotropia observada no universo é, assim, um enigma da causalidade. Do ponto de vista do modelo do *big bang*, a isotropia é simplesmente um acidente.

É claro que é possível ajustar as condições iniciais no modelo matemático do universo primitivo de tal maneira que, quando partes do universo causalmente desligadas entram em contato pela primeira vez, o resultado seja um universo homogéneo e isotrópico. É isto que se supõe no modelo padrão do *big bang*. Mas tal «ajuste preciso» das condições iniciais do universo para produzir os resultados observados não explica o porquê da uniformidade. Na realidade, somos nós — os construtores de modelos — que, artificialmente, lá colocámos desde o início a uniformidade.

A uniformidade do universo não é senão um dos problemas suscitados pelo modelo padrão do *big bang*. Outro enigma é que espaço seja quase plano por todo o universo. Inicialmente, poderíamos pensar que isso fosse natural. Mas, de acordo com a teoria da relatividade geral, o espaço do universo inteiro pode encurvar-se, estando essa curvatura relacionada com o parâmetro cósmico Ω, a razão entre a densidade média de matéria e uma

densidade crítica conhecida. Se for igual a 1, então o espaço será plano; se Ω for maior do que 1, então o espaço terá uma curvatura convexa, positiva; se Ω for menor do que 1, o espaço terá uma curvatura côncava, negativa.

Os astrónomos estimam que a densidade média da matéria coloca os valores de Ω entre $^{1}/_{10}$ e 2; portanto, as observações não são suficientemente rigorosas para permitirem decidir qual das três curvaturas possíveis existe realmente. Mas o que é notável, e muitas vezes se menospreza, é que o valor de Ω é muito próximo da unidade: mesmo que o espaço seja curvo, é muito aproximadamente plano, caso em que Ω é igual a 1. Por que razão é isto assim? Teoricamente, o parâmetro Ω pode tomar qualquer valor. Podia ser tão grande como 10.000, caso em que o universo teria entrado em colapso há muito, ou podia ser tão pequeno como $^{1}/_{10.000}$, caso em que a matéria estaria tão diluída que as galáxias não poderiam existir.

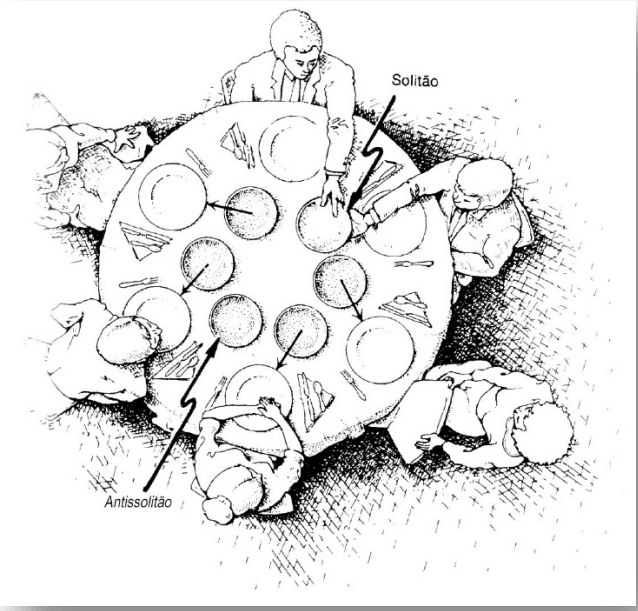
O enigma do espaço quase plano é complicado pelo facto de só mediante um «ajuste preciso» das condições iniciais do universo, de uma parte em milhões de biliões, se conseguir que o valor de Ω se situe entre os valores que os astrónomos hoje estimam — valores próximos da unidade. Corresponderia a «ajustar» o disparo inicial de uma bola de tal maneira que esta acabasse por pousar suavemente na mão de uma criança. Tal «ajuste», embora possível, dificilmente fornece uma explicação satisfatória do que se observa no universo.

Sucede ainda que uma terceira propriedade do nosso universo contemporâneo é igualmente intrigante do ponto de vista do modelo do *big bang*: a ausência de monopólos magnéticos, solitões topológicos, porções torcidas de energia do campo, na descrição das teorias de campos de padrão unificados — as GUTs. De acordo com elas, quando o universo primitivo tinha apenas 10^{-35} segundos de idade, a matéria consistia num gás de quarks, leptões e gluões a temperatura muito elevada, interagindo simetricamente, sem distinção entre as diferentes interações. À medida que o universo se expandiu e arrefeceu, a temperatura baixou, a simetria quebrou-se e as várias interações tornaram-se distintas. Atingida certa temperatura crítica, a interação eletromagnética tornou-se distinta, dado que a simetria de unificação se quebrou. Quando isso aconteceu, os monopólos magnéticos podiam ter sido criados no universo primitivo. Não é difícil compreender porquê.

Recordemos a nossa ilustração da quebra espontânea da simetria — um grupo de pessoas sentadas à volta de uma mesa de jantar redonda e instadas a pegarem no prato de salada à sua esquerda ou à sua direita, quebrando, assim, a simetria direita-esquerda. Se uma das muitas pessoas sentadas pegar no prato da direita e outra pessoa, sentada num outro lugar qualquer, pegar no prato da esquerda, segue-se que, pelo menos, um prato será reclamado por duas pessoas e outro prato por nenhuma. Esta confirmação da simetria quebrada contém, pelo menos, duas torções topológicas, uma localizada no prato reclamado por duas pessoas (o solitão) e a outra no prato reclamado por ninguém (o antissolitão). Curiosamente, se dissermos a uma das duas pessoas que reclamam o mesmo prato de salada para alterar a escolha e o

vizinho fizer o mesmo (e assim sucessivamente), a posição do prato reclamado por duas pessoas (o solitão) mover-se-á em torno da mesa até encontrar o antissolitão, aniquilando-o. Então todos à mesa ficarão com um prato de salada. A simetria direita-esquerda está ainda quebrada, mas no mesmo sentido a toda a volta da mesa.

Para desenvolvermos a analogia entre a quebra de simetria à mesa de jantar e a quebra de simetria no universo primitivo precisamos de qualquer coisa análoga à temperatura elevada na mesa de jantar. O efeito da temperatura elevada seria análogo ao de uma confusão agitada em que cada pessoa alterasse continuamente a escolha do prato de salada entre a direita e a esquerda. Então, em média, nem a escolha da direita nem a da esquerda seriam preferidas e a situação seria simétrica: um exemplo da restauração da simetria a temperatura elevada.



Um solitão e um antissolitão à mesa de jantar. Se uma das pessoas que reclamam o mesmo prato escolher, em vez disso, o prato adjacente, o solitão começará a mover-se à volta da mesa até encontrar o antissolitão e aniquilá-lo.

A descida da temperatura corresponde a uma diminuição da agitação: cada pessoa decide-se agora por um prato de salada. A simetria fica então quebrada. Todavia, quando isto acontece, algumas pessoas podem apanhar o prato da direita, outras o da esquerda, o que, como vemos, produz as torções topológicas — os solitões. O mesmo sucede no universo primitivo; à medida que a temperatura desce, as simetrias são quebradas. Mas, se se quebram diferentemente em diferentes regiões do universo, então podem formar-se solitões topológicos, os monopólos magnéticos. Quantos destes monopólos e antimonopólos seriam produzidos?

Suponhamos que a mesa de jantar redonda é tão grande que não cabe numa só sala, mas intersecta várias salas, de maneira que ficam paredes a separar grupos de pessoas sentadas à mesma mesa. Em qualquer das salas todas as pessoas pegariam num prato de salada à medida que ficassem menos agitadas e, uma vez que podem comunicar umas com as outras, todas concordariam em pegar no prato que lhes ficasse do mesmo lado. Contudo, dado que não é possível a comunicação entre salas, as pessoas em salas adjacentes teriam 50% de probabilidade de pegarem em pratos diferentes. Se agora as paredes fossem retiradas, estimaríamos que existiria um número de torções igual a metade do número de salas.

O universo primitivo imita esta imagem. As diferentes salas correspondem a regiões do universo causalmente desligadas que, à medida que o universo evolui com o tempo, entram em contato

causal. Uma vez que os campos de simetria quebrada podem ser alinhados diferentemente em diferentes regiões, da mesma maneira que a escolha dos pratos de salada pode ser diferente em salas diferentes, produzem-se monopólos magnéticos quando as regiões entram em contato umas com as outras. O universo está a expandir-se com rapidez suficiente para que os monopólos e os antimonopólos se afastem rapidamente uns dos outros e não se aniquilem em grande número. Os físicos calcularam o número de monopólos produzidos desta maneira e estimaram que a densidade de massa do universo atual deveria ser dominada por esses monopólos magnéticos. Isto é um resultado absurdo, especialmente se nos lembrarmos de que nenhum monopólo foi jamais detetado na Terra e de que existem limites severos para a sua abundância no universo. Deste modo, o modelo do *big bang*, aliado a algumas ideias acerca da quebra de simetria dos campos, coloca ainda um terceiro enigma: que aconteceu a todos os monopólos magnéticos?

Todos os três enigmas — isotropia, espaço plano e monopólos — ficarão esclarecidos se admitirmos a existência de uma nova Era *pré-big bang* da evolução cósmica, chamada «universo inflacionário». A principal característica desta breve, mas importante, Era é que o espaço do universo nela sofreu uma enorme expansão, muitíssimo maior do que a muito mais lenta expansão posterior ocorrida durante a Era do *big bang*. O tempo de vida do universo inflacionário durou talvez dos 10^{-35} segundos aos 10^{-33} segundos, muito antes, portanto, do primeiro nanossegundo. É difícil imaginar como algo que durou tão pouco tempo e aconteceu tão cedo na história térmica do universo possa resolver os enigmas

da isotropia, do espaço plano e da ausência de monopólos. No entanto, e curiosamente, a existência de uma Era inflacionária muito primitiva consegue-o precisamente.

A história da descoberta teórica do universo inflacionário começa na Primavera de 1979. Alan Guth e Henry Tye, dois jovens físicos teóricos da Universidade de Cornell, começaram a estudar matematicamente o enigma colocado pela ausência cosmológica de monopólos magnéticos. Guth, como muitos outros físicos das partículas dessa altura, estava cético quanto ao uso da cosmologia para impor restrições às teorias das partículas quânticas, mas Tye, já convertido a esta ideia, persuadiu-o de que essa abordagem tinha mérito. Em Setembro Guth e Tye, nessa altura na divisão de teoria do Stanford Linear Accelerator Center, escreveram um artigo sugerindo que, se o universo primitivo tivesse atravessado uma nova fase «superarrefecedora» anterior ao *big bang*, durante a qual a temperatura baixara abruptamente, então o problema da abundância dos monopólos magnéticos podia ser resolvido. A inflação não fazia, por enquanto, parte do cenário.

No Outono desse ano Guth interrogou-se pela primeira vez acerca do problema chave: como é que uma tal fase de superarrefecimento afetaria a evolução do universo? Descobriu que a resposta era a seguinte: antes do *big bang* quente o universo deveria ter passado por uma fase inflacionária, um período durante o qual o espaço sofreu uma expansão exponencial. No modelo usual do *big bang* o espaço do universo também se expande, mas a um ritmo muito mais lento. No espírito de Guth, o universo superarrefecido e a inflação estavam agora ligados logicamente.

Mas admitir uma fase inflacionária tão estranha parecia-lhe na altura uma ideia forçada e não explicava nada, exceto a ausência de monopólos.

Felizmente, Guth lembrou-se de uma exposição informal (*talk*) feita pelo físico Robert Dicke, de Princeton, durante uma visita a Cornell. Este falara acerca do trabalho que ele e P. James E. Peebles tinham realizado e que punha em evidência que o facto de o universo ser plano não tinha explicação no âmbito do modelo padrão do *big bang*. Em Novembro Guth compreendeu que a sua nova ideia da inflação resolvia este incómodo enigma. Bastou admitir a existência de uma época inflacionária para prever um valor de $\Omega=1$ (um espaço plano) e para tornar o ajuste preciso desnecessário. Este extraordinário resultado tornou a hipótese da inflação mais atraente. Mas nessa altura Guth nem sequer sabia do enigma da isotropia-causalidade.

Em Dezembro, ou nos princípios de Janeiro de 1980, Guth e outros teóricos do Stanford Center estavam a discutir um artigo de Anthony Zee que punha em destaque o enigma da causalidade. Guth compreendeu quase imediatamente que o seu cenário inflacionário resolvia também esse enigma cosmológico e que estava, portanto, no bom caminho.

Obteve então resultados pormenorizados importantes, elaborando um modelo cósmico explícito baseado na GUT SU(5), modelo este que acarretava a existência de uma Era inflacionária anterior ao *big bang*. O artigo ficou terminado no fim do Verão de 1980 e foi publicado em 1981. A ideia do universo inflacionário tinha

nascido, a primeira ideia realmente fresca em várias décadas de cosmologia. Curiosamente, Guth tivera a resposta — a Era inflacionária do universo — antes de saber a que pergunta ela respondia.

Refletindo sobre a sequência destes acontecimentos alguns anos mais tarde, Guth acha espantoso que ninguém o tenha precedido na descoberta, tanto mais que todas as peças para resolver o *puzzle* estavam à vista. Wolfgang Pauli, um teórico da geração anterior, queixando-se de falta de criatividade nos seus últimos anos, dizia: «Eu sei demasiado.» Guth, um novato em cosmologia, não «sabia demasiado» e, ao fazer a análise da relação entre a cosmologia e a física das partículas quânticas, encontrou uma nova maneira de ver o universo primitivo.

Para visualizarmos o universo inflacionário empregaremos uma imagem simples. Em vez de um espaço curvo, tridimensional, do universo real (que é difícil de visualizar), imaginemos o espaço do universo como uma linha a uma dimensão, quer como uma linha infinita para um universo «aberto», quer como uma circunferência para um universo «fechado». Suponhamos, para maior exatidão, que o universo é fechado e que o seu espaço é uma circunferência. Então a expansão do universo é representada pelo raio da circunferência, que aumenta, e pela circunferência, que «estica» como uma tira de borracha perfeitamente elástica. Isto é o modelo «tira de borracha» do universo.

Podemos ainda imaginar as ondas de luz propagando-se ao longo deste espaço a uma dimensão, como pequenas vibrações elásticas

numa tira de borracha. A importante característica física de uma onda de luz é ser a maneira mais rápida de enviar um sinal energético de um ponto a outro do espaço vazio. Logo, os raios de luz estabelecem se um acontecimento pode ou não influenciar causalmente um segundo acontecimento — um raio de luz resultante do primeiro evento deve ser capaz de alcançar o segundo.

Imaginemos a seguir que a circunferência está em expansão. O raio da circunferência pode expandir-se tão rapidamente quanto queira; não está limitado pela velocidade da luz, porque a expansão não transporta energia. Por exemplo, no modelo padrão o raio do universo cresce mais rapidamente do que a velocidade da luz; de facto, cresce com uma potência do tempo decorrido.

Se um acontecimento físico tiver lugar num tal espaço em expansão num dado momento, então podemos imaginar que os raios de luz são emitidos por esse acontecimento em todas as direções do espaço. Se um segundo acontecimento tiver lugar algures no espaço antes que qualquer raio de luz, partindo do primeiro, o tenha atingido, então o segundo acontecimento está «causalmente desligado» do primeiro — não existe processo pelo qual o primeiro acontecimento possa ter influenciado o segundo. No futuro, os raios de luz que partem de um acontecimento acabarão por atingir o outro, pondo-os então em contato pela primeira vez.

Isto é o que acontece no modelo padrão. Muitas regiões do espaço que hoje vemos na nossa galáxia correspondem a acontecimentos que estiveram causalmente desligados de nós (e

uns dos outros) no universo primitivo. Imaginemos que no universo primitivo, cujo raio era muito pequeno, existiam duas minúsculas flutuações que cresceram até, finalmente, constituírem galáxias. Presumimos que estes dois acontecimentos estão correlacionados com precisão, de tal maneira que em tempos posteriores, quando entrarem em contato causal, as duas galáxias parecerão a mesma. Isto é «ajuste preciso» do universo primitivo para o ajustar aos factos posteriores. Mas suponhamos que estabelecemos as condições iniciais diferentemente, de tal maneira que um dos acontecimentos foi uma pequena flutuação e o outro foi nada. Certamente que podemos fazer isto, dado que os dois acontecimentos estão causalmente desligados. Então esta condição inicial pode resultar num universo atual no qual exista uma grande região do céu sem nenhuma galáxia, um universo extremamente anisotrópico. O facto de no modelo padrão o universo correntemente observado consistir num número incontável de milhares de milhões de regiões que estavam causalmente desligadas nos estádios iniciais do *big bang* e eram, portanto, independentes umas das outras, sendo hoje, porém, muito parecidas, logo aparentemente correlacionadas, constitui o enigma da isotropia-causalidade.

No contexto do modelo simples da tira de borracha vemos agora como a inflação resolve não somente este incómodo problema da causalidade, mas também os problemas do espaço plano e da abundância de monopólos. Recapitulando, a ideia básica do universo inflacionário é a de que o espaço do universo sofreu uma enorme expansão num estádio inicial anterior ao *big bang*. Durante este breve período inflacionário, que durou, provavelmente, dos 10^{-}

^{35} segundos de idade até aos 10^{-33} segundos, o raio do universo aumentou exponencialmente com o tempo, em contraste com o ritmo muito mais lento do crescimento durante a expansão do *big bang*. Para termos uma ideia da intensidade da inflação em causa, imaginemos um universo tira de borracha do tamanho de um anel expandindo-se durante o período inflacionário até chegar a ter uma circunferência do tamanho do universo observado atualmente — um fator de 10^{30}. Mesmo esta intensa inflação é muito pequena, comparada com o fator de 10^{50} necessário para que a inflação resolva os enigmas cosmológicos. Após esta intensa inflação ter cessado, o universo continuou a expandir-se, mas ao ritmo muito mais lento do *big bang*.

É fácil ver que a inflação resolve o problema da causalidade, pois qualquer pequena região do espaço causalmente ligada aquando do começo da inflação é agora «esticada» até estar transformada numa imensa região do espaço ao terminar a Era inflacionária. Todos os acontecimentos nessa imensa região do espaço estão agora correlacionados porque foram originados a partir de uma única região extremamente pequena e causalmente ligada.

De acordo com este cenário, todo o universo hoje visível e todo o universo visível num futuro inimaginavelmente longo tiveram origem numa região espacial única e causalmente ligada antes de a inflação ter começado. Terminada a inflação, novos eventos e processos podem ter lugar na imensa região do espaço causalmente ligada, por exemplo, à formação e evolução das galáxias, eventos estes que *não* estão causalmente ligados de forma direta. Mas todos estes acontecimentos subsequentes foram postos em movimento

por acontecimentos mais antigos, anteriores à inflação, que *estavam* causalmente ligados. Estes acontecimentos receberam a sua «ordem de marcha» antes da inflação; mesmo que tenham vindo a perder o seu contato causal, sabem o que devem fazer.

Do ponto de vista do universo inflacionário, deixa, pois, de constituir surpresa que, ao dirigirmos o olhar para as extensões longínquas do espaço, vejamos galáxias que não parecem ser muito diferentes da nossa. A nossa saudação, ao vermos essas galáxias, não deve ser, portanto, «Olá!», mas «Olá, novamente!», pois há muito tempo — antes da inflação — estivemos todos juntos sob a forma de uma flutuação de densidade inicial.

A inflação resolve também o problema do espaço plano, o enigma que consiste no facto de Ω ser tão próximo da unidade. A enorme região do espaço que hoje observamos corresponde apenas a um minúsculo arco de uma imensa circunferência inflacionada. Este espaço, tem a aparência de uma linha, um espaço plano, e não de uma circunferência, da mesma forma que a Terra parece plana quando se observa uma pequena região dela. Uma previsão precisa do cenário do universo inflacionário é, portanto, que Ω=1: o universo observado deve ser hoje exatamente plano, independentemente da curvatura que tinha antes da inflação.

A previsão precisa de que Ω=1 pode ser um problema para a imagem inflacionária, dado que observações correntes apontam para um valor de Ω de cerca de um décimo. Sucede ainda que, se admitirmos simplesmente que Ω=1, então os cálculos das abundâncias relativas dos elementos criados durante o *big bang*

(que são sensíveis ao valor de Ω) não parecem concordar com as abundâncias observadas.

Apesar destes conflitos potenciais entre as observações e o cenário inflacionário, a maior parte dos cosmólogos pensam que a previsão de Ω=1 acabará por ser confirmada. Num encontro de cosmólogos e de físicos de partículas em São Francisco no ano de 1984 levou-se a referendo qual seria o valor de Ω. Não obstante provas experimentais em contrário, o consenso geral foi que Ω=1. Embora dificilmente represente um método científico para determinar Ω, este referendo reflete quão profundamente enraizada está a ideia da inflação.

A inflação resolve também o terceiro enigma — o problema da abundância de monopólos. Recordemos que os monopólos foram produzidos como torções topológicas quando regiões do espaço causalmente desligadas (nas quais os campos de quebra de simetria têm orientações diferentes) estiveram em contato umas com as outras. Mas, se a totalidade do universo hoje observado teve origem numa única região causalmente ligada, e não em várias, então nenhum monopólo deve existir no universo atual. Quaisquer monopólos produzidos na fase pré-inflacionária do universo foram simplesmente «apagados» — tornaram--se tão diluídos que não existem para efeitos práticos.

Vemos aqui uma primeira propriedade da inflação: dilui o universo. Quaisquer partículas quânticas específicas produzidas antes ou durante a inflação serão «apagadas» — a casa do universo é mantida limpa pela inflação.

O universo inflacionário resolve então os três enigmas cosmológicos colocados pelo modelo do *big bang*. Mas qual terá sido a causa deste notável comportamento do universo inicial, esta enorme inflação? E por que cessou ela? Por enquanto nada disse acerca disto.

O comportamento dinâmico da geometria do universo é governado pelas equações de Einstein que relacionam a curvatura do espaço com as propriedades da matéria nesse espaço. No modelo do *big bang* a matéria foi representada por um gás de todas as partículas quânticas, não tendo as propriedades deste gás, de acordo com o modelo padrão, como consequência uma fase inflacionária.

Nos modelos mais simples do universo inflacionário, como o que primeiro foi considerado por Guth, a matéria do universo está representada por um único campo com *spin* zero, que é uniforme no espaço, mas cuja amplitude varia no tempo. Um tal campo com *spin* zero, que pode ser um campo de Higgs de quebra de simetria, com todas as propriedades devidas, está presente em muitas das teorias de campo unificado. Admitir a existência deste campo, embora seja uma «ideia louca», não é extremamente descabido. A sua presença produz uma densidade de massa-energia e pressão uniformes no universo (no nosso modelo de tira de borracha, ao longo do fio) e, se pudermos determinar essa densidade e pressão, as equações de Einstein dar-nos-ão a curvatura do espaço. Como a amplitude do campo varia no tempo, a densidade de massa-energia e a pressão no espaço do universo variam também no tempo, o mesmo sucedendo, portanto, com a curvatura.



Desenho de A. Linde, físico que contribuiu para desenvolver a ideia do universo inflacionário. Uma propriedade desejável da inflação é que ela varra da casa do universo os objetos indesejáveis (porque não se observam), como os monopólos magnéticos, os gravitinos (o superparceiro do gravitão, o quantum de gravidade) e as «paredes de domínio», que podem formar-se entre regiões do espaço com configurações de campo topologicamente distintas.

As equações matemáticas que descrevem a variação no tempo do campo com *spin* nulo revelam-se idênticas às que descrevem uma bola rolando por uma colina a baixo; façamos, pois, uma analogia entre o campo e a bola. Dado que esta analogia é matematicamente precisa, até os físicos teóricos discutem a dinâmica da inflação em termos de bola e colina, sabendo que podem traduzir estes conceitos para a linguagem dos campos. Nesta analogia a coordenada horizontal da bola na colina é igual à amplitude do campo com *spin* zero e a altura vertical da bola acima da base da colina — a sua energia potencial *V* — é igual à densidade de massa-energia do campo. A energia cinética do movimento da bola, à medida que rola colina a baixo, é igual à soma da densidade de massa-energia e da pressão (medida em certas unidades)

produzida pelo campo com *spin* nulo no espaço do universo. Donde se segue que para calcular a densidade de massa-energia e pressão no universo à medida que evolui no tempo basta determinar como uma bola rola pela vertente de uma colina com determinada forma — um problema simples da física clássica.

A seguir, temos de dizer como a pressão e a densidade de massa-energia afetam a variação da curvatura do universo. No nosso modelo simples da tira de borracha isso equivale a determinar a variação do raio da circunferência com o tempo. Se o universo está em expansão, então a massa-energia que nele existe abrandará a expansão, da mesma maneira que a massa da Terra, por influência da gravidade, retarda o movimento de uma pedra atirada para cima.

O efeito da pressão é mais complicado. Poderia pensar-se que uma pressão positiva no universo favoreceria a expansão, da mesma maneira que o aumento de pressão no interior de um balão aumenta o tamanho do balão. Mas a pressão no universo não se exerce contra nenhuma parede — o universo não tem paredes — e a analogia com a pressão do balão está errada. De facto, um efeito da pressão no universo é produzir uma contribuição adicional para a densidade de massa-energia do universo. Daí que um aumento de pressão retarda efetivamente a expansão, em vez de a acelerar. Inversamente, se a pressão diminui, a taxa de expansão aumenta. Uma vez que não há nenhuma regra que proíba ao universo ter pressão negativa, é possível imaginar valores negativos da pressão, o que implica uma taxa de expansão ainda mais elevada. Se a pressão toma um valor precisamente simétrico do da densidade de

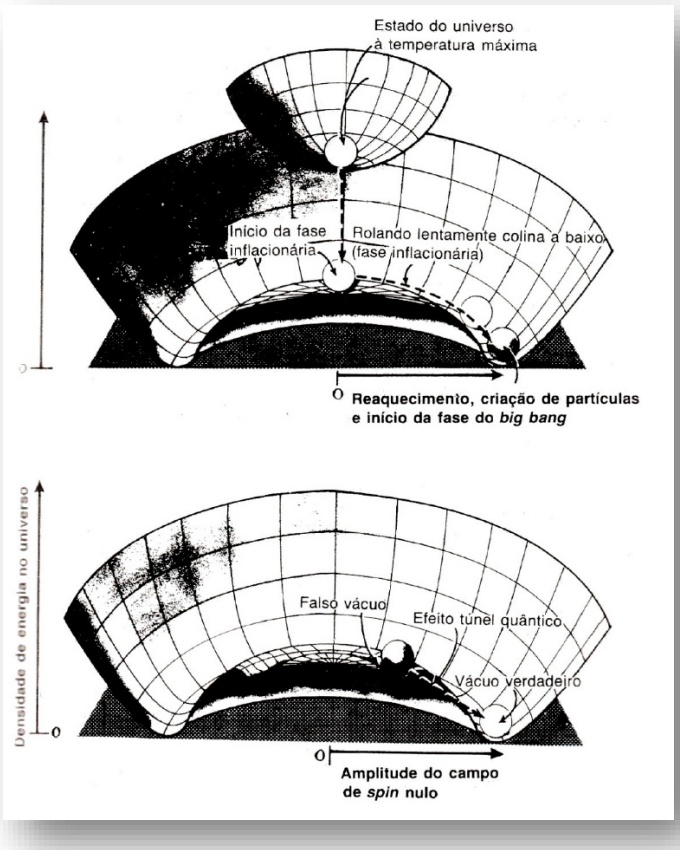
massa-energia (que se supôs não ser zero), então as equações de Einstein preveem que algo de notável acontece: o raio do universo expande-se exponencialmente com grande rapidez — uma expansão verdadeiramente inflacionária.

Agora que compreendemos, por um lado, a relação quantitativa entre a densidade de massa-energia e a pressão e, por outro, a taxa de expansão do universo, voltemos à bola rolando pela vertente da colina.

Neste caso, a energia do movimento da bola é igual à soma da densidade de massa-energia com a pressão. Assim, se a bola rolar muito lentamente, de tal maneira que a energia de movimento seja desprezível, então a soma da densidade de massa-energia com a pressão é efetivamente zero. A pressão é então simétrica da densidade de energia, precisamente a condição para que haja um crescimento inflacionário exponencial do raio do universo. A existência de uma expansão inflacionária do universo está, assim, associada a um «lento rolar» do movimento da bola. Que espécie de forma da colina produzirá um rolar tão lento?

Antes de responder a esta questão temos de considerar uma outra propriedade da colina, que é importante e complica as coisas: a forma da colina depende da temperatura do universo e varia, à medida que a temperatura desce, segundo uma lei calculável. Antes de a inflação começar, o universo está extremamente quente e a forma da colina de potencial é um poço fundo (conforme a figura). A posição estável da bola é no fundo do poço. À medida que o universo se expande, a sua temperatura baixa e, de acordo com

certos modelos específicos da teoria de campos, a colina de potencial toma a forma de um planalto central plano, com uma encosta gradualmente descendente até um «penhasco» que desce até zero para depois subir abruptamente. Consequentemente, à medida que a temperatura baixa, a bola, representando o estado da matéria do universo, começa por estar no cimo do planalto — situação não estável. Os físicos referem-se a este estado da bola como a um «estado de falso vácuo», porque representa um universo que não é estável. A posição estável da bola é no sopé da colina — o «estado de vácuo verdadeiro ».



Uma bola rolando colina a baixo — uma analogia matemática para a dinâmica do universo inflacionário. O estado da bola representa o estado do universo: a altura é a densidade de energia do universo; a posição é a amplitude do campo com *spin* zero, que é motor da inflação; a energia de movimento é proporcional à soma da pressão e da densidade de energia no universo. A forma da colina depende da temperatura do universo. A uma temperatura muito alta a bola repousa no fundo de um poço. À medida que a temperatura baixa, o poço toma a forma da colina que

a figura ilustra e a bola começa a rolar pela vertente a baixo (a fase inflacionária) até ficar em repouso no vale (a fase do *big bang*). A figura de baixo mostra o que pode acontecer se estivermos num estado de falso vácuo. Num tal estado, o verdadeiro vale não foi ainda alcançado. O nosso universo decairia então para o verdadeiro vácuo por meio do efeito túnel quântico.

Flutuações quânticas aleatórias no campo com *spin* nulo darão início ao lento rolar da bola planalto a baixo. Uma vez que a energia do movimento da bola é desprezível durante este período de lento rolar, a pressão é simétrica da densidade de massa-energia, o que implica que o universo cresça por inflações até atingir dimensões enormes. Logo, a existência do longo planalto plano é crucial para que a ideia da inflação funcione. Quando a bola começa, por fim, a descer rapidamente pelo penhasco, a energia do movimento da bola já não é desprezível e a fase inflacionária termina. Inicia-se uma nova «fase de reaquecimento».

O campo com *spin* zero que a bola representa interage com os outros campos quânticos — os gluões, os quarks, os leptões e os campos de Higgs — e esta interação pode ser representada por uma contribuição para a força de «atrito» que se exerce sobre a bola à medida que esta cai do penhasco. Como resultado do atrito e do reaquecimento, a bola oscila para a frente e para trás no vale, perdendo energia, até acabar por parar no fundo do vale — o vácuo verdadeiro estável. A densidade de energia e a pressão associadas ao campo com *spin* zero são agora nulas; este já não domina a matéria do universo. O que aconteceu foi que a imensa energia originalmente encerrada no campo com *spin* zero (quando a bola estava no planalto) foi usada para criar a bola de fogo primordial do *big bang*, com todas as partículas quânticas. Durante a inflação o universo estava relativamente frio, mas a força de atrito

reaqueceu o universo agora inflacionado até uma temperatura elevadíssima, e a verdadeira fase do *big bang* começou. O campo com *spin* zero, uma vez cumprida a missão de inflacionar o universo, torna-se inativo, tendo dado origem à bola de fogo irradiante das partículas quânticas. A inflação gerou também a enorme entropia que hoje vemos na radiação de micro-ondas — é a verdadeira «morte térmica» do universo. Depois do reaquecimento o universo continua a expandir-se ao ritmo mais lento associado ao modelo padrão do *big bang* como um gás irradiante de partículas quânticas. O *big bang* quente começou.

Em resumo, o cenário do universo inflacionário começa com o universo num estado extremamente quente em que a densidade de matéria é dominada por um único campo com *spin* zero. A temperatura baixa rapidamente, mas o universo vai parar a um «estado de falso vácuo», com as dimensões a crescerem exponencialmente. A inflação cessa depois de o «lento rolar» ter acabado e a enorme densidade de energia do campo com *spin* zero reduz-se a zero, desembocando no «vácuo verdadeiro». Durante a fase de aquecimento a energia do campo cria o gás quente de partículas quânticas do *big bang*. O universo tem agora 10^{-33} segundos de idade.

A tarefa principal dos físicos teóricos que trabalham na ideia do universo inflacionário é descobrir um modelo da teoria quântica relativista dos campos que tenha um campo com *spin* zero e para a qual a forma calculada da colina seja a adequada a que o cenário funcione em todos os pormenores.

A versão original do universo inflacionário de Guth baseou-se na GUT SU(5), a qual, embora implique inflação, tinha, além disso, a infeliz propriedade que consistia em esta não durar o suficiente para produzir o universo homogéneo que hoje vemos. Guth estava ciente deste problema e expô-lo. Posteriormente, Andrei Linde, na União Soviética e, independentemente, Andreas Albrecht e Paul Steinhardt, na Universidade da Pensilvânia, descobriram que outros modelos da teoria de campos possuíam também a colina, com um longo planalto, que possibilita o lento rolar e a prolongada inflação necessária. Desde então os físicos inventaram muitos outros modelos da teoria de campos baseados nas GUTs e nas ideias da supersimetria, tudo com a intenção de melhorar o cenário inflacionário e de o derivar de primeiros princípios. Muitos teóricos estão convencidos de que o universo inflacionário, com uma duração que vai dos 10^{-35} segundos até cerca de 10^{-33} segundos, é uma exigência lógica das teorias de campo que comportam as «ideias loucas». A ideia do universo inflacionário é o principal resultado do recente matrimónio da cosmologia com a teoria quântica dos campos.

Os teóricos que aceitam o sucesso da ideia do universo inflacionário aplicam-na a outros enigmas cosmológicos. Por exemplo, a inflação e a fase de reaquecimento subsequente fornecem a chave para a solução do enigma da assimetria matéria-antimatéria. Recordemos que um dos requisitos para que esta assimetria se produza é que o universo primitivo esteja num estado de não equilíbrio, durante o qual a taxa de expansão do universo é rápida, comparada com a taxa de colisões entre partículas quânticas. A fase inflacionária satisfaz este requisito crucial. Os

teóricos que estimam a assimetria matéria-antimatéria produzida durante a fase inflacionária e de reaquecimento encontram, aproximadamente, a assimetria observada. Sem tal fase de não equilíbrio (que ocorre naturalmente no cenário inflacionário, mas não na fase posterior do *big bang*) é difícil compreender a origem da assimetria matéria-antimatéria.

Os físicos estão também entusiasmados por o cenário do universo inflacionário poder fornecer a solução para o mistério da origem das galáxias. Mesmo antes de o cenário inflacionário ter sido descoberto, sabia-se que as galáxias podiam ter crescido a partir de minúsculas flutuações na densidade de matéria no universo primitivo. Mas a origem de tais flutuações foi sempre um enigma.

Uma solução proposta por muitos físicos é que estas flutuações foram flutuações quânticas no gás de partículas quânticas. O que é atraente nesta solução é que tais flutuações quânticas estão sempre presentes — não há necessidade de explicar donde vieram. Infelizmente, cálculos pormenorizados mostraram que estas flutuações eram milhares de milhões de vezes demasiado pequenas para produzir galáxias no modelo do *big bang*. Quando o cenário do universo inflacionário foi descoberto, os físicos constataram que existia uma nova maneira de examinar o problema. Vulgares flutuações aleatórias no campo com *spin* zero, que produziu a inflação, poderiam, durante o período inflacionário, produzir as flutuações de densidade necessárias para, posteriormente, engendrarem as galáxias.

Em 1982 vários físicos que assistiam à conferência de Nuffield, em Inglaterra, calcularam estas flutuações de densidade devidas ao campo com *spin* zero e chegaram todos à mesma conclusão: as flutuações eram agora excessivamente grandes por um fator de 1000. Nessa altura até este resultado errado pareceu constituir progresso, uma vez que todas as tentativas anteriores de calcular a amplitude das flutuações tinham produzido estimativas excessivamente pequenas. Posteriormente, físicos inteligentes encontraram modelos da teoria de campos que forneciam a amplitude exata para as flutuações. Pela primeira vez modelos teóricos algo simples baseados na teoria de campos podiam explicar a origem e evolução das flutuações a partir das quais as galáxias acabaram por se formar. A inflação não explica só a uniformidade geral do universo; determina também corretamente os desvios dessa uniformidade — a granulosidade que hoje vemos na forma de galáxias. Se esta ideia estiver correta, então as magnificentes galáxias nasceram das quase impercetíveis flutuações de campo ocorridas nos primeiros 10^{-33} segundos de vida do nosso universo.

Além de oferecer novas perspetivas para o problema da origem das galáxias e da assimetria matéria-antimatéria, o cenário inflacionário chama a atenção para um desastre potencial que poderá esperar o nosso universo. Depois da inflação o campo com *spin* zero torna-se inativo no vale do sopé da colina — o «vácuo verdadeiro». Mas como sabemos que este é realmente o «vácuo verdadeiro» e que não existe um vale mais fundo do outro lado da colina? A resposta é que não sabemos.

É possível imaginar e construir modelos matemáticos segundo os quais estaremos hoje a viver, não no verdadeiro vácuo, mas ainda noutro falso vácuo. Se assim for, então, de acordo com a teoria quântica, existe uma pequena probabilidade de, após muitos anos de vácuo falso, o nosso universo cair no vácuo verdadeiro, mais abaixo — a bola pode atravessar a colina que separa os dois vales por efeito túnel. Sidney Coleman e Frank de Luccia, na Universidade de Harvard, estudaram esta possibilidade e concluíram que depois de «atravessar o túnel» para o verdadeiro vácuo o universo viveria menos do que alguns microssegundos antes de entrar em colapso. Refletindo sobre isto, fizeram notar o seguinte:

> Isto é desencorajador. A possibilidade de estarmos a viver num falso vácuo nunca foi animadora. O decaimento no vácuo é a derradeira catástrofe ecológica; num novo vácuo existem novas constantes da Natureza; depois deste decaimento não somente a vida, tal como a conhecemos, é impossível, como o é também a química. Podíamos, porém, extrair um conforto estoico da possibilidade de no decorrer do tempo o novo vácuo vir a conservar, se não a vida, tal como a conhecemos, pelo menos, algumas estruturas capazes de conhecer a alegria. Esta possibilidade foi agora eliminada.

É decerto a cosmologia, e não a economia, a «ciência lúgubre».

A inflação anterior ao *big bang* e a posterior fase de reaquecimento resolvem muitos antigos enigmas — a origem da isotropia e o facto de o espaço ser plano, a ausência de monopólos magnéticos e, possivelmente, a assimetria matéria-antimatéria e a origem das galáxias. Ideias loucas como as GUTs, que primeiro promoveram o cenário do universo inflacionário, podem não

sobreviver ao teste do tempo. Mas a ideia central da inflação — que o universo sofreu em determinado momento uma expansão imensa — pode sobreviver, dada a inexistência atual de soluções alternativas para os formidáveis enigmas cosmológicos postos pelo universo observável.

Se admitirmos que o cenário inflacionário é válido, podemos então imaginar o que poderia ter tido lugar *antes* da inflação, a origem mesma do universo? A inflação expande o universo a um grau quase inimaginável e dilui toda a matéria e estrutura que contém. Na verdade, uma das maravilhosas propriedades deste cenário é que a inflação pode levar o universo de quase não importa que condições iniciais ao modelo padrão do *big bang*; o «ajuste preciso» das condições iniciais não é necessário. Mas esta mesma característica significa que nenhuma estrutura pode sobreviver à inflação e fornecer-nos pistas acerca do universo pré-inflacionário, da mesma maneira que a uniformidade e o carácter plano do espaço nos forneciam pistas acerca da existência de uma inflação.

A existência do universo inteiro, com as suas três dimensões espaciais e uma temporal, constitui a única pista para a sua origem. O facto de o universo existir carece de explicação. No próximo capítulo examinaremos as modernas especulações acerca do velho enigma de como começou o mundo.

CAPÍTULO 17

Antes da Inflação – A Origem do Universo

Estou propenso a acreditar que a origem e as propriedades do nosso universo possam ser explicáveis dentro do quadro da ciência convencional.

EDWARD P. TRYON, 1973

Como começou o universo? Há séculos que os homens pensantes enfrentam esta questão. Alguns acreditam que a resposta deve situar-se fora do domínio da ciência, sentem que a criação do universo foi um ato divino. Outros repudiam inteiramente a questão e mantêm que o universo nunca começou e que existiu desde sempre — uma visão que, mais recentemente, é expressa pelo modelo do estado estacionário do universo. No entanto, todos os dados astronómicos corroboram o simples facto de que o nosso universo era muito diferente no passado remoto e que teve uma origem precisa. É concebível que o universo seja infinitamente periódico, sofrendo, alternadamente, expansões e contrações. Mas esta periodicidade, a ser verdadeira, não pode ser estabelecida com base nas observações atuais. Um universo cíclico é uma possibilidade; não somos, porém, compelidos a adotar esta conceção e, por simplicidade, admitimos que a origem do nosso universo é um evento único.

Se olharmos para o universo em concordância com os modelos cosmológicos habituais, a temperatura e a densidade da matéria continuarão a crescer sem limites à medida que recuamos no tempo. Acabaremos por encontrar a singularidade do espaço-tempo onde as leis da física deixam de fazer sentido. Esta circunstância levou algumas pessoas a adotar uma atitude a que chamo «misticismo da singularidade» — a ideia de que até os cientistas devem desistir da tentativa de compreender racionalmente a origem do universo. Lembram-me aqueles que nas primeiras décadas deste século pensavam que a física tinha encontrado o merecido castigo ao tentar compreender o átomo. Todavia, o átomo foi posteriormente compreendido por meio da nova teoria quântica. Analogamente, nada barra o caminho a uma descrição racional da própria origem do universo e um dia tal será conseguido. A singularidade no começo do tempo deve ser vista como um enigma provocante, e não como um sinal de que devemos desistir.

Existem, no entanto, pessoas razoáveis que se opõem à ideia da compreensão da origem do universo por ela estar para além de tudo o que pode ser sujeito a observação ou teste. Mais ainda, argumentam que, se o cenário inflacionário estiver correto, então qualquer característica pré-inflacionária do universo teria sido apagada e não teria restado nenhuma pista observável para a sua origem.

Contudo, tais objeções não são válidas. Certamente a própria existência do universo inteiro e o *big bang* constituem provas de que houve alguma espécie de origem. Existem outras características do nosso universo que podem fornecer informação

sobre a sua origem, embora possamos, à primeira vista, não as tomar como pistas. Por exemplo, o cenário inflacionário requer que antes da inflação o universo tenha estado imensamente quente e sido muito denso, requisitos que deviam decorrer, logicamente, de uma teoria da própria origem. Ainda outro exemplo é a mais dramática de todas as propriedades do universo que sobreviveram à inflação: a dimensionalidade três-mais-um do espaço-tempo. Uma outra característica do universo pré-inflacionário é que apresenta um elevado grau de simetria, e também isto deve ser explicado por qualquer teoria da génese.

Ao tentarmos compreender a própria origem do universo, vale a pena recordarmos o «método postulacional de Einstein». Este método consiste em adivinhar intuitivamente um postulado físico (que não pode ser testado diretamente), deduzir logicamente as suas consequências e comparar posteriormente estes resultados com a experiência. Se o teste falhar, o postulado admitido deve também ser rejeitado.

Os cientistas terão certamente de adivinhar o modelo físico correto da origem do universo e até as leis físicas que governam tal acontecimento. No entanto, e curiosamente, o nosso conhecimento atual das leis da física e as características observadas do universo restringem severamente o nosso jogo de adivinhas.

A possibilidade de as novas ideias que provêm da física quântica serem usadas para elaborar modelos matemáticos definidos sobre a própria origem do universo que evitem uma verdadeira singularidade física entusiasma alguns cientistas. Tais modelos,

como as «ideias loucas», em que muitas vezes se apoiam, não têm suporte experimental preciso. Mas, neste estádio muito inicial de construção imaginativa de modelos, a falta de suporte experimental não inquieta os físicos, porque o que é notável acerca destes modelos não é tanto o poderem vir a revelar-se verdadeiros ou falsos, mas o facto de tais modelos racionais e matemáticos da origem do universo serem possíveis. Parece que o universo, apesar da sua imensidade e da sua estranha origem, será um dia compreendido pela razão, como entidade física que é.

A origem do universo teve lugar antes do período inflacionário. À medida que o universo se contrai, torna-se mais quente e mais denso e, de acordo com a relatividade geral, acaba por reduzir-se à singularidade do espaço-tempo. Porém, esta imagem puramente clássica do colapso deve ser modificada se tivermos em conta a teoria quântica. Os físicos sabem que a descrição clássica da geometria do espaço-tempo deixa de ser válida à escala de comprimentos de Planck, antes de se atingir a singularidade. A geometria do universo torna-se então um mar fervilhante, descrito como «espuma do espaço-tempo», e a influência da gravidade quântica torna-se dominante. Uma vez que o espaço e o tempo são os conceitos mais fundamentais na física — análogos ao uso das palavras nas frases —, é difícil dizer o que resta das leis convencionais da física neste estranho estádio do universo. Seria como se as palavras tivessem perdido o respetivo significado. No entanto, os físicos descobriram que podem servir-se de uma nova linguagem de configurações de campo para descrever a origem do universo. Como pensam alguns físicos acerca deste acontecimento?

Primeiro, devo deixar claro que os físicos e os cosmólogos não estão de acordo acerca de nenhum modelo «padrão» para a origem do universo. Todos os modelos que aqui iremos referir devem ser vistos como muito preliminares e a abandonar logo que surjam modelos melhores. O que os físicos procuram é um modelo preciso deste acontecimento, como os modelos precisos dos interiores das estrelas ou dos últimos estádios do *big bang*, em que se dá a síntese do hélio.

Um critério que qualquer destes modelos deve satisfazer é não deixar em aberto interrogações sem resposta acerca de um estádio do universo «anterior à origem». De outro modo, não se trataria verdadeiramente de uma teoria da origem do universo. Por exemplo, podemos querer acreditar que o universo começou como «qualquer coisa» — uma semente inicial. Mas depois ficamos a braços com a pergunta seguinte: donde veio essa semente? A ideia alternativa de que o universo começou do «nada», uma criação *ex nihilo*, satisfaz o critério de não deixar nenhuma questão sem resposta sobre o estádio preexistente do universo. Mas o que é «nada»?

O que os físicos habitualmente consideram «nada», aquilo a que chamam «estado de vácuo», é o estado correspondente ao valor mais baixo possível da energia total para um determinado sistema físico. Tal sistema físico poderia ser o sistema solar, uma galáxia ou o universo inteiro. O espaço plano e vazio encaixa nesta descrição do «estado de vácuo». Mas, se pusermos qualquer coisa nesse estado de vácuo, como um eletrão ou um fotão, a sua energia total aumenta e deixa de ser um estado de nada; não está vazio.

Esta definição do estado de vácuo — tão razoável como parece — depende da nossa capacidade de definir com precisão o que se entende por «energia total» de um sistema físico. Uma maneira de o fazer é notando que a energia é equivalente à massa e a massa é uma fonte de um campo gravitacional. Se nos afastarmos de um sistema físico e medirmos o campo gravitacional que produz, poderemos determinar a sua massa total e, portanto, a sua energia total. Contudo, só se pode atribuir um significado não ambíguo a este campo gravitacional produzido pelo sistema se o espaço for plano a grandes distâncias do sistema. Se o espaço longínquo não for plano, mas curvo, então a curvatura pode ser pensada como representando a presença de um campo gravitacional adicional que não pode ser desligado do que foi produzido pela matéria. Isto significa que não podemos determinar o campo gravitacional devido à matéria e, portanto, a energia do sistema.

A «energia total» de um sistema físico, um conceito com significado no espaço plano, perde-se se admitirmos que o espaço tenha uma curvatura arbitrária. Uma vez que o espaço do universo inteiro se pode encurvar, a energia total do universo deixa, assim, simplesmente, de ter significado. Esta conclusão de que os conceitos de energia total e de conservação da energia total não se aplicam ao universo inteiro é algo surpreendente, mas verdadeira. Logo, se quisermos definir um «nada» — o estado de vácuo — aplicável ao universo como um todo, devemos procurar outras propriedades do vácuo que não envolvam conceito de «energia total».

Uma característica do estado de vácuo é que deveria ser eletricamente neutro, não deveria ter carga elétrica, nem qualquer das outras espécies de cargas estudadas pelos físicos das partículas quânticas, pois, se o vácuo possuísse uma carga total absolutamente conservada, então nunca poderíamos livrar-nos de todas as partículas que transportam a carga e o vácuo não seria «nada» — seria alguma coisa.

Se definirmos vácuo como o estado para o qual todas as grandezas conservadas fisicamente são zero, então concluímos, curiosamente, que o universo inteiro poderia ser equivalente a nada. Uma primeira reação a esta sugestão é que é certamente um disparate, porque o universo inteiro é tudo, não é «nada». No entanto, se examinarmos esta sugestão de perto, descobrimos que, de facto, o universo poderia ser equivalente a um estado de nada e, portanto, é possível que o nosso universo tenha tido origem no vácuo.

O primeiro a expressar esta ideia foi Edward Tryon, um antigo aluno de Steven Weinberg e hoje um físico do Hunter College, na cidade de Nova Iorque. Ed e eu encontrámo-nos de tempos a tempos em encontros científicos ou nas ruas do Upper West Side de Manhattan, onde ambos vivemos[1]. Durante um destes encontros, muitos anos atrás, Ed falou-me desta ideia sobre a origem do universo — que ele podia ter começado a partir do vácuo —, o que me estimulou a pensar acerca do problema. Num artigo de 1973, intitulado «É o universo uma flutuação de vácuo?», Tryon sublinha que a soma de todas as cargas conservadas, como a carga elétrica, pode ser zero se considerarmos a totalidade do universo, o qual

pode, portanto, ter sido criado a partir do vácuo. Nenhuma lei da física impede uma criação *ex nihilo*.

Tryon sugere também que o universo teve origem numa flutuação quântica do vácuo — uma flutuação minúscula que se transformou no *big bang*. Tal como as partículas quânticas podem ser espontaneamente criadas a partir de um vácuo, também o universo pode ter sido criado de um vácuo. Todavia, embora uma flutuação quântica espontânea num vácuo possa momentaneamente produzir uma partícula e uma antipartícula, estas são rapidamente destruídas, porque a produção de partículas reais, com energia positiva, acarreta uma violação da conservação de energia. Não obstante a conservação da energia poder ser violada em conformidade com a relação quântica de incerteza, só o pode ser durante um curto intervalo de tempo. Como pode então o universo inteiro emergir de uma flutuação quântica do vácuo se nem sequer duas partículas o podem fazer?

A razão pela qual as partículas reais não brotam hoje vindas do espaço vazio é porque o nosso espaço é muito plano e num tal espaço a lei de conservação da energia proíbe semelhante processo. Mas no universo primitivo o espaço era altamente curvo e, portanto, a conservação da energia total não tinha sentido. Uma flutuação quântica do vácuo, criadora de partículas reais, *pode* ter lugar se o espaço for altamente curvo. Tal flutuação poderia «desembestar», criando a grande quantidade de partículas quânticas que hoje identificamos com o *big bang*.

Apesar de parecer hoje original e intrigante, o artigo de Tryon teve pouco impacto no pensamento cosmológico até há pouco tempo, provavelmente porque o autor não elaborou um modelo matemático que concretizasse as especulações. No princípio da década de 70 Y. Zel'dovitch e A. Starobinski, na União Soviética, desconhecendo o trabalho de Tryon, sugeriram que as flutuações quânticas na geometria do espaço-tempo durante a Era de Planck podiam produzir partículas e antipartículas. Estas novas partículas retirariam energia gravitacional da geometria flutuante, acabando por amortecer as flutuações e por produzir um universo quente, como no modelo do *big bang*.

A partir de 1978, R. Brout, P. Englert, E. Gunzig e, mais tarde, P. Spindel, da Universidade de Bruxelas, produziram uma série de trabalhos que descrevem o primeiro modelo do universo originado a partir do nada através de um processo quântico. Para eles o estado do vácuo, o estado de nada a partir do qual o universo começa, é o espaço plano vazio; mostram como, se uma flutuação quântica produzir algumas partículas num tal espaço, as interações gravitacionais mútuas entre elas fazem com que o espaço se torne curvo. Segue-se então uma produção de partículas em cascata e o espaço torna-se curvo: um universo aberto e expandindo-se, cheio de matéria, é criado a partir do espaço plano vazio. Os autores concluem:

> Numa palavra, mostramos que as leis da mecânica quântica formuladas no quadro da relatividade geral são perfeitamente compatíveis com a criação espontânea de toda a matéria e radiação no universo, a qual tem, no princípio, uma origem arbitrária, no espaço-tempo.

Contudo, um enigma permanece: logo que uma tal flutuação do vácuo se desencadeia, pode manter-se a si própria; mas como começa?

A origem do universo foi claramente um acontecimento muito violento. Trabalhando na Universidade Rockefeller, em 1981, David Atkatz e eu interessámo-nos por este enigma. Desenvolvemos um modelo matemático no qual a origem do universo é um evento do tipo efeito túnel, da mecânica quântica, semelhante ao decaimento de um núcleo atómico, cujas partículas atravessam a barreira (de potencial) nuclear que ordinariamente as confina. A origem do universo pode ser vista como um evento semelhante, mas com consequências mais dramáticas. A ideia base é que o estado inicial de «nada» é um estado de falso vácuo, como uma esfera colocada atrás de uma barreira que tem probabilidade pequena, mas não nula, de a atravessar para um estado de energia mais baixa. Ao fazer isto, o vácuo falso decai para o estado de *big bang* das partículas interagentes.

De acordo com o nosso modelo, o universo teve de ser espacialmente fechado — um universo compacto —, o que significa que o espaço inicial do universo — «nada» — era um espaço que consistia numa minúscula «esfera» perfeita, sem nenhuma matéria dentro dela. Apesar de o espaço de uma esfera perfeita ser um estado de «nada» tão bom como o espaço plano e vazio, ficámos desapontados por (pelo menos, no nosso modelo) o espaço plano não poder decair num universo com *big bang*. Tínhamos esperado encontrar uma maneira de explicar o cenário de Brout-Englert-Gunzig-Spindel — a criação a partir do espaço plano. Todavia, um

universo que comece como uma minúscula esfera compacta sugere-nos algo mais:

> O que perspetivamos é que o universo começou como [um espaço Kaluza-Klein multidimensional e compacto]. Um subespaço quadridimensional deste [espaço multidimensional] passa então por efeito túnel para a configuração de bola de fogo, sendo as dimensões restantes as simetrias internas que observamos.

Do nosso ponto de vista, o universo começa como um espaço multidimensional com um elevado grau de simetria. Mas o universo com esta geometria pode ser instável e sofrer um decaimento por efeito túnel. Como consequência, as quatro dimensões tornam-se «grandes» dimensões e expandem-se rapidamente em tamanho, enquanto as restantes — as dimensões pequenas — permanecem pequenas e são hoje representadas pelas simetrias «internas» das partículas quânticas.

Não demos qualquer explicação da razão pela qual são quatro as dimensões que crescem, em vez de sete ou onze. Essa explicação pertence ao futuro. Contudo, se esta ideia estiver correta, então a origem do universo é o acontecimento que estabelece o número hoje observado das dimensões espácio-temporais. Este número, é importante que o recordemos, é uma pista para a origem do universo, que sobrevive ao período inflacionário subsequente — a dimensionalidade do espaço-tempo não se dilui.

J. Richard Gott III, astrofísico da Universidade de Princeton, inventou ainda um outro modelo, desta vez para criar universos

abertos. Gott admite a existência de um espaço que se expande exponencialmente — semelhante ao do período inflacionário —, chamado «espaço de, De Sitter», do nome do físico holandês que primeiro explorou as suas propriedades matemáticas. Stephen Hawking mostrara em 1974 que se dá produção de partículas quânticas — a radiação de Hawking — em horizontes de acontecimentos como a superfície de um buraco negro. No espaço de, De Sitter, tais horizontes de acontecimentos são omnipresentes. Gott pensou então que as partículas devem ser produzidas por toda a parte nesse universo, enchendo-o de matéria quente, e mostrou que dentro deste «superespaço» se poderiam formar bolhas de espaço aberto ordinário, que cresceriam essencialmente à velocidade da luz, as quais se encheriam de radiação através do processo de Hawking. À medida que uma bolha cresce, torna-se mais diluída e à sua rápida expansão desacelera-se. O mundo dentro de tal bolha parece-se muito com o nosso próprio mundo durante o estádio do *big bang*. Enormes quantidades de bolhas poderiam ter-se formado no «superespaço» de, De Sitter, correspondentes a enormes quantidades de universos, nenhum dos quais pode jamais comunicar com qualquer outro. Segundo Gott, a radiação de fundo de micro-ondas a 3 K que podemos observar hoje é o fóssil da radiação produzida pelo processo de Hawking quando o nosso universo era uma minúscula bolha.

Todos os modelos da origem do universo que até agora apresentei pressupõem a existência prévia de algum tipo de espaço vazio — o vácuo donde tudo começa. O modelo de Brout, Englert, Gunzig e Spindel pressupunha um espaço plano e vazio a quatro dimensões. Atkatz e eu pressupusemos um espaço fechado, talvez

com mais de quatro dimensões. Gott pressupõe a existência prévia de um espaço de, De Sitter, a quatro dimensões. Alex Vilenkin, físico teórico da Universidade de Tufts, não ficou satisfeito com nenhuma destas noções de «nada». «O espaço é ainda qualquer coisa», fez-me notar um dia Alex, «e penso que o universo devia realmente começar como nada. Nem espaço nem tempo — nada.» Quando Alex me falou pela primeira vez nesta possibilidade, perguntei-lhe: «Que é que entende por "nada"?» Ele encolheu os ombros e declarou enfaticamente: «Nada é nada!»

Em 1983 escreveu um artigo, intitulado «O nascimento dos universos inflacionários», no qual construiu um modelo matemático em que o «nada» — nenhum espaço ou tempo — passa, por efeito túnel, a uma geometria de espaço e tempo, donde poderia partir um estado inflacionário para o universo. O seu trabalho foi mais longe do que algumas ideias semelhantes desenvolvidas em Moscovo em 1982 por P. Grishuk e Y. B. Zel'dovitch que, contudo, não propuseram um mecanismo definido. Mas como podem o espaço e o tempo ser criados a partir do «nada»?

Descrevi anteriormente uma visualização do espaço de um universo inflacionário fechado como um anel perfeitamente elástico — o modelo da tira de borracha. Usemos agora esta imagem para explicar a criação do universo (desta vez um universo a uma dimensão). À medida que o tempo recua, o anel encolhe-se até se transformar num ponto, desaparecendo em seguida. Do nosso ponto de vista, no espaço a três dimensões o anel está mergulhado no nosso espaço e desaparece no nosso espaço. Mas, do ponto de vista de um ser que podemos imaginar vivendo no espaço

unidimensional do anel, o espaço não desaparece noutro espaço de dimensão mais elevada; pura e simplesmente, desaparece. Da mesma forma, o espaço-tempo real a quatro dimensões do nosso universo pode, pura e simplesmente, desaparecer no nada absoluto. Se pode desaparecer no nada, também pode ser criado a partir do nada.

Uma maneira de visualizar a criação do espaço e do tempo, segundo Vilenkin, é imaginar uma esfera suspensa sobre uma folha plana infinita. O espaço físico do universo, o anel a uma dimensão, é descrito pela intersecção da superfície da esfera com a folha, que é nada, nem sequer espaço. A esfera «desce» então em direção ao plano e toca-o — o espaço do universo é agora um ponto. À medida que a esfera penetra o plano e o atravessa, o ponto cresce até se tornar uma circunferência, representando o nosso universo de anel a uma dimensão. Então, à medida que a esfera continua a mover-se através do plano, o anel cresce primeiro e depois entra em colapso até só restar um ponto, desaparecendo seguidamente no nada. Um espaço circular a uma dimensão brotou do nada e tornou a mergulhar nele. No modelo de Vilenkin a criação do espaço-tempo é um processo de efeito túnel e o universo não começa por ser um ponto. Em vez disso, parte como um anel vindo do nada por efeito túnel; brota do nada e depois cresce.

Um defeito desta imagem de uma esfera descendo através de uma folha é que o tempo existe sempre. Mas o tempo pode ser visto como uma dimensão adicional e podemos construir uma imagem semelhante para a criação da dimensão temporal a partir do nada. Finalmente, deitemos fora a esfera, o plano e o espaço a três

dimensões no qual imaginámos estarem mergulhados a esfera e o plano para alcançar a ideia de uma verdadeira criação do espaço e do tempo a partir do nada — absolutamente nada. Vilenkin conclui assim:

> As vantagens do cenário aqui representado são de [um] carácter estético. O cenário fornece um modelo cosmológico que não tem uma singularidade no *big bang* (existem ainda singularidades finais) e não requer nenhumas condições iniciais ou de fronteira. A estrutura e a evolução do(s) universo(s) são totalmente determinadas pelas leis da física.

Sinto-me atraído pela noção que consiste em combinar a ideia de Vilenkin com o cenário da criação perspetivado por Atkatz e por mim. O espaço e o tempo criados a partir do nada absoluto pelo mecanismo de efeito túnel de Vilenkin são precisamente o espaço-tempo multidimensional, que é o ponto de partida por nós sugerido. Uma vez que este espaço multidimensional exista, então, dado que é instável, passa, por efeito túnel, para um universo a três dimensões espaciais e uma temporal. Estas dimensões são inflacionadas até um tamanho imenso para se tornarem o espaço-tempo do nosso universo físico.

Stephen Hawking, da Universidade de Cambridge, e James Hartle, da Universidade da Califórnia, em Santa Bárbara, levaram estas ideias mais longe. Num artigo de 1984, intitulado «A função de onda do universo», abordaram diretamente a questão da perturbadora singularidade no princípio dos tempos. Tais singularidades, para as quais as quantidades físicas, como a densidade de matéria, se tornam infinitas, aparecem muitas vezes

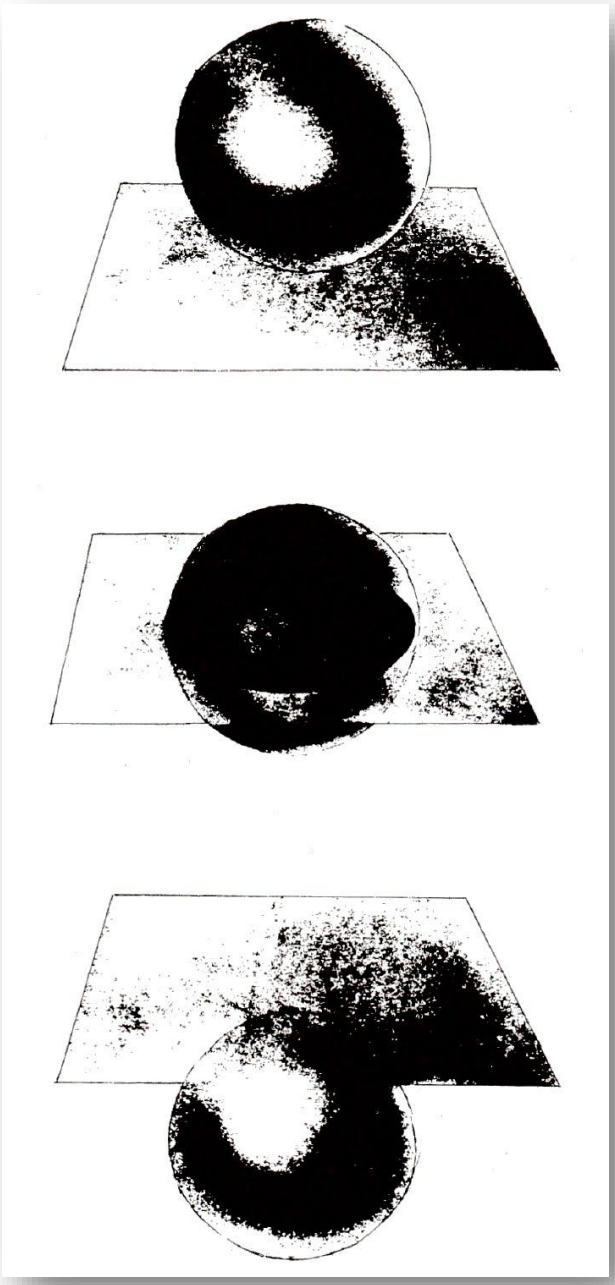
como soluções das equações da física clássica. Por exemplo, a equação clássica de um eletrão carregado negativamente movendo-se no campo elétrico do protão (positivamente carregado) num átomo de hidrogénio implica que, se o eletrão se precipitar no protão, libertar-se-á uma quantidade infinita de energia. No entanto, quando o átomo de hidrogénio é examinado à luz da teoria quântica, não se encontra semelhante singularidade infinita. De acordo com a teoria quântica, nenhuma infinidade realmente aparece porque o eletrão tem probabilidade finita de se precipitar sobre o protão. A singularidade é, assim, um artifício de uma descrição clássica e, se o problema for tratado corretamente por intermédio da teoria quântica (na qual o átomo de hidrogénio é descrito por uma função de onda de probabilidade de Schrödinger), a singularidade não está presente.

Segundo a teoria quântica, todos os corpos materiais têm uma função de onda de probabilidade de Schrödinger associada, que descreve o respetivo estado com precisão. Para a maioria dos objetos macroscópicos, como mesas e cadeiras, tal descrição, em termos de onda de probabilidade, não é muito esclarecedora, porque os objetos grandes podem ser descritos adequadamente pela física clássica. Para os objetos microscópicos, como átomos ou eletrões, a descrição ondulatória torna-se necessária, porque as propriedades quânticas de corpos tão pequenos são importantes. Normalmente, não pensamos o universo como um objeto microscópico para o qual as propriedades quânticas sejam importantes. Todavia, se recuarmos no tempo, o universo contrai-se até ser, ele também, um objeto microscópico, para o qual os efeitos quânticos são importantes. Para descrever o universo é

necessário calcular a função de onda do universo da mesma forma que vulgarmente calculamos a função de onda de um eletrão. E, tal como a função de onda de um eletrão especifica a probabilidade de encontrar um eletrão, também a função de onda do universo especifica a probabilidade de criar um universo.

Hawking e Hartle propõem uma definição para a função de onda do universo, em particular a função de onda específica, quando está no «estado fundamental», essencialmente o estado de vácuo. Se esta descrição ondulatória do universo inteiro fizer sentido, então mostraram que a singularidade no princípio dos tempos desaparece, tal como a descrição de uma função de onda remove a singularidade no átomo de hidrogénio. Calcularam, em seguida, a probabilidade de o universo emergir de um estado de «nada», como no modelo de Vilenkin, para um estado de «alguma coisa». Hawking concluiu um artigo posterior, datado de 1984, a que deu o título «O estado quântico do universo», com a seguinte observação: «Pode muito bem ser, portanto, que o universo observado deva a sua existência a efeitos gravitacionais quânticos.»

Embora especulativa, como todas as ideias acerca da origem do universo, a análise de Hawking-Hartle sugere que a origem do universo pode ser tratada como qualquer outro evento quântico. Trata-se de uma invulgar aplicação da teoria quântica, uma vez que envolve a criação do espaço e do tempo, em vez da criação de algumas partículas quânticas, como sucede no decaimento de um núcleo. Contudo, em princípio, não existe razão para que a origem do universo não possa ser sujeita à análise racional e ao cálculo no âmbito da teoria quântica.



Uma imagem da criação do espaço e do tempo — aqui o espaço de um «universo em anel» — a partir do nada. O espaço-tempo do universo é a intersecção da superfície esférica com o plano. Começa como «nada». Torna-se um ponto quando a esfera toca o plano e no final contrai-se novamente até «nada».

O nada «antes» da criação do universo é o vazio mais completo que se possa imaginar — nem espaço, nem tempo, nem matéria existem. É um mundo sem lugar, sem duração ou eternidade, sem número — é aquilo a que os matemáticos chamam «o conjunto vazio». Porém, este vazio impensável converte-se em existência

plena — uma consequência necessária das leis físicas. Onde estão essas leis escritas nesse vazio? Que é que «diz» ao vazio que está prenhe com um universo possível? Parece que até o vazio está sujeito à lei, a uma lógica que existe antes do espaço e do tempo.

É fácil criticar cada uma destas ideias acerca da origem do universo com base em argumentos técnicos. Há até quem objete contra toda esta investigação com o argumento de que, se a origem do universo é um evento não observável, situa-se fora do domínio da ciência, mas que, contudo, teria também objetado contra o cálculo da abundância relativa dos elementos no *big bang*, pois este não é observável. Objeções destas têm pouco mérito.

Outros críticos mais sérios apontam o facto de todos estes modelos envolverem a generalização das teorias e conceitos correntes bem para além das condições em que foram testados, não passando, portanto, de palpites imaginativos. Podem muito bem ter razão. Toda uma comunidade de cientistas muito inteligentes pode ter sido levada até uma teoria do universo primitivo que no futuro (com a vantagem da posterioridade) seria vista como uma fantasia baseada em informações incompletas e extrapolações imaginativas. A construção das teorias, conquanto crie um quadro para o pensamento, nunca é substituta da experiência e da observação. Os novos aceleradores de altas energias e os telescópios atualmente em fase de projeto irão dizer-nos muito sobre se estas ideias estão corretas ou não.

Por vezes, ocorre-me desejar que este livro acerca das ideias atuais da física e da cosmologia pudesse ser publicado como um

bloco de apontamentos com folhas que se pudessem soltar. Desta maneira, algumas páginas poderiam ser arrancadas e substituídas por outras que descrevessem melhores ideias logo que estas aparecessem. Muito do nosso pensamento científico corrente acerca da física microscópica, as «ideias loucas» e a cosmologia estão, provavelmente, errados e terão de ser abandonados. No futuro talvez se dê uma grande revolução na física que modifique toda a nossa ideia da realidade. Recordaremos então as tentativas correntes para compreender a origem do universo como desesperançosamente inadequadas, como as tentativas dos filósofos medievais para tentarem compreender o sistema solar antes das revelações de Copérnico, Kepler, Galileu e Newton. O que hoje olhamos como «a origem do universo» pode ser o limiar temporal de mundos para lá da nossa imaginação. Mas também é possível que estejamos próximos do fim da nossa busca. Ninguém sabe.

Os homens adoraram outrora o Sol, maravilhados pelo seu poder e beleza. Agora que os astrofísicos compreendem a física do Sol e das estrelas e a fonte da sua energia, não são mais os mistérios que dantes eram. Na nossa cultura já não adoramos o Sol, nem o vemos como uma presença divina, como os nossos antepassados fizeram. Mas muitas pessoas, nossas contemporâneas, associam ainda os seus sentimentos mais profundos ao universo como um todo e veem a sua origem como algo de misterioso. O tamanho, o esplendor e a glória do universo provocam ainda o sentimento de um ser transcendente e eterno.

Algum dia (e esse dia ainda não está próximo) a origem física e a dinâmica de todo o universo serão tão bem compreendidas como hoje o são as estrelas. A existência do universo não guardará mais mistério para aqueles que escolherem compreendê-lo do que a existência do Sol. Stephen Hawking, que tanto contribuiu para a moderna compreensão do universo, espantou muitos colegas na palestra inaugural da Plumean Professorship, em Cambridge, intitulada «Estará à vista o fim da física teórica?», ao afirmar que os principais mistérios do universo serão esclarecidos dentro de algumas décadas.

Fazendo uma comparação que reflita o meu próprio otimismo, talvez os cientistas estejam agora num estádio de compreensão do universo semelhante àquele em que Eddington se encontrava da compreensão das estrelas quando escreveu *The Internal Constitution of Stars* na década de 20. Com grande poder conceptual, Eddington aplicou as leis da física então conhecidas à compreensão das estrelas e sugeriu até que a energia nuclear (na altura não compreendida ainda) era a fonte da sua energia. Tinha razão nas linhas gerais e até em muitos pormenores, mas a moderna teoria das estrelas tinha de esperar pelo desenvolvimento da experimentação e da teoria nucleares. Se tal comparação for válida, então é possível que já compreendamos, em linhas gerais, a própria origem do universo.

Muitos cientistas razoáveis discordarão desta apreciação otimista. Mas, pondo de lado as nossas opiniões pessoais relativamente à questão de quão perto estamos de atingir o nosso

objetivo de compreender o universo, todos os cientistas concordam que o importante é continuar a trabalhar nesse sentido.

A nossa crescente familiaridade com estas novas ideias não deve impedir-nos de constatar quão estranhas elas eram até há uns anos atrás. Quando os historiadores da ciência examinarem as décadas de 70 e 80 registarão que pela primeira vez os cientistas construíram modelos matemáticos racionais baseados nas leis da física que descrevem a criação do universo a partir do nada, o que assinala o início de uma nova perspetiva da criação de tudo o que existe.

---

[1] O autor faleceu em 23 de Julho de 1988. (*N. do T.*)

PARTE QUATRO

Reflexões

Na altura em que Dendid criou todas as coisas,

criou o Sol,

e o Sol nasce, e morre, e volta de novo;

criou a Lua,

e a Lua nasce, e morre, e volta de novo;

criou as estrelas,

e as estrelas nascem, e morrem, e voltam de novo;

criou o homem,

e o homem nasce, e morre, e nunca mais volta.

CANÇÃO AFRICANA

CAPÍTULO 18

O Computador Cósmico

Existe um número infinito de universos possíveis e, como só um deles pode ser real, deve haver razão suficiente para a escolha de Deus, que O levou a decidir-se por um em detrimento dos outros.

G. W. LEIBNIZ, A Monadologia, 1714

Mesmo quando os cientistas fornecem uma explicação precisa e coerente do universo, podem sempre ser postas outras questões interpretativas acerca das suas descobertas. Os filósofos, poetas e escritores criaram metáforas que interpretam o universo, metáforas que muitas pessoas, incluindo cientistas, conservam no inconsciente ao refletirem sobre a realidade. Por exemplo, antes do nascimento da ciência moderna muitas pessoas viam o universo como um organismo ou uma revelação divina. Depois de Copérnico, de Kepler, de Galileu e de Newton a imagem do universo mudou radicalmente — era um grande relógio, posto inicialmente em movimento por mão divina, mas agora funcionando por si próprio e completamente determinado quanto ao futuro. Com a descoberta da teoria quântica e da natureza estatística dos acontecimentos atómicos, esta imagem de um relógio determinista tornou-se inadequada para descrever o universo. Estas descobertas trouxeram a necessidade de novas metáforas.

Uma característica importante da teoria quântica, que não é partilhada com a antiga teoria clássica, é que a informação que obtemos acerca do mundo por meio de medições depende de como decidimos obter essa informação — o método de medição. A teoria quântica coloca ênfase na informação, na sua representação e transformação. Uma vez que os computadores também transformam a informação, uma imagem interessante do universo quântico é a de que é um computador gigante — um sistema de processamento da informação.

Nesta metáfora do universo como computador cósmico os corpos materiais do universo, as partículas quânticas, são o *hardware*. As regras lógicas a que estas partículas obedecem, as leis da Natureza, são o *software*. A evolução do universo pode ser vista como a execução do «programa» especificado pelas leis da Natureza, embora não se trate de um programa determinista, como os dos computadores digitais. Resta determinar qual será a saída última deste computador cósmico. Mas sabemos já que o programa deu origem a «sub-rotinas» complexas, que podemos identificar com a vida, as quais são tão complicadas que parecem ter vida própria, independentemente do computador cósmico.

É divertido explorar esta metáfora do universo como um computador e examinar qual a relação deste computador cósmico com outro computador, a máquina «mole» que jaz na nossa cabeça, o cérebro, que não é certamente um computador digital (nem o universo o é), mas um órgão que transforma informação. Os nossos computadores cerebrais são parte do computador cósmico e estão também a tentar compreender a sua estrutura última. A dificuldade

que experimentamos ao fazê-lo pode ser vista como um «problema de interface», de ligação dos computadores um com o outro. Facto notável é que, a despeito deste problema de interface, parecemos ser capazes de compreender o universo.

Inicialmente, isto parece paradoxal: como pode a parte compreender o todo? Mas o cérebro é um computador muito mais complexo do que o universo macroscópico. E é certamente possível a um computador complexo construir um modelo matemático das suas partes mais simples e das regras a que obedecem. Não existem problemas de autorreferência nem paradoxos até que um computador tente examinar os próprios programas internos — e não é isso o que estamos a fazer quando examinamos o *hardware* do universo?

Uma razão pela qual poderemos talvez aceder à completa compreensão do universo — conhecer os seus *hardware* e *software* básicos — é que o universo pode ser macroscópica e microscopicamente finito. Um tema recorrente na história da física é que, sempre que aparecem infinitos nas nossas descrições matemáticas das grandezas físicas, isso é um sinal seguro de que ainda não compreendemos o que se está a passar. Quando se alcança um conhecimento melhor, os infinitos desaparecem. A Natureza evita os infinitos.

Um exemplo é a descoberta de Max Planck da quantificação da troca de energia. A lei clássica da distribuição dos comprimentos de onda da luz emitida por um carvão ao rubro implicava que a energia total de radiação fosse infinita — uma conclusão absurda. Planck

modificou a lei clássica, admitindo que a troca de energia radioativa era quantizada — vinha em pequenos pacotes. Isto removia o infinito absurdo, condizia com a distribuição dos comprimentos de onda observados e abria a porta à subsequente teoria quântica.

Não há muito tempo, muitas pessoas acreditavam no modelo estático do universo, que obrigava a que ele fosse infinito, possivelmente no espaço e certamente no tempo. Mas este modelo não é apoiado pela observação astronómica que, em vez disso, favorece o modelo do *big bang*, no qual até o tempo é finito e teve um «começo». O nosso universo pode ser finito; pelo menos, não existe grande prova em contrário.

Se o universo, o nosso computador cósmico, é finito, embora extremamente grande, então a nossa cosmologia começa a ter semelhanças com a conceção medieval do mundo como um lugar circunscrito.

Ironicamente, apesar das enormes descobertas científicas, o que estaremos talvez a apreender serão as condições, em última análise finitas e limitadoras, da existência do nosso universo. Contudo, desta vez não será uma ideia provinciana de Deus e do universo a impor restrições ao que é possível; serão as próprias leis da Natureza.

Se os cientistas alcançarem a curto prazo uma compreensão das leis básicas da criação e da evolução posterior do universo, então, como sublinhou Stephen Hawking na sua palestra Plumean, estará

à vista o fim da física teórica. Muitos físicos discordam desta opinião e apontam a sua semelhança com a sustentada por cientistas clássicos eminentes no fim do século XIX, os quais pensaram igualmente ter a física chegado ao fim. Estes físicos creem que a física é inesgotável e que existem incontáveis estruturas e novas leis que estão para ser descobertas. Quem tem razão?

Ninguém pode dizê-lo ao certo. Todavia, é certo que a resposta a esta questão não virá das especulações teóricas, por mais atraentes que pareçam. A única pedra de toque para a verdade empírica é a experiência e a observação. Algumas experiências, ou novas observações, poderiam destruir a imagem do universo desenvolvida até agora, ou poderiam, por outro lado, conferir-lhe mais solidez.

Suponhamos, para efeitos de discussão, que os cientistas descobrem a lei mestra do universo, o *software* básico do computador cósmico. Então, embora a física, tal como a conhecemos, desaparecesse, restariam, pelo menos, duas fronteiras de pesquisa. Uma é a «fronteira da complexidade» — a fronteira do conhecimento acerca dos processos complexos pelos quais a matéria se organiza em formas vivas e não vivas. Uma coisa é conhecer as leis básicas da física, outra é deduzir as consequências complexas dessas leis.

A outra fronteira pode ser chamada «a fronteira da simplicidade». A lei mestra da física será, provavelmente, bastante simples (de outra maneira não poderemos encontrá-la). Será difícil resistir à tendência intelectual de tentar compreender por que é que

essa lei particular se aplica, e não outra qualquer, igualmente simples. Aqui a metáfora do computador cósmico pode ajudar-nos. Encarando as leis do universo como *software*, prevejo a possível futura fusão da física com a ciência da informação, um ramo da matemática. A ciência da informação poderá talvez dizer-nos se os nossos cérebros e o computador cósmico estão ligados de tal maneira que só uma lei seja consistente com a nossa capacidade de compreender a ordem aparente do universo.

Existem enigmas levantados pela imagem do universo como um computador. As leis da Natureza, como a lei da gravitação de Newton, estão normalmente representadas no *hardware* do universo, como sucede com o sistema solar. As leis da física quântica estão mesmo representadas no espaço plano «vazio» pelas partículas e antipartículas quânticas que, ora existem, ora deixam de existir. Mas suponhamos que o universo começa do nada, esse completo vazio que anteriormente apresentei e no qual nem o espaço nem o tempo existem. No vazio completo não existe *hardware*. Sem *hardware* não pode haver *software*. Como estão as leis do universo representadas nesse vazio? O que é que «diz» ao vazio para executar o programa de se converter em existência plena?

De acordo com os modelos matemáticos correntes da origem do universo a partir do nada, existe qualquer coisa presente nesse vazio, uma probabilidade descrita pela «função de onda do universo». A teoria quântica impõe que para todas as entidades físicas, quer sejam um átomo, quer uma bola, exista uma «função de onda» associada que especifica a probabilidade dos resultados

das medições levadas a cabo sobre essa entidade. Uma importante propriedade desta descrição, usando a «função de onda», é que devemos separar conceptualmente (e também na prática) a entidade física — o «observado» — do instrumento de medição — o «observador».

A função de onda não é, de maneira alguma, um corpo material; apenas especifica as probabilidades dos acontecimentos materiais. Pode estabelecer-se uma analogia com o lançamento de dados — um evento material —, cujos resultados podem ser descritos por uma distribuição de probabilidades. A distribuição probabilística é como uma «mão invisível» que parece influenciar os eventos materiais.

Inicialmente parece desnecessário considerar a «função de onda do universo», porque o universo é tão grande que é válida uma descrição clássica com certezas, em vez de probabilidades. Contudo, se recuarmos no tempo até ao *big bang*, o universo contrai-se até às dimensões de uma partícula quântica e então a descrição ondulatória é importante. Mas onde está a separação entre o observador e o observado? O universo inclui tudo, até mesmo o observador; não pode haver tal distinção.

Este problema sugere a alguns físicos que a ideia usual de uma função de onda não é aplicável ao universo inteiro. O físico John Wheeler é da opinião de que a evolução dos observadores posteriores devida ao ato de observar o universo cria uma «textura de significado» que se torna o próprio universo — um cenário a que chama o «universo participante». Outros físicos, como Stephen

Hawking, James Hartle e Alex Vilenkin, pensam que não existe nenhum problema de fundo e tratam de inventar prescrições matemáticas para calcular a função de onda do universo.

Se aceitarmos a ideia de que o universo quântico é descrito por uma tal função de onda, então ela especifica a probabilidade de criar um universo de espaço-tempo e matéria a partir do nada. Num certo sentido, até o nada absoluto está prenhe do pleno: existe uma probabilidade de o nada se tornar alguma coisa. Mas o que é «alguma coisa»? Que espécie de universo se cria? Esta questão acerca da origem do universo também inquietou Einstein, que se interrogava se Deus teria tido alguma alternativa ao criar o universo do modo que o fez. Einstein tinha alguma coisa em mente quando fez esta pergunta: se as leis da Natureza deixariam, no fim, os valores de algumas constantes físicas por especificar. Se todas as constantes fundamentais estiverem especificadas, então Deus não teve alternativa. Se algumas constantes não estão, porém, especificadas, então podem ser criados outros universos possíveis — o computador cósmico pode executar outros programas.

A possibilidade de existirem outros universos sempre intrigou as pessoas que se interrogavam sobre o porquê da existência do nosso universo. Era claramente esta a ideia do filósofo G. W. Leibniz quando em 1714 escreveu na *Monadologia*:

> Existe um número infinito de universos possíveis e, como só um deles pode ser real, deve haver razão suficiente para a escolha de Deus, que O levou a decidir-se por um em detrimento dos outros. E esta razão pode ser encontrada

somente na conveniência ou no grau de perfeição que estes mundos possuem.

As implicações do raciocínio de Leibniz não escaparam aos seus contemporâneos. O filósofo satírico Voltaire criou em *Candide* a personagem do Dr. Pangloss, que mais não era do que um Leibniz mal disfarçado. O Dr. Pangloss, sempre otimista, reagia aos constantes desastres que lhe aconteciam e aos seus companheiros lembrando a toda a gente que «este é o melhor dos mundos possíveis» — fraco conforto para as vítimas.

Leibniz pensava que o nosso universo era especial porque era «perfeito». Hoje alguns cosmólogos e físicos pensam que o nosso universo é especial porque contém a vida. Estes cientistas pensam que outros universos possíveis poderiam ser inóspitos para a vida e, portanto, não teriam filósofos, cosmólogos e físicos para sequer pensar por que é que existiam. O nosso universo, pelo contrário, acolheu os seres que podem pensar acerca do universo dentro do qual nasceram e interrogar-se acerca das propriedades deste. Parece que o atributo da vida seleciona do conjunto de todos os universos possíveis um pequeno subconjunto (talvez um só) que pode observar-se a si próprio.

Para desenvolverem tal argumento estes cientistas supõem que os outros universos são caracterizados por diferentes valores das constantes físicas fundamentais, como a constante da gravitação ou o valor da carga do eletrão. Isto influenciará as estruturas físicas desses universos imaginários, estruturas que eventualmente não permitirão a evolução da vida. O nosso universo real pode,

portanto, ser obrigado a não ser muito diferente do que é, porque, de outra maneira, simplesmente, não estaríamos aqui para o observar. Nas palavras do cosmólogo britânico John D. Barrow, «as observações dos parâmetros cosmológicos feitas pelos astrónomos são as vítimas de um efeito de seleção que tudo envolve — a nossa própria existência». A esta ideia de que a existência da vida introduz um critério seletivo, uma «razão suficiente» para universos fisicamente possíveis, foi dado um nome, o «princípio antrópico».

O princípio antrópico parece-se menos com um princípio físico e mais com um princípio biológico, lembrando o princípio da seleção natural de Darwin, aplicado aqui a todo o universo. Vinte anos atrás, quando era estudante graduado em física e pela primeira vez ouvi invocar este raciocínio antrópico para explicar por que é que o nosso universo é como é, fiquei intrigado. Aqui estava uma forma de raciocínio completamente estranha ao caminho usual que os físicos teóricos seguem na sua tarefa de pesquisar as leis matemáticas da Natureza. E este princípio seletivo pode, no entanto, fazer luz sobre essas mesmas leis.

Mas, quanto mais pensava sobre o princípio antrópico, menos me parecia que se tratasse de um grande princípio seletivo darwinista e mais ele se me afigurava como uma explicação rebuscada para aquelas características do universo que os físicos ainda não sabem explicar. Os físicos e os cosmólogos que apelam para o raciocínio antrópico parecem estar a abandonar gratuitamente o programa feliz da ciência física convencional, que consiste em compreender as propriedades quantitativas do nosso universo com base em leis físicas universais. Talvez que ao

tentarem encontrar uma descrição completa e quantitativa dos parâmetros cósmicos que caracterizam o nosso universo real se tenham deixado vencer pelo desespero e pela frustração.

É claro que existem cientistas eminentes que não partilham a minha opinião negativa acerca do princípio antrópico. Poderíamos debater os seus méritos e deméritos durante muito tempo. Mas tal debate interminável é um sintoma do que está mal com o princípio antrópico que, ao contrário dos princípios da física, não oferece nenhum caminho para se saber se está certo ou errado, não há maneira de poder testá-lo. Ao contrário dos princípios físicos convencionais, o princípio antrópico não está sujeito a refutação experimental — o sinal seguro de que não é um princípio científico. Não é possível qualquer conclusão empírica acerca da sua veracidade e um debate para averiguar se é verdadeiro ou falso poderia durar para sempre.

A maioria dos físicos e dos astrofísicos ativos em investigação, simplesmente, não prestam atenção ao princípio antrópico. Esforçam-se por compreender as propriedades básicas da matéria e do universo em termos dos princípios físicos expressos na linguagem precisa da matemática, princípios estes que conduzem a conclusões testáveis. Esta tarefa, notavelmente, tem conhecido grandes êxitos.

Em contraste, a influência do princípio antrópico sobre o desenvolvimento dos modelos cosmológicos contemporâneos tem sido estéril: não explicou nada, teve até influência negativa, como foi evidenciado pelo facto de o valor de certas constantes, como a

razão das abundâncias de fotões e partículas nucleares, para a qual o raciocínio antrópico foi uma vez invocado como explicação, poder agora ser explicado por novas leis físicas. Se compararmos o progresso feito na compreensão das constantes fundamentais através de novas teorias físicas com o progresso feito usando o princípio antrópico, rapidamente verificaremos que não se ganhou qualquer conhecimento pela adoção do raciocínio antrópico. Eu optaria pela rejeição do princípio antrópico, por estabelecer uma confusão desnecessária no reportório conceptual da ciência.

A minha visão pessoal é que, embora ainda não conheçamos as leis fundamentais, a possibilidade da vida no universo, governada por aquelas leis, estará inscrita nelas quando e se as encontrarmos. A existência da vida no universo não é um princípio seletivo que atue sobre as leis da Natureza, é antes uma consequência delas. A vida é parte do programa executado pelo computador cósmico e, tanto quanto sabemos, é a sua mais complexa sub-rotina. Os cientistas da informação desenham hoje computadores com sub-rotinas capazes de alterar o programa do próprio computador. Será possível que a vida, ou o que quer que ela se torne, possa alterar o programa do computador cósmico, mudando o curso do seu destino?

Será preciso algo mais do que uma metáfora para responder a esta importante questão; será necessária uma compreensão mais profunda da vida e do cosmos do que aquela que hoje temos. No entanto, o desejo de conhecer a resposta a tais interrogações acerca do nosso destino nunca desaparecerá. E esse desejo é talvez o programa mais profundo do nosso computador cósmico.

CAPÍTULO 19

Ciência na Primeira Pessoa

> Posso viver com a dúvida e a incerteza. Penso que é muito mais interessante viver não sabendo do que ter respostas que podem estar erradas.
>
> *RICHARD FEYNMAN, 1981*

Isaac Newton, o descobridor da mecânica clássica, era um unitário — uma crença que manteve secreta enquanto viveu na cultura trinitária cristã da Universidade de Cambridge. Embora estivesse profundamente envolvido com ideias religiosas, manteve-as fora dos seus escritos científicos. Einstein, que sustentava opiniões elaboradas, se não pouco usuais, e que tinha a própria visão da impersonalidade do universo, sempre separou estas perspetivas da investigação científica. A separação entre crenças e opiniões privadas e o trabalho científico de cada um é mais do que simples profissionalismo — está na essência do empreendimento científico.

Os cientistas, individualmente, trazem muitas motivações pessoais para a condução das suas investigações. Podem defender conceções peculiares da realidade e filosofias políticas não convencionais. Podem roubar ideias e usar o talento científico como veículo para aumentar a sua influência na sociedade. No entanto, a intenção da investigação nas ciências naturais é descobrir a ordem

do mundo, as regras a que o universo obedece, e transcrever o que é descoberto, de maneira que a verdade possa ser restabelecida por todos os indivíduos competentes, independentemente da sua cultura, opção política, raça ou sexo. É uma notável propriedade do nosso universo que esta intenção possa ser concretizada.

À imagem do mundo que resulta de tal investigação científica gosto de chamar «ciência na terceira pessoa». Esta é a ciência que vemos publicada em revistas especializadas e que ouvimos ter sido exposta em conferências e seminários. É claro que esta ciência está «ligada à cultura» (especialmente na escolha de assuntos adequados à investigação) e está também sujeita à falibilidade dos cientistas que a fazem. Contudo, a sua intenção, ainda que imperfeitamente realizada, é que este conhecimento seja verdadeiro para todos nós. Se a ciência falhar nesse intento, falha igualmente no seu compromisso com a verdade, por provisório que seja.

A ciência na terceira pessoa mostra-nos a ordem material do mundo. Certamente não pedimos que essa ordem fosse da maneira que é e certamente não fomos nós que a inventámos. Não obstante, essa ordem está agora a ser revelada e é como é. É claro que a imagem do mundo da ciência está longe de estar completa — existe uma vasta fronteira desconhecida que hoje desafia os cientistas. Em apenas alguns séculos as descobertas científicas alteraram radicalmente a conceção do universo que as pessoas instruídas tinham. E, como as pessoas acreditam que estas descobertas são verdadeiras, não podem deixar de lhes reagir intelectual e emocionalmente. Os cientistas não são exceção. Einstein, cujo

trabalho tanto contribuiu para criar a nossa visão corrente do cosmos, expressava muitas vezes a sua opinião sobre como podíamos relacionar-nos com o universo. Escreveu uma vez:

> Um ser humano é parte de um todo a que chamamos «universo»; uma parte limitada no tempo e no espaço. Tem consciência de si mesmo, dos seus pensamentos e sentimentos, como qualquer coisa separada do resto — uma espécie de ilusão óptica da sua consciência. Esta ilusão é como que uma prisão para nós, restringindo-nos aos nossos desejos pessoais e à afeição por algumas pessoas mais próximas de nós. A nossa tarefa deve ser libertar-nos desta prisão pelo alargamento de todo o círculo de compaixão, de modo a abraçar todas as criaturas vivas e toda a Natureza na sua beleza. Ninguém é capaz de o conseguir completamente, mas a luta por tal objetivo é, em si mesma, uma parte da nossa libertação e um fundamento para a nossa segurança interior.

Isto é um exemplo daquilo a que chamaria «ciência na primeira pessoa» — os pensamentos pessoais de um indivíduo, interpretando e reagindo à realidade do mundo descoberta pela ciência. A particularidade da ciência na primeira pessoa é que ela reflete a sensibilidade pessoal e a consciência da pessoa que escreve. A intenção destas reflexões não é que elas sejam verdadeiras ou façam sentido para todas as pessoas, como na ciência na terceira pessoa, mas que possam ser verdadeiras apenas para a própria experiência de cada um de nós, uma experiência que desejamos partilhar. Nem a ciência na primeira pessoa nem a ciência na terceira pessoa podem ser substitutas uma da outra — a intenção da investigação e das reflexões é diferente em cada caso.

Nos seus escritos na «primeira pessoa» os cientistas não são de modo algum privilegiados. Os acidentes da sua história pessoal influenciam a sua experiência da realidade, tal como acontece com as outras pessoas. Gostaria de examinar agora alguns exemplos de ciência na primeira pessoa.

A descoberta recente da imensa idade e dimensões do universo e a visão da natureza efémera das nossas vidas individuais, que estas descobertas suscitam, fazem com que desperte nas pessoas uma variedade de atitudes — as de uma pessoa racional, de uma pessoa espiritual, do construtor pragmático e do explorador por divertimento. A maior parte de nós experimenta a realidade, não como uma ilusão, mas como qualquer coisa absoluta e concreta. A nossa experiência da realidade é apaixonada e a maior parte de nós projeta no universo as suas esperanças ou medos. Mas tais projeções da nossa imaginação impedem, de facto, que exploremos essa mesma realidade com a abertura essencial à descoberta.

Podemos ser tentados a tomar a atitude de *Homo sapiens* — a pessoa que segue unicamente a razão. Aqueles que seguem a inexorável lógica da razão concluem que, uma vez que no fim toda a vida será destruída, não veem qualquer saída; a esperança, insistem, é uma ilusão. Bertrand Russell, o lógico, filósofo e humanista inglês, exprimiu eloquentemente esta posição no seu livro de 1923 *A Free Man's Worship*:

> Que o homem é o produto de causas que não conheciam o fim a que se destinavam; que a sua origem, crescimento, esperanças e medos, amores e crenças não são mais do que

> o resultado de arranjos fortuitos de átomos; que nenhum fogo, nenhum heroísmo ou intensidade de pensamento ou de sentimento podem preservar a vida para lá do túmulo; que todos os esforços humanos ao longo dos tempos, toda a devoção, toda a inspiração, todo o brilho refulgente do génio humano, estão destinados à extinção na vasta morte do sistema solar; que todo o edifício das conquistas do homem será inevitavelmente enterrado sob os escombros de um universo em ruínas. Todas estas afirmações, se não indiscutíveis, são, porém, quase certas e nenhuma filosofia que as rejeite pode perdurar. Só dentro do andaime destas verdades, só sobre os firmes alicerces do desespero inabalável, pode a habitação da alma ser construída em segurança.

Muitos concordariam com isto, argumentando que a razão conduz unicamente a esta conclusão. Steven Weinberg, físico teórico laureado com o prémio Nobel, falou por muitos cientistas quando escreveu:

> Quanto mais sabemos acerca do universo, mais se torna evidente que ele não tem objetivo nem sentido.

Às escalas de tempo do nosso universo, todos os nossos atos parecem vazios. É como se nunca tivéssemos existido. Os atos simples da vida de todos os dias, o heroísmo do esforço, as ideias de ordem política, o sofrimento incompreensível dos indivíduos e das multidões, a esperança da religião, as descobertas e explorações da arte e da ciência, a loucura, a glória e alegria da nossa espécie estão destinados a um esquecimento tão completo que nem sequer o ato da sua destruição será lembrado. Esta não é a conclusão de uma religião pessimista, nem o desvario de um

filósofo infeliz, mas a única conclusão racional que emerge da nossa visão científica do cosmos.

Mas será que Russell tem razão? Não estou convencido disso. Em primeiro lugar, é logicamente possível que a nossa perspetiva atual do universo esteja seriamente errada ou extremamente limitada, o que já aconteceu antes e pode vir a acontecer de novo. Todas as conceções correntes do universo, como a de Russell, baseiam-se, em última análise, em modelos cosmológicos, que são, eles próprios, invenções do pensamento e imaginação humanos. Presentemente, nenhum modelo cosmológico corrente considera importante, por exemplo, o efeito da vida inteligente sobre o futuro do universo. Talvez a vida venha a influenciar o cosmos de uma maneira imprevista. Pode a vida existir indefinidamente? Ou está sujeita à extinção total? Ninguém sabe responder a esta pergunta com segurança com base no conhecimento atual.

No entanto, mesmo que Russell tenha razão — que toda a vida esteja condenada à aniquilação —, poderá daqui deduzir-se que podemos construir a nossa filosofia só «sobre os firmes alicerces do desespero inabalável»? Não vejo a lógica disto. Será que o objetivo da vida, se é que ela tem algum, é promover a sua continuidade eterna? Duvido. Julgo que Russell encontrou uma maneira de justificar o próprio desespero — e todos nós parecemos precisar, por vezes, de tal justificação —, projetando-o na sua previsão da morte última do universo.

Um risco que os racionalistas impenitentes correm é o de se tornarem, por vezes, propensos a atitudes espirituais ou

excêntricas à medida que envelhecem. Sou da opinião de que dentro de cada super-racionalista existe um espiritualista lutando para sair cá para fora. Muitas vezes, quando o espiritualista sai do armário, surge com uma forma bizarra e pouco usual, consequência do seu longo confinamento.

Tomemos o caso de Auguste Comte, o fundador do positivismo científico no século XIX. Comte, um super-racionalista segundo a tradição francesa, amava profundamente a governanta, que era também sua amante. Quando ela morreu, Comte ficou emocionalmente destruído. Criou um culto à volta dela e exigiu aos seus alunos que a adorassem. É claro que, na sua perspetiva, esta adoração era diferente da das religiões convencionais e era perfeitamente racional. Comte nunca percebeu o que lhe acontecera.

Aqueles que Russell afastaria por estarem iludidos poderiam responder que a morte — mesmo a morte do universo — não deve ser vista como um vazio sem sentido, mas como uma forma de transformação. Tais indivíduos acreditam numa consciência cósmica, ou Deus — a atitude do *Homo spiritualis*. Tal ser transcendente não pode ser examinado pelas ciências naturais e pode ser visto, quer positivamente, como o fundamento da existência (é muitas vezes o caso na teologia ocidental), quer negativamente, como o grande vazio (típico de muitas religiões orientais). Estas perspetivas espirituais do mundo são profundamente atraentes. Comparado com tais perspetivas, o naturalismo científico parece superficial e pobre.

A ubíqua necessidade humana de descobrir um fundamento último para a existência parece-me ser de origem orgânica ou biológica, como o impulso sexual. Se assim for, negar tais impulsos espirituais resulta apenas em distorções da nossa humanidade. As formas que tais impulsos espirituais tomam em diferentes pessoas podem, como o impulso sexual, ser superficiais ou profundas, bizarras ou belas, ou quase inexistentes.

O paleontologista e filósofo jesuíta belga Pierre Teilhard de Chardin exprime esse ponto de vista. Encarou a essência do universo, tal como Platão, como um ser vivo em processo de transformação. Nas suas visões místicas do futuro viu toda a vida e matéria interligadas naquilo a que chamou a «noosfera», convergindo para um ponto ómega, a realização final de toda a existência:

> O universo é um coletor e um conservador, não de energia mecânica, como supusemos, mas de pessoas. A toda a nossa volta, uma por uma, como uma exalação contínua, as «almas» libertam-se, levando para cima a sua carga incomunicável de consciência. Uma a uma, embora não isoladamente, pois para cada uma delas, pela própria natureza do ómega, só pode existir um ponto possível de emersão definitiva — o ponto no qual, sob a ação sintetizante da união personalizadora, a noosfera apresentará coletivamente o seu ponto de convergência — no fim do mundo.

Em 1924 concluiu o seu livro *O Futuro do Homem* com estas palavras:

> Como uma vasta maré, o Ser terá dominado o estremecimento de todos os seres. A extraordinária aventura do mundo terá terminado no seio de um oceano tranquilo, do qual cada gota estará, porém, ainda consciente de ser ela própria. O sonho de todos os místicos terá encontrado a própria e completa plenitude.

Estão estas visões místicas do futuro certas ou erradas? Será isso importante? Elas servem o fim de manter a esperança, permitindo àqueles que as têm continuar a trabalhar para a sobrevivência da espécie. As palavras de Teilhard podem mitigar o sentimento do nosso próprio sofrimento ao sugerir que ele tem um significado mais profundo do que normalmente imaginamos. Tal como Simone Veil fazia notar no seu *Caderno de Apontamentos*, «devemos identificar-nos com o próprio universo. Tudo o que é menos que o universo está sujeito ao sofrimento».

Outros há que não se envolvem em pensamentos elevados acerca do fim e destino do universo. A necessidade de criar condições para a sobrevivência humana — um emprego, uma família, uma comunidade, uma nação — arrasta-os para uma vida de atividade prática. A importância que dão aos problemas imediatos da vida reflete a perspetiva do *Homo faber* — construtor ou fabricante. Andrew Carnegie, o grande capitalista e filantropo americano que consolidou a indústria do aço nos Estados Unidos, expressava esta visão em *The Gospel of Wealth*[1] em 1990:

> Sobre as especulações quanto ao futuro [...] parece insensato falar muito. Penso que não temos nada a ver com aquilo que pode acontecer daqui a mil ou um milhão de anos; aliás, nenhum de nós sabe o que está para vir. Os

> nossos deveres situam-se no presente, nos nossos dias e na nossa geração, e mesmo estes são, por vezes, bastante difíceis de discernir. A raça trabalha lentamente, elevando-se degrau a degrau; tem mesmo de construir cada degrau antes de se colocar sobre ele, porque [...] [se] tenta saltar sobre o espaço vazio, em direção a um ideal qualquer, não subirá mais, caindo, em vez disso, nas profundezas. Não posso, portanto, considerar essas especulações senão uma perda de tempo — de valioso tempo —, que é imperativamente necessário para lidar com o novo degrau possível no caminho ascendente.

O objetivo da vida, como Carnegie o viu, é trabalhar para criar as condições para uma vida melhor. Muitos socialistas, apesar da sua oposição ao capitalismo, concordariam com a mensagem essencial de Carnegie: o que importa é melhorar a vida das pessoas que agora vivem e a vida daquelas que irão viver no futuro.

Em concordância com a sua atraente robustez e vitalidade, a atitude do *Homo faber* desdenha a especulação intelectual — o pragmatismo está no cerne desta filosofia. As crenças, na perspetiva das pessoas práticas, só são importantes na medida em que influenciam o comportamento individual e social. Estes pragmáticos estão desinteressados da questão da verdade última, que insistem ser provisória.

É o *Homo faber*, e não o *Homo sapiens* ou o *Homo spiritualis*, que move o mundo moderno. Alguns são, porém, incapazes de ignorar aquilo que o ativista desdenha — a possibilidade de toda esta atividade e edificação vir a ser totalmente destruída — e abraçam também o ponto de vista filosófico, expresso por Sócrates, de que «não vale a pena viver a vida não examinada».

Ainda outra atitude diferente em relação à realidade cósmica é a do *Homo ludens* — o que brinca. Muitas vezes encaramos as brincadeiras das crianças ou dos animais jovens como uma preparação para a vida adulta. Mas por que razão tantos de nós param de brincar à medida que nos tornamos mais velhos? A criança que existe em nós, sempre em exploração, é a fonte da nossa força criadora.

Uma característica que observei nos indivíduos altamente criativos é a sua libertação dos estereótipos sociais e conceptuais — criam a sua própria exploração da realidade e mesmo as suas formas próprias de cultura. Têm um espírito radicalmente aberto, e a brincadeira e o humor, que são dos mais elevados atributos da vida cultural, constituem a chave para esta abertura de espirito.

A atividade do *Homo ludens* não deve restringir-se à brincadeira das crianças, mas deve incluir também a brincadeira dos adultos, que muitas vezes se parece com trabalho. Os adultos podem jogar o jogo da vida com bastante seriedade, sem, no entanto, deixarem de o ver como um jogo. O desporto e o teatro são atividades do *Homo ludens* que se tornam absorventes ao ponto de se perder de vista o facto de se tratar de jogos e brincadeiras.

Alcançar o sentimento de liberdade nas brincadeiras não é, contudo, fácil; a aparente ausência de esforço é uma consequência do treino e da disciplina.

Richard Feynman, físico teórico do Cal Tech, tipifica muitas das características do *Homo ludens* com os seus embustes, gracejar permanente e respostas inesperadas. Este cientista trabalhava duramente nos problemas da física devido ao prazer que isso lhe trazia. Nunca se servindo de um conceito físico sem lhe imprimir uma marca pessoal, ofereceu-nos novas perspetivas inventivas. Numa entrevista, dada em 1981, Feynman recordou como voltara à investigação física depois de ter trabalhado em Los Alamos, no projeto da bomba atómica, durante a Segunda Guerra Mundial:

> Um dia pensei comigo próprio: ainda não fiz nada importante e nunca o farei. Mas costumava gostar das coisas da física e da matemática. Nunca era nada importante, mas fazia as coisas pela graça de as fazer. Então decidi: vou fazer coisas só pela graça que há em fazê-las.
>
> Nessa tarde, enquanto almoçava, um miúdo atirou um prato [ao ar] para a cafetaria. O prato tinha no centro um medalhão azul. O miúdo atirou o prato ao ar, que, ao cair, rodava e oscilava, e eu interroguei-me sobre qual seria a relação entre os dois movimentos. Para mim isto era uma brincadeira, não tinha a mínima importância. Então comecei também a brincar com as equações do movimento de objetos em rotação e descobri que, quando a oscilação era pequena, a rotação do desenho azul era duas vezes mais rápida do que a oscilação. Tentei então descobrir por que era assim, diretamente, a partir das leis de Newton, em vez de usar as equações complicadas, e trabalhei neste problema apenas para me divertir[2].
>
> Fui então procurar Hans Bethe e disse-lhe: «Olha, vou mostrar-te uma coisa divertida.» Expus-lhe o problema e ele retorquiu: «Sim, é muito divertido e interessante, mas para que serve?» Disse-lhe então: «Isso não interessa. Não tem qualquer utilidade. Estou apenas a fazê-lo por graça.» Continuei a brincar com esta rotação, que me levou ao problema semelhante da rotação do *spin* de um eletrão

> segundo a equação de Dirac, e isso fez-me voltar à eletrodinâmica quântica, o problema em que trabalhara. Continuei a brincar com o problema de maneira descontraída, como inicialmente, e foi quase como se tirasse a rolha a uma garrafa — tudo acabou por sair. Daí a muito pouco tempo fiz as descobertas pelas quais viria a ganhar o prémio Nobel.

A atitude brincalhona de Feynman em relação aos enigmas da física leva-nos a deduzir que ele tinha pouco interesse pelo misticismo ou por qualquer conceção rígida da realidade. Disse ele ainda:

> Se esperarmos que a ciência nos dê todas as respostas às maravilhosas interrogações sobre o que somos, para onde vamos, qual é o significado do universo, etc., então poderemos facilmente ficar desiludidos e procurar uma resposta mística. De que maneira pode um cientista aceitar uma resposta mística, não sei. Não posso acreditar nas histórias que têm sido elaboradas acerca da nossa relação com o universo, porque parecem ser demasiado simples, demasiado coerentes, demasiado provincianas. As pessoas perguntam-me se a ciência é verdadeira. Digo que não, não sabemos o que é verdadeiro. Tentamos descobrir, e tudo pode estar errado.
>
> Posso viver com a dúvida e a incerteza. Penso que é muito mais interessante viver não sabendo do que ter respostas que podem estar erradas. Tenho respostas aproximadas, crenças possíveis e diferentes graus de certeza acerca de várias coisas, mas não estou absolutamente certo de nada. Existem muitas coisas das quais nada sei, como, por exemplo, se tem algum significado perguntar por que estamos aqui. Mas não tenho de saber a resposta. Não me sinto assustado por não saber nada, por estar perdido num universo misterioso sem nenhum fim, que, tanto quanto

> posso dizer, é o que realmente se passa. Mas isto não me assusta.

Feynman, perpetuamente curioso, tentará quase tudo para explorar a realidade. Conta ele a seguinte história: encontrava-se dentro de um tanque de privação de sentidos e passou por uma experiência exossomática — sentiu que saía «para fora do corpo», que viu jazendo diante dele. Para testar a realidade da experiência tentou mexer o braço e, de facto, viu que o braço se movia no corpo. Ao descrever isto, disse ter ficado preocupado com o facto de poder ficar fora do próprio corpo e decidiu voltar para ele. Terminada a história, perguntei-lhe que conclusão tirava desta experiência estranha. Feynman respondeu, com a precisão observacional de um verdadeiro cientista: «Não vi *nenhuma* lei da física ser violada.» Na realidade, as descrições críveis de tais experiências que tenho lido, e também a minha própria, confirmam a observação de Feynman: as experiências «fora do corpo» não violam as leis físicas mais do que a experiência do sonho.

A intuição e a abertura de espírito de Feynman evocam a admiração de muitos cientistas criativos. Como o leitor, provavelmente, já adivinhou, as minhas simpatias vão mais para as atitudes do *Homo faber* e do *Homo ludens*, quanto mais não seja porque se libertaram da obsessão da certeza e da esperança. No entanto, todas as atitudes que até aqui considerámos — as do *Homo sapiens*, do *Homo spiritualis*, do *Homo faber* e do *Homo ludens*, bem como muitas outras — vivem dentro de cada um de nós. São vozes dirigindo o curso das nossas vidas, chamando-nos para percorrer caminhos diferentes.

As pessoas perguntam muitas vezes como podem articular os conceitos sobre a realidade, quer sejam místicos, quer pragmáticos, com as descobertas da ciência. A maior parte de nós tem uma forte tendência para unificar a nossa experiência da realidade numa única entidade — uma verdade universal para o universo — e passar parte das nossas vidas a procurá-la. O problema de procurar uma qualquer espécie de verdade última é que podemos pensar que a encontrámos — uma forma de certeza que nos fecha a porta a novas experiências acerca da realidade.

«Mas», pode alguém perguntar, «não procuram os cientistas a verdade e a certeza? Que dizer das descobertas da ciência — a existência da molécula genética do ADN[3], da estrutura dos átomos, das propriedades da luz? Não estão os cientistas certos da verdade destas descobertas? E que dizer acerca das verdades da matemática — não são elas absolutamente certas?» Podemos ter confiança na verdade provisória do conhecimento científico. Mas devemos distinguir cuidadosamente entre o conhecimento científico e aquelas perspetivas da realidade que, embora coerentes com o nosso conhecimento científico, representam extrapolações que vão para além dele. Não pode haver certeza universal sobre a verdade de qualquer visão do mundo, que não é capaz de ter, nem deve ter, um fundamento racional. Tais visões do mundo estão baseadas em crenças, que refletem a pluralidade das nossas experiências e a variedade das nossas necessidades. Separar o nosso conhecimento das nossas crenças, a nossa ciência da nossa fé, é uma tarefa que não devemos destruir pela simples fusão destas duas componentes com o propósito da unificação.

Talvez exista uma verdade última para o universo — não sei. No entanto, a suspensão de tais crenças abre-nos novos caminhos de exploração. Mais tarde podemos comparar os nossos novos conhecimentos e crenças com os antigos. Muitas vezes tais comparações envolvem contradições, mas estas, por seu turno, geram novas perspetivas criadoras acerca da ordem da realidade. A capacidade de tolerar a complexidade e aceitar a contradição, e não a necessidade da simplificação e da certeza, são os atributos do explorador.

Séculos atrás, quando algumas pessoas suspenderam as pesquisas sobre a verdade absoluta e começaram, em vez disso, a perguntar como funcionavam as coisas, nasceu a ciência moderna. Curiosamente, foi abandonando a busca da verdade absoluta que a ciência começou a fazer progressos, abrindo o universo material à exploração humana. Foi só por ser provisório e aberto à mudança, mesmo radical, que o conhecimento científico começou a evoluir. E, ironicamente, a sua vulnerabilidade à mudança é a fonte da sua força.

O universo assombra-me. Este sentido do oceano insondável e maravilhoso do que existe arrastou-me para a ciência. O universo infunde-me temor respeitoso, intriga-me e, por vezes, irrita-me com uma ordem natural que traz tanta dor e sofrimento. No entanto, qualquer emoção ou sentimento que tenha em relação ao cosmos parece-me obter, como resposta, não a benevolência, nem a hostilidade, mas somente o silêncio. O universo parece-me um ecrã perfeitamente neutro, no qual posso projetar qualquer paixão ou atitude com igual resultado.

E onde estou eu? Estou no presente, este momento imperfeito, tentando permanecer vulnerável à sua intensa especificidade. Não existe para mim outro tempo, nem outro lugar, nem qualquer consciência cósmica ou fácil misticismo, sob o qual me possa proteger. Para que este momento possa ser visto como o fulcro da existência, nenhum pormenor, nenhuma imperfeição, nenhum impedimento de culpa ou ressentimento, pode ficar por reconhecer. Sou a testemunha desta realidade — a racha na chávena, a dor nos meus membros, o som de fundo de uma voz, o amor que me rodeia, estes mesmos pensamentos e sentimentos que enchem a minha consciência. Nenhum espírito que seja curioso, nenhuma alma que queira libertar-se da necessidade da certeza, nenhum coração que esteja aberto a novas experiências, pode voltar as costas à terrível especificidade do mundo — o próprio fundamento da existência. Quando perguntaram a um estudante por que razão viajara de tão longe para estudar com um certo rabi, ele respondeu: «Não é para o ouvir falar sobre a sabedoria do Talmude que vim, mas para o ver atar os atacadores dos sapatos.» No amor do mundano, na abertura à exploração, no jogo da imaginação, na sublimação da agressão em atividade criativa, na necessidade de comunicar com e amar outras pessoas, reside a fonte de toda a grande poesia, arte e ciência e a minha esperança pessoal da libertação da espécie.

---

[1] Literalmente, *O Evangelho da Riqueza*. (*N. do T.*)

[2] A «oscilação» a que Feynman se refere é o movimento de nutação do prato. (*N. do T.*)

[3] Ácido desoxirribonucleico. (*N. do T.*)

Bibliografia

Extraída de abundante literatura, segue-se uma lista incompleta de livros que podem interessar ao leitor comum que deseje aprofundar os seus conhecimentos de cosmologia e astronomia. A maior parte destes livros, mas não todos, expressam o ponto de vista da cosmologia clássica anterior aos excitantes desenvolvimentos que agora relacionam a cosmologia com a teoria quântica dos campos. Contudo, têm aparecido artigos em revistas que descrevem o estudo do universo primitivo. Recomendo, em especial, os que apareceram em *Scientific American*, *American Scientist e Sky and Telescope* nos últimos dez anos, bem como artigos de revisão e relatos na *Science e na Nature*.

ABBOTT, E. A. — *Flatland*, Dover Publications, Nova Iorque, 1952.

ABELL, George — *Realm of the Universe*, 3.º ed., Saunders College Publishing, Filadélfia, 1984.

ABETTI, Giorgio — *The History of Astronomy*, Abeland-Schuman, Londres e Nova Iorque, 1952,

ALLEN, David A. — *Infrared*, *the New Astronomy*, Keith Reid, Ltd., Londres, 1975.

ALPHER, R. A., e HERMAN, R. — «Reflections of big bang cosmology», in *Cosmology, Fusion and Other Matters* — George Gamow Memorial Volume, ed. F. Reines, Colorado Associate University Press, Boulder, 1972.

ASIMOV, Isaac — *Eyes on the Universe*, Houghton Mifflin Co., Boston, 1975, The Collapsing Universe, Walker and Co., Nova Iorque, 1977 (O Colapso do Universo, Círculo de Leitores, Lisboa, 1982), e *The Universe from Flat Earth to Black Holes and Beyond*, Walker and Co., Nova Iorque, 1980.

BERENDZEN, R., HART, R., e SEELEY, D. —*Man Discovers the Galaxies*, Science History Publications, Nova Iorque, 1976.

BERGMAN, P. G. — *The Riddle of Gravitation* — From Newton to Einstein to Today's Exciting Theories, Charles Scribner's Sons, Nova Iorque, 1968.

BOK, Bort J., e BOK, Priscilla F. — *The Milky Way*, Harvard University Press, Cambridge, Mass., 1974.

BONDY, H. — *Cosmology*, Cambridge University Press, Cambridge, 1960, e *Relativity and Common Sense*, Doubleday, Anchor Books, Nova Iorque, 1964.

BOVA, Ben — *The New Astronomers*, St. Martin's Press, Nova Iorque, 1972.

BURGER, D. — Sphereland — A Fantasy about Curved Spaces and An Expanding Universe, Thomas Y. Crowell, Apollo, Nova Iorque, 1969.

CHAISON, Eric — *Cosmic Dawn*, Atlantic-Little, Brown, Boston, 1981.

CHARON, J. — *Cosmology — Theories of the Universe*, McGraw-Hill, Nova Torque, 1969.

DAVIES, P. C. W. — *Space and Time in the Modern Universe*, Cambridge University Press, Cambridge, 1977, *The Runaway Universe*, J. M. Dent, Londres, 1978, *The Forces of Nature*,

Cambridge University Press, Cambridge, 1979, e *Superforce — The Search for a Grand Unified Theory of Nature*, Simon and Schuster, Nova Iorque, 1984 (Superforça — Em Busca de Uma Teoria Unificada da Natureza, Gradiva, Lisboa, 1988).

DE VAUCOULEURS, Gerard — *Discovery of the Universe*, The McMillan Co., Nova Iorque, 1957.

DICKSON, F. B. — *The Bowl of Night — The Physical Universe and Scientific Thought*, Massachusetts Institute of Technology Press, Cambridge, Mass., 1968.

EDDINGTON, A. S. — *The Expanding Universe*, Cambridge University Press, Cambridge, 1933, reimp. na University of Michigan Press, Ann Arbor, 1958.

FEINBERG, Gerald — What's the World Made Of? Atoms, Leptons, Quarks and Other Tantalizing Particles, Doubleday, Anchor Books, Nova Iorque, 1977.

FERRIS, Timothy — *The Red Limit — The Search for the Edge of the Universe*, William Morrow and Co., Nova Iorque, 1977, e *Galaxies*, Sierra Club Books, São Francisco, 1980.

FRITZSCH, Harald — *Quarks*, Simon and Schuster, Nova Iorque, 1983. *Frontiers in Astronomy*, «Readings from Scientific American», introd. por Owen Gingerich, W. H. Freeman and Co., São Francisco, 1970.

GAMOW, G. — *The Birth and Death of the Sun*, Viking Press, Nova Iorque, 1952, *The Creation of the Universe*, Viking Press, Nova Iorque, 1952, e *Gravity*, Doubleday, Anchor Books, Nova Iorque, 1962.

GARDNER, M. — *The Relativity Explosion*, Random House, Nova Iorque, 1976.

GEROCH, Robert — *General Relativity from A to B*, University of Chicago Press, Chicago, 1978.

GINGERICH, O. (ed.) — *Cosmology + 1*, W. H. Ffeeman, São Francisco, 1977.

GLASBY, John S. — *Boundaries of the Universe*, Harvard University Press, Cambridge, Mass., 1971.

GREENSTEIN, George — *Frozen Star*, Freundlich Books, Nova Iorque, 1983.

GRIBBIN, John — *Our Changing Universe*, E. P. Dutton and Co., Nova lorque, 1976.

HARRISON, Edward — *Cosmology*, Cambridge University Press, Cambridge 1981.

HARWIT, Martin — *Cosmic Discovery*, Basic Books, Nova Iorque, 1981

HEY, J. Stanley — *The Evolution of Radio Astronomy*, Science History Publications, Nova Iorque, 1973.

HODGE, Paul — *The Revolution of Astronomy*, Holiday House, Nova Iorque, 1970.

HOSKIN, M. A. — *William Herschel and the Construction of the Heavens*, Science History Publications, Nova Iorque, 1963.

HOYKE, Fred — *Frontiers of Astronomy*, Mentor Books, Nova Iorque, 1955, *The Nature of the Universe*, Mentor Books, Nova Iorque, 1955, e *Galaxies, Nuclei and Quasars*, *Harper and Row*, Nova Iorque, 1965.

HUBBLE, E. — *The Realm of the Nebulae*, Yale University Press, New Haven, 1936, reimp. por Dover Publications, Nova Iorque, 1958.

INGLIS, Stuart J. — *Planets, Stars and Galaxies*, John Wiley and Sons, Nova Torque, 1976.

ISLAM, J. N. — *The Ultimate Fate of the Universe*, Cambridge University Press, Cambridge, 1983.

JAMMER, M. — *Concepts of Space — The History of the Theories of Space in Physics*, Harvard University Press, Cambridge, Mass., 1954, reimp. Por Harper and Row, Nova Iorque, 1960.

JASTROW, Robert — *The Exploration of Space*, McMillan and Co., Nova lorque, 1960, *Until the Sun Dies*, W. W. Norton and Co., Nova Iorque, 1977, e *Red Giants and White Dwarfs*, W. W. Norton and Co., Nova Iorque e Londres, 1979 (*A Arquitetura do Universo*, Edições 70, Lisboa, 1981 e 1989).

JEANS, J. H. — *Astronomy and Cosmology*, Cambridge University Press, Cambridge, 1929, e *The Mysterious Universe*, Cambridge University Press, Cambridge, 1930, reimp. por McMillan, Londres, 1937.

KAUFMANN, William J. — *Black Holes and Warped Spacetime*, W. H. Freeman, São Francisco, 1979.

KILMISTER, Clive — *The Nature of the Universe*, Thames and Hudson, Londres, 1971.

KING, H. C. — *Exploration of the Universe*, Signet, Nova Iorque, 1964.

KIPPENHAHN, Rudolf — *700 Billion Suns*, Basic Books, Nova Iorque, 1983.

LAURIE, John (ed.) — *Cosmology Now*, BBC, The Broadwater Press, Inglaterra, 1936.

LEMAÎTRE, G. — *The Primeval Atom*, Van Nostrand, Nova Iorque, 1951.

MENZEL, Donald H., WHIPPLE, Fred L., e DE VAUCOULEURS, Gerard — *Survey of the Universe*, Prentice-Hall, Englewood Cliffs, N. J., 1970.

MITTON, S. — *Exploring the Galaxies*, Charles Scribner's Sons, Nova Iorque, 1976.

MORRISON, Philip — *Powers of Ten*, W. H. Freeman, São Francisco, 1983

MORTZ, L. — *The Universe — Its Beginning and End*, Charles Scribner's Sons Nova Iorque, 1975.

NORTH, J. D. — *The Measure of the Universe — A History of Modern Cosmology*, Oxford University Press, Clarendon Press, Oxford, 1965.

PAGE, Thornton, e PAGE, Lou Williams — *Beyond the Milky Way*, The McMillan Co., Collier-McMillan Canada, Ltd., Toronto, 1969.

PAGE, Thornton, e PAGE, Lou Williams (ed.) — *Starlight*, The McMillan Co., Nova Iorque, 1967, e *The Evolution of Stars,* The McMillan Co., Nova lorque, 1968.

PAGELS, Heinz — *The Cosmic Code*, Bantam Books, Nova Iorque, 1983 (*O Código Cósmico*, Gradiva, Lisboa, 1986).

PAIS, Abraham — *Subtle is the Lord...*, Oxford University Press, Nova lorque, 1982.

«Report of the panels», in Astronomy and Astrophysics for the 1980's, vol. 2, National Academy Press, Washington, D. C., 1983.

ROMAN, Colin A. — *The Astronomers*, Evans Brothers, Londres, 1964.

ROOD, Robert T., e TREFIL, James S. — *Are We Alone?*, Charles Scribners's Sons, Nova Iorque, 1981.

RUCKER, R. B. — *Geometry, Relativity and the Fourth Dimension*, Dover Publications, Nova Iorque, 1977.

SASLAW, W., e JACOBS, K. (eds.) — *The Emerging Universe — Essays on Contemporary Astronomy*, Cambridge University Press, Cambridge, 1971.

SCHATZMAN, E. — *The Structure of the Universe*, McGraw-Hill Book Co., Nova Iorque, 1968.

SCHRAMM, David N., e WAGONER, Robert V. — *«What can deuterium tell us?», in Physics Today*, Dezembro de 1974.

SCHRAMM, David N,., e TURNER, Michael S. — «*Cosmology and elementary particle physics*», *in Physics Today*, Setembro de 1979.

SCIAMA, D. W. — *The Unity of the Universe,* Faber and Faber, Londres, 1959, e Modern Cosmology, Cambridge University Press, Cambridge, 1971.

SHAPLEY, Harlow — Of Stars and Men — The Human Response to An Expanding Universe, Beacon Press, Boston, 1958, View from a Distant Star — Man's Future in the Universe, Basic Books, Nova Iorque, 1963, Through Rugged Ways to the Stars, Charles Scribner's Sons, Nova Iorque, 1969, e Galaxies, rev. por Paul W. Hodge, Harward University Press, Cambridge, Mass., 1972.

SHIPMAN, Harry L. — *Black Holes, Quasars and the Universe*, Houghton Mifflin Co., Boston, 1980.

SHKLOVSKI, I. S., e, SAGAN, C. — *Intelligent Life in the Universe*, Holden-Day, Nova Iorque, 1966 (*A Vida Inteligente no Universo*, Publicações Europa-América, Mem Martins, 1985).

SILK, Jospeh – «*The evolution of the universe – A layperson's guide to the big bang*», in Griffith Observer, Julho de 1955, e The Big Bang, W. H. Freeman, São Francisco, 1979.

SILK, Joseph, e BARROW, John D. — *The Left Hand of Creation*, Basic Books, Nova Iorque, 1983 (*A Mão Esquerda da Criação*, Gradiva, Lisboa, 1989).

STRUVE, O., e ZEBERGS, V. — *Astronomy in the 20th Century*, McMillan, Nova Iorque, 1962.

SULLIVAN, Walter — *We Are not Alone*, McGraw-Hill Book Co., Nova Iorque, 1964, e *Black Holes*, Warner Books, Nova Iorque, 1980.

TAYLOR, E. F., e WHEELER, J. A. — *Spacetime Physics*, W. H. Freeman, São Francisco, 1966.

TOULMIN, S., e GOODFIELD, J. — The Fabric of the Heavens — The Development of Astronomy and Dynamics, Harper and Row, Nova Iorque, 1961.

TREFIL, James — From Atoms to Quarks — An Introduction to the Strange World of Particle Physics, Charles Scribner's Sons, Nova Iorque, 1980, e The Moment of Creation — Big Bang Physics from Before the First Millisecond to the Present Universe, Charles Scribner's Sons, Nova Iorque, 1983.

VERSCHUR, Guerrit L. — *The Invisible Universe*, The English Universities Press, Londres, 1974.

WAERDEN, B. L. van der — *Science Awakening*, Noordhoft, Groningen, Holanda, 1954.

WEINBERG, Steven — The First Three Minutes — A Modern View of the Origin of the Universe, Basic Books, Nova Iorque, 1977 (Os Três Primeiros Minutos, Gradiva, Lisboa, 1988).

WHITNEY, Charles A. — *The Discovery of Our Galaxy*, Alfred A. Knopf, Nova Iorque, 1971.

WHITROW, G. T. — *The Structure and Evolution of the Universe*, Hutchinson, Londres, 1959.

WIGNER, E. F. — *Symmetries and Reflections*, Indiana University Press, Bloomington, 1967.

WOOD, Harley — Unveiling the Universe — The Aims and Achievements of Astronomy, Angus and Robertson, Londres, 1967.

WRIGHT, H. — Explorer of the Universe — A Biography of George Ellery Hale, E. P. Dutton, Nova Iorque, 1966.

WRIGHT, H., WARNOW, J., é WEINER, C. (eds.) — *The Legacy of George Ellery Hale*, The MIT Press, Cambridge, Mass., 1972.