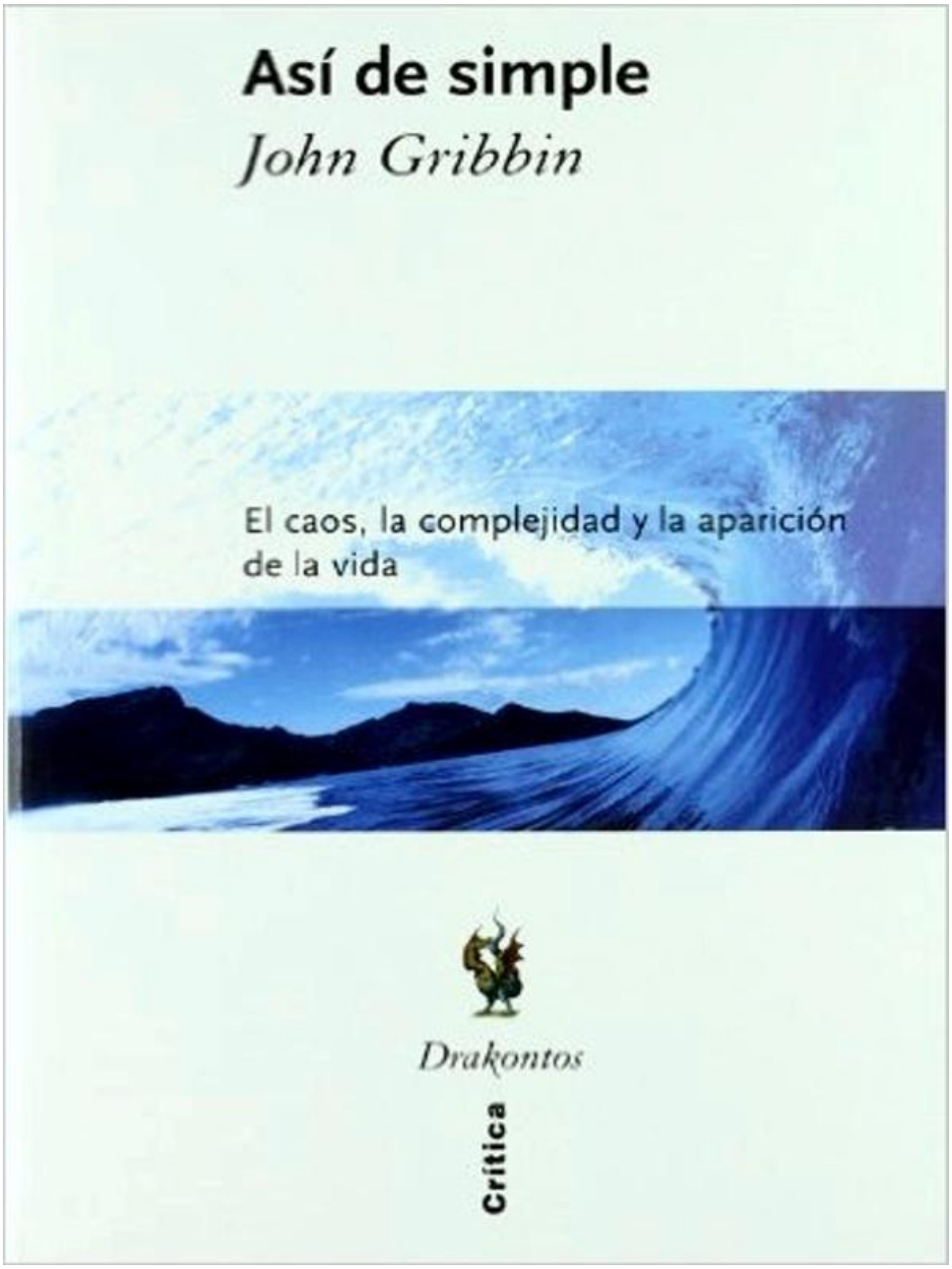
Así de simple
John Gribbin
El caos, la complejidad y la aparición de la vida
Drakontos
Crítica

A Jim Lovelock

Siempre me ha intrigado que, cuando se trata de aplicar las leyes tal como las entendemos actualmente, una computadora necesite hacer un número infinito de operaciones lógicas para efectuar cálculos relativos a lo que sucede en cualquier zona diminuta del espacio y en cualquier porción insignificante de tiempo. ¿Cómo puede suceder todo eso en un espacio diminuto? ¿Por qué se necesita una cantidad infinita de operaciones lógicas para averiguar lo que va a pasar en un fragmento diminuto de espacio/tiempo? A menudo he formulado la hipótesis de que en última instancia la física no necesitará una expresión matemática, ya que al fin se descubrirá la maquinaria y las leyes llegarán a ser sencillas, como un juego de ajedrez con todas sus aparentes complejidades.

Richard Feynman

The Character of Physical Law

La sencillez de la naturaleza no se puede calibrar por la de nuestras ideas. Siendo infinitamente variada en sus efectos, la naturaleza sólo es sencilla en sus causas, y su economía consiste en producir un gran número de fenómenos, a menudo muy complicados, mediante un pequeño número de leyes generales.

Pierre-Simon Laplace

1749-1827

Agradecimientos

Como en todos mis libros, Mary Gribbin ha desempeñado en éste un papel importante en cuanto a garantizar que las ideas que aquí se presentan resulten inteligibles para alguien que carezca de formación científica. Aparecería con su nombre como coautora en la portada del libro, si no fuera por el prejuicio que tiene nuestro editor de Penguin contra las autorías conjuntas.

Nuestro agradecimiento a la Universidad de Sussex por su apoyo al permitimos acceder a una excelente biblioteca y a una conexión de alta velocidad con Internet. Lo mismo deseamos expresar a la Alfred C. Munger Foundation por su contribución económica a nuestros gastos de investigación y viajes.

El autor y la editorial agradecen la autorización para reproducir las siguientes figuras: figura 2.1 de Edward Lorenz, The Essence of Chaos, University of Washington Press, 1993; figuras 2.3, 3.1, 3.4, 3.5, 3.7, 3.8 y 3.9 de J. C. Sprott en http://sprott.physics.wisc.edu/fractals/chaos; figuras 2.2 y 3.10 de James Gleick, Chaos, Heinemann, 1988; figuras 4.1 y 5.7 de Per Bak, How Nature Works, Oxford University Press, 1997; figura 4.2 de James Murray, Mathematical Biology, Springer Verlag, 1993; figura 5.2 de J. Feder, Fractals, Plenum, 1988; figura 5.4 de Philip Ball, The Self-Made Tapestry, Oxford University Press, 1999; figuras 5.8 y 5.9 de Stuart Kauffman, At Home in the Universe, Oxford University.

Introducción

La sencillez de la complejidad

El mundo que nos rodea parece ser un lugar complicado. Aunque hay algunas verdades sencillas que parecen eternas (las manzanas caen siempre hacia el suelo y no hacia el cielo; el Sol se levanta por el este, nunca por el oeste), nuestras vidas, a pesar de las modernas tecnologías, están todavía, con demasiada frecuencia, a merced de complicados procesos que producen cambios drásticos y repentinos. La predicción del tiempo atmosférico tiene todavía más de arte que de ciencia; los terremotos y las erupciones volcánicas se producen de manera impredecible y aparentemente aleatoria; las fluctuaciones de la bolsa siguen ocasionando prosperidad y bancarrota sin una pauta obvia. Desde la época de Galileo (en números redondos, a comienzos del siglo XVII) la ciencia ha hecho progresos enormes, ignorando en gran medida estas complejidades y centrándose en cuestiones sencillas, intentando explicar por qué las manzanas caen al suelo y por qué el Sol se levanta por el este. Los avances fueron de hecho tan espectaculares que hacia mediados del siglo XX ya se había dado respuesta a todas las cuestiones sencillas. Conceptos tales como la teoría general de la relatividad y la mecánica cuántica explicaron el funcionamiento global del universo a escalas muy grandes y muy pequeñas respectivamente, mientras el descubrimiento de la estructura del ADN y el modo en que éste se copia de una generación a otra hizo que la propia vida, así como la evolución, parecieran sencillas a nivel molecular. Sin embargo, persistió la complejidad del mundo a nivel humano, al nivel de la vida. La cuestión más interesante de todas, la que planteaba cómo la vida pudo haber surgido a partir de la materia inerte, siguió sin respuesta.

No es de extrañar que sea precisamente a escala humana donde se den las características más complejas del universo, las que se resisten más a rendirse ante los métodos tradicionales de la investigación científica. Realmente, es posible que seamos lo más complejo que hay en el universo. La razón es que, a escalas más reducidas, entidades tales como los átomos se comportan individualmente de un modo relativamente sencillo en sus interacciones mutuas, y que las cosas complicadas e interesantes surgen cuando se unen muchos átomos de maneras

complicadas e interesantes, para formar organismos tales como los seres humanos. Pero este proceso no puede continuar indefinidamente, ya que, si se unen cada vez más átomos, su masa total aumenta hasta tal punto que la gravedad aplasta toda la estructura importante y la aniquila. Un átomo, o incluso una molécula tan simple como la del agua, es algo más sencillo que un ser humano, porque tiene poca estructura interna; una estrella, o el interior de un planeta, es también algo más sencillo que un ser humano porque la gravedad aplasta cualquier estructura hasta aniquilarla. Esta es la razón por la cual la ciencia puede decirnos más sobre el comportamiento de los átomos y el funcionamiento interno de las estrellas que sobre el modo en que las personas se comportan.

Cuando los problemas sencillos se rindieron ante el empuje de la investigación, fue algo natural que los científicos abordaran rompecabezas más complicados que iban asociados con sistemas complejos. Como veremos, aunque anteriormente algunos individuos valientes habían atacado ya estos rompecabezas, para que por fin fuera posible comenzar a comprender el funcionamiento del mundo a una escala humana más compleja, hubo que esperar hasta la década de 1960, que fue cuando aparecieron los poderosos y rápidos (para lo que se estilaba en aquella época) ordenadores electrónicos. Estos nuevos inventos empezaron a ser conocidos por un público más amplio entre mediados y finales de la década de 1980, primero con la publicación del libro, ahora convertido en un clásico, Order out of Chaos, de Iya Prigogine e Isabelle Stengers, y luego con Chaos, de James Gleick[1]*.*

Por aquellos tiempos me encontraba ocupado escribiendo sobre los grandes triunfos de la ciencia en épocas anteriores y, aunque de manera ocasional intenté abordar las ideas del caos y de la complejidad, constaté que tales intentos me daban dolor de cabeza, por lo que mayormente me mantuve a una cierta distancia. Sin embargo, después de diez años esperando en vano que la teoría del caos se esfumara o que alguien escribiera un libro explicándola con un lenguaje que yo pudiera entender, decidí que, si nadie más iba a explicarla de una manera clara, lo tendría que hacer yo. Esto significaba que tendría que leer todo lo que pudiera encontrar sobre el tema e intentar comprenderlo por mi cuenta. Mientras lo hacía, descubrí que, después de todo, no era tan difícil. También la teoría de la relatividad

[1] Hay traducción al castellano: James Gleick, *Caos: la creación de una ciencia*, Seix Barral, Barcelona, 1998, trad. de Juan Antonio Gutiérrez-Larraya.

y la mecánica cuántica se consideraron, cuando eran nuevas, como unas ideas demasiado difíciles para que cualquiera las entendiera, salvo los expertos, pero ambas se basan en conceptos sencillos que son inteligibles para cualquier persona lega en la materia, siempre que esté dispuesta a aceptar su parte matemática con los ojos cerrados.

No tendría que haberme sentido sorprendido al descubrir que sucede lo mismo con el caos y la complejidad, pero así me sentí, y recuerdo claramente el momento en que por fin capté de qué se trataba. Tal como yo lo entendí, lo que en realidad importa es simplemente que algunos sistemas («sistema» no es más que una palabra de la jerga científica para designar cualquier cosa, como un péndulo que oscila, o el sistema solar, o el agua que gotea de un grifo) son muy sensibles a sus condiciones de partida, de tal modo que una diferencia mínima en el «impulso» inicial que les damos ocasiona una gran diferencia en cuanto a cómo van a acabar, y existe una retroalimentación, de manera que lo que un sistema hace afecta a su propio comportamiento. Parecía demasiado bueno para ser verdad, demasiado sencillo para ser cierto. Por eso, pregunté a la persona más inteligente que conozco, Jim Lovelock, si yo iba por el buen camino. ¿Era realmente verdad, le pregunté, que todo este asunto del caos y de la complejidad se basaba en dos ideas sencillas, la sensibilidad de un sistema a sus condiciones de partida, y la retroalimentación? Sí, me respondió, eso es todo lo que hay sobre el tema.

De algún modo, esto es como decir que «todo lo que hay» sobre la teoría especial de la relatividad es que la velocidad de la luz es la misma para todos los observadores. Es cierto, es sencillo, y es fácil de captar. Sin embargo, la complejidad de la estructura que se levanta sobre este hecho sencillo resulta asombrosa y requiere algunos conocimientos matemáticos para poder apreciarla plenamente. Pero yo sabía por experiencia que podía explicar lo esencial de la relatividad a personas carentes de formación científica y el hecho de darme cuenta de que había unas verdades sencillas similares en la base del caos y la complejidad me dio la confianza necesaria para intentar hacer lo mismo en este campo. El resultado está ahora en manos de los lectores: este libro es un intento de dar el paso necesario para tratar de explicar el caos y la complejidad de una manera sencilla, desde la base y para todos. La idea principal es que el caos y la

complejidad obedecen leyes sencillas, en esencia las mismas leyes sencillas que descubrió Isaac Newton hace más de trescientos años.

Lejos de echar abajo cuatro siglos de esfuerzos científicos, como nos permitirían creer algunos textos, estos nuevos avances nos muestran cómo los conocimientos científicos basados en leyes sencillas, y firmemente arraigados durante tanto tiempo, pueden explicar (aunque no predecir) el comportamiento aparentemente inexplicable de los sistemas meteorológicos, los mercados bursátiles, los terremotos, e incluso los seres humanos. Esperando convencer al lector, diremos que el estudio del caos y la deducción de la complejidad a partir de sistemas sencillos están ahora a punto de lograr explicar el propio origen de la vida. Según una frase atribuida a Murray Gell-Mann, pero que es un eco de la valoración de Richard Feynman citada al principio de este libro, el complicado comportamiento del mundo que vemos alrededor de nosotros, incluso el mundo de los seres vivos, no es más que «una complejidad superficial que surge de una profunda sencillez».[2] Precisamente la sencillez es la base de la complejidad, que es el tema de este libro.

[2] Citado, por ejemplo, en *Complexity,* de Roger Lewin.

Capítulo 1
El orden surgido del caos

Antes de la revolución científica del siglo XVII, el mundo parecía estar regido por el caos, entendiendo esta palabra de un modo bastante diferente a como la utilizan los científicos hoy en día. En absoluto se sugería que pudiera haber leyes sencillas y ordenadas que sustentaran la confusión del mundo, y lo más cerca que se llegó de ofrecer una razón que explicara el comportamiento del viento y el clima, la existencia de hambrunas o las órbitas de los planetas fue decir que eran resultado del capricho de Dios o de los dioses. Allí donde se percibía orden en el universo, este orden se atribuía a la respuesta que daban los objetos físicos a una necesidad de que se preservaran la armonía y el orden siempre que fuera posible, se suponía que las órbitas de los planetas y la del Sol alrededor de la Tierra (de la que se pensaba que estaba en el centro del universo) eran círculos, porque los círculos eran perfectos; los objetos caían hacia el suelo porque el centro de la Tierra estaba en el centro de todo, que era el centro de simetría del universo, y por lo tanto, el lugar más deseable en que uno se podía encontrar. Incluso el filósofo Aristarco de Samos, que vivió en el siglo II a.C., pensaba en una órbita circular cuando se atrevió a sugerir que la Tierra se movía alrededor del Sol.

Estos ejemplos ilustran una diferencia absolutamente crucial entre la ciencia de los antiguos y la de los tiempos posteriores a Galileo. Los antiguos griegos fueron unos matemáticos excelentes, y en particular unos geómetras de primera, que conocían muy bien las relaciones entre los objetos estacionarios. Por supuesto, esta geometría tenía sus raíces en culturas aún más antiguas, siendo fácil imaginarse cómo esta ciencia primitiva pudo haber surgido a partir de cuestiones prácticas de la vida cotidiana durante el desarrollo de las sociedades agrícolas de la prehistoria, mediante problemas asociados con la construcción de edificios y el trazado de las ciudades, así como, a medida que la sociedad se hizo más complicada, con la necesidad de dividir la tierra en parcelas. Pero los antiguos no entendían en absoluto cómo se movían los objetos, es decir, no conocían las leyes del movimiento. Basta con pensar lo asombrados que estaban con las famosas paradojas de Zenón, como la del soldado al que nunca puede matar una flecha. Si

huye, en el momento en que la flecha llega a la posición en que estaba, él ya se ha desplazado; en el tiempo que tarda la flecha en cubrir ese desplazamiento, él ha podido desplazarse un poco más lejos; y así sucesivamente.

A pesar de la existencia de personas como Aristarco, el universo centrado en la Tierra siguió siendo la imagen aceptada (lo que los científicos llamarían ahora un «modelo»), incluso después de que Nicolás Copérnico publicara en 1543 su modelo de universo centrado en el Sol (que seguía basado en los círculos). Su libro, *De Revolutionibus Orbium Coelestrum* quedó terminado en lo esencial en 1530, y buena parte de sus contenidos fueron tema de discusión antes de ser publicados, lo cual indujo a Martín Lutero a comentar en 1539:

«Este loco desea volver del revés toda la astronomía; pero las Sagradas Escrituras nos dicen que Josué ordenó al Sol que se detuviera, no a la Tierra».

Respondiendo a críticas similares, Galileo replicó más tarde:

«La Biblia muestra la manera de llegar al cielo, no la manera en que se mueven los cielos».

Fue un contemporáneo de Galileo, Johannes Kepler, quien, utilizando las observaciones minuciosamente recopiladas por Tycho Brahe, señaló, para aquellos que tuvieran los ojos abiertos, que el planeta Marte no sólo se movía alrededor del Sol, sino que lo hacía recorriendo una órbita elíptica, dejando así sin base la idea de que la perfección circular, la preferida por los antiguos griegos, era la que gobernaba el cosmos.

Incluso para las personas que saben poco de ciencia, o de la historia de la ciencia, Galileo (que vivió desde 1564 hasta 1642) es famoso hoy en día como el hombre que enfocó los cielos con uno de los primeros telescopios, halló pruebas para justificar el modelo heliocéntrico copernicano y tuvo con la Iglesia Católica una bronca que terminó con su condena por herejía y la prohibición de sus libros en los países católicos, cosa que (por supuesto) hizo que se vendieran como rosquillas calientes en cualquier otro lugar. Pero Galileo llegó mucho más lejos. Fue él sobre todo quien sentó los principios del método científico de investigación, que incluía teorías (o modelos) de tipo comparativo donde se utilizaban los resultados de

experimentos y observaciones, y también fue Galileo el primero que abordó la cuestión del movimiento de una manera científica.

La clave de los trabajos de Galileo sobre el movimiento es algo que descubrió en 1583, cuando era estudiante de medicina en Pisa. Durante un aburrido sermón que se pronunciaba en la catedral de esta ciudad, observó una lámpara de araña que se movía hacia un lado y otro, y cronometró la oscilación con su propio pulso. De esta manera, Galileo constató que el tiempo que tardaba la lámpara en realizar una oscilación completa era siempre el mismo, tanto si recorría un amplio arco como si describía uno pequeño. Experimentos posteriores demostraron que el tiempo que invierte un péndulo en realizar una oscilación completa dependía de su longitud, no de la amplitud de su oscilación.

Este es el fundamento del reloj de péndulo, pero, aun sin llegar a construir un reloj (diseñó uno, que luego construyó su hijo), posteriormente Galileo pudo utilizar un péndulo como cronómetro preciso cuando realizó experimentos para estudiar el comportamiento de unas bolas que rodaban hacia abajo por una rampa. Estos experimentos aportan más información sobre el pensamiento de Galileo y el método científico. Deseaba estudiar la caída de los objetos para investigar el efecto que producía la gravedad sobre los cuerpos en movimiento. Pero al caer, las bolas se movían demasiado rápido para que él pudiera seguirles la pista. Por este motivo, decidió dejarlas rodar por una rampa inclinada, ya que se dio cuenta de que este procedimiento le proporcionaba una versión prolongada y retardada del modo en que unas bolas caen por efecto de la gravedad. Mediante estos experimentos Galileo desarrolló el concepto de aceleración.

La velocidad de un objeto nos dice qué distancia recorre en un tiempo determinado, por ejemplo, un segundo. Una velocidad constante de 9,8 metros por segundo significa que cada segundo el objeto en movimiento cubre una distancia de 9,8 metros. Pero Galileo descubrió que los objetos que caen (o las bolas que bajan rodando por una rampa) se mueven cada vez más rápido, con una velocidad que aumenta cada segundo. Lo más importante de aquellos experimentos fue ver que la velocidad aumentaba siempre en una misma cantidad cada segundo. Esto se llama aceleración uniforme, y una aceleración uniforme de 9,8 metros por segundo *cada segundo* significa que, partiendo de una situación de reposo, en el primer segundo

el objeto alcanza una velocidad de 9,8 metros por segundo; pasados dos segundos se mueve a 19,6 metros por segundo, y su velocidad después del tercer segundo es de 29,4 metros por segundo, y así sucesivamente.

He elegido concretamente este ejemplo porque 9,8 metros por segundo cada segundo es la aceleración que produce la gravedad sobre un objeto que cae hacia la superficie de la Tierra; dado que el tiempo aparece dos veces en el cálculo, se llama a esto un efecto de segundo orden, siendo la velocidad un efecto de primer orden. El círculo se cierra diciendo que esta aceleración debida a la gravedad explica por qué los péndulos se comportan como lo hacen.

Galileo hizo algo más, algo fundamental para la historia que contamos en este libro. Constató que las bolas que descienden rodando por planos inclinados se frenaban un poco a causa del rozamiento. De hecho, lo que él midió no era una aceleración perfectamente uniforme.

Sin embargo, dio un salto radical y con consecuencias, asombroso para su época, cuando extrapoló a partir de sus observaciones reales el modo en que las bolas se moverían sin el efecto del rozamiento, es decir, sobre una pendiente ideal, perfectamente deslizante. Este tipo de extrapolación formaría parte de la base de la investigación científica mundial durante los cuatro siglos siguientes. Cuando los científicos, en particular los físicos, intentaron explicar el mundo en términos de leyes matemáticas, formularon estas leyes para describir el comportamiento de objetos míticos, tales como esferas perfectamente rígidas, que rebotaban unas contra otras sin deformarse y rodaban sobre superficies sin sentir el rozamiento, y otros objetos similares.

Sin embargo, a diferencia de los filósofos de la antigua Grecia, sabían que aquella imagen de perfección no representaba el mundo real. Partiendo de estas ecuaciones, podían intentar introducir términos suplementarios, factores de corrección, para tener en cuenta las imperfecciones del mundo real, incluyendo así, por ejemplo, el efecto de la resistencia del aire que actúa sobre un cuerpo que cae. La resistencia del aire explica por qué en la Tierra un martillo y una pluma caen a velocidades diferentes, mientras que en la Luna, que carece de atmósfera, caen a la misma velocidad, como demostraron los astronautas del Apolo.

Todo esto ayudó a Galileo a eliminar de la ciencia otro aspecto de la perfección

geométrica que sus predecesores habían imaginado para el mundo real. Antes de Galileo se pensaba que cuando un cañón disparaba una bala con un cierto ángulo de elevación sobre la horizontal, la trayectoria de la bala sería una línea recta al salir de la boca del cañón, luego describiría un arco de circunferencia perfecto durante cierto tiempo y finalmente caería verticalmente al suelo.

En este movimiento sólo aparecía la perfección imaginaria de las líneas rectas y las circunferencias. Aplicando su descubrimiento de que la gravedad produce una aceleración de frenado constante sobre la bala del cañón, y suponiendo que ésta tiene una velocidad inicial al salir de la boca del cañón, Galileo demostró que la trayectoria de la bala debía ser, en realidad, una única línea con una suave curvatura, un arco parabólico, hasta llegar a su objetivo. Los mismos cálculos demostraron que el alcance máximo del cañón (suponiendo siempre la misma carga de pólvora y el mismo peso de la bala) se lograría al disparar con un ángulo de elevación de 45 grados sobre la horizontal.

Se trataba de cuestiones prácticas que tuvieron gran importancia en la época turbulenta en que vivió Galileo, y este tipo de trabajos militares contribuyeron a crear su temprana fama. Independientemente de lo que los filósofos y los teólogos pudieran decir sobre la perfección, los ejércitos que estaban en el campo de batalla no tenían tiempo de razonar sutilmente sobre la idea de que el movimiento circular era deseable; todos querían saber hacia dónde tenían que apuntar sus armas para conseguir el máximo efecto, y Galileo se lo dijo.

Fue una combinación del descubrimiento de las órbitas elípticas por parte de Kepler, y de las teorías de Galileo sobre la aceleración y el método científico, lo que preparó el camino para el mayor descubrimiento científico del siglo XVII, y quizá de todos los siglos: la ley de la gravitación universal de Isaac Newton. Newton había nacido en 1642 y falleció en 1727. Su gran obra *Philosophiae Naturalis Principia Mathematica* o, dicho abreviadamente los *Principia,* se escribió entre los años 1684 y 1687 y se publicó en 1687, pero se basaba en ideas que Galileo había desarrollado veinte años antes, cuando era un joven graduado de Cambridge a punto de convertirse en miembro del Trinity College, y se había visto obligado a pasar varios meses en la casa que tenía su madre en Lincolnshire, ya que la universidad se cerró por temor a la peste.

Al igual que Galileo, Newton subrayó la importancia de comparar las teorías y los modelos con los resultados de los experimentos y las observaciones del mundo real, y, siempre que le fue posible, realizó por sí mismo importantes experimentos para comprobar sus teorías. En la actualidad, este procedimiento está tan asumido como parte del método científico que puede parecer obvio incluso a los que no son científicos y es difícil valorar hasta qué punto, incluso en el siglo XVII, muchos filósofos harían especulaciones sobre la naturaleza del mundo físico de una manera abstracta, sin mancharse las manos haciendo experimentos. El ejemplo clásico es la cuestión de si dos pesos diferentes que dejamos caer desde la misma altura y al mismo tiempo llegarán juntos al suelo, un asunto que estuvo dando vueltas durante décadas entre los pensadores abstractos, después de que un ingeniero flamenco, Simón Stevin, hiciera realmente los experimentos con pesos de plomo que dejaba caer desde una altura de unos 10 metros y, tras descubrir que caían a la misma velocidad, publicara los resultados en 1586.[3]

También Newton adoptó y perfeccionó la idea de Galileo, valorando de manera positiva los modelos deliberadamente simplificados (como los planos sin rozamiento) para utilizarlos en la descripción de aspectos concretos del mundo real. Por ejemplo, una característica fundamental de los trabajos de Newton sobre la gravedad y las órbitas es el hecho de que, en sus cálculos relativos a los efectos de la gravedad, consideró objetos tales como Marte, la Luna o una manzana, como si toda su masa estuviera concentrada en un solo punto, y de esta manera, siempre que nos encontremos en el exterior del objeto en cuestión, su influencia gravitatoria se mide en función de nuestra distancia a dicho punto, que es el centro de masa del objeto (y asimismo el centro geométrico, si el objeto es una esfera). En sus *Principia,* Newton demostró que éste es el caso de los objetos esféricos. Él sabía que la Tierra no es exactamente esférica (de hecho, supo calcular cuánto se pandea la Tierra en el ecuador a causa de su rotación); sin embargo, consideró razonable suponer que, en una primera aproximación, la Tierra era esférica (y también el Sol, Marte, etc.) y calcular las órbitas a partir de esta suposición. De hecho, ciertos cálculos posteriores han determinado que, siempre que estemos suficientemente

[3] Dicho sea de paso, no hay pruebas de que el propio Galileo realizara este experimento desde la torre inclinada de Pisa, un mito que parece haberse originado a partir de la pintoresca biografía de Galileo escrita por Vincenzo Viviani, que sólo tenía diecisiete años cuando se convirtió en el encargado de escribir con veneración los textos del gran hombre, cuando éste se quedó ciego a una edad avanzada.

alejados de ellos, incluso objetos cuya forma es muy irregular actúan, a efectos de la gravedad, como si toda su masa estuviera concentrada en un punto, pero esto no reduce la importancia que tiene la idea de utilizar aproximaciones idealizadas de la realidad cuando es necesario, o práctico, para efectuar los cálculos de una manera más directa.

Sin embargo, en esta historia hay más que lo que vemos a simple vista. En sus *Principia*, Newton demostró la validez de considerar la influencia gravitatoria de un objeto esférico como si toda su masa estuviera concentrada en el centro, utilizando técnicas geométricas que habrían entendido los antiguos griegos y que ciertamente les resultaban familiares a los contemporáneos de Newton. Esta demostración es muy difícil. No obstante, sabemos actualmente que, mucho antes de escribir sus *Principia,* Newton había desarrollado (o descubierto) la técnica matemática que hoy en día se denomina cálculo y que, utilizando esta técnica, la demostración resulta muy fácil. Algunos expertos sospechan que, en realidad, Newton resolvió el problema empleando primero el cálculo, y que luego continuó aplicando un proceso minucioso que consistía en traducir todo a la terminología clásica, para asegurarse de que sus contemporáneos lo entenderían. Si fue así, es posible, en cierto sentido, que se pillara los dedos, porque, al guardar silencio sobre su nueva técnica matemática, preparó el camino para una dura disputa con el alemán Wilhelm Leibniz, que inventó la técnica por su cuenta de manera independiente (y le dio el nombre por el que todavía se conoce). A Leibniz se le ocurrió la idea un poco más tarde que a Newton, pero tuvo el buen sentido de publicar sus trabajos y ésta es, en parte, la razón por la que hubo duras disputas sobre la prioridad (éstas surgieron en parte también porque ambos protagonistas se negaban a hacer concesiones y Newton, en particular, era una persona bastante desagradable que creía con arrogancia en sus propias habilidades y reaccionaba con malevolencia ante cualquiera que él considerara como su oponente). Pero aquí la cuestión de la prioridad nos preocupa poco. Lo importante es que el cálculo constituye una técnica con la que es posible descomponer los problemas en componentes muy pequeñas que pueden manipularse de manera matemática y los resultados se combinan para dar solución a un problema global. Por ejemplo, en el caso de la influencia gravitatoria de un objeto esférico, la esfera se puede tratar como si estuviera

dividida (diferenciada) en un número infinito de trozos de materia infinitamente pequeños (infinitesimales) y es posible escribir una ecuación que describa la influencia gravitatoria que ejerce un fragmento de materia de este tipo en función de su posición en la esfera. El efecto global de todos los trozos de materia cuando funcionan juntos se puede sumar, o integrar, a partir de esta ecuación.[4]

El mismo tipo de procedimiento se puede aplicar al tiempo. Por ejemplo, el vuelo de una flecha se puede diferenciar en una descripción de cómo se mueve la flecha en cualquiera de los infinitos puntos de su trayectoria, y lo mismo se puede hacer con la ruta de un corredor que intenta escapar de la flecha. Al realizar la integración, las dos ecuaciones diferenciales nos dirán exactamente dónde alcanza la flecha al corredor, sin paradojas. Cuando el cálculo llegó a ser una técnica ampliamente conocida, se tuvo la impresión de que Newton y Leibniz habían conseguido dominar el tiempo, haciendo que fuera posible describir el comportamiento de los objetos en movimiento con la misma precisión que los antiguos griegos habían logrado aplicar a la descripción de las relaciones entre objetos estáticos. Al menos, en principio, lo habían conseguido. Es fácil el cálculo cuando se trata del caso aislado de un planeta que describe una órbita en tomo a una estrella, o un corredor que intenta huir de una flecha; pero, como todo, la dificultad empieza cuando manejamos sistemas más complicados, aunque se aplican las mismas reglas básicas. En general, se puede escribir de manera inmediata un conjunto de ecuaciones diferenciales para describir un sistema de este tipo; lo que puede resultar problemático es resolver (integrar) las ecuaciones.

Independientemente de cuáles fueran las técnicas matemáticas que utilizara Newton de manera privada en la década de 1680, en los *Principia* demostró, utilizando procedimientos que sus contemporáneos conocían, que para que los planetas describieran órbitas elípticas alrededor del Sol (y, por consiguiente, para que las observaciones concordaran con la teoría), la gravedad debía cumplir una ley en la que apareciera la proporcionalidad inversa al cuadrado de la distancia. Concretamente, la fuerza de atracción entre dos masas situadas a una cierta distancia (midiendo dicha distancia desde el centro de masa de cada objeto) es

[4] Una copia única de los manuscritos de Arquímedes, que había estado perdida durante siglos, se tradujo a finales del siglo XX y reveló que el gran matemático griego había desarrollado lo esencial de este tipo de integración hace más de dos mil años. Por supuesto, Newton y Leibniz desconocían esto, pero el hecho en sí hace que su disputa sobre prioridades se vea con otra perspectiva.

proporcional al producto de ambas masas y se divide por el cuadrado de la distancia que separa la una de la otra (por esto se habla de una ley de la «inversa del cuadrado»). La constante de proporcionalidad que aparece en la fórmula es la constante gravitatoria, un número que expresa hasta qué punto es intensa la fuerza de la gravedad y se escribe mediante la letra *G.* Utilizando los símbolos habituales,

$$F = G\frac{Mm}{r^2}$$

Este hallazgo no fue del todo una sorpresa para los contemporáneos de Newton. Científicos como Robert Hooke, Edmond Halley y Christopher Wren, todos ellos miembros de la Royal Society, sabían ya que una ley de la inversa del cuadrado daría como resultado órbitas elípticas. El triunfo de Newton fue demostrar que *sólo* una ley de la inversa del cuadrado daría tal resultado. Pero fue aún más lejos. De su trabajo se deducía que se trataba de una ley *universal.* No sólo era aplicable a las manzanas que caían de los árboles, o a la Luna cuando orbitaba alrededor de la Tierra, o a los planetas cuando describían sus órbitas alrededor del Sol. Y siempre era igual, no sólo se cumplía en verano, o en los meses con «r». Se cumplía para la fuerza de atracción que ejercía cualquier trozo de materia, en cualquier lugar del universo, sobre cualquier otro fragmento de materia, en cualquier otra parte del universo. Newton hizo que éste pareciera un mundo ordenado, sin espacio para las interferencias de dioses caprichosos[5]. Aportó además otras tres leyes: las leyes del movimiento que describen el comportamiento de objetos que se mueven en el laboratorio, o en el universo a gran escala, o en el sistema solar y más allá de éste. Estas leyes han de considerarse también como leyes universales, que se cumplen siempre y en cualquier lugar.

Se trata de las leyes del movimiento de Newton, que constituyen la base de trescientos años de ciencia, pero que pueden resumirse de una forma muy sencilla y que marcan el desarrollo del modo científico de observar el mundo. La primera ley de Newton dice que todo objeto permanece en reposo o continúa desplazándose en

[5] Quizá valga la pena señalar que, en una fecha tan reciente entonces como 1609, Johannes Kepler, tras constatar que debía existir algo que mantuviera los planetas en sus órbitas alrededor del Sol, lo había llamado la «fuerza del Espíritu Santo», y nadie se había reído.

línea recta, salvo que se le aplique una fuerza o se encuentre sometido a ella. Lo de que «permanece en reposo» coincide con lo que nos dice el sentido común (como tal entendemos lo que percibimos a diario en la Tierra). Los objetos no se mueven a menos que algo les haga moverse. Pero el asunto del movimiento constante en línea recta choca frontalmente con este sentido común. Cuando ponemos algo en movimiento aquí en la Tierra, si lo abandonamos a su suerte, cae al suelo y se detiene. El punto de vista de Newton, que ampliaba las teorías de Galileo, era que, si los objetos caen, es sólo a causa de una fuerza exterior (la gravedad), y que únicamente dejan de moverse porque actúan sobre ellos otras fuerzas (rozamiento). Sin haber visto las imágenes, que actualmente nos resultan familiares, de objetos que obedecen las leyes de Newton en el entorno de gravedad cero y casi totalmente carente de rozamiento que existe en el interior de una nave espacial, Newton supo imaginar lo que sería el mundo si se pudiera ignorar el rozamiento y otras fuerzas exteriores.

La segunda ley nos dice en qué medida se ve afectado el movimiento de un objeto cuando se le aplica una fuerza. Afirma que una fuerza aplicada a una masa produce una aceleración que viene dada por la sencilla ecuación

$$F = ma$$

o

$$a = F/m$$

Es esta ley, junto con la ley de la gravedad, la que explica o describe la órbita que recorre un planeta alrededor del Sol. Ambas leyes juntas, la segunda ley y la ley de la gravedad, explican el comportamiento de los objetos que caen que Galileo observó y describió. Si la Tierra es la masa M, entonces la fuerza gravitatoria que actúa sobre cualquier masa situada en la superficie de la Tierra es proporcional a su propia masa, *m,* ya que

$$F = G\frac{Mm}{r^2}$$

Pero la aceleración que produce esta fuerza se obtiene dividiendo *F* entre *m.* Por lo tanto, la masa *m* se elimina de la ecuación, con lo que todos los objetos que caen por efecto de la gravedad tienen la misma aceleración cuando se encuentran cerca de la superficie de la Tierra. Todos ellos tienen otra aceleración diferente (aunque es la misma para todos) en la superficie de la Luna.

La tercera ley dice que toda fuerza que actúa sobre un objeto produce una fuerza igual en sentido opuesto, o, en palabras de Newton, a cada acción se opone siempre una reacción igual, pero de sentido contrario. Un ejemplo sencillo de este fenómeno es lo que sucede cuando la fuerza que se produce en un rifle para enviar una bala a toda velocidad en una trayectoria determinada produce una reacción que se percibe como el golpe del rifle contra el hombro del tirador. Si alguien golpea una mesa con el puño, podrá sentir la reacción de una manera igualmente obvia. Otro caso, algo menos obvio, es el del Sol cuando atrae a un planeta mediante la gravedad y éste atrae al Sol con una fuerza de la misma magnitud y sentido opuesto, como si ambos estuvieran unidos por una goma elástica tensada. Incluso la manzana que cae de un árbol está atrayendo la Tierra hacia ella, aunque lo hace con una fuerza cuya magnitud es insignificante (lo es porque, aunque se aplique la misma fuerza a la manzana y a la Tierra, la aceleración que éstas experimentan es proporcional a la fuerza, pero *inversamente* proporcional a la masa, y la Tierra posee una masa mucho mayor que la de la manzana). Por lo que respecta a las órbitas, esto significa que, en vez de ser la Luna la que describe una órbita alrededor de la Tierra, en realidad tanto la Luna como la Tierra orbitan alrededor de su centro de masa mutuo (el cual, debido a que la Tierra tiene una masa mucho mayor que la de la Luna, se encuentra bajo la superficie terrestre).

Estas leyes se cumplen igualmente, por una parte, para los planetas y satélites que se mueven en el espacio, y, por otra (teniendo en cuenta el rozamiento), para los sucesos que tienen lugar en la Tierra, donde, por ejemplo, en una mesa de billar se puede ver una buena aproximación del modo en que actúan las leyes de Newton. Pero, por ahora, nos centraremos en el comportamiento de los planetas, porque

éste fue el caso en que comenzó a haber problemas con la explicación del mundo que había dado Newton. Durante siglos, en gran medida, se optó por esconder estos problemas bajo la alfombra, pero, como es sabido, fue en el ámbito de la mecánica de los cuerpos celestes donde surgieron finalmente los primeros indicios del caos, en el sentido en que lo entiende la ciencia moderna.

Parte del problema es que, aunque las leyes de Newton (incluida la ley de la gravedad) pueden utilizarse perfectamente para calcular las órbitas de dos objetos que giran uno alrededor del otro en un universo que, por lo demás, se considera vacío (la Luna alrededor de la Tierra, o la Tierra alrededor del Sol, etc.), estas leyes no pueden dar respuestas exactas para cálculos que se refieran a tres o más objetos en movimiento sometidos a una atracción gravitatoria mutua (como la Tierra, la Luna y el Sol en su movimiento conjunto a través del espacio). Esto es lo que se llama el «problema de los tres cuerpos», aunque se aplique a cualquier número de «cuerpos» mayor que dos; hablando de una forma más general, los físicos aluden a veces al «problema de *N* cuerpos», donde *N* puede ser cualquier número mayor que dos. Las ecuaciones que describen estos sistemas pueden escribirse, pero no resolverse, no son integrables y se dice que no tienen soluciones analíticas (lo cual significa que las técnicas del análisis matemático no funcionan con estas ecuaciones). De las que pueden resolverse analíticamente a veces se dice que son deterministas; la ecuación relativa a la órbita de un planeta alrededor del Sol es determinista y posee soluciones analíticas en forma de elipses. Sin embargo, es importante señalar que la falta de soluciones del problema de tres cuerpos no tiene como causa nuestras deficiencias humanas como matemáticos; el problema está construido según las leyes de las matemáticas.

No obstante, a menudo hay maneras de soslayar el problema. Se trata de aplicar aproximaciones. Concretamente para tres cuerpos que describen órbitas uno alrededor del otro es posible llevar a cabo los cálculos en pasos repetidos, suponiendo primero que uno de los cuerpos se encuentra inmóvil, mientras se realiza el cálculo para los otros dos, y así sucesivamente. Las respuestas que se obtienen nunca son perfectas, porque en realidad los tres cuerpos están moviéndose al mismo tiempo. Pero, si se hacen suficientemente pequeños los pasos temporales de que consta este proceso repetitivo, en muchos casos las órbitas que

calculamos estarán muy cerca de las reales. En el sistema solar, el Sol es mucho mayor que cualquiera de los pianolas (o, de hecho, mayor que todos los planetas juntos), por lo que su influencia gravitatoria es predominante y, en una primera aproximación, se pueden ignorar las influencias de los planetas. Por ejemplo, la órbita de Marte se calcula suponiendo que los demás planetas no existen y el resultado es una elipse perfecta. La órbita real de Marte difiere ligeramente de esta solución analítica de las ecuaciones del movimiento, pero las diferencias pueden explicarse reconociendo los efectos perturbadores de otros planetas, especialmente los de los gigantes Júpiter y Saturno.

De manera similar, la órbita que recorre la Luna en torno a la Tierra puede calcularse, en una primera aproximación, ignorando los efectos del Sol, que está muy alejado, y añadiéndolos luego como una corrección del cálculo. Si estamos dispuestos a llevar a cabo una serie de cálculos laboriosos para cada una de las sucesivas correcciones en las que intervienen todos los planetas y sus influencias mutuas (cosa que hoy en día es más bien fácil utilizando ordenadores de alta velocidad), se puede conseguir una aproximación muy buena para la predicción de la órbita realmente observada del planeta o del satélite en cuestión. Pero *nunca* podremos obtener una solución matemática analítica exacta que nos permita predecir para siempre cómo se moverá en su órbita un planeta o un satélite. Siempre hay un cierto margen de error.

Además, el problema de los tres cuerpos es del todo inabordable si los tres objetos tienen más o menos el mismo tamaño (es decir, la misma masa) y están a aproximadamente la misma distancia uno de otro. La ausencia de soluciones analíticas significa que la propia naturaleza no «sabe» exactamente cómo cambiarán (evolucionarán) las órbitas a medida que transcurra el tiempo. Incluso en el caso de los planetas del sistema solar, siempre cabe la posibilidad de que las órbitas no permanezcan para siempre tal como son hoy en día, ya que pueden experimentar cambios literalmente impredecibles.

Newton era consciente de esto, pero era profundamente religioso (aunque sus creencias no coincidían totalmente con la ortodoxia de la Iglesia de entonces) y resolvió el problema, al menos así lo pensaba él, sugiriendo que, si los planetas comenzaban a salirse de sus órbitas habituales (quizá desplazándose en espiral

hacia el Sol, o perdiéndose en las profundidades del espacio), siempre podría intervenir Dios para colocarlos de nuevo en su camino. Esto provocó una respuesta cortante por parte de Leibniz, que comparó el universo ordenado y determinista de Newton con un reloj, afirmando con sarcasmo que el Dios de Newton debía de ser un relojero bastante torpe si era incapaz de hacer un reloj que marcara siempre la hora correcta, pues para que funcionara bien tenía que intervenir cada vez que se estropeaba.

El problema continuó sin tener solución hasta finales del siglo XVIII, cuando el matemático francés Pierre Laplace (1749-1827) aparentemente puso orden en el sistema solar. Se centró primero en calcular (utilizando el laborioso procedimiento de las correcciones sucesivas paso a paso que hemos esbozado con anterioridad) las órbitas de Júpiter y Saturno, que son los planetas más grandes del sistema solar y ejercen la máxima influencia gravitatoria el uno sobre el otro, y cada uno de ellos sobre los planetas menores, después de la que ejerce el Sol.

Laplace descubrió que actualmente la órbita de Júpiter se está expandiendo de forma lenta, mientras que la de Saturno se está reduciendo, justo el tipo de efecto que preocupaba a Newton. Pero Laplace vio que estas modificaciones están ligadas a un cambio rítmico en la influencia gravitatoria que ejercen ambos planetas el uno sobre el otro, y que dicho cambio se produce porque Saturno describe dos órbitas alrededor del Sol mientras Júpiter recorre su órbita cinco veces.

Esto significa que los dos planetas gigantes se encuentran a distancia mínima uno del otro cada cincuenta y nueve años. Utilizando las leyes de Newton y la técnica de correcciones reiteradas paso a paso, Laplace calculó que el efecto producido por todo esto era la inversión de los cambios globales observados en las órbitas de los dos planetas cada 929 años, después de 929 años durante los cuales la órbita de Júpiter se expande, mientras la de Saturno se contrae, hay otro intervalo de 929 años durante el cual la órbita de Júpiter se contrae, mientras la de Saturno se expande, y así sucesivamente. Laplace pensó que así había restablecido el orden en el sistema solar y, según la famosa anécdota, le comentó a Napoleón que él «no necesitaba» la «hipótesis» de Dios, ya que:

Las irregularidades de estos dos planetas parecían antes imposibles de explicar mediante la ley de la gravitación universal, pero actualmente constituyen una de

sus pruebas más concluyentes.

Como veremos, sólo tenía razón hasta cierto punto. Sin embargo, demostró que en otros planetas se daba el mismo tipo de estabilidad y, por consiguiente, en el conjunto del sistema solar, de tal modo que, desde principios del siglo XIX en adelante, las leyes de Newton parecían implicar realmente que el sistema solar y el universo en toda su extensión funcionaban con la estabilidad y fiabilidad de un reloj perfecto, sin necesidad de intervenciones externas para mantenerlos a la hora exacta.

El éxito de las leyes de Newton permitió a los físicos resolver muchos problemas, siendo el fundamento sobre el cual se construyó toda la ciencia moderna. Aunque los físicos de los siglos XIX y XX fueron conscientes de que hay muchas situaciones en las cuales no existen soluciones analíticas para las ecuaciones correspondientes, pudieron hacer unos progresos tan enormes resolviendo las ecuaciones en los casos en que eran deterministas, y aplicando técnicas de aproximación en otros casos, que los problemas más inabordables quedaron en gran medida ignorados.

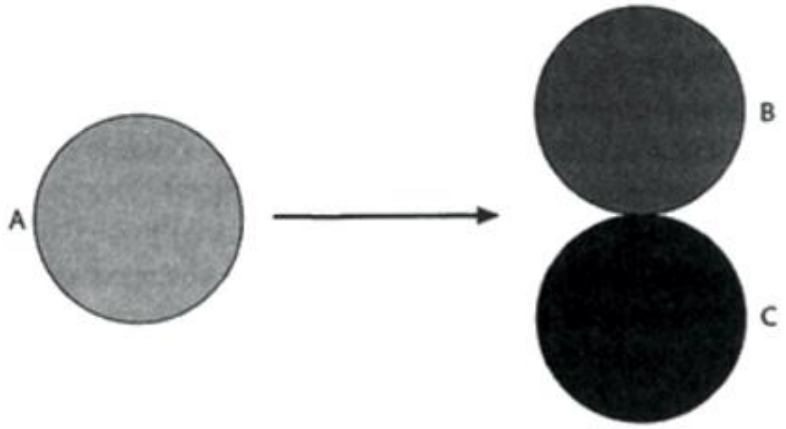

Figura 1.1. Según las leyes de Newton, si una esfera perfectamente elástica golpea simultáneamente dos esferas perfectamente elásticas que se encuentran en contacto, es imposible predecir adónde irán las tres esferas.

Después de todo, no deja de ser natural que se resuelvan primero todos los problemas fáciles, antes de preocuparse por los difíciles. Pero sólo unas pocas

personas se interesaron por los problemas que quedan más allá del alcance de las ecuaciones de Newton (al menos hasta el punto de indicar la existencia de tales problemas), y otro ejemplo del problema de los tres cuerpos nos ayudará a hacernos una idea de lo limitadas que son en realidad las ecuaciones de Newton.

Se trata de una situación que se puede visualizar recurriendo a los choques entre tres bolas en una mesa de billar. Si una bola en movimiento golpea a otra bola que está en reposo (o incluso si ambas bolas están en movimiento), entonces las leyes de Newton pueden utilizarse para averiguar exactamente cómo se mueven las dos bolas después de producirse la colisión, siempre que conozcamos las masas de las bolas, sepamos la velocidad a la que se desplaza la bola en movimiento, y tengamos en cuenta el rozamiento en la forma habitual (o lo ignoremos).

Pero, si dos bolas inmóviles están situadas de tal forma que hay contacto entre ambas, y la bola en movimiento las golpea lateralmente, chocando simultáneamente con ambas bolas inmóviles, las leyes de Newton no pueden decirnos, en general, cómo se separaran después de la colisión las dos bolas que están en contacto una con otra.[6]

Si una de las dos bolas en contacto recibe el golpe antes que la otra, aunque la diferencia sea sólo una fracción de segundo, las leyes nos dicen adonde van a ir las tres bolas.

Sin embargo, si se produce una colisión auténticamente simultánea, esto no es así. Podríamos intentar soslayar esta dificultad argumentando que en la práctica es altamente improbable que la colisión sea del todo simultánea, pero sigue siendo preocupante que existan situaciones sencillas en las que los resultados son imprevisibles, a pesar de que las leyes funcionan satisfactoriamente cuando se trata de explicar hechos que van desde la oscilación de un simple péndulo hasta el vuelo de una nave espacial dirigida a la Luna.

Realmente, a nadie le preocupaba esto en el siglo XIX. Estaban todos demasiado ocupados aplicando tanto las leyes de Newton como el método científico desarrollado por Galileo y New- ton, para diseñar una imagen global de un universo ordenado y determinista. La imagen que prevalecía era la que Laplace había explicado en 1814 en su *Essai philosophique sur les probabilités:*

6 Se puede decir lo mismo si las tres bolas en movimiento chocan entre sí simultáneamente, pero en la práctica, esto es aún más difícil de conseguir.

Imaginemos una mente inteligente que, en todo instante, pudiera tener conocimiento de todas las fuerzas que controlan la naturaleza y también, de las condiciones en que se encuentran en cada momento todas las entidades de que consta la naturaleza. Si esta mente tuviera una inteligencia suficiente para analizar todos estos datos, podría abarcar con una sola fórmula los movimientos de los cuerpos de mayor tamaño del universo y los de los átomos más ligeros; para ella nada sería incierto; el futuro y el pasado estarían ambos presentes ante sus ojos.

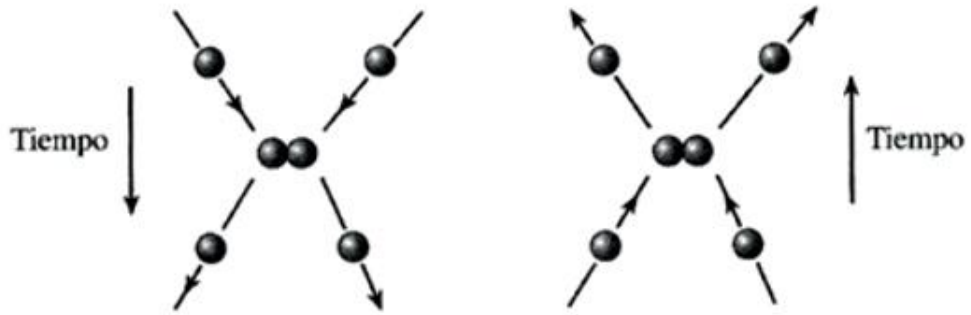

Figura 1.2. Según las leyes de Newton, las colisiones entre pares de objetos son perfectamente reversibles. La imagen resulta igualmente plausible con independencia de cómo tracemos la «flecha del tiempo».

El equivalente moderno de esta mente sería un superordenador que conociera las posiciones y las velocidades de todas las partículas del universo, y pudiera utilizar las leyes de Newton y las que describen las fuerzas de la naturaleza (como la gravedad y el electromagnetismo), no solo para predecir la trayectoria futura de cada partícula, sino para averiguar toda la historia de su procedencia, porque en las leyes de Newton no hay nada que revele la dirección del tiempo y funcionan de la misma manera si éste transcurre en sentido contrario, como podemos ver fácilmente si nos imaginamos el proceso inverso del choque entre dos bolas de billar, o si invertimos el movimiento orbital de todos los planetas del sistema solar. No hay una flecha del tiempo en las leyes de Newton y, según Laplace y muchos otros, estas leyes parecen describir un mundo completamente determinista en el cual el pasado y el futuro están fijados de una manera rígida y no hay lugar para el

libre albedrío. Lo que ninguno de estos científicos parece haber observado es que el argumento fundamental se desploma si, en cualquier momento y lugar del universo, se produce una colisión simultánea entre tres partículas, aunque la valoración de si esto sería suficiente para restablecer el libre albedrío es una cuestión cuya discusión prefiero dejar a los filósofos.

Este mismo problema relativo al tiempo se planteó a partir de uno de los mayores triunfos de la física del siglo XIX: la investigación de la naturaleza de la luz y de otras formas de radiación electromagnética, que tuvo su momento culminante en la obra del escocés James Clerk Maxwell (1831-1879). La explicación dada por Maxwell sobre la radiación electromagnética se basa en la obra de Michael Faraday, que vivió entre 1791 y 1867, y propuso la definición de los «campos» eléctrico y magnético que surgen en torno a los imanes o a objetos que poseen carga eléctrica. Un campo de este tipo puede hacerse visible mediante el clásico experimento de diseminar virutas de hierro sobre una hoja de papel colocada horizontalmente sobre un imán; cuando se le da unos golpecitos suaves al papel, las virutas de hierro se colocan por sí mismas formando unas líneas que unen los dos polos magnéticos del imán, dibujando las llamadas líneas de fuerza asociadas al campo magnético.

Fue Faraday el primero en sugerir que la luz podría estar producida por algún tipo de vibración de las líneas de fuerza asociadas con imanes y partículas cargadas, que vibrarían como lo hacen las cuerdas de un violín al ser pulsadas. Pero Faraday carecía de los conocimientos matemáticos necesarios para desarrollar esta idea hasta conseguir un modelo perfectamente configurado, siendo Maxwell quien terminó el trabajo en la década de 1860. Demostró que todos los fenómenos eléctricos y magnéticos conocidos en aquella época, incluido el comportamiento de la luz, podían describirse mediante un conjunto de sólo cuatro ecuaciones, que actualmente se denominan ecuaciones de Maxwell. Estas ecuaciones eran para las radiaciones y los campos electromagnéticos lo que las leyes de Newton para la materia sólida, en principio, lo describían todo y hacían posible el cálculo de las consecuencias de cualquier interacción electromagnética, aunque esto podía requerir una gran cantidad de paciencia si la situación era complicada. Las ecuaciones de Maxwell describen el modo en que funcionan las dinamos y los motores eléctricos, la razón por la cual la aguja de una brújula señala el norte, la

magnitud que tiene la fuerza que ejercen mutuamente dos cargas eléctricas situadas a cierta distancia (ya se sabía que tanto la electricidad como el magnetismo, al igual que la gravedad, obedecen leyes de la inversa del cuadrado), y mucho más.

Entre los dos, Newton y Maxwell aportaron el conjunto de herramientas matemáticas necesarias para controlar todo lo que la física conocía a mediados del siglo XIX. Por otra parte, lo más maravilloso de las ecuaciones de Maxwell era que, sin que se hubiera pedido, proporcionaban una descripción de la luz, las ecuaciones se crearon para describir otros fenómenos electromagnéticos, pero incluían en sí mismas una solución que describía las ondas electromagnéticas que se desplazaban por el espacio a cierta velocidad. Esta velocidad es exactamente la de la luz (que ya había quedado bien determinada en la década de 1860 y pronto podría medirse con una precisión aún mayor), no dejando lugar a dudas de que la luz se desplaza como una onda electromagnética.

Sin embargo, las ecuaciones de Maxwell tienen dos características curiosas: una de ellas pronto tendría un profundo impacto en la física, y la otra fue considerada hasta tiempos muy recientes sólo como una rareza de menor importancia. La primera característica innovadora de estas ecuaciones es que dan la velocidad de la luz como un valor constante, independientemente de cómo se mueva la fuente de luz con respecto a la persona (o aparato) que mide su velocidad. Según las ecuaciones, si alguien me ilumina con una linterna, ambos medimos la velocidad de la luz que parte de ella como un valor constante igual a c. Esto está bien si ambos estamos inmóviles. Pero si me muevo con respecto a quien me está iluminando, o él se mueve con respecto a mí, aunque lo hagamos a gran velocidad, seguiremos midiendo un valor *c* para la velocidad de la luz que emite la linterna. Esto es cierto cuando yo me acerco a la persona que tiene la linterna (se podría pensar, por sentido común, que yo tendría que medir la velocidad *c* más mi propia velocidad con relación a esa persona) y también cuando me alejo de ella (aunque el sentido común dice que en este caso tendría que medir la velocidad *c* menos mi propia velocidad con relación a ella). Esta condición de las ecuaciones de Maxwell de que la velocidad de la luz sea constante para todos los observadores, independientemente de su movimiento, fue lo que llevó a Albert Einstein a desarrollar su teoría especial

de la relatividad, publicada en 1905, y a avanzar hacia la teoría general de la relatividad, que se publicó una década más tarde.

Sin embargo, la teoría de la relatividad no tiene espacio en este libro, salvo para reconocer que existe y que explica el comportamiento de los objetos en movimiento y la gravedad de una manera más completa que la teoría newtoniana. No echa abajo, ni sustituye, a la física newtoniana, sino que engloba en sí misma los conceptos de Newton y extiende la descripción del mundo físico a condiciones para las cuales la teoría newtoniana resulta inadecuada, en particular, a los objetos que se mueven a velocidades muy grandes y a aquellos que se encuentran en campos gravitatorios muy fuertes. Sin embargo, como descripción del movimiento de la Tierra alrededor del Sol, por ejemplo, la física newtoniana tiene una exactitud de hasta una cienmillonésima, y es incluso más precisa para explicar los fenómenos que se producen a escala humana.

Las ecuaciones de Maxwell tienen también sus limitaciones, especialmente en la descripción de fenómenos que se producen a escalas muy pequeñas, tales como el comportamiento de los átomos y de las partículas que los componen. En este caso, es preciso modificar tanto la descripción clásica de las interacciones electromagnéticas (Maxwell), como la descripción clásica de las interacciones entre partículas (Newton), fenómenos en los cuales se cumplen las reglas de la física cuántica. Los efectos cuánticos, o, al menos, un efecto cuántico concreto, aparecerán en nuestra historia más tarde, aunque sólo desempeñarán un papel relativamente modesto. Casi todo lo que tratamos aquí se puede describir (y quizá entender) utilizando la mecánica newtoniana. Incluso podríamos decir que, si vale la pena mencionar las ecuaciones de Maxwell, es por otra característica suya que resulta un tanto extraña.

Esta característica extraña de las ecuaciones de Maxwell es que, como las ecuaciones de Newton, no contienen la flecha del tiempo. Esto no tiene por qué ser tan preocupante si pensamos en

cosas tales como las partículas dotadas de carga eléctrica que se mueven en los campos magnéticos e imaginamos que se invierte el sentido del tiempo. Seguirían cumpliéndose las leyes de la física descritas por las ecuaciones de Maxwell y la teoría de New- ton. Sin embargo, según estas ecuaciones, no existe diferencia

alguna entre una onda luminosa que se propaga desde una lámpara y el fenómeno correspondiente con inversión del tiempo: una onda luminosa que converge en una lámpara. Esto sería para nosotros tan extraño como ver que sobre una mesa de billar las bolas dispersas se reúnen formando claramente un triángulo, utilizando para ello la energía obtenida por rozamiento, y la bola blanca se dispara y llega al extremo del taco del jugador, dándole un golpecito. No obstante, ambas modalidades son posibles cuando se combinan las leyes de Newton y las ecuaciones de Maxwell. Está claro que sucede algo extraño con la flecha del tiempo.

Lo que fue durante mucho tiempo la explicación habitual de la razón por la que vemos una dirección predominante del tiempo surgió a partir de otro gran triunfo de la física del siglo XIX: la descripción de la relación entre calor y movimiento (termodinámica). Esto tuvo una importancia práctica fundamental en el mundo industrial cuando se utilizaba la fuerza de las máquinas de vapor. Desde nuestro punto de vista, la importancia de la termodinámica reside en el hecho de que proporciona a los físicos un modo de explicar el comportamiento de gran número de objetos, en especial, partículas de gas, que, en cierto sentido, funcionan juntos en un sistema complejo. Esto incluye el uso de promedios y estadísticas, pero se basa en gran medida en la idea de que un gas está constituido por una cantidad innumerable de partículas diminutas (átomos y moléculas) que no cesan de rebotar y chocar entre sí y con las paredes del recipiente que las contiene, cumpliendo las leyes del movimiento de Newton. Esta teoría cinética de los gases fue un ejemplo importante del modo en que las leyes universales de la física ponían orden en el caos.

La palabra «gas» fue acuñada por el físico flamenco Joannes (originalmente Jan) van Helmont a partir de la palabra griega que significa «caos»; este término apareció impreso por primera vez en el libro de Van Helmont titulado *Ortus Medicinae,* publicado cuatro años después del fallecimiento de Joannes, en 1648, por su hijo Franciscus. La idea de que los gases eran como un caos se consideró acertada durante trescientos años, hasta que Maxwell, desde Gran Bretaña, y su contemporáneo Ludwig Boltzmann, desde Viena, consolidaron la teoría cinética (que hasta entonces había sido sólo una especulación), dándole una firme base científica fundamentada en las leyes de Newton. Por ejemplo, la presión que ejerce un gas

sobre las paredes del recipiente que lo contiene se explica en términos de acción y reacción (tercera ley de Newton), fenómeno que se produce cuando las partículas del gas chocan con la pared y rebotan alejándose de ella e internándose de nuevo en el volumen del gas.

Esto produce lo que se percibe como una presión constante, sencillamente porque es tal la cantidad de partículas, que el número de colisiones por segundo resulta enorme. En números redondos, la cantidad de moléculas de gas que caben en un recipiente del tamaño de una caja de cerillas es aproximadamente un billón de billones (la unidad seguida de 24 ceros). Este número es tan grande que realmente no hay mucha diferencia si pensamos en una caja diez veces mayor o diez veces menor. Una molécula contenida en el aire, que esté en la atmósfera al nivel del mar y a una temperatura de 0 °C, experimenta cerca de cuatro millones de colisiones por segundo[7]. Todo esto es lo que nos da la sensación de que el aire es un medio continuo y significa también que, dejando a un lado lo que Laplace pudiera haber imaginado, sería completamente inútil intentar determinar el comportamiento de un gas calculando las trayectorias de cada molécula según las leyes de Newton, estamos ante un problema de *N* cuerpos donde *N* es igual a un billón de billones. Aquí es donde la estadística entra en escena.

La disciplina que trata la aplicación de las leyes de la mecánica de una manera estadística para describir el comportamiento de grandes números de moléculas y átomos contenidos en un gas (o algún otro sistema) se conoce como mecánica estadística. La fuerza que impulsó su desarrollo en el siglo XIX fue la necesidad de encontrar una explicación científica de los fenómenos termo- dinámicos observados, no sólo en los experimentos, sino también en la vida cotidiana, por ejemplo, el hecho constatado de que sin cierto tipo de intervención exterior el calor fluye siempre desde un objeto más caliente a otro más frío. En el lenguaje coloquial se dice que los objetos calientes se enfrían. En el mundo real, si colocamos un cubito de hielo sobre una superficie caliente, el hielo se funde a medida que se calienta; nunca vemos que el agua en estado líquido se convierta de manera espontánea en cubos de hielo y emita calor mientras lo hace, aunque, como en el ejemplo de las bolas de billar que se colocan ellas mismas formando un triángulo, por lo que

[7] Se utiliza el billón entendido como un millón de millones (10^{12}).

respecta al movimiento de átomos y moléculas individuales, parece que esto pudiera estar justificado por las leyes de Newton. La realidad cotidiana lleva instalada una flecha del tiempo y parece estar estrechamente vinculada con las leyes de la termodinámica y la mecánica estadística.

Estas leyes se desarrollaron a partir de observaciones del modo en que suceden las cosas en el mundo real. Esto equivale a comenzar a partir de la observación de que las manzanas siempre caen hacia abajo desde los árboles, nunca hacia arriba, y, adoptar esto como una ley de la naturaleza, para medir luego a qué velocidad descienden, cómo se aceleran, extendiendo las observaciones a otros objetos (como la Luna y Marte), y hallando finalmente las ecuaciones matemáticas que describen este comportamiento, en este caso la ley de la gravedad. Para la termodinámica, el punto de partida fue una investigación sobre el modo en que el calor fluye a través de objetos sólidos, realizada por Joseph Fourier en Francia durante los años anteriores a 1811, que es cuando ganó un premio de la Academia Francesa de las Ciencias por este trabajo.

Fourier descubrió una sencilla ley matemática que explicaba esta transferencia de calor, el flujo de calor es proporcional a la diferencia de temperaturas, fluyendo siempre dicho calor (por supuesto) desde el extremo más caliente de un objeto hasta su extremo más frío. En el desarrollo de la termodinámica, esta ley tan sencilla ocupa un lugar equivalente al que ocupa en el desarrollo de la teoría de la gravedad la ley de Galileo relativa a la aceleración de un cuerpo en caída libre. Además, al igual que la ley de la gravedad, la ley de Fourier es universal, de hecho, no sólo se cumple para los sólidos, sino también para líquidos y gases, siendo independiente de la materia de que estén hechos los objetos, aunque (a diferencia de la aceleración debida a la gravedad) la constante de proporcionalidad que interviene en la relación es diferente para las distintas sustancias químicas. El calor fluye con mucha mayor rapidez a través de los metales que, por ejemplo, a través de la madera, algo que han aprendido de primera mano la mayoría de las personas a partir de experiencias dolorosas.

Aparte de su importancia directa para el estudio de los flujos de calor, el descubrimiento de esta sencilla ley por parte de Fourier proporcionó un conocimiento profundo de la naturaleza del mundo físico. A nivel de átomos y

moléculas, sería un esfuerzo inútil intentar predecir las propiedades globales de un objeto lo suficientemente grande como para ser manipulado a escala humana (los llamados objetos macroscópicos); las interacciones entre todas las partículas son tan complicadas que desafían cualquier análisis directo y representan (tal como precisó Van Helmont) algo muy próximo a lo que se suele llamar vulgarmente caos.

Sin embargo, con miles y miles de millones de moléculas que participan en interacciones mutuas unas con otras, el caos en cierto modo desaparece, o al menos se aminora, y aparece (o reaparece) un orden que está asociado con leyes sencillas. Nadie sabía cómo podía suceder esto. Era conocido el hecho de que unas pocas partículas que cumplieran las leyes de Newton podían describirse de una manera sencilla (incluso cuando la resolución de las ecuaciones supusiera en la práctica la aplicación de unas tediosas aproximaciones técnicas), y también era sabido que un número enorme de partículas que actuaran conjuntamente podían ser descritas asimismo sin entrar en complicaciones. Pero lo que no se sabía era cómo pasar de un nivel a otro, y esto tampoco tuvo importancia mientras se pudo desarrollar la termodinámica como ciencia aplicada en la era de la máquina de vapor.

Se tardó muchas décadas, la mayor parte del resto del siglo XIX, en completar este desarrollo, pero se fue logrando una descripción del modo en que se comportan los objetos en lo relativo a sus características macroscópicas, tales como la temperatura, la presión, la densidad, e incluso su comportamiento químico, y también por lo que respecta a cómo responden estas características ante los cambios impuestos al sistema desde el exterior, como un aumento de la presión o un descenso de la temperatura. No se intentó predecir estos cambios en términos de comportamiento de las moléculas y los átomos, aunque la interpretación de lo que sucedía sí reconocía la existencia de dichas partículas y utilizaba los promedios de sus características de una manera estadística.

Un importante trabajo precursor del de Fourier fue el realizado por el conde Rumford en Baviera durante la década de 1790. Rumford, que había nacido en lo que entonces era la colonia británica de Massachusetts en 1753 y vivió hasta 1814, se llamaba en principio Benjamín Thompson, pero recibió el título de conde de manos de Jorge III de Baviera, al que prestó servicios como consejero durante una

pintoresca carrera como soldado, estadista, espía y benefactor público. Fue mientras trabajaba en Baviera, supervisando la fabricación de un nuevo cañón para el ejército, cuando planteó la idea de que el calor es una forma de trabajo. El cañón fue perforado con una herramienta accionada por la fuerza de unos caballos y, cuando los caballos trabajaban más, aumentaba la temperatura de la herramienta de perforar y del hueco que se estaba perforando en el cañón. Una máquina de vapor convierte el calor en trabajo; el proceso de perforación de un cañón convertía el trabajo en calor, como subproducto de la fabricación del cañón.

La gestación de estas ideas duró mucho tiempo. Un paso decisivo fue el que dio James Joule en Gran Bretaña durante la década de 1840. Lo hizo todavía mejor que Rumford, ya que realizó experimentos precisos para medir la cantidad de trabajo necesario para producir un determinado ascenso de la temperatura en una cierta cantidad de agua, uno de estos experimentos, concebido de una forma maravillosamente sencilla, consistía en calentar agua en un recipiente revolviéndola con una especie de rueda de paletas, que se accionaba mediante una cuerda atada a un peso que caía. Este trabajo y los de algunos contemporáneos de Joule, tales como Hermann Helmholtz en Alemania, desembocaron en el principio de la conservación de la energía. Éste dice que la energía no puede crearse ni destruirse, sólo puede transformarse pasando de una forma a otra.

La energía que el caballo necesita para hacer su trabajo procede del heno que come, de tal forma que el combustible químico que éste le proporciona se combina con el oxígeno para dar potencia a los músculos; la energía almacenada en el heno en forma de sustancia química procede en última instancia de la luz del Sol; y así sucesivamente. El principio de conservación de la energía llegó a conocerse también como primer principio de la termodinámica y dice que en un sistema cerrado (un sistema que no interacciona en modo alguno con el mundo exterior, lo cual es otra idealización de los físicos, como el plano sin rozamiento) la energía total permanece constante. Pero, como muestra el ejemplo de la perforación del cañón, ninguna transformación de trabajo en energía es perfecta, ya que el calor siempre se disipa como un subproducto, de tal forma que hay algo de energía que desaparece de la circulación. Dado que el calor siempre fluye de un lugar más caliente a otro más frío (segundo principio de la termodinámica), finalmente, en cualquier sistema cerrado

toda la energía acaba convirtiéndose en calor, y todas las diferencias de temperatura se irán nivelando hasta dejar un sistema templado y sin características especiales donde no sucede nada interesante.

Lo que actualmente se conoce como segundo principio de la termodinámica se puede expresar de muchas formas diferentes, pero su primer enunciado se debe al físico británico William Thomson (quien fuera posteriormente lord Kelvin) en 1852. La cuestión esencial sobre la que Thomson llamó la atención es esta idea de disipación, que, aunque el modo en que funciona el mundo natural se puede describir como un gran motor que convierte el calor en trabajo (o en movimiento, que viene a ser lo mismo), debe haber siempre algo de calor que se disipa durante el proceso, aunque realmente no se pierde, sino que se propaga por todo el universo, haciendo que su temperatura global suba una pizca, una cantidad casi imperceptible.

Esto va más allá del principio, o ley, de conservación de la energía (el primer principio de la termodinámica), porque en este caso, aunque la cantidad total de energía del mundo (expresión con la que los Victorianos se referían a lo que actualmente llamaríamos el universo) se mantiene siempre igual, la cantidad de energía *útil* siempre está disminuyendo. Esto implica que los físicos necesitaban un método para cuantificar la cantidad de energía útil existente en un sistema cerrado, o en el mundo (el universo en toda su amplitud), de tal modo que pudieran tenerla en cuenta y manejarla en sus ecuaciones. Esto indujo a Rudolf Clausius a proponer el concepto de entropía, lo cual hizo en Alemania a mediados de la década de 1860.

El modo más sencillo de entender lo que mide la entropía es pensar en términos de la cantidad de orden que hay en un sistema, y el ejemplo clásico consiste en imaginarse una caja que está dividida en dos mitades mediante una pared separadora móvil. Una mitad de la caja está llena de gas y la otra se encuentra inicialmente vacía, es el vacío. Tenemos así un sistema que posee una cierta cantidad de orden, o de estructura, porque hay una diferenciación entre las dos mitades del recipiente.

Si se introduce al azar un robot consistente en una sonda microscópica, nos podrá decir en qué lado de la pared separadora se encuentra, comprobando si está rodeado por el gas o por el vacío. Imaginemos ahora que abrimos esa pared

separadora. Todos sabemos por experiencia lo que va a suceder. El gas se propaga hasta llenar la caja de manera uniforme.

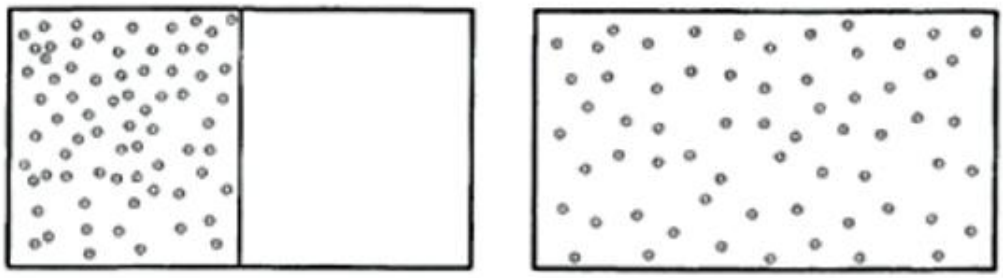

Figura 1.3. Cuando observamos un gran número de partículas, como por ejemplo los átomos de un gas que hay en el interior de una caja, la dirección del tiempo resulta obvia. Si se elimina una separación, el gas se propaga hasta llenar toda la caja, y no tenemos dificultad alguna para distinguir cuál es el «antes» y cuál el «después», aunque no se represente una flecha del tiempo.

Entonces habrá en el sistema menos orden (o, si se quiere, más desorden), porque desde el interior de la nube de gas es imposible decir en qué mitad de la caja estamos. La temperatura, de hecho, también habrá descendido al expandirse el gas cuando entra en el vacío. Ahora, ciertamente, podemos restablecer el estado de orden original del gas que hay en la caja, empujando todo el gas hacia un extremo de la caja mediante un pistón. Esto calentará el gas, que recuperará su temperatura original al verse comprimido. Sin embargo, para lograr esto, el pistón debe realizar un trabajo, y durante el proceso es inevitable que se desprenda calor a causa del rozamiento. Por lo tanto, aunque la caja, que es un sistema cerrado, vuelva a su estado original, como resultado se ha perdido más calor en la totalidad del universo, por lo que el mundo habrá cambiado.

Ampliando ligeramente esta analogía, un tablero de ajedrez pintado con cuadrículas alternativamente negras y blancas tiene un cierto orden. Si se mezcla exactamente la misma cantidad de pintura y se utiliza para pintar el tablero de un tono gris uniforme, habrá menos orden. Podemos imaginar que existe una máquina lo suficientemente hábil como para deshacer la mezcla de moléculas de pintura y separar de nuevo la pintura blanca y la pintura negra con sus colores puros, pero,

una vez más, la máquina no puede ser eficiente al 100 por cien y tiene que emitir calor al universo mientras trabaja.

La entropía mide la cantidad de orden que hay en un sistema y, si el desorden aumenta, también lo hace la entropía. Sabiendo que en el mundo real el desorden crece en todo sistema cerrado (las cosas se desgastan) a medida que pasa el tiempo, el inevitable aumento de la entropía define una dirección del tiempo, una flecha que parte del pasado ordenado y apunta hacia el futuro desordenado. Dado que este proceso parecía inevitable y universal, los especialistas en termodinámica de la era victoriana preveían un destino último del universo en el que toda la energía útil se habría convertido en calor y todo sería una mezcla templada de materia a temperatura uniforme, una situación desoladora que llamaban la «muerte térmica» del universo[8].

La vida, por supuesto, parece desafiar este proceso creando orden y estructuras a partir de materiales desordenados (o, en todo caso, menos ordenados). Una planta construye su estructura, y puede fabricar flores de gran belleza, a partir de dióxido de carbono, agua y unos pocos restos de otros productos químicos. Pero sólo puede hacerlo con ayuda de la luz solar, es decir, con energía procedente de una fuente externa. La Tierra, y en particular la vida que se desarrolla sobre ella, no es un sistema cerrado. Es posible demostrar, utilizando las ecuaciones desarrolladas por Thomson, Clausius y sus contemporáneos, que, en cualquier lugar del universo donde aparece un foco de orden, esto se hace a costa de que se produzca *más* desorden en algún otro lugar.

Cuando en un frigorífico utilizamos el compartimento del congelador para producir cubos de hielo, desafiando aparentemente la segunda ley de la termodinámica, si podemos hacerlo es gracias a que está funcionando una máquina que bombea fluidos alrededor del frigorífico y genera más calor que el que «pierde» el agua al congelarse. El proceso de enfriamiento, que tiene lugar en los tubos situados en el interior de la nevera, es esencialmente igual que la dilatación y el enfriamiento que

[8] Este tenebroso pronóstico ha quedado ya descartado. El hecho de que el universo se expande (que no se descubrió hasta finales de la década de 1920) altera todo el contexto de esta predicción, y la constatación de que la gravedad tiene de hecho energía negativa, que data de la década de 1940, descarta en esencia el tipo de muerte térmica que se imaginaron los Victorianos. Estas curiosas ideas se comentan en mi libro *In the Beginning* (Penguin, 1994) [Hay traducción al castellano: John Gribbin, *En el principio: el nacimiento del universo viviente,* Alianza Editorial, Madrid, 1994, [trad. Jesús Unturbe] y en el de Paul Davies, *The Cosmic Blueprint* (Heinemann, 1987) [Hay traducción al castellano: Paul Davies, *Proyecto cósmico: nuevos descubrimientos acerca del orden del universo,* Ediciones Pirámide, 1989, trad. Manuel Martínez!; de lodo esto hablaremos con detalle más adelante.

se producen en el gas de nuestra caja imaginaria cuando se suprime la separación en compartimentos; el proceso de calentamiento que se produce en los conductos exteriores del frigorífico, en su parte posterior, es esencialmente el resultado de la compresión que se aplica al gas para restaurar su volumen inicial y que hace que este gas se caliente. Después, el calor se dispersa en el aire antes de que el fluido regrese a los conductos interiores del frigorífico.

Si se deja un frigorífico con la puerta abierta y funcionando en una habitación sellada con paredes perfectamente aisladas, la habitación, en vez de enfriarse, se calentará cada vez más a causa del trabajo que realiza el motor para mantener en marcha el proceso de enfriamiento dentro de la nevera.

A escala macroscópica, según unas leyes deducidas a partir de experimentos y observación siguiendo procedimientos científicos aprobados, ensayados y comprobados, el universo actúa de un modo irreversible. Nunca se puede hacer que las cosas vuelvan a ser como solían. Pero precisamente en nuestro sencillo y clásico ejemplo de irreversibilidad termodinámica, la entropía y la flecha del tiempo, la caja de gas con la pared separadora móvil, vemos con gran claridad la dicotomía aparente entre el mundo macroscópico y el mundo microscópico. A nivel de los átomos y las moléculas que componen el gas (en realidad un nivel sub-microscópico, pero nadie lo tiene en cuenta), toda colisión es, según las leyes de Newton, perfectamente reversible, como los choques entre las bolas de una mesa de billar. Podemos imaginamos que retiramos la pared separadora, dejando que el gas llene toda la caja, y luego movemos una pared mágica para invertir el movimiento de cada átomo y cada molécula del gas. En este modelo de movimiento inverso no habría nada que estuviera prohibido por las leyes de Newton. Obedeciendo ciegamente estas leyes, los átomos y las moléculas recorrerían su camino inverso para volver a quedarse en una de las mitades del recinto, con independencia del número de colisiones que sufrieran durante este proceso. Sin embargo, en el mundo real nunca vemos que los sistemas actúen de esta manera, por ejemplo, que todo el gas que hay en una habitación se desplace de repente a uno de sus extremos. Esta dicotomía entre las leyes que se cumplen a gran escala y las que se cumplen a pequeña escala fue un enorme quebradero de cabeza para la física a finales del siglo XIX.

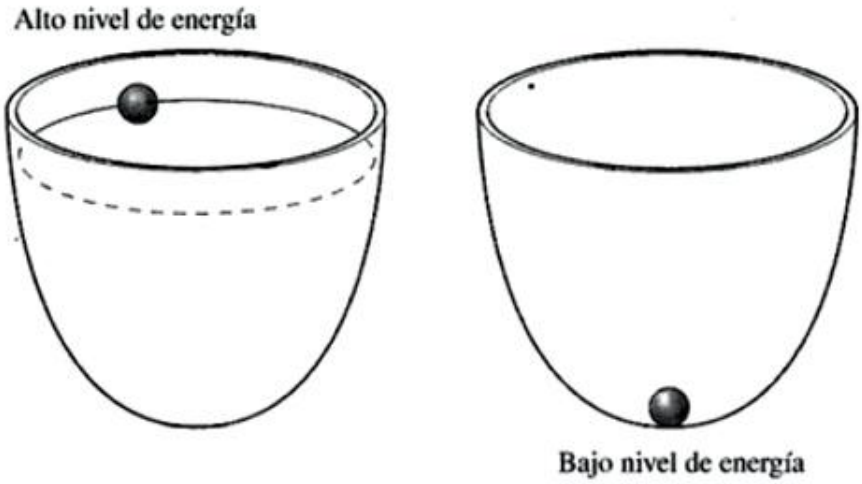

Figura 1.4. Si se les deja actuar por sí mismos, los sistemas (incluso un «sistema» tan simple como una canica y un cuenco) tienden a descender a un estado de energía mínima y entropía máxima, siempre que no haya aportación de energía desde el exterior.

Los que intentaron resolver este enigma tuvieron que desarrollar un nuevo lenguaje (como ya hemos empezado a ver) y un nuevo modo de reflexionar sobre el mundo físico. Uno de los conceptos clave, fundamental para la historia que hemos de contar aquí, es el de *atractor.* Si se libera un gas en el interior de una caja, ya sea desplazando una pared separadora, tal como hemos explicado, o bien mediante un orificio en una pared de la caja, el resultado final es que el gas se distribuye de una forma equilibrada, se expande uniformemente por la caja.

Este estado corresponde a la entropía máxima que puede conseguirse en el sistema, y es el mismo con independencia del modo en que el gas entró en la caja, no importa, por ejemplo, cuál es la pared de la caja en la que está el orificio que permitió la entrada del gas, o en qué parte de la pared se perforó. Se dice que el estado final de equilibrio (que corresponde también al estado de energía mínima) es un atractor, porque el sistema actúa como si se sintiera atraído hacia dicho estado. Una vez que se alcanza, no hay modo de decir cómo ha llegado el sistema a ese estado, en el estado final de equilibrio no está escrita su historia.

Otros sistemas que nos resultan más familiares pueden describirse también de esta

manera. Si se deja rodar una canica dentro de un cuenco, después de dar unas vueltas y realizar unas cuantas oscilaciones la canica acabará inmóvil en el fondo del cuenco; ese estado es un atractor para este sistema.

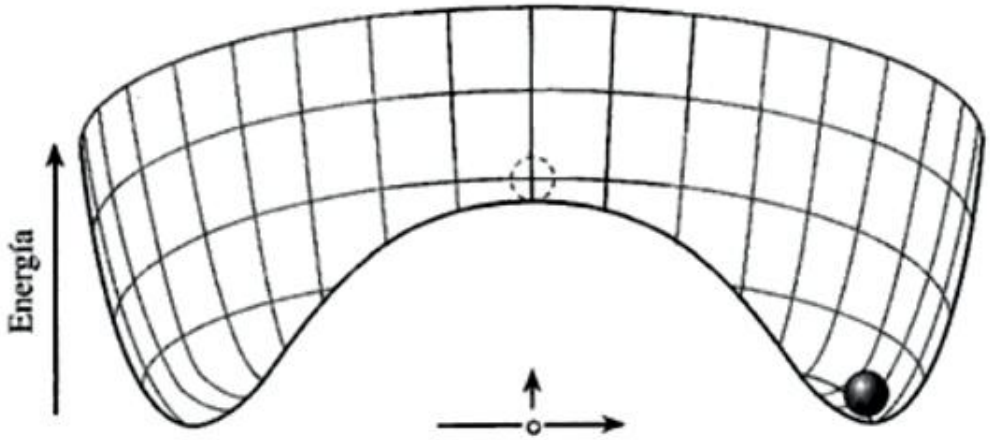

Figura 1.5. El estado en que se estabilizan los sistemas se denomina un atractor. En el ejemplo que se muestra en la página anterior, el atractor es un solo punto en el fondo del cuenco. Sin embargo, un atractor puede ser también toda una amplia región, como se puede ver en esta ilustración. La canica que se encuentra en la parte más alta está abocada a caer al valle, pero todo lo que está en el fondo del valle posee el mismo poder de atracción.

Sin embargo, incluso en un sistema así de simple, el atractor no tiene por qué ser tan sencillo. Si se hace rodar la canica en el ala vuelta hacia arriba de un sombrero de ala ancha, terminará parándose en algún lugar del valle que forma el ala vuelta hacia arriba, pero todos los puntos situados en un círculo que recorre el fondo de dicho valle forman parte del atractor (llamado a veces el «atractor del sombrero mejicano»), ya que todos ellos corresponden al mismo estado de energía mínima.

Para un péndulo perfecto imaginario, sin rozamiento (lo que se llama un péndulo ideal), oscilar a un lado y a otro es un atractor; para un péndulo real, en el que el rozamiento gasta energía, la oscilación se hará cada vez más lenta y finalmente se parará, quedando el péndulo colgado verticalmente; también esto puede considerarse un estado atractor. No importa cómo iniciemos la oscilación del péndulo (de izquierda a derecha, adelante y atrás, o describiendo círculos); si

esperamos, se detendrá en el mismo estado final, sin nada que indique cómo llegó allí. En el lenguaje de la termodinámica se dice que cuando un sistema alcanza el equilibrio, ha olvidado sus condiciones iniciales. Lo único que importa es dónde se encuentra entonces.

Pero ¿cuándo alcanza un sistema el equilibrio? En el mundo real no existe un aislante perfecto, por lo que nuestra caja de gas siempre estará intercambiando calor con el mundo exterior. El péndulo oscilante, que chirría para detenerse, podría ser uno de los mejores ejemplos de un sistema próximo al equilibrio, pero hasta este sistema se ve perturbado por las moléculas de aire que chocan con él, e incluso si lo situamos en un recinto donde se ha hecho previamente el vacío, habrá una cierta interacción con el mundo exterior a través de la cuerda del péndulo, que ha de estar atada a algo. La verdad es que no existe nada que pueda considerarse un sistema cerrado (salvo la totalidad del universo), y ningún sistema se encuentra en un equilibrio perfecto. Puede llegar muy cerca de esa situación de equilibrio, tanto como se quiera, si se espera el tiempo suficiente, pero el equilibrio nunca es del todo perfecto.

No se trata de mera pedantería. Tomando un ejemplo del libro de Ilya Prigogine e Isabelle Stengers[9], si se unen dos recipientes mediante un tubo relativamente estrecho y lleno de una mezcla de hidrógeno y ácido sulfhídrico, ambos en estado gaseoso, entonces, en situación de equilibrio, si está todo a la misma temperatura, habrá una mezcla uniforme de ambos gases en cada recipiente. Sin embargo, si uno de los recipientes se mantiene ligeramente más caliente que el otro, la mezcla gaseosa comienza a separarse, concentrándose las moléculas de hidrógeno, que son más ligeras, en el recipiente que está a mayor temperatura, mientras que las moléculas de ácido sulfhídrico, que son más pesadas, se concentran en el recipiente que está más frío (algo similar sucede con cualquier par de gases de diferentes pesos moleculares, no sólo con el hidrógeno y el ácido sulfhídrico). En este caso, una pequeñísima desviación con respecto a la situación de equilibrio produce orden a partir del caos. Cuando se pierde el equilibrio, un flujo de energía puede crear orden de manera espontánea, siempre que se den las circunstancias adecuadas.

[9] *Entre le temps et Véternité,* Fayard, París, 1988. [Flay traducción al castellano: *Entre el tiempo y la eternidad,* Alianza Editorial, Madrid, 1994, [trad. de Javier García Sanz.] Prigogine falleció en 2003, justo cuando se comenzaba a imprimir este libro.

Esto nos da una idea crucial de lo que sucede con nuestra propia existencia, ya que no se puede negar que somos criaturas ordenadas, y hay pruebas abrumadoras de que el universo comenzó a existir en un estado de mucho menos orden.

En general, un sistema que esté cercano al equilibrio será atraído hacia un estado en el cual *es mínima la velocidad a la que se produce la entropía*, pero esta idea no llegó a estar del todo clara hasta bien entrado el siglo XX. Por la misma razón que los físicos, después de Newton, se centraron primero en resolver problemas sencillos que podían tratarse utilizando ecuaciones y leyes del movimiento newtonianas (porque eran los problemas que resultaban más fáciles de resolver), los expertos en termodinámica se centraron al principio en el estudio de sistemas en equilibrio y en el desarrollo de leyes y ecuaciones que sirvieran para describirlos (como la segunda ley de la termodinámica).

Fue este planteamiento el que condujo al desarrollo de una forma de entender la termodinámica en términos de probabilidad, ampliando este enfoque bajo la denominación «mecánica estadística».

Una gran parte de estos trabajos se iniciaron a partir de los intentos de describir de forma matemática el comportamiento de los gases, porque éstos son sistemas relativamente simples en los que se producen colisiones entre moléculas siguiendo más o menos las leyes newtonianas, con lo que había una cierta esperanza de poder resolver las ecuaciones relativas a estos casos. Lo que llegaría a conocerse como teoría cinética de los gases (porque la palabra «cinética» implica movimiento, en este caso de las moléculas que componen el gas) se desarrolló gracias a los esfuerzos de varias personas cuyas ideas se enriquecieron mutuamente durante la segunda mitad del siglo XIX.

Uno de los principales argumentos de esta historia comienza con Clausius, que planteó en 1858 la idea de la «trayectoria libre media», que es la media de las distancias que recorre una molécula entre las colisiones que sufre dentro de un gas a una temperatura y una presión determinadas. También propuso la idea de que las moléculas tienen un «radio efectivo», de tal modo que se comportan como unas pequeñas bolas rígidas que poseen ese radio. Estas ideas fueron asumidas y desarrolladas por Maxwell, que fue aún más lejos que Clausius al considerar que hay toda una gama de velocidades entre las moléculas que intervienen en las

colisiones. Demostró cómo, si se hace una elección adecuada de todas estas propiedades, muchas de las características observadas en el comportamiento de los gases se podrían describir mediante unas ecuaciones relativamente sencillas. A su vez, la obra de Maxwell proporcionó la inspiración para posteriores creaciones desarrolladas por Ludwig Boltzmann en Alemania. Boltzmann fue el que compaginó todo en la primera versión de la mecánica estadística. Al otro lado del Atlántico, en Estados Unidos, Josiah Willard Gibbs realizó también una importante contribución a esta nueva rama de la ciencia.

No necesitamos entrar aquí en todos los detalles de la mecánica estadística, pero podemos ofrecer al lector una idea de lo que es el concepto clave de probabilidad, tal como lo aplicó Boltzmann, utilizando un ejemplo muy sencillo (un «modelo» mental) referido al gas que está en el interior de una caja, del que ya hemos hablado antes. Recordemos que el enigma es por qué, si las leyes newtonianas de la mecánica no lo prohíben, nunca vemos cómo se desplaza todo el gas para entrar en una mitad de la caja. ¿Qué es lo que tiene de especial el estado en el cual un gas se reparte más o menos uniformemente para llenar la caja?

Con el fin de hacernos una idea de lo que está pasando, imaginemos primero que tenemos que poner en la caja sólo dos partículas. Se desplazan según las leyes de Newton, rebotando la una contra la otra y también contra las paredes de la caja. Si tomamos una instantánea del interior de la caja en un momento dado, ¿cómo podrían estar situadas las dos partículas? Hay dos maneras de tener una partícula en cada mitad de la caja, la partícula A en la mitad de la izquierda y la B en la derecha, o B en la izquierda y A en la derecha. Y hay dos modos en los que ambas partículas pueden estar en una sola mitad, las dos a la izquierda o las dos a la derecha.

Por consiguiente, sólo hay cuatro estados posibles en los que puede encontrarse la caja, utilizando este criterio. Si cada estado tiene el mismo número de probabilidades (y no hay razón para pensar que pueda ser de otra manera), entonces hay un 25 por ciento de probabilidades de ver ambas partículas a la izquierda, un 25 por ciento de verlas a la derecha, y un 50 por ciento (2 x 25 por ciento) de ver cada una en su lado. Con estas probabilidades, al echar un vistazo rápido a la caja, no nos sorprenderemos de ver alguna de estas posibilidades en ese

preciso momento.

Ahora realizaremos el mismo proceso con cuatro partículas. Para que sea más sencillo, nos centraremos esta vez en una sola de las dos mitades de la caja. Hay cuatro maneras de tener una partícula a la izquierda y todas las demás a la derecha (tomando cada una de las partículas A, B, C y D en cada caso como la partícula que queda aislada). Sin embargo, hay seis maneras de tener dos partículas en la parte izquierda de la caja (AB, AC, AD, BC, BD, CD).

Es ligeramente más probable, con una probabilidad de 4:6, que veamos el mismo número de partículas en cada mitad de la caja, que ver sólo una partícula, y sigue habiendo un solo modo de no tener partícula alguna a la izquierda. La probabilidad aumenta radicalmente a favor de una distribución uniforme cuando aumenta el número de partículas.

El lector podrá hacer los cálculos si lo desea, pero para sólo ocho partículas sigue habiendo un único modo de que todo el «gas» esté en la parte derecha de la caja, y sin embargo hay setenta maneras de distribuir las partículas uniformemente entre las dos mitades. Boltzmann argumentó que nos resulta más fácil ver cómo se dispersa el gas para llenar una caja, que imaginarlo apiñándose en un rincón, porque es mucho más probable el primer caso que el segundo. Esto último no es imposible, sólo es bastante improbable.

El sencillo esquema que acabamos de esbozar realmente no hace justicia a la obra de Boltzmann, que incluye la deducción de una ecuación matemática en la que se relaciona la entropía con la probabilidad[10], y sitúa la mecánica estadística sobre una adecuada base cuantitativa. Pero esperamos que este esbozo sirva para dar una idea de lo que sucede.

Por supuesto, la cuestión fundamental es cuántas partículas intervienen en las colisiones que se producen en una pequeña caja llena de gas, y lo improbable que es que, como resultado, acaben todas en un rincón de la caja. No es casualidad que el primer conocimiento real de las cifras correspondientes surgiera también durante la década de 1860, ya que todas estas ideas se alimentaban mutuamente. El hilo del asunto se remonta a la obra del italiano Amadeo Avogadro, quien sugirió en 1811 que volúmenes iguales de gas a la misma temperatura y presión contienen un

[10] Boltzmann estaba tan orgulloso de su ecuación, $S = k \log P$, que ésta aparece grabada en su lápida.

mismo número de partículas. Fue el autor de la idea de que una caja llena de cualquier gas a una temperatura y una presión concretas siempre contiene el mismo número de pequeñas esferas rígidas (actualmente las equipararíamos con moléculas) que rebotan por todas partes y chocan entre sí y con las paredes de la caja. No importa que el gas sea oxígeno, dióxido de carbono, o cualquier otro, o incluso, como el aire que respiramos, una mezcla de varios gases; lo único que importa en este contexto es el número de moléculas que intervienen en el asunto.

Prácticamente no sorprende que tardara largo tiempo en aparecer alguien que planteara una cifra precisa para determinar el número de moléculas necesarias en este modelo, en todo caso, lo sorprendente es que no tardara más. En realidad hay varias maneras de abordar el problema, pero aquí sólo expondremos la que utilizó el austríaco Joseph Loschmidt a mediados de la década de 1860 para calcular el número de moléculas contenidas en un centímetro cúbico de gas a una temperatura de 0 °C y a la presión atmosférica normal (el resultado se conoce como número de Loschmidt).

Partió de la hipótesis de que hay mucho espacio vacío dentro de un gas y cada molécula ocupa un volumen especificado en términos de su radio efectivo. La presión se podía calcular a partir de la teoría cinética, utilizando la trayectoria libre media, la velocidad media de las moléculas, y cosas por el estilo. Si en el gas hubiera unas pocas moléculas grandes, la trayectoria libre media tendría un valor alto; si hubiera muchas moléculas de menor tamaño, sería menor el valor de la trayectoria libre media. Para que el modelo concordara con el modo en que cambia la presión en un gas real, sólo valdría un valor de la trayectoria libre media y, por consiguiente, un solo valor del número total de partículas presentes. El valor que se da actualmente al número de Loschmidt es $2{,}687 \times 10^{19}$ (10^{19} es la unidad seguida de 19 ceros o diez millones de billones).

Recordemos que éste es el número de moléculas que contiene cada centímetro cúbico de aire al nivel del mar un frío día de heladas, lo cual nos da también una idea de lo pequeños que son los átomos y las moléculas. Para calibrarlo mejor, pensemos que hay alrededor de 300.000 millones (300×10^9) de estrellas brillantes en nuestra galaxia de la Vía Láctea, y el número de galaxias brillantes que resultan visibles mediante nuestros telescopios es aproximadamente el mismo. Por lo tanto,

multiplicando este número por sí mismo, resulta que hay alrededor de 90.000 x 10^{IX} o 9.000 x 10^{19} estrellas brillantes en todo el universo visible. Dividiendo esta cantidad por el número de Loschmidt, obtenemos que en números redondos 450 centímetros cúbicos de aire frío, menos de medio litro[11], contienen tantas moléculas como estrellas hay en el universo.

Estas moléculas están tan cerca unas de otras que la trayectoria libre media entre colisiones, si las moléculas se mueven a una velocidad media de 460 metros por segundo, mide sólo 13 millonésimas de metro. Aplicada de manera adecuada a unos números tan grandes de moléculas, la estadística nos dice que tendríamos que esperar un tiempo mucho mayor que la edad del universo para tener una buena oportunidad de ver apiñado en un extremo todo el gas que contiene un recipiente del tamaño de una caja de cerillas.

Sin embargo, esto no significa que un suceso así sea imposible. Boltzmann pensó que había encontrado un modo de demostrar que es imposible y de introducir la flecha del tiempo en la termodinámica. Pero resultó que estaba equivocado. Boltzmann planteó unas ecuaciones que describían el comportamiento global de una caja de gas, utilizando una media estadística adecuada de las propiedades de todas y cada una de las moléculas, en vez de intentar seguir el rastro de las interacciones de cada molécula contenida en la caja.

Traduciendo esto al lenguaje cotidiano, se puede decir que sus ecuaciones explican el modo en que las moléculas de gas de la caja alcanzan un promedio a medida que transcurre el tiempo. Por ejemplo, cuando el gas entra por primera vez en la caja, algunas moléculas estarán moviéndose a mayor velocidad que otras, lo cual se corresponde con el hecho de que tienen una temperatura más elevada, de tal modo que el gas que se encuentra en un rincón de la caja puede estar más caliente que el resto. Pero, cuando una molécula de movimiento rápido choca con otra cuyo movimiento es más lento, la energía del movimiento de las dos moléculas (su energía cinética) tiende a repartirse entre ambas, de tal manera que al final tenemos dos moléculas que se mueven a una velocidad media (pensemos en una bola de billar de movimiento rápido que choca con otra bola que se encuentra en reposo, sin realizar giro alguno).

[11] Son mil los centímetros cúbicos que hay en un litro, no cien.

Por consiguiente, las moléculas más calientes se enfriarán y las moléculas más frías se calentarán, hasta que todo el gas de la caja esté a una temperatura uniforme, logrando así el equilibrio termodinámico. Éste sería un proceso rápido y eficiente, ya que toda molécula que se encuentra en el aire a 0 °C experimenta algo menos de cuatro millones de colisiones por segundo. Las ecuaciones planteadas por Boltzmann mostraban procesos naturales aleatorios que producían el equilibrio térmico en dirección única.

Así pues, ¿por qué vivimos en un mundo ordenado? Boltzmann llegó finalmente mucho más lejos que cualquier otro al formular la hipótesis de que la muerte térmica del universo en toda su extensión ya se había producido, y que, de esta manera, el universo debía estar en equilibrio. Entonces, ¿cómo podía existir un lugar como la Tierra, que estaba lejos de alcanzar el equilibrio y que sustentaba la vida? Su respuesta fue que todo el espacio que podemos ver e investigar con nuestros telescopios (que, a finales del siglo XIX, no era más que la zona local de las islas de estrellas que llamamos Vía Láctea) debe representar (a escala del universo) una pequeña fluctuación local que se aparta del equilibrio.

En nuestra hipotética caja de gas, cuando éste se encuentra en equilibrio, sólo por casualidad, podría haber ocasiones en que son más las moléculas que se mueven de izquierda a derecha, que aquellas que lo hacen de derecha a izquierda, incrementándose de manera breve la densidad del gas que se encuentra en la parte derecha de la caja, antes de que los procesos termodinámicos vuelvan a igualarlo. Boltzmann afirmó que la parte visible del universo era equivalente a aquella fluctuación temporal y local, pero en un cosmos mucho más grande. En su obra *Lectures on Gas Theory*[12], escribió:

Podemos elegir entre dos tipos de explicación. O bien suponemos que todo el universo se encuentra en el momento actual en un estado muy improbable, o que los *eones*[13] durante los cuales perdura este estado improbable, y la distancia de aquí a Sirio [una de las estrellas más cercanas], son diminutos comparados con la edad y el tamaño de todo el universo. En este universo, que en conjunto está en equilibrio térmico y, por consiguiente, muerto, se encuentran aquí y allá regiones

[12] Existe una reedición de la University of California Press, Berkeley, 1964; originalmente se publicó en alemán en dos volúmenes, en 1896 y 1898.
[13] Del griego cacov (eternidad), «período inmensamente largo». *(N. de la t.)*

relativamente pequeñas, del tamaño de nuestra galaxia; unas regiones (a las que podríamos llamar «mundos») que se desvían significativamente del equilibrio térmico durante intervalos relativamente cortos de esos *«eones»* de tiempo. Entre estos mundos las probabilidades de su estado (es decir, la entropía) aumentarán a medida que ellos disminuyen.

En el conjunto del universo las dos direcciones del tiempo son indistinguibles, del mismo modo que en el espacio no hay hacia arriba y hacia abajo. Sin embargo, al igual que en un determinado lugar de la superficie terrestre podemos llamar «hacia abajo» a la dirección que apunta hacia el centro de la Tierra, también un organismo vivo que se encuentra en uno de estos mundos durante un cierto intervalo de tiempo puede definir la «dirección» del mismo como un vector que va desde el estado menos probable al más probable (el primero será el «pasado» y el segundo el «futuro») y, en virtud de esta definición, descubrirá que su pequeña región, aislada del resto del universo, siempre está «inicialmente» en un estado improbable. Me parece que este modo de ver las cosas es el único que nos permitirá comprender la validez de la segunda ley, y la muerte térmica de cada mundo aislado, sin invocar un cambio unidireccional de todo el universo desde un estado inicial definido hasta un estado final.

Es desacertado decir que Boltzmann eligiera la gravedad como ejemplo para definir una manera de obtener una dirección preferente en el espacio, ya que, como hemos mencionado con anterioridad, es la influencia de la gravedad en el universo, tomado en sentido amplio, lo que a escalas tan grandes pone patas arriba los conceptos tradicionales relativos a la entropía. Resulta paradójico que tuviera que descartar la idea de «un cambio unidireccional de la totalidad del universo desde un estado inicial determinado hasta un estado final», ya que dicha idea es un sucinto resumen del modelo del big bang, que es el que normalmente se prefiere.

Sin embargo, y por encima de todo, las ideas de Boltzmann parecen anticuadas porque actualmente podemos sondear el espacio mucho más allá de Sirio (que está a una distancia de sólo 8,6 años luz) para investigar las propiedades del universo a una escala de diez mil *millones* de años luz, sin encontrar rastro del cosmos uniforme y muerto que preconiza Boltzmann. Pero, tengamos cuidado con reírnos demasiado de sus ideas pintorescas y pasadas de moda. Hoy en día hay reputados

cosmólogos que afirman que la totalidad del universo visible no es más que una de las muchas burbujas en expansión que existen en un cosmos mucho más amplio, quizá infinito, y más o menos uniforme, listas ideas no están basadas en los principios de la termodinámica, pero sirven como recordatorio de que, si bien podemos reírnos de Boltzmann por ser tan pueblerino como para creer que 8,6 años luz era una gran distancia, en un cosmos verdaderamente infinito incluso diez mil millones de años luz pueden considerarse el equivalente cósmico de un paseo hasta la farmacia de la esquina.

Sin embargo, en sentido estricto, nada de esto es importante para lo que aquí queremos contar. La razón es que, como ya hemos indicado, hay una fisura en la argumentación de Boltzmann, no en sus reflexiones filosóficas, sino en las matemáticas que utilizó para describir el comportamiento de los gases. Boltzmann partió del supuesto de que las moléculas que están a punto de chocar unas con otras no saben nada las unas de las otras, sus movimientos no están en correlación y se encuentran sometidos a lo que a veces se llama «caos molecular», una expresión en la que el término caos se utiliza en el sentido que tiene en la vida cotidiana, no en el sentido con que se va a utilizar más adelante en este libro.

Pero, por supuesto, después del choque, las trayectorias que siguen las moléculas están correlacionadas, ya que han intercambiado energía y momento cinético en su colisión. Por lo tanto, Boltzmann, sin ser consciente de ello, había incorporado a sus cálculos desde el principio una microscópica flecha del tiempo, con lo cual no era sorprendente que al final de dichos cálculos apareciera una flecha del tiempo macroscópica. Sin embargo, no hay flecha del tiempo en las leyes de Newton, que sí se utilizan para describir las colisiones, las leyes funcionan igual de bien cuando todo el sistema se recorre «hacia atrás en el tiempo». No puede haber correlaciones que funcionen en un sentido y no en el otro. O, para expresarlo de otro modo, como Laplace habría dicho, las posiciones y las velocidades de todas las moléculas del gas en cualquier instante han almacenado en sí mismas una «memoria» de toda la historia pasada del gas, de tal modo que el comportamiento de cada molécula está correlacionado con el de cualquier otra.

Fue Loschmidt quien señaló la fisura que presentaba la argumentación de Boltzmann, y lo hizo a mediados de la década de 1870, poco después de que se

publicara la formulación original de la teoría de Boltzmann[14]. Este problema fue retomado y desarrollado por Henri Poincaré (1854-1912), un matemático y físico francés. Poincaré llevó la argumentación de Laplace a su conclusión lógica. Utilizando el máximo rigor matemático, demostró que, si tenemos una caja de gas que contiene un número determinado de partículas (tantas como se quiera, con tal de que no sea una cantidad realmente infinita) y éstas cumplen estrictamente las leyes del movimiento de Newton, entonces, después de un intervalo de tiempo suficientemente largo, la distribución de las partículas dentro de la caja debe volver a su estado original, desplazándose cada una de ellas en la misma dirección y con la misma velocidad que al principio.

Podemos plantear una analogía muy esclarecedora con lo que sucede al barajar un mazo de cartas que contiene una baraja completa. Cualquiera que sea el estado del que partimos (cualquiera que sea el orden de las cartas), en principio, después de barajar, se romperá la pauta inicial. Sin embargo, si continuamos barajando durante un período de tiempo suficiente, y esto es verdaderamente un proceso realizado al azar, llegará un momento en que el mazo vuelve a estar en su orden inicial. Dado que el instante de partida es arbitrario, esto significa que el mazo de cartas o la caja llena de gas deben pasar una y otra vez por todos los estados posibles, incluidos aquellos casos raros en que todo el gas se encuentra en un extremo de la caja, y los muy numerosos en que está disperso de manera uniforme por toda la caja, pero con las partículas situadas en posiciones ligeramente diferentes. Todos los estados posibles del gas dentro de la caja se repiten una y otra vez, con una periodicidad característica que recibe el nombre de tiempo de recurrencia de Poincaré, o tiempo cíclico de Poincaré.

Si la entropía aumenta durante un lapso de tiempo, inevitablemente debe disminuir de nuevo más tarde, para hacer que el gas vuelva a su estado original (la base del universo de Boltzmann). Este comportamiento recurrente y cíclico es un resultado directo de la aplicación estricta de las leyes de Newton, en las que el pasado y el futuro tienen la misma categoría.

Pero las escalas de tiempo que intervienen aquí están más allá de lo que podemos comprender. En el caso de una caja de gas aislada que contenga *N* moléculas, el

[14] Como consecuencia de estas críticas, Boltzmann formuló posteriormente la teoría de que nuestro mundo es una fluctuación local dentro de un universo muerto, intentando así salvar algo del naufragio.

tiempo que transcurre entre recurrencias será, según Poincaré, 10^N segundos. Recordemos que un solo centímetro cúbico de aire al nivel del mar y a 0 °C contiene más de 10^{19} moléculas. El tiempo que tendremos que esperar para ver que el gas contenido en un recipiente de este tamaño recorre todo su ciclo es 10 *elevado a* 10^{19} segundos, y el resultado de esta potencia es la unidad seguida de diez millones de billones de ceros. La edad total del universo desde el big bang es sólo de unos 10^{17} segundos.

Si dividimos 17 entre 10^{19}, nos haremos una idea de lo insignificante que es la probabilidad de ver que el gas de la caja se desvía del equilibrio termodinámico durante toda la vida del universo hasta la fecha actual, en la medida en que dicha caja esté realmente aislada del mundo exterior. Incluso en el caso de que la caja contuviera sólo 52 partículas, que es el número de cartas que hay en una baraja, la duración del ciclo sería de 10^{52} segundos, que es 10^{35} veces la edad del universo[15].

A partir de todo esto, la termodinámica, y en particular la idea de una entropía creciente, se convirtieron en teorías estadísticas, en el mundo real, dado el número de partículas existentes, hay una probabilidad abrumadoramente grande de que surja la entropía, pero es posible (en el sentido de que no está prohibido por las leyes de la física) que la entropía descienda.

Al principio, esto no era algo a lo que todos los físicos pudieran acostumbrarse fácilmente. Dos siglos después de que Newton hubiera publicado los *Principia* y setenta y cinco años después de Laplace, se les decía que el mundo, después de todo, no es determinista en el sentido más simple del término, y que se han de tener en cuenta el azar y la probabilidad cuando se intente describir o calcular el comportamiento de sistemas macroscópicos. Sin embargo, de una manera gradual, se acostumbraron a la idea y las generaciones siguientes se educaron en el hábito de considerarla natural, incluso obvia. Pero, mientras estas teorías se incorporaban a la corriente principal de la ciencia, Poincaré avanzaba hacia la resolución de otros problemas. A finales de la década de 1880, descubrió algo tan sorprendente sobre el funcionamiento de las propias leyes de Newton, que las implicaciones permanecieron ampliamente ignoradas durante unos setenta años, como dice la Biblia «tres veintenas y diez años más» de una vida humana.

[15] Para que la duración del ciclo fuera igual a la edad del universo, lógicamente la caja tendría que contener sólo 17 partículas.

En esencia descubrió que Newton había tenido razón al preocuparse por la estabilidad de las órbitas de los planetas, y que Laplace se había equivocado al pensar que había resuelto el enigma. Poincaré demostró que el problema de los tres cuerpos (por no hablar del de *N* cuerpos) a menudo no se puede resolver ni siquiera mediante las conocidas técnicas de aproximación, y que ciertas órbitas en apariencia sencillas que obedecen las conocidas leyes newtonianas de la mecánica y la gravedad pueden comportarse de un modo caótico y verdaderamente impredecible.

Capítulo 2
El regreso del caos

La técnica reiterativa que se utiliza para obtener «soluciones» en casos como el problema de los tres cuerpos tiene un inconveniente. A veces no funciona, y no siempre podemos decir a priori si va a funcionar o no. La técnica que se aplica para «resolver» las ecuaciones diferenciales pertinentes (recordemos que no se pueden resolver analíticamente) implica realizar aproximaciones sucesivas, en las cuales, como es sabido, el primer paso del proceso de cálculo sólo da una solución aproximada; el segundo paso añade (con un poco de suerte) una corrección para obtener una aproximación más precisa de la realidad; el tercer paso nos da una aproximación aún mejor, y así sucesivamente hasta que nos parezca que la aproximación es lo suficientemente buena para el objetivo que nos hayamos propuesto. Pero nunca podremos conseguir *con exactitud* la «respuesta» que encaja a la perfección con el comportamiento de los objetos del mundo real en los que se centra nuestro interés.

Lo que estamos haciendo es sumar una serie de números, en principio, una serie de números infinitamente larga. A los matemáticos les interesan estas series infinitas para sus propios objetivos, independientemente de la importancia que puedan tener para los estudios del comportamiento de cosas tales como los planetas que orbitan alrededor del Sol, y conocen una gran cantidad de series infinitas cuyas sumas se comportan lo suficientemente bien como para ofrecer una aproximación cada vez mejor de un número concreto.

Un buen ejemplo lo constituye uno de los procedimientos que se utilizan habitualmente para calcular un valor aproximado del número **π**, el cociente entre la circunferencia de un círculo y su diámetro. Se puede calcular realmente el valor de π/4 con tanta precisión como se desee, sumando la serie numérica:

$$1 - 1/3 + 1/5 - 1/7....$$

Esto nos da una primera aproximación del valor de *π* que sería 4, (4 x 1), que no es muy brillante; una segunda aproximación cuyo valor sería 2,6666... (4 x 2/3), que

es algo mejor, y que, curiosamente, se encuentra al otro lado de la respuesta «correcta»; una tercera aproximación que sería 3,46666..., y así sucesivamente. Estas aproximaciones van resultando cada vez mejores y convergen en el verdadero valor de **π**, en este caso concreto desde ambos lados.

Pero el proceso es tedioso, la suma del primer millón de términos de la serie nos da para π un valor de 3,1415937, que sólo es correcto en sus cinco primeras cifras decimales. No obstante, se puede calcular π de este modo hasta el grado de precisión que se desee (hasta alguna cifra de los decimales), si tenemos la paciencia necesaria.

Sin embargo, muchas series infinitas no convergen de este modo. Por ejemplo, si planteamos la suma de todos los números enteros positivos

$$(1 + 2 + 3 + 4 + \ldots)$$

se obtiene una serie que simplemente se hace cada vez mayor, es decir, tiende a infinito; algo más sorprendente, a primera vista, es lo que se obtiene sumando todas las fracciones que son los inversos de los números enteros positivos:

$$(1 + 1/2 + 1/3 + 1/4 + \ldots)$$

es una serie que también tiende a infinito, aunque cada término de la serie sea menor que el anterior. Además, hay series cuya suma oscila hacia arriba y hacia abajo a medida que añadimos más términos, como sucedía con la aproximación de π/4, pero se aleja cada vez más de un valor concreto, en vez de converger en un valor específico. Tomemos la serie

$$1 - 2 + 3 - 4 + 5 \ldots.$$

En este caso, las sucesivas «iteraciones» nos dan los valores 1, -1, 2, -2, 3, -3, y así sucesivamente. Cuantos más términos tengamos en la serie, mayor será la oscilación a cada lado del cero en las aproximaciones sucesivas.

Y hay series cuyas sumas no siguen en absoluto una pauta regular que resulte

evidente. De todas las series que no convergen en un número finito concreto se dice que divergen y, del mismo modo que se utilizan los números, puede haber series que contengan variables. A estas últimas se les llama series de potencias, como en el caso de la serie

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \ldots$$

que nos da la posibilidad, si tenemos destreza matemática, de describir cosas tales como las órbitas de los astros. Estas series infinitas han constituido una parte importante de las matemáticas desde los tiempos de Isaac Newton, que ideó diversas técnicas para aplicarlas en la resolución de problemas y llegó a desarrollar el cálculo a través de estos trabajos.

En algunos casos, es posible demostrar matemáticamente que una serie de este tipo debe converger, incluso sin calcular millones de términos de la serie. En otros se puede probar que la serie es divergente. Sin embargo, en muchas ocasiones no hay modo de averiguar si la serie es convergente o divergente. En tales casos, aunque podamos estar trabajando con lo que parece ser una serie claramente convergente, con independencia de cuántos términos hayamos sumado ya, no podemos estar seguros de que la iteración siguiente no vaya a producir algo inesperado que cambie de manera significativa, quizá radicalmente, la respuesta que habíamos calculado con anterioridad.

Esta es exactamente la situación en que se encontraban a mediados del siglo XIX los astrónomos que utilizaban técnicas de aproximación para calcular las órbitas de los planetas del sistema solar. Parecía que las aproximaciones funcionaban bien y daban soluciones satisfactorias.

Tal como había descubierto Laplace, describían órbitas bien definidas alrededor del Sol, que sólo cambiaban de maneras regulares y predecibles cuando se tenían en cuenta las complicaciones que ocasionaban los efectos de las influencias gravitatorias que ejercían entre sí los planetas. Pero nadie había sido capaz de probar que las series utilizadas en aquellos cálculos debían converger siempre. A los astrónomos esto no les preocupaba demasiado, ya que las aproximaciones parecían

funcionar bien; pero los matemáticos lo vieron como un desafío. Nadie pensaba realmente que el sistema solar no fuera estable, pero habría sido estupendo tener una prueba de que lo era.

En la década de 1850, uno de los más destacados expertos en teoría de números era Lejeune Dirichlet, un matemático alemán que trabajaba en la Universidad de Gotinga y tenía un interés especial por el comportamiento de las series infinitas. En 1858 Dirichlet habló a su discípulo Leopold Kronecker sobre una nueva técnica que había desarrollado para resolver algunas familias de ecuaciones diferenciales que no tenían soluciones analíticas. Aunque no mencionó los detalles, también señaló que había utilizado esta técnica para demostrar que las soluciones aproximadas en términos de series infinitas para las ecuaciones que describían las órbitas de los planetas (las ecuaciones de la mecánica celeste) eran convergentes.

Por desgracia, Dirichlet falleció en mayo de 1859, a los cincuenta y cuatro años de edad, sin contar a nadie los detalles de la demostración que afirmaba haber realizado. Kronecker informó al resto del mundo matemático sobre la afirmación de Dirichlet, pero ni él, ni ningún otro, fueron capaces de utilizar las nuevas técnicas para probar que las series en cuestión realmente convergían. No obstante, Dirichlet tenía una reputación tan excelente que nadie dudó de que hubiera llevado a cabo lo que había dicho, sobre todo porque todo el mundo deseaba creer que las órbitas de los planetas eran realmente estables.

Sin embargo, hacia finales de la década de 1880, los matemáticos empezaron a inquietarse por el hecho de no haber sido todavía capaces de reproducir lo que Dirichlet pretendía haber realizado. Fue entonces cuando surgió algo así como la oportunidad de abordar el problema de manera concertada. Unos matemáticos de la Universidad de Estocolmo, en Suecia, plantearon la idea de contribuir a las celebraciones del sexagésimo cumpleaños de su rey, Oscar II, organizando un concurso internacional, con un premio importante en metálico para un trabajo matemático original que debería referirse a una de las cuatro cuestiones que plantearía un comité.

Una de estas cuestiones, inspirada en la pretensión de Dirichlet (y por el hecho de que, pasados casi treinta años, nadie hubiera logrado reconstruir la supuesta demostración), pedía una demostración de que las series de potencias que

aparecían en los cálculos de las órbitas planetarias debían converger siempre. En pocas palabras, la cuestión podía resumirse de manera concisa, diciendo: ¿Es estable el sistema solar? Entre los que se inscribieron para este concurso figuraba Henri Poincaré, que aceptó en concreto este desafío y en el desarrollo de su trabajo inventó las técnicas que todavía se utilizan hoy en día para estudiar el comportamiento de sistemas dinámicos.

La primera contribución crucial, la idea del espacio de las fases, surgió a partir del trabajo del matemático irlandés William Hamilton (1806-1865), a principios del siglo XIX[16]. Hamilton halló un modo de reformular las leyes newtonianas de la física basándose en dos propiedades de las partículas: su posición y su momento (que es igual a la masa de la partícula multiplicada por su velocidad).

Describió interacciones basadas en variaciones del momento y no en fuerzas. Desde un punto de vista físico, ambas cosas son exactamente equivalentes[17], pero este planteamiento proporcionaba un modo de estudiar un sistema completo de partículas que ejercen interacciones mutuas, utilizando para esta descripción un conjunto de ecuaciones diferenciales (ecuaciones de Hamilton) que se pueden manipular siguiendo ciertas reglas.

La posición y el momento son las variables fundamentales de esta descripción y, desde un punto de vista matemático, aparecen en pie de igualdad; pero cada una de ellas viene dada por su propio sistema de ecuaciones diferenciales, que contiene información sobre el modo en que cambia la magnitud en cuestión con el paso del tiempo. El estado de una partícula concreta en cualquier instante viene dado por su posición y su momento.

Aquí es donde entra en escena el espacio de las fases. En una representación gráfica bidimensional, la posición de cualquier punto puede definirse mediante dos números, las coordenadas *x* e *y,* que nos dicen a qué distancia se encuentra el punto sobre la horizontal de izquierda a derecha, y a qué altura está desde la parte superior hasta la inferior. De manera similar, la posición de una partícula en el espacio tridimensional ordinario puede representarse mediante tres números, que son las coordenadas que indican su distancia a un punto cero elegido previamente.

[16] El término «fase» se utiliza por razones históricas y no tiene relevancia actualmente, siendo tan sólo un nombre como otro cualquiera.

[17] Para los lectores aficionados a las matemáticas, diremos que la fuerza es la velocidad con que varía el momento, o la primera derivada de esta magnitud.

Con un esfuerzo imaginativo relativamente pequeño es posible trasponer esta idea a un espacio imaginario que representa todos los momentos que puede tener una partícula. Dado que el momento es proporcional a la velocidad, y ésta es una propiedad tridimensional (incluye también la dirección, y no sólo la rapidez del movimiento), un solo punto de este espacio imaginario de las velocidades representará las componentes de la velocidad de una partícula en cada una de las tres direcciones, que pueden elegirse de manera que formen entre sí ángulos rectos. Los números pueden representarse sobre los ejes x, *y, y z, y* es posible combinar las tres componentes para indicar el momento global de la partícula.

Ahora llegamos al paso que sólo daría un matemático. ¿Por qué no combinar estos dos conjuntos de información, es decir, tres dimensiones que representan el espacio y otras tres que informan sobre el movimiento? El resultado es un espacio imaginario de seis dimensiones, en el que un solo punto nos da tanto la posición como el momento de la partícula en un instante determinado. Éste es el espacio de las fases de una partícula.

Afortunadamente, no es necesario que seamos capaces de visualizar este espacio de las fases. Las ecuaciones que describen lo que está sucediendo pueden ampliarse de dos a tres, y a muchas dimensiones, de una manera relativamente sencilla. Mientras podamos manejar las ecuaciones, no necesitamos imaginarnos lo que está sucediendo, salvo (si eso nos ayuda) valiéndonos de una equivalencia mediante espacios tridimensionales, una especie de corte transversal a través de un espacio de las fases en seis dimensiones.

Esto es especialmente importante porque no podemos detenernos aquí. Son necesarias seis dimensiones para describir el estado de una sola partícula. Si hay dos partículas que ejercen interacciones mutuas en una caja que, por lo demás, está vacía, necesitamos un espacio de las fases de doce dimensiones para describir el sistema de esta manera, y así sucesivamente. El estado de una caja llena de gas a presión normal y temperatura ambiente se puede representar también mediante un solo punto en un espacio de las fases, pero ahora este espacio tiene una dimensión igual a seis veces el número de partículas que hay en la caja, un número que de por sí ya es enorme, como hemos comentado antes. Por supuesto, el gran número de ecuaciones diferenciales utilizadas para describir un sistema así no

puede resolverse en la práctica, ni siquiera cuando las matemáticas nos dicen que, en principio, las ecuaciones son solubles.

Sin embargo, la estadística que se utiliza en la mecánica estadística clásica incluye el análisis de las probabilidades asociadas con la distribución de puntos en el espacio de fases que describe en estos términos un sistema como el que constituye una caja llena de gas.

Por ejemplo, es mayor el número de puntos (una cantidad cuantificable) del espacio de las fases correspondiente a una distribución uniforme de partículas distribuidas por toda la caja, que el de los puntos que representan a todas las partículas que se encuentran en un extremo de la caja; también es mayor el número de puntos correspondientes a una distribución uniforme de los momentos de las partículas, que el de los puntos que representan una división de momentos cuando la mitad de las partículas se mueve muy rápido y la otra mitad muy lentamente. Esto se debe a que el espacio de fases que contiene todos los puntos del primer caso ocupa un volumen mucho mayor, y el que contiene los puntos que representan el segundo caso ocupa un volumen mucho menor.

Podemos imaginar el espacio de las fases como un paisaje, con valles ondulados, simas profundas, colinas y montañas. Las ecuaciones de Hamilton permiten a los matemáticos analizar, de un modo general, la manera en que este sistema global cambia a medida que transcurre el tiempo, sin resolver las innumerables ecuaciones diferenciales. Por ejemplo, si nos imaginamos que vertemos agua sobre nuestro paisaje del espacio de las fases, ésta Huirá por los valles, invirtiendo más tiempo en fluir por los grandes valles que en pasar por las pequeñas hendiduras del paisaje; se almacenará en las profundas simas y se precipitará desde las cimas de las montañas. Las ecuaciones de Hamilton nos dicen cómo se mueven los sistemas reales «a través» del espacio de las fases y nos indican las regiones hacia las cuales se verán atraídos los valles profundos y las simas. Podemos incluso extender la analogía para incluir en ella los equivalentes a la evaporación y la lluvia, que corresponderían a trayectorias relativamente raras que sacan partículas del «río» y las transportan a las «cimas de las montañas», desde las cuales fluyen de nuevo inevitablemente hacia abajo. El desplazamiento de una *sola* partícula a través del espacio de las fases (su trayectoria) representa el modo en que el sistema en

conjunto varía con el transcurso del tiempo, y la cantidad de tiempo que pasa la partícula en cada zona del espacio de las fases es proporcional al volumen de esa zona.

No se puede decir con exactitud adónde irá la partícula, al igual que no es posible predecir la trayectoria exacta de una molécula de agua en un río caudaloso; pero sí se puede decir que hay una probabilidad enorme de que siga cierto tipo de trayectoria a través del espacio de las fases, del mismo modo que la molécula de agua tiene pocas probabilidades de ir a parar a algún lugar situado más allá de las orillas del río. Sería mucho menor la probabilidad de que siguiera una de las trayectorias raras y poco transitadas que ascienden hacia la cima de la montaña y vuelven a bajar de nuevo.

Retomando un ejemplo muy sencillo, a un péndulo perfecto y sin rozamiento que oscila hacia uno y otro lado sólo le corresponde una dimensión del espacio real y una del espacio de las velocidades, de tal modo que el espacio de las fases es bidimensional y se puede trazar sobre una hoja de papel plana. Optaremos por representar su posición horizontalmente a través de la hoja y su velocidad hacia arriba o hacia abajo sobre la vertical, de tal modo que el sentido ascendente corresponda al movimiento hacia la derecha y el descendente al movimiento hacia la izquierda. Cuando la lenteja del péndulo se mueve de izquierda a derecha y luego en sentido contrario, su velocidad parte de cero en un extremo del recorrido, alcanza el máximo en una dirección en medio de la oscilación, cae a cero en el otro extremo, y luego repite este proceso a la inversa.

Combinemos esto con su posición cambiante en el espacio real y resultará que describe un círculo en el espacio de las fases. Si aceptamos que haya rozamiento, el péndulo irá frenándose gradualmente y se detendrá, describiendo una espiral en el espacio de las fases, siendo el punto situado en el centro de dicha espiral un atractor de este sistema concreto.

Esta idea de imaginarnos el espacio de las fases utilizando montañas, valles y cosas por el estilo constituye una versión muy simple de la topología, una materia en la que fue pionero Poincaré, quien la aplicó, en el contexto del espacio de las fases, al problema de demostrar que el sistema solar es estable. En esencia lo que hizo fue convertir un problema de mecánica y dinámica en un problema geométrico.

Recordemos que un solo punto del espacio de las fases corresponde a la totalidad del estado del sistema, en el caso del gas contenido en una caja, este punto único (o las ecuaciones de Hamilton correspondientes a él) representa la posición y el momento de todas y cada una de las partículas de este gas. Una trayectoria que atraviesa el espacio de las fases expresa el modo en que varía el estado a medida que transcurre el tiempo. Pero, si esta trayectoria pasa alguna vez por un punto por el que ya había pasado antes, esto significa que el sistema en su totalidad ha vuelto *exactamente* a un estado anterior.

En este caso, según las leyes de la mecánica newtoniana, debe proceder de ese estado exactamente del mismo modo que lo hizo con anterioridad, redescribiendo su trayectoria a través del espacio de las fases. Con respecto a las cajas llenas de gas, esto entra de lleno en la idea de Poincaré de ciclos y recurrencia; en relación con las órbitas, significa que, si una trayectoria trazada a través de un espacio de las fases que representa los estados posibles de, por ejemplo, tres cuerpos regresa al mismo punto del espacio en que ha estado antes, entonces las propias órbitas, con independencia de lo complicadas que puedan parecer, deben repetirse con periodicidad, y los cuerpos que las recorren no se alejarán de repente unos de otros, ni chocarán entre sí.

Podríamos deducir también que si la trayectoria que atraviesa el espacio de las fases vuelve a pasar muy cerca de un punto por el que ya había pasado con anterioridad, el comportamiento subsiguiente del sistema será muy parecido al que había tenido antes. Sin embargo, como veremos más adelante, no deberíamos dar por supuesto ninguno de estos hechos.

En el trabajo que presentó al concurso, Poincaré ni siquiera intentó describir topológicamente el comportamiento del sistema solar en su totalidad, sino que se centró en hacer una representación geométrica del espacio de las fases para las órbitas asociadas con un caso muy sencillo, conocido como el problema «restringido» a tres cuerpos. Como ya hemos visto, en un universo que contenga sólo dos cuerpos que gravitan, no hay dificultad alguna para describir sus órbitas. Como Newton demostró, se mueven el uno en torno al otro con una trayectoria totalmente regular y periódica que Poincaré y sus sucesores pudieron describir hablando de un bucle cerrado en el espacio de las fases. Sin embargo, si añadimos

otro cuerpo, como ya hemos visto, el movimiento se vuelve demasiado complicado para hacer cálculos analíticos.

El problema restringido a tres cuerpos intenta salvar esta dificultad imaginando dos cuerpos más o menos del mismo tamaño y un tercer cuerpo mucho menor que los otros dos, de tal modo que, a la hora de realizar cálculos, se pueda ignorar el efecto que produce sobre cuerpos de mayor tamaño, aunque éstos ejerzan un efecto poderoso sobre el cuerpo pequeño (que a veces recibe el nombre de «partícula de polvo»).

Esta situación aún no puede resolverse analíticamente, pero todo lo que se intentará hacer será calcular la órbita de la partícula de polvo, sumando las series adecuadas, en el marco de las posiciones cambiantes de los otros dos cuerpos. Se trata de una situación imposible, ya que la partícula de polvo debe ejercer siempre alguna influencia gravitatoria sobre sus compañeras, pero la esperanza de algunos matemáticos, como Poincaré, era que se podría conseguir así una aproximación razonable a lo que estaba sucediendo en el sistema solar si pensamos en las interacciones mutuas entre, por ejemplo, el Sol, Júpiter y la Tierra, o el Sol, la Tierra y la Luna[18].

Uno de los grandes supuestos simplificadores de Poincaré consistió en tener en cuenta sólo una pequeña parte del correspondiente espacio de las fases, de hecho, un corte transversal del espacio de las fases. Dicho corte representaba una superficie por la cual tenía que pasar la trayectoria que se estaba investigando. Actualmente este corte se conoce como sección de Poincaré. Si una trayectoria que atraviesa el espacio de las fases comienza en un punto de una sección de Poincaré y regresa luego a ese punto, entonces sabemos que es exactamente periódico, con independencia de lo complicado que pueda haber sido el comportamiento del sistema entre estas dos intersecciones con la sección de Poincaré.

Con el fin de llevar a los jueces suavemente hacia sus nuevas teorías, Poincaré comenzó abordando el problema de un modo convencional, lo cual significa que seguía teniendo que resolver ecuaciones diferenciales para intentar obtener la ansiada demostración de que las órbitas del sistema solar (o, al menos las órbitas de una partícula de polvo en el problema restringido a tres cuerpos) eran

[18] De hecho, este procedimiento lo utilizó con cierto éxito el astrónomo americano George Hill a finales del siglo XIX para calcular la órbita de la Luna.

periódicas, y, dado que las ecuaciones no podían resolverse analíticamente, era necesario sumar series, y no terminó del todo el trabajo. Halló las ecuaciones adecuadas y dedujo las series infinitas correspondientes, pero en realidad no demostró que las series *debían* converger; sólo probó que eran soluciones adecuadas que *podían* converger. A continuación, utilizó su nuevo planteamiento topológico para demostrar que (según creía él) aquellas tenían que ser las soluciones correctas, ya que las correspondientes trayectorias que cruzaban el espacio de las fases volvían a atravesar la sección de Poincaré por el mismo punió del que habían partido.

Poincaré escribió más de doscientas páginas para expresar todo esto, y gran parte de todo ello resultó nuevo para los jueces. Pero este trabajo les dio la respuesta que deseaban; era obviamente un planteamiento inteligente, y tenían que terminar de valorarlo a tiempo para el cumpleaños del rey, que se celebraba el 21 de enero de 1889. Poincaré ganó el premio. Sin embargo, cuando se publicó su trabajo y otros matemáticos tuvieron tiempo para estudiarlo con detenimiento, descubrieron que Poincaré había cometido un error.

Su «demostración» no se sostenía. Después de un arranque de trabajo intenso para responder a las críticas de sus colegas, Poincaré corrigió los puntos débiles de su razonamiento y publicó en 1890 un trabajo revisado, que actualmente se considera como un clásico de las matemáticas de todos los tiempos. Para su sorpresa, al corregir sus propios errores descubrió que su nuevo planteamiento demostraba que, dicho con el lenguaje antiguo, incluso en el caso de la órbita de la partícula de polvo del problema restringido a tres cuerpos, la serie correspondiente es *típicamente* divergente, la inestabilidad es normal, y las órbitas que son estables de manera permanente constituyen la excepción.

Utilizando los términos de su nuevo planteamiento geométrico, descubrió que, aunque hay órbitas periódicas cuya trayectoria en el espacio de las fases vuelve a pasar repetidas veces por un punto determinado de la sección de Poincaré, si la trayectoria corta dicha sección a tan sólo una pequeñísima distancia del punto, entonces el sistema puede seguir una pauta de comportamiento totalmente distinta, con trayectorias que toman caminos muy diferentes a través del espacio de las fases y cruzan repetidamente la sección de Poincaré sin pasar en ningún caso dos

veces por el mismo punto.

Puede haber una cantidad infinita de puntos de corte en la sección de Poincaré, lo cual permite que la trayectoria que atraviesa el espacio de fases haga recorridos infinitamente complicados y variados sin volver jamás a su punto de partida[19].

Lo bueno es que, al menos en el caso del problema restringido a tres cuerpos, y de forma más general para los planetas del sistema solar, las órbitas se pueden calcular, en principio, con tanta precisión como se desee aplicando las laboriosas técnicas de aproximación y realizando las sumas de las series correspondientes con un número suficientemente grande de términos, aunque no sean, en sentido estricto, perfectamente periódicos.

Como veremos, estos cálculos muestran que hay situaciones relativamente simples en las que los planetas pueden describir en esencia las mismas órbitas durante intervalos de tiempo que son muy largos en comparación con cualquier escala de tiempo humana, aunque no son estrictamente periódicos a escalas de tiempo que son largas si se comparan, por ejemplo, con la edad del Sol.

Pero la cuestión clave que se desprende de la obra de Poincaré es la constatación de que, en determinadas circunstancias (no en todas, aunque tampoco han de ser necesariamente raras), los sistemas que parten de estados casi iguales pueden evolucionar muy pronto de maneras completamente diferentes.

Hay dos procedimientos que nos pueden ayudar a hacernos una idea de todo esto. El primero es la ya mencionada analogía del espacio de las fases como un paisaje por el que fluye agua. La trayectoria de una sola molécula de agua a través del espacio de las fases representa el estado variable de todo un sistema, que puede ser tan sencillo como el problema restringido a tres cuerpos, o tan complicado como todo el universo. Imaginemos que hay un ancho río que fluye cruzando el paisaje y que representa una región en la cual es altamente probable que se encuentren las ecuaciones hamiltonianas correspondientes al sistema.

En algún punto, el río se rompe y diverge hacia una complicada red de canales, que forman un delta como el del Ganges. Una sola molécula que se desplace con el flujo

[19] En realidad es aún más complicado. Recordemos que el espacio de las fases obliga a utilizar dimensiones más altas que aquellas a las que estamos acostumbrados. Por lo tanto, una sección de Poincaré no es (en general) una simple superficie bidimensional. Esto hace que aparezca un enredo aún más complicado, con infinitas posibilidades, llamado un enredo homoclino. Pero, para lo que aquí nos interesa, es suficiente con saber que las posibilidades son realmente infinitas.

del río puede ir hacia la izquierda, entrando en una rama de los canales que se bifurcan, mientras que otra molécula de su entorno inmediato se desplaza hacia la derecha, yendo a parar a otra rama de la red. Por consiguiente, siguen caminos bastante distintos. Dicho de otro modo, dos trayectorias que atraviesan el espacio de las fases y que transcurren durante un largo trecho una al lado de la otra, porque corresponden a estados del sistema casi idénticos, pueden divergir hacia estados muy diferentes. De manera similar, imaginemos una gota de agua que cae sobre una cadena de montañas de cimas altas y afiladas. Si cae a un lado de una cresta, descenderá fluyendo en una dirección, hacia un profundo océano que se comporta como un fuerte atractor para ese sistema; si cae por el otro lado de la cresta, fluirá en dirección opuesta, hasta otro profundo océano, que es un atractor igualmente poderoso para el sistema. Sin embargo, estas dos trayectorias, que conducen a estados finales tan diferentes, pueden comenzar en puntos infinitamente cercanos.

También nos puede ayudar un ejemplo relacionado con la vida cotidiana. En nuestra experiencia diaria, los cambios se producen, en su mayoría, de una forma lineal. Hace falta el doble de fuerza para levantar un paquete de azúcar de dos kilos que para levantar uno de un kilo, y así también en otros casos. Si damos dos pasos por la calle, recorremos una distancia que es el doble de la que recorreríamos si diéramos sólo uno. Pero, supongamos que caminar no es un proceso lineal, de tal modo que, después de dar el primer paso, cada paso sucesivo nos llevara a una distancia que fuera el doble de la recorrida con el paso anterior. Pongamos que el primer paso nos lleva a un metro del punto de partida. Entonces, el segundo paso nos haría avanzar dos metros más, el tercero, cuatro metros, y así sucesivamente.

Andando de forma lineal, once pasos nos harían avanzar once metros por la calle. Sin embargo, andando de esta forma no lineal, sólo con el undécimo paso recorreríamos una distancia de 1.024 metros, que sería por sí mismo un recorrido superior en un metro a todo lo que habríamos andado con los diez pasos anteriores. En números redondos, con once pasos no avanzaríamos once metros por la calle, sino unos *dos kilómetros.*

Los procesos no lineales hacen que los objetos se alejen muy rápidamente de su punto de partida y, si dos trayectorias comienzan de un modo no lineal en

direcciones ligeramente diferentes a través del espacio de las fases, van a divergir no menos rápidamente la una con respecto a la otra. Poincaré descubrió que muchos sistemas del mundo real son muy sensibles a sus condiciones iniciales (como la gota de lluvia que cae sobre la cresta de la montaña) y se alejan de dichas condiciones de una manera no lineal.

Lo importante es que esto limita nuestra capacidad de predecir el comportamiento de tales sistemas. Cuando se trata de un sistema lineal, si cometemos un pequeño error al medir o estimar alguna propiedad inicial del sistema, este error se transmitirá a lo largo de nuestros cálculos y producirá un pequeño error al final de los mismos. Pero, si se trata de un sistema no lineal, un pequeño error cometido al principio de los cálculos producirá un error muy grande en el resultado final. Un sistema lineal es más o menos igual a la suma de sus partes; un sistema no lineal puede ser mucho más o mucho menos que la suma de sus partes. Según constató Poincaré, esto significa que, en algunas circunstancias, no podemos calcular cómo variará un sistema con el paso del tiempo, porque la información de que disponemos sobre sus condiciones iniciales no es lo suficientemente precisa. En 1908 escribió en su obra *Science et Méthode*[20] lo siguiente:

Una causa muy pequeña que escapa a nuestra percepción determina un efecto considerable que no podemos dejar de ver, y luego decimos que este efecto se debe al azar. Si conociéramos con exactitud las leyes de la naturaleza y la situación del universo en su momento inicial, podríamos predecir exactamente la situación de ese mismo universo en un momento posterior. Pero, incluso si se diera el caso de que las leyes naturales no fueran ya un secreto para nosotros, sólo podríamos conocerla situación inicial de una manera aproximada. Si esto nos diera la posibilidad de predecir una situación posterior con la misma aproximación, esto es todo lo que necesitamos, y diríamos que se ha logrado predecir el fenómeno, que está gobernado por las leyes. Pero no siempre es así; puede suceder que unas pequeñas diferencias en las condiciones iniciales produzcan unas diferencias muy grandes en los fenómenos finales. Un pequeño error en lo anterior causará un enorme error en lo posterior. La predicción se hace imposible, y lo que tenemos es un fenómeno fortuito.

[20] Hay traducción al castellano. El texto completo *(Ciencia y método)* está incluido en *Sobre la ciencia* y *su método.* Círculo de Lectores, S. A., Barcelona, 1997, trad. de M. García Miranda, L. Alonso, A. B. Besio y J. Banfi.

Dicho de otro modo, aunque en principio Laplace tenía razón, su universo determinista nunca se pudo predecir en la práctica, ni hacia adelante ni hacia atrás en el tiempo.

En el mismo volumen, Poincaré elaboró su teoría sobre este tema utilizando el ejemplo concreto de la previsión meteorológica:

¿Por qué tienen los meteorólogos tantas dificultades para pronosticar el tiempo con cierta seguridad? ¿Por qué parece que las lluvias e incluso las tormentas llegan por azar, de tal modo que mucha gente considera bastante normal rezar para que llegue la lluvia o el buen tiempo, aunque les parecería ridículo hacerlo para pedir un eclipse? Vemos que las grandes perturbaciones se producen generalmente en zonas donde la atmósfera se encuentra en equilibrio inestable.[21] *Los meteorólogos ven claramente que el equilibrio es inestable, que se formará un ciclón en algún lugar, pero no están en situación de decir con exactitud dónde; una décima de grado más o menos en cualquier punto dado y el ciclón se desencadenará aquí y no allí, causando estragos en zonas que en otro caso no se habrían visto afectadas. Si se hubieran dado cuenta de que existía esta diferencia de una décima de grado, habrían podido saber de antemano lo que pasaba, pero las observaciones no son suficientemente amplias o precisas, y ésta es la razón por la cual todo parece ser fruto del azar.*

En una visión retrospectiva, este ejemplo puede considerarse especialmente anticipatorio de lo que se conocería en el futuro. Poincaré formuló ideas muy avanzadas para su tiempo cuando intentaba explicar cómo las leyes ordenadas de la física pueden dar lugar a comportamientos aparentemente caóticos en la vida cotidiana, y pasarían muchas décadas antes de que estas teorías se recuperaran y pasaran a formar parte de la corriente principal de la ciencia.

[21] Un equilibrio inestable es como un lápiz que se balancea sobre su punta. Poincaré dice que en tal situación sabemos que el lápiz se va a caer, pero no podemos predecir en qué dirección lo hará.

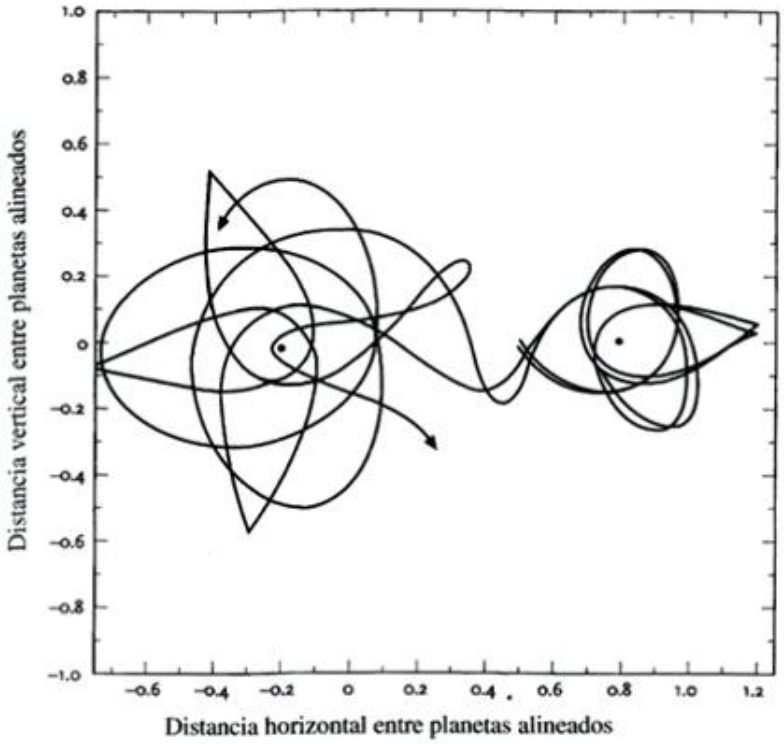

Figura 2.1. El problema de los tres cuerpos. Si un pequeño «satélite» describe una órbita en torno a dos «planetas» de mayor tamaño, incluso un pequeñísimo cambio en las condiciones iniciales de la trayectoria del satélite produce un gran cambio en su órbita. Dado que nunca podremos conocer las condiciones iniciales con exactitud, la órbita es impredecible. Esta imagen muestra dos de estas trayectorias, en un marco de referencia que rota, de tal modo que los dos planetas aparecen estacionarios.

Sin embargo, cuando esta recuperación se produjo, tuvo lugar, en gran medida, gracias a los esfuerzos de los meteorólogos por mejorar sus previsiones del tiempo y por comprender por qué dichas previsiones fallaban tan a menudo. La cuestión clave es en qué medida han de ser amplias y precisas las observaciones iniciales. Un buen punto de partida es abordar el problema a través de la obra pionera del científico inglés Lewis Fry Richardson, que fue también un hombre avanzado para su tiempo y halló durante la segunda década del siglo XX el método para hacer una previsión meteorológica científica. Lo hizo mucho antes de que existieran los aparatos necesarios para este trabajo (en particular, ordenadores electrónicos de

alta velocidad).

Richardson, que nació en 1881 y falleció en 1953, tenía intereses muy variados y no era en absoluto un meteorólogo estereotipado[22]. Siendo cuáquero e hijo de un granjero, Richardson llegó a ser superintendente del observatorio meteorológico de *Eskdalemuir*, Escocia, cuando estalló la Primera Guerra Mundial. A causa de sus creencias religiosas no tomó parte en la contienda, sino que se presentó voluntario para conducir ambulancias, prestando servicios muy cerca del frente.

Allí, en su tiempo libre, se pasaba horas realizando tediosos cálculos aritméticos, en un intento por conseguir la primera «predicción» meteorológica numérica, que no fue realmente una predicción en el sentido habitual del término, ya que se había llevado consigo un montón de datos meteorológicos auténticos relativos a una hora determinada de un día concreto, e intentaba utilizar esta información para calcular las amplias líneas generales de las pautas meteorológicas en Europa seis horas después del momento al que se referían los datos utilizados.

Incluso si llegaba a dar resultado, llevaría meses realizar una «predicción» para seis horas más tarde; pero lo que le interesaba a Richardson durante la segunda década del siglo XX era la posibilidad de demostrar que las aproximaciones matemáticas y las leyes de la física podían utilizarse de este modo.

La inspiración para emprender este proyecto le llegó de la obra del noruego Vilhelm Friman Koren Bjerknes que era diecinueve años mayor que Richardson y había sugerido, una década antes, que aquella técnica de predicción del tiempo era, en principio, posible.

Fue el primero en plantear que las ecuaciones que describían las condiciones atmosféricas se conocerían lo suficiente como para que la predicción fuera posible, si las condiciones iniciales para comenzar los cálculos se conocían con la precisión necesaria. La idea que subyace a este planteamiento de la predicción meteorológica, y que es fundamental en la moderna meteorología, es la medición de las propiedades importantes de la atmósfera, como son la temperatura y la presión, en una red de puntos que está situada sobre la superficie terrestre y se prolonga hacia arriba en el aire. Cuanto más próximos se encuentren los puntos de esta red, más preciso es el modelo matemático del estado de la atmósfera en ese

[22] Fue también tío del actor Ralph Richardson, aunque esto no tiene importancia para nuestra historia.

momento. A continuación se aplican las leyes de la física para averiguar cuál es el modo en que variarán las condiciones en cada punto de la red al estar dicho punto sometido a las influencias de los puntos vecinos (cuando el calor fluye desde los puntos más calientes a los más fríos, el viento sopla desde las zonas sometidas a altas presiones hacia las zonas en las que la presión es baja, se produce la convección, etc.)[23].

Es evidente la similitud que existe entre el tipo de técnicas necesarias para resolver el enigma de las órbitas de los planetas, y el sentido común nos dice que es mucho mayor la esperanza de éxito cuando se utiliza una red más cerrada, en la cual los puntos están más próximos. Las condiciones que se dan en los lugares situados entre los puntos de la red se calculan sencillamente por interpolación, tomando una media adecuada de las propiedades de los puntos circundantes de la red.

Lo que se esperaba era que, si aquella técnica llegaba a funcionar, se podría hacer, en principio, tan precisa como quisiéramos utilizando más y más puntos de una red entretejida de tal manera que fuera cada vez más fina. Resultó que las predicciones de Richardson eran desesperadamente imprecisas, pero a él esto no le sorprendió, ya que era consciente de que se veía obligado a utilizar para los datos una red de puntos muy tosca (tanto por las limitaciones en cuanto a datos disponibles, como por el tiempo que requerían los cálculos), en la cual no era realista esperar que pudieran representarse todas las sutilezas del tiempo atmosférico en Europa. Lo que importaba en aquella época era que Richardson había utilizado la técnica para calcular cómo podría cambiar el tiempo, aunque los resultados de sus cálculos no se correspondieran con la evolución en el mundo real.

Richardson se sintió lo suficientemente entusiasmado con su técnica como para empezar a trabajar en la redacción de un libro, *Weather Prediction by Numerical Process,* cuando todavía estaba de servicio en el frente occidental.

En medio de la confusión que generó un desplazamiento rápido ocasionado por las incidencias de los combates, se perdió en 1917 la única copia de su manuscrito, pero apareció unos meses más tarde bajo un montón de carbón. Nadie sabe cómo llegó allí, pero, cuando recuperó el manuscrito, Richardson pudo terminar el libro,

[23] Podemos comparar este planteamiento con la predicción tradicional «sinóptica», que se basa en el estudio de las pautas del tiempo en un mapa y en la utilización de un conocimiento básico de la física necesaria, además de una buena intuición, que se desarrolla con la experiencia, para adivinar cómo se desplazarán los anticiclones y las borrascas durante las horas siguientes.

que se publicó en 1922 y contiene su más famosa visión del futuro de la predicción meteorológica. Consciente de que no resulta práctico (con independencia de lo gratificante que pueda ser para los matemáticos) hacer una predicción para las seis horas siguientes si se tarda meses (o incluso si se tardara sólo siete horas) en efectuar los cálculos, Richardson pedía a sus lectores que se imaginaran una «factoría de predicciones meteorológicas» en la que 64.000 «computadoras» humanas, cada una de ellas equipada con una calculadora mecánica (una especie de gloriosa máquina de sumar, precursora de las calculadoras electrónicas de bolsillo), trabajarían encargándose cada una de ellas de un aspecto concreto del problema[24], sentadas en un anfiteatro del tamaño de un estadio de fútbol y bajo la batuta de una especie de director de orquesta matemático subido a un escenario situado en medio del centro de operaciones, que se comunicaría con ellas mediante destellos de luz o mensajes enviados por tubos neumáticos. Se trataba, por supuesto, de una fantasía científica, nunca sería posible coordinar debidamente la actividad de tantas personas. El propio Richardson escribió:

Quizá algún día, en un futuro lejano, será posible avanzar los cálculos más rápido de lo que avanza el tiempo atmosférico, con un riesgo para la humanidad inferior al del salvamento, gracias a la información obtenida. Pero esto es un sueño.

No obstante, fue un sueño que casi llegó a ver realizado. Tres décadas después, antes de que Richardson hubiera muerto, el invento del ordenador electrónico abrió la posibilidad de realizar prácticamente el mismo trabajo con máquinas, con *una sola* máquina que sustituiría a los 64.000 ordenadores humanos de Richardson y (finalmente) a muchos más. La primera predicción numérica del tiempo que pudo realizarse con éxito se hizo mediante una de estas máquinas en 1950, aunque en aquella época el aparato todavía tardaba en formular la predicción más tiempo que el que tardaba la atmósfera real en evolucionar.

El propio Richardson, sin duda consciente de que había llegado con la predicción numérica todo lo lejos que se podía llegar en la década de 1920, pasó a dedicarse a otras cosas, como dirigir el departamento de física del Westminster Training College de Londres desde 1920 hasta 1929, llegando luego a ser decano del Paisley

[24] Una posibilidad es que cada una de estas personas trabajara con los da-los correspondientes a un punto de la red.

Technical College de Glasgow, hasta su jubilación en 1940.

Paralelamente estudió psicología y, tanto antes como después de jubilarse, escribió varios libros sobre las causas y la psicología de los conflictos armados. Teniendo en cuenta esto, y sus creencias cuáqueras, resulta paradójico que el enorme progreso de las técnicas computacionales, que en vida de Richardson llegó a hacer realidad la predicción numérica del tiempo, se viera estimulado por el trabajo de los decodificadores durante la Segunda Guerra Mundial.

Sin embargo, aunque a principios de la década de 1950 parecía que el sueño de Richardson iba a hacerse realidad y que todo lo que hacía falta para hacer mejores predicciones era disponer de ordenadores más rápidos y de redes de observación más detalladas, diez años más tarde un hombre dio con la prueba que dejaba sin fundamento esta hipótesis y demostraba que incluso Poincaré había sido tremendamente optimista al suponer que el conocimiento del estado de la atmósfera con una precisión de décimas de grado podría explicar por qué un huracán se producía *aquí* y no *allí.*

En 1959 Edward Lorenz era un meteorólogo matemático de treinta y dos años que trabajaba en el Massachusetts Institute of Technology (MIT). Aunque no era capaz de escribir un conjunto completo de ecuaciones para describir el estado de la atmósfera en cualquier momento, tuvo la idea de que, en principio, esas ecuaciones podían describir cualquier estado de la atmósfera como condiciones iniciales, aunque no fuera de una manera realista o plausible, pero que, si se realizaba una simulación correspondiente a dicho estado, entonces los procesos naturales llevarían el modelo de atmósfera a ser uno de los relativamente pocos estados estables, lo que llamaríamos ahora atractores, aunque no era éste el lenguaje que utilizaban los meteorólogos en la década de 1950.

Algunos pioneros de la predicción meteorológica matemática habían utilizado técnicas lineales, por la sencilla razón de que tenían que empezar por algún sitio y las ecuaciones lineales son más fáciles de manejar que las ecuaciones no lineales. Lorenz fue aún más lejos cuando desarrolló un modelo computacional para una versión muy simplificada de la atmósfera (los físicos a veces llaman a estas simplificaciones modelos «de juguete») basada en un conjunto de doce ecuaciones no lineales. El modelo tenía que ser muy sencillo porque los ordenadores de que se

disponía eran también muy sencillos, Lorenz utilizó una máquina del tamaño de un gran mesa de trabajo, que tenía una memoria de tan solo 4k y 32 bits, mucho menos potente que el chip que podemos encontrar actualmente dentro del reloj digital o de la lavadora. Pudo demostrar que las técnicas lineales de predicción estaban «lejos de ser perfectas»[25] y, a continuación, se puso a estudiar el comportamiento de sus propias simulaciones.

Así como la potencia de los ordenadores que existían entonces era limitada, también lo eran los instrumentos que se utilizaban para reflejar los resultados, no había impresoras láser en color, ni monitores con pantalla plana. El resultado que producía el ordenador de Lorenz era sencillamente una lista con números impresos, y él programaba aquel ordenador para redondear en la tercera cifra decimal con el fin de obtener un conjunto completo de doce números en cada línea del papel donde aparecían los resultados.

El modelo avanzaba a pasos de seis horas virtuales, y simulaba la situación atmosférica correspondiente a un día en un minuto. Con una sofisticación realmente sorprendente, Lorenz programó la máquina para imprimir los detalles correspondientes a sólo una o dos variables en cada línea, de tal modo que la distancia del símbolo «retorno de página» representaba el valor de la variable; entonces pudo dibujar a mano una línea que unía estos puntos para expresar en una gráfica el valor cambiante de la variable (la dirección del viento, por ejemplo). Fue entonces cuando surgió la feliz casualidad.

Lorenz decidió que deseaba volver a echar un vistazo a una parte de la simulación que ya había realizado, llevándola más allá en el tiempo futuro y, en vez de comenzar desde el principio y esperar a que el ordenador llegara al intervalo que le interesaba, tecleó como parámetros iniciales el conjunto de números de aquella parte intermedia de lo ya realizado, leyéndolos del papel, y luego se fue a tomar un café mientras la máquina traqueteaba. Después de aproximadamente una hora, durante la cual el aparato había simulado dos meses de situaciones atmosféricas de mentirijillas, Lorenz volvió a comprobar lo que se estaba haciendo y descubrió que los números que estaba imprimiendo no guardaban relación alguna con los de los «días» correspondientes del proceso anterior. Al principio pensó que el ordenador

[25] Véase Edward Lorenz, The Essence of Chaos. [La esencia del caos, Editorial Debate, Barcelona, 2000, trad. de Francisco Páez de la Cadena.]

funcionaba mal, pero, cuando comparó detenidamente los dos resultados impresos, descubrió que el segundo proceso comenzaba como la sección correspondiente del primero, aunque se desviaba de éste cada vez más, multiplicándose los errores por dos cada cuatro días de la simulación, es decir, se trataba de un crecimiento no lineal.

Lorenz no tardó en darse cuenta de qué era lo que había sucedido. Los números que había introducido eran los de tres cifras decimales tomados del resultado que le había dado la impresora; pero dentro del ordenador los números se calculaban con seis cifras decimales. Por ejemplo, donde Lorenz había tecleado 0,506, el número utilizado en el primer proceso de cálculo había sido en realidad 0,506127. El modelo era tan sensible a las condiciones iniciales que la diferencia, un cuarto de una décima del uno por ciento, hacía que los dos procesos divergieran completamente el uno con respecto al otro al cabo de un breve tiempo. Si la atmósfera real era tan sensible a sus condiciones iniciales, no había ninguna esperanza de poder utilizar las técnicas de predicción numérica para pronosticar el tiempo con más de unos pocos días de antelación.

Lorenz anunció su descubrimiento sin hacer mucho ruido en un congreso científico que se celebró en Tokio en 1960 y continuó desarrollando estas teorías durante los años siguientes pero, como veremos más adelante, no se apreció toda la importancia de su trabajo hasta mucho más tarde. Por lo que respecta a la predicción del tiempo atmosférico, Lorenz consiguió demostrar que la atmósfera real puede ser en efecto muy sensible a las condiciones iniciales. Utilizó un sencillo modelo matemático (numérico) de convección, que sólo alude a la complejidad del tiempo atmosférico real, pero que posee una característica fundamental de sensibilidad a pequeños cambios. Utilizando nuestro modelo del espacio de las fases, nos podemos imaginar esto como un paisaje inundado de agua, imaginemos dos profundos pozos en el espacio de las fases, que representan dos atractores igualmente potentes, separados por una barra de arena cubierta por una capa de agua muy poco profunda que une los dos pozos.

Las típicas trayectorias del espacio de las fases circundan repetidas veces un pozo u otro, pero unas pocas trayectorias cruzan sobre la barra de arena y entran en el otro pozo, al que circundan varias veces antes de volver a pasar por la barra para

regresar al primer pozo. El quid del descubrimiento de Lorenz está en que, sin conocer la localización precisa de un punto en el espacio de las fases, es imposible predecir con exactitud cuándo podrá cruzar al otro lado la trayectoria que pasa por ese punto, por lo que los cambios de un estado a otro parecen producirse aleatoriamente.

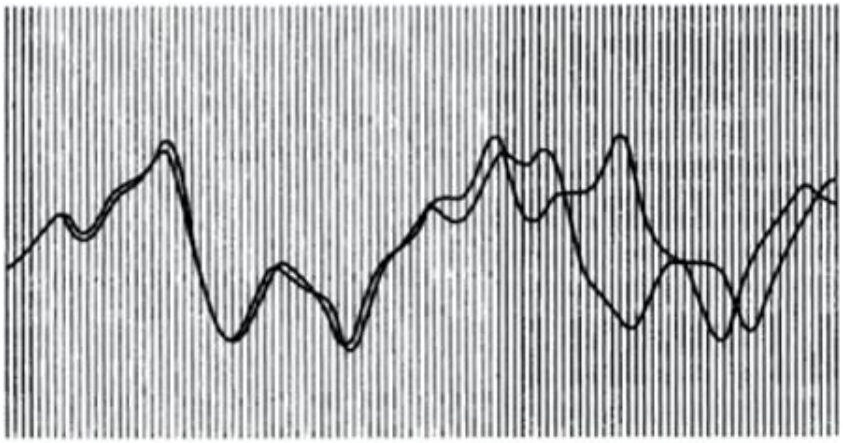

Figura 2.2. Edward Lorenz descubrió que, tal como sucede con la órbita de un satélite en el problema de tres cuerpos, las «predicciones» obtenidas mediante computadora con respecto a las propiedades del tiempo atmosférico (por ejemplo, la temperatura) divergen ampliamente aunque partan de condiciones iniciales casi idénticas.

O bien, desde un punto de vista ligeramente distinto, las trayectorias que pasan sobre la barra de arena son muy sensibles a perturbaciones muy ligeras, y un pequeño empujón dado desde fuera puede ser todo lo que se necesita para que el sistema cambie de una trayectoria que regresa a uno de los pozos a otra que se dirige al otro pozo. Este tipo de funcionamiento es el que hace que el límite de exactitud en la predicción del tiempo sea de entre diez y catorce días, aproximadamente, y que el tiempo atmosférico real pueda asimismo pasar de un estado estable a otro también estable de un modo básicamente impredecible.

La cuestión de la determinación *precisa* de las condiciones iniciales es el tema en el que Lorenz centró su atención y constituye el núcleo central de la manera moderna de entender el caos. Esta sensibilidad del tiempo atmosférico (y de otros sistemas complejos) a las condiciones iniciales se denomina a veces «efecto mariposa»,

según el título de una publicación que presentó Lorenz en un congreso celebrado en Washington DC en 1972 «*Does the Flap of a Butterfly's Wings in Brazil Set off a Tornado in Texas?*»[26].

La analogía no debe tomarse al pie de la letra, ya que a esta escala en el mundo real intervienen tantos procesos, dejando a un lado las contribuciones individuales, que resulta improbable que el batir de alas de una mariposa en Brasil pueda identificarse como la causa de un tornado específico en Texas (o, al revés, que el batir de alas de otra mariposa en China evite que se forme un tornado en Texas).

Pero no deja de ser una impactante metáfora del caos. Se da la coincidencia de que, cuando el atractor de dos polos descubierto por Lorenz se traza gráficamente sobre el papel (o sobre la pantalla de un ordenador), forma un dibujo bastante parecido a las alas de una mariposa; se conoce habitualmente como el atractor de Lorenz (aunque el propio Lorenz lo llama atractor de mariposa), y se ha convertido en la imagen visual más familiar e impactante del caos, incluso para quien no sabe qué representa en realidad esa imagen.

Resulta que el tiempo meteorológico unas veces es más caótico que otras. Cuando hoy en día los meteorólogos realizan simulaciones numéricas, no se limitan a tomar los datos exactos de los puntos de la red que representan las observaciones como materia prima para una única predicción. En vez de eso, para averiguar si los errores y las imprecisiones inherentes a las observaciones tienen una importancia significativa, habitualmente realizan el proceso correspondiente a cada predicción varias veces con ligeras variaciones en las condiciones iniciales. Si todas las predicciones resultan más o menos iguales, saben que se puede confiar en la pauta global de la predicción (o de las predicciones), podríamos decir que el sistema del tiempo meteorológico circunda uno de los pozos profundos del espacio de las fases. Pero, a veces, cuando realizan el proceso para la misma predicción con condiciones iniciales ligeramente distintas, obtienen unas «predicciones» muy diferentes para el tiempo que va a hacer unos pocos días más tarde.

Esto les indica que el tiempo meteorológico es un estado caótico, equivalente a las trayectorias que pasan sobre la barra de arena en el espacio de las fases, y no se puede confiar en ninguna de las predicciones.

[26] «¿Una mariposa que bate sus alas en Brasil puede desencadenar un tornado en Texas?» Esta publicación está incluida en su libro *The Essence of Chaos.*

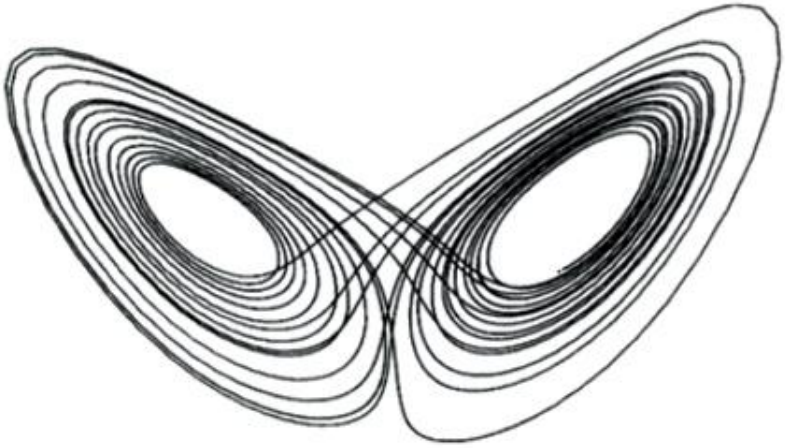

Figura 2.3. El atractor (la «mariposa») de Lorenz.

Esta es la razón por la cual los que hacen el pronóstico meteorológico en televisión parecen estar algunas veces más acertados con sus previsiones y otras veces no tanto. Como uno de ellos reconocía con aire pesaroso, «podemos predecir el tiempo con exactitud, siempre que éste no haga algo inesperado»[27].

Podemos ver cómo funciona el caos mediante la calculadora de bolsillo, utilizando la técnica de la iteración, que es tan importante en este tipo de trabajo numérico.

Tomemos una expresión sencilla, como es $2x^2 - 1$. Demos *a x* un valor comprendido entre 0 y 1, con unas cuantas cifras decimales, por ejemplo, 0,2468.

Utilicemos este número como *x* e introduzcámoslo en el cálculo para la próxima iteración. Obtendremos lo que parece ser una sucesión de números aleatorios, producto de un proceso completamente determinista que obedece a una ley sencilla. A continuación, intentémoslo de nuevo con un número que difiere del primero sólo en la última cifra decimal, por ejemplo, 0,2469.

Si tenemos la paciencia necesaria para seguir dando a las teclas de la calculadora (o la capacidad mental para confeccionar un sencillo programa informático que haga el trabajo por nosotros) descubriremos que se obtiene una sucesión completamente diferente de números aparentemente aleatorios después de unas pocas iteraciones. Ahora bien, en ambos casos cada número está determinado con exactitud por el

[27] Citado por Ian Stewart en *Does God Play Dice?* [Hay traducción al castellano: *¿Juega Dios a los dados?: la nueva matemática del caos*, Editorial Crítica, Barcelona, 1991 y 2001 (edición revisada), trad. de Miguel Ortuño, Jesús Ruiz Martínez, Rafael García Molina.]

número anterior en este proceso de retroalimentación, la pauta de los números es totalmente determinista. Esto es precisamente lo que Lorenz observó, y si hacemos la prueba de realizar cualquier proceso de este tipo con distintas calculadoras, no nos ha de sorprender que obtengamos «respuestas» diferentes, porque cada calculadora redondea los números de manera diferente al realizar sus cálculos internos.

También podemos ver un tipo de comportamiento diferente. Tomemos la expresión x^2 - 1, que tiene un aspecto muy similar a la expresión anteriormente mencionada. A continuación, para cualquier valor de *x* con el que comencemos (entre 0 y 1), después de un cierto intervalo de tiempo la pauta se asienta en un es- lado estable, oscilando entre 0 y -1. Se dice que es una pauta «periódica de período 2», porque, una vez que el sistema se ha asentado en esta pauta, con independencia de dónde comencemos, sólo hay que realizar dos pasos para volver a la situación inicial.

Asimismo podemos tener un comportamiento periódico en otros sistemas en los cuales sean necesarios más pasos para volver a la situación en la que habíamos comenzado, pero el número de pasos siempre será el mismo en un sistema determinado, con independencia de la situación de partida. Una sencilla ley establece un comportamiento periódico que converge en un atractor; otra ley, también sencilla y aparentemente muy similar a la primera, nos da un comportamiento «aleatorio» que es muy sensible a las condiciones iniciales. Si somos aficionados a este tipo de cosas, podemos seguir explorando[28]; descubriremos que algunas iteraciones convergen en un único valor y se quedan clavadas en él (un atractor simple, un sistema del que se dice a veces que tiene «período uno» porque un solo paso reiterativo nos lleva a la situación en que comenzamos), algunas son aparentemente aleatorias y otras resultan periódicas, siendo el caso de período dos solamente un caso sencillo que sugiere otras posibilidades mucho más complicadas.

Todas estas pautas de comportamiento se ven en el mundo real, donde pueden corresponder a cosas tan diversas como el goteo de un grifo, el modo en que varían las poblaciones de animales salvajes, o las oscilaciones del mercado bursátil. Pero en esencia hemos descubierto ya la sencillez subyacente de la que emergen el caos

[28] La contribución de Franco Vivaldi a *The New Scientist Guide to Chaos* (editada por Nina Hall) nos puede dar algunas ideas.

y la complejidad, leyes sencillas, procesos no lineales, sensibilidad a las condiciones iniciales y a la retroalimentación son los factores que hacen funcionar el mundo. Sin embargo, antes de ir a cosas realmente complicadas, parece conveniente terminar la historia de las órbitas planetarias, que es la cuestión a partir de la cual empezó todo esto.

Al igual que en el caso de la predicción meteorológica, la clave para conocer las propiedades a largo plazo que caracterizan al sistema solar surgió con la llegada de ordenadores electrónicos razonablemente rápidos, capaces de efectuar integraciones numéricas a una velocidad relativamente grande. Poincaré había demostrado que el sistema solar estaba en principio sometido a lo que actualmente llamamos caos, pero entonces, como ahora, *estaba* bastante claro que, en la práctica, los planetas similares a la Tierra describían órbitas estables durante mucho tiempo, de otro modo no estaríamos aquí para cavilar sobre estas cosas.

El modo de abordar la cuestión del caos dentro del sistema solar era examinar sistemas más bien parecidos al problema restringido a tres cuerpos, donde un pequeño objeto (o una serie de pequeños objetos) se mueve sometido a la influencia gravitatoria de dos objetos grandes. Los de mayor tamaño del sistema solar son el Sol y el planeta Júpiter, existiendo además toda una familia de pequeños objetos sometidos a su influencia gravitatoria, los miles y miles tic cuerpos rocosos conocidos como asteroides que describen sus órbitas a lo largo de un cinturón situado alrededor del Sol entre la órbita de Marte y la de Júpiter.

Los planetas del sistema solar forman dos familias, separadas por este cinturón de asteroides. Más cerca del Sol hay cuatro pequeños planetas rocosos (Mercurio, Venus, la Tierra y Marte); más alejados se encuentran cuatro grandes planetas gaseosos (Júpiter, Saturno, Urano y Neptuno)[29].

La razón básica para establecer esta división se ve con bastante claridad, cuando los planetas se estaban formando, el calor de aquel Sol joven se llevaba la mayor parte del material gaseoso sacándolo de la parte interior del sistema solar; pero en las regiones exteriores hacía el frío suficiente para que el gas se acumulara en los cuatro planetas gigantescos.

[29] Por razones históricas, el objeto conocido como Plutón todavía se suele llamar planeta, aunque, de hecho, sencillamente es el miembro de mayor (amaño perteneciente a un cinturón de objetos helados que se encuentran más allá de la órbita de Neptuno y que son los restos sobrantes de la formación del sistema solar.

Tampoco el origen del cinturón de asteroides es hoy en día un gran misterio. Los modelos y estudios realizados por ordenador con respecto a unos polvorientos discos de materia que se encuentran en tomo a estrellas jóvenes sugiere que los planetas se formaron «desde un principio» como granos de polvo que se adherían luego unos a otros para formar primero granos de mayor tamaño, luego rocas, y así sucesivamente. Sin embargo, en la zona situada entre Marte y Júpiter, la influencia gravitatoria de este último alteró este proceso antes de que pudiera finalizar y evitó que se formara un planeta completo. Las simulaciones más recientes sugieren que para entonces Júpiter había crecido lo suficiente para empezar a dominar con su influencia gravitatoria esta parte del sistema solar, y que quizá se habían formado seis o siete objetos del tamaño de Marte (que tiene una décima parte de la masa de la Tierra), pero, al estar sometidos a la influencia de Júpiter, chocaron entre sí violentamente y, en vez de adherirse unos con otros para constituir un solo planeta grande, se rompieron y fragmentaron para dar lugar al cinturón de asteroides, quedando Marte como único superviviente.

Pero este modelo pormenorizado de la formación de los asteroides se desarrollaría a finales de la década de 1970, cuando un estudiante de doctorado de Caltech, Jack Wisdom, decidió estudiar las órbitas actuales de los asteroides utilizando una nueva técnica numérica y el mejor ordenador que se podía conseguir entonces. Los asteroides no habían sido descubiertos hasta el siglo XIX, porque eran pequeños y de brillo muy apagado. Pero, a medida que crecía el número de asteroides observados, pronto se vio claramente que sus órbitas no están espaciadas de manera uniforme por el cinturón, sino que hay huecos correspondientes a órbitas vacías. Estos huecos se llaman lagunas de Kirkwood, porque fue el astrónomo estadounidense Daniel Kirkwood quien llamó la atención sobre su existencia a finales de 1860. Desde el principio se vio que había algo especial en relación con estas órbitas, corresponden a ciertos fenómenos llamados «resonancias» con la órbita de Júpiter. Sin embargo, hasta que llegó Wisdom (con su ordenador), no estuvo del todo claro por qué se mantenían vacías aquellas órbitas de resonancia.

Todos hemos experimentado y utilizado la resonancia cuando, de niños, jugábamos en un columpio. La resonancia es una manera de conseguir una gran repercusión a cambio de un esfuerzo relativamente pequeño, haciendo este esfuerzo justo en el

momento adecuado e impulsando un sistema del modo en que éste «desea» funcionar. La pequeña «inversión» de impulso que hacemos en el columpio debe realizarse en el momento preciso para que el balanceo describa un arco cada vez más amplio.

Algo del mismo tipo sucede cuando Júpiter le da un codazo a un asteroide que está en su órbita, cada vez que el asteroide pasa al lado de este planeta, entre él y el Sol. En los casos de la mayoría de las órbitas, estos codazos son más o menos aleatorios y ocurren en diferentes momentos en cada órbita, como si estuviéramos en un columpio y nos meneáramos sin un propósito determinado, por lo que no se produce un gran efecto. Pero las órbitas de resonancia son diferentes. Supongamos que un asteroide describe una órbita alrededor del Sol cuya duración es exactamente el doble de larga que la de la órbita de Júpiter. Entonces, el asteroide siempre recibe el codazo de Júpiter en la misma zona de su órbita. La influencia del planeta gigante puede contribuir a perturbar la órbita, del mismo modo que, impulsando el columpio en el momento adecuado, podemos hacer que oscile cada vez más apartado de la vertical. Esto no importaría mucho si la órbita fuera estable en el sentido tradicional, y una pequeña perturbación hiciera sencillamente que se desviara un poco hacia un lado y luego de nuevo hacia el otro (como los cambios a largo plazo que tienen lugar en las órbitas de Júpiter y Saturno, y que fueron estudiados por Laplace).

No obstante, si la órbita es del tipo de las que son sensibles a las perturbaciones, la resonancia podría desplazar rápidamente el asteroide de su órbita más o menos circular alrededor del Sol, haciéndolo pasar a una órbita mucho más elíptica, pero que todavía vuelve con la misma regularidad al punto de resonancia con Júpiter. Esto es exactamente lo que Wisdom descubrió, y publicó en 1982, el caos funciona en el cinturón de asteroides, en particular para las órbitas que están próximas a la resonancia 3:1 con Júpiter.

Pero ni siquiera esto resolvía el enigma de las lagunas de Kirkwood, porque una perturbación que puede desplazar a los asteroides sacándolos de estas lagunas, podría actuar fácilmente a la inversa, sacando a los asteroides de las órbitas elípticas para mandarlos de nuevo a llenar las lagunas (recordemos el atractor de dos lóbulos, como el que descubrió Lorenz).

Wisdom señaló que la solución de esta pieza del rompecabezas llegó con la constatación de que los asteroides expulsados de las lagunas de Kirkwood como resultado del caos, bajo la influencia de Júpiter, se trasladaban a órbitas que cruzaban las de los planetas interiores, incluido el planeta Tierra. Muchos de estos asteroides, durante la larga historia del sistema solar, han salido totalmente de lo predicho por las ecuaciones correspondientes, como resultado de su colisión con uno de los planetas interiores, por lo que no hay una reserva de la que se puedan extraer asteroides para cubrir las lagunas. Todos los planetas interiores del sistema solar (y también nuestra Luna) muestran en sus magullados rostros las cicatrices de muchos de estos impactos, y está ampliamente difundida la creencia de que uno de estos impactos, hace sesenta y cinco millones de años, puso fin en la Tierra a la era de los dinosaurios y abrió el camino para la irrupción de los mamíferos, incluidos nosotros mismos. Esto implica que debemos nuestra existencia de forma directa a los efectos del caos existente en el cinturón de asteroides, y trae consigo el espeluznante corolario según el cual el final de la civilización puede producirse de la misma manera.

No obstante, por si sirve de algo, nos puede proporcionar cierto consuelo el saber que el final de la civilización *no* llegará por el hecho de que la Tierra salte repentinamente a una órbita que la lleve a ser engullida por el Sol o a alejarse perdiéndose en las profundidades del espacio. Durante los años transcurridos desde 1982, los ordenadores han ido haciéndose mejores y más rápidos, y los matemáticos han ideado programas más ingeniosos, por lo que han podido proyectar la mirada a gran distancia hacia el futuro (o los posibles futuros) del sistema solar.

Como sus colegas de las ciencias del espacio, los investigadores parecen deleitarse construyendo unos acrónimos tortuosos o haciendo juegos de palabras para denominar sus proyectos, dos de los cuales se conocieron como LONG-STOP (que viene de «Longterm Gravitational Study of the Outer Planets» = Estudio gravitatorio de los planetas exteriores a largo plazo) y *Digitary Orrery* (un «orrery» es un modelo mecánico, a veces de relojería, que simula el sistema solar).

Algunos de estos estudios contemplan hasta cientos de millones de años en el futuro y muestran que, aunque las órbitas de los planetas son técnicamente

caóticas, a diferencia de la situación de los asteroides en las lagunas de Kirkwood hay sólo una probabilidad pequeñísima de que alguna de las órbitas de los planetas, incluida la de la Tierra, vaya a sufrir un cambio drástico durante el tiempo de vida que le quede al Sol en su forma actual, es decir, durante los próximos cinco mil millones de años, más o menos. Sin embargo, estos cálculos no predicen «la» órbita de la Tierra, o de cualquier otro planeta.

Son sensibles a las condiciones iniciales, aunque de un modo muy restringido. Por ejemplo, un modelo desarrollado por Jacques Laskar en el Bureau des Longitudes de París demostraba que la Tierra describirá esencialmente la misma órbita durante al menos 200 millones de años (el límite de esa integración numérica concreta). Pero también demostraba que una diferencia de sólo 15 metros al especificar la posición de la Tierra al principio de la integración no se mantiene como un «error» de 15 metros en todos los pasos de la iteración, sino que, debido a la no linealidad, crece hasta una inexactitud tan grande al cabo de 100 millones de años que no podemos decir en qué parte de su órbita estará nuestro planeta en ningún momento concreto después de tanto tiempo.

El error crece hasta un valor tan grande como el de la circunferencia de la órbita de la Tierra, es decir, casi 050 millones de kilómetros. Los errores de redondeo inherentes a cualquier integración numérica, la decisión sobre el número de cifras decimales con las que vamos a hacer los cálculos, ya sea utilizando lápiz y papel o un ordenador electrónico, afecta al resultado del cálculo prácticamente del mismo modo en que se veía afectado el sencillo modelo del tiempo meteorológico de Lorenz.

Una consecuencia de esto es que, si utilizamos un modelo así para calcular cuál va a ser la órbita de la Tierra (o cualquier otra órbita) durante los próximos 100.000 años, y luego invertimos el cálculo, no volveremos al punto de partida. Si el sistema es razonablemente estable, volveremos a alguna situación próxima a la de partida, en la misma zona del espacio de las fases, pero quizá todo lo que podamos decir sea que el planeta en cuestión permanece más o menos en la misma órbita alrededor del Sol. Las trayectorias calculadas no son reversibles, aunque hayamos llegado a pensar que los objetos que obedecen las leyes de Newton deberían seguir trayectorias reversibles.

Mientras esto sea sólo un artificio para realizar los cálculos, no cambia mucho el modo en que pensamos sobre el universo. En el caso de planetas como la Tierra, por encima de un cierto número de lugares decimales, la órbita que calculamos (o la trayectoria a través del espacio de las fases) no cambia mucho (aunque *sí* cambia), por más que cambiemos el proceso de redondeo en el ordenador. Obtenemos una órbita diferente que es casi exactamente la misma que la resultante del cálculo anterior. Cada vez que realizamos el cálculo con una ligera modificación de las condiciones iniciales obtenemos una órbita similar.

Todas las órbitas ocupan prácticamente la misma zona del espacio real y la misma del espacio de las fases, por lo cual, aunque la órbita se desplace caóticamente de una posibilidad a otra, no se sale de esta gama de posibilidades en ninguna escala temporal que sea importante para los seres humanos. Se trata de una especie de caos restringido. A veces, se hace una analogía con el comportamiento de la bola de una ruleta, después de haber sido lanzada en la rueda que está girando y antes de que se encaje en el hueco correspondiente a un número. La bola rebota de una manera esencialmente caótica, pero sus rebotes están todos ellos restringidos a «órbitas» que quedan dentro del borde de la rueda (suponiendo que el croupier haya hecho su trabajo debidamente).

Se podría establecer una analogía más casera con alguien que se evita el lío mañanero de elegir un par de calcetines comprándose solamente calcetines verdes. Así, sabe que puede tomar al azar dos calcetines del cajón y que ambos serán verdes, aunque haya calcetines diferentes con propiedades distintas. Incluso con los ojos cerrados, nunca sacará del cajón un calcetín rojo.

Esto plantea una cuestión a la que hay que referirse, aunque es más bien un aspecto colateral del hilo principal de nuestra historia. Hay personas que creen que los planetas interiores del sistema solar, en particular Venus, han cambiado sus órbitas radicalmente durante los últimos milenios y que dichos cambios se mencionan en varios mitos y leyendas de la antigüedad. Esta creencia se basa en una interpretación muy imaginativa de aquellos mitos antiguos y en un conocimiento muy pobre de la dinámica orbital y fundamentos tales como la ley de la conservación del momento; sin embargo, y por desgracia, cuando algunos astrónomos respetables hablan sobre el caos en el sistema solar, lo que hacen es

echar grano al molino de tales fantasías, y podemos ver cómo se sugiere que la teoría del caos «explica» o «demuestra» que Venus no se ha desplazado a su órbita actual hasta tiempos muy recientes.

Nada de esto. Los cálculos pueden realizarse tanto hacia atrás como hacia adelante en el tiempo, y muestran siempre el mismo tipo de pauta, no podemos decir *exactamente* dónde estaba Venus en su órbita, por ejemplo, el 4 de julio del año 5.000.000 a.C., pero podemos decir con gran seguridad en qué clase de órbita estaba entonces: una que era como la actual (un calcetín verde, no uno rojo). Es totalmente nula la probabilidad de que Venus, o cualquiera de los otros planetas, haya experimentado un cambio orbital drástico en algún momento durante los últimos cinco millones de años, que cubren de sobra el tiempo transcurrido desde que nuestra línea genealógica se separó de la de los otros monos africanos. Es seguro que ningún ser humano ha visto jamás a Venus vagando por el cielo mientras cambiaba de órbita.

Sin embargo, antes de dejar el tema de los planetas, merece la pena mencionar que el caos no sólo afecta a las órbitas. Los planetas giran sobre sus ejes (la Tierra lo hace una vez cada veinticuatro horas), y oscilan mientras giran, a causa de la influencia gravitatoria del Sol (de manera similar, la peonza que hace bailar un niño oscila mientras gira, a causa de la influencia gravitatoria de la Tierra).

Entre el período del giro y el de la oscilación pueden producirse resonancias, que en la zona correspondiente del espacio de las fases pueden llevar a un cambio repentino y caótico en el ángulo de inclinación del planeta (su oblicuidad). Esto podría suceder, por ejemplo, si el giro del planeta se frena gradualmente y, como resultado, el sistema entra en una región sensible del espacio de las fases. La inclinación de la Tierra con respecto a la vertical (relativa a una línea imaginaria que uniría la Tierra y el Sol) es actualmente de unos 23 grados, y ésta es la causa del ciclo estacional del clima. Pero resulta que la presencia de nuestra gran Luna actúa como un estabilizador que evita variaciones bruscas en la inclinación del planeta. No es éste el caso de Marte, o de los demás planetas interiores, donde no hay un gran satélite que actúe como estabilizador (las dos lunas diminutas de Marte, de las que se cree que son asteroides capturados, son totalmente inadecuadas para esta tarea).

Resulta que la inclinación media de Marte es de unos 24 grados, pero las simulaciones realizadas por ordenador muestran que esto puede cambiar drásticamente en una gama muy amplia de valores, al menos en 20 grados por encima o por debajo de esta media. También hay evidencia directa de que Marte ha experimentado cambios climáticos radicales en el pasado, en forma de lo que parece ser lechos fluviales que se han secado y que erosionan la superficie actualmente árida del planeta rojo; estos acontecimientos climáticos extremos pueden atribuirse a épocas pasadas en las cuales la helada región polar se inclinó hacia el Sol, produciéndose una situación extrema de calores estivales, y se calentó lo suficiente como para hacer que el casquete polar de agua o dióxido de carbono helados se evaporara o se fundiera.

Probablemente el mismo tipo de oscilación se haya producido en Venus y Mercurio, pero no hay huellas visibles sobre la superficie de estos planetas, porque en Mercurio no hay atmósfera y en Venus la superficie parece haber sido removida por una intensa actividad volcánica.

Esto excluye a la Tierra, que se mantiene hasta ahora felizmente estable. Sin embargo, a muy largo plazo, dado que la Luna se aleja de la Tierra como resultado de la actividad de fuerzas mareomotrices, su influencia se debilitará y, en un futuro lejano, la inclinación de nuestro planeta variará también caóticamente, aumentando quizá repentinamente hasta en 90 grados, con lo cual en verano hará tanto calor en un polo como actualmente en el ecuador, mientras el otro polo padece un duro y oscuro invierno, invirtiéndose la situación seis meses más tarde.

Esto no será en absoluto lo ideal para la vida tal como la conocemos ahora, por lo tanto, aunque el caos en el sistema solar pueda haber sido responsable de nuestra existencia (por el impacto del asteroide que hizo desaparecer los dinosaurios), es la *ausencia* de caos en la cambiante oblicuidad de la Tierra lo que, gracias a la presencia de la Luna, ha permitido que la vida en la Tierra evolucione en circunstancias climáticas más o menos estables durante miles de millones de años.

Sin embargo, hay otro lance imprevisto en este relato. La mejor hipótesis de que disponemos para justificar la presencia de una Luna tan grande asociada con la Tierra es que su aparición se produjo durante las primeras etapas de la historia del sistema solar, cuando un objeto del tamaño de Marte y procedente del cinturón de

asteroides se vio desviado de su órbita por electo del caos y colisionó con la Tierra, lanzando por electo del choque una cantidad de material fundido que luego formó la Luna en el espacio.[30]

Hay un último asunto, quizá el más importante de todos, sobre el que queremos llamar la atención antes de dejar la investigación sobre el sistema solar. Un poco antes hemos señalado que, si utilizamos una integración numérica para calcular la órbita de un planeta y luego invertimos el cálculo, no volvemos al punto de partida. Las trayectorias calculadas no son reversibles. Ya dijimos que, en la medida en que esto es sólo un efecto artificial debido al modo en que hacemos los cálculos, no cambia mucho nuestra manera de concebir el universo, ya que las leyes de Newton son, en principio, reversibles. Pero, ¿cómo «hace» el universo sus cálculos? Al igual que Laplace, podríamos imaginarnos una inteligencia perfecta (viva o electrónica) que sería capaz de almacenar información exacta sobre todas las propiedades importantes (como la posición y el momento de cada partícula) y luego llevar a cabo los cálculos correspondientes de una manera perfecta. En este caso seguramente serían reversibles. Pero, ¿cuántas cifras decimales necesitaríamos para conseguir la perfección? ¿Cuánta memoria tendría que tener un ordenador perfecto?

A primera vista, esto parece un logro imposible. Sin embargo, resulta que tiene una respuesta muy sencilla, relacionada con la naturaleza de los números. La mayoría de la gente, cuando habla de números, piensa en números naturales, como 1, 2, 27, 44, 196, etc. También estamos acostumbrados a fracciones sencillas: 1/2, 1/3, 3/4, y otras por el estilo. Para muchos usos de la vida cotidiana no pasamos de esto.

Incluso cuando manejamos decimales, suele ser únicamente para referirnos a algo como el dinero, y lo habitual es utilizar sólo un par de decimales, por lo que nos resulta familiar la idea de que £17,46 significa «diecisiete libras y cuarenta y seis peniques». Pero hay una cantidad literalmente infinita de números que nunca se tiene en cuenta en la vida cotidiana. Aún peor, hay una infinidad de números entre cualquier par de números que podamos pensar. Esto es sumamente obvio por lo que respecta a los decimales. Tomemos los enteros 1 y 2. Entre el uno y el dos debe haber un conjunto de números que tienen una cifra decimal (1,1; 1,2;... 1,8;

[30] Véase *Vire on Earth,* de John y Mary Gribbin.

1,9), y entre cada dos de estos números hay otro conjunto de números que tienen una cifra más en su parte decimal (1,11; 1,12; 1,13... 1,18; 1,19), y entre cada par de estos números..., ya se hace el lector una idea.

Con independencia de cuántas sean las cifras decimales que haya en los dos números de que partimos, el mismo razonamiento nos dice que hay una cantidad infinita de números entre ambos, incluso en el caso de números como 247,8503468295667 y 247,8503468295668, o en el de números con suficientes cifras decimales como para llenar todo este libro y que difieren sólo en el último dígito, o números lo bastante grandes como para llenar de libros todo el universo y que sólo tienen diferente el último dígito.

Esto no tendría gran importancia si fuera meramente una curiosidad matemática. Sabemos, por ejemplo, que algunos infinitos se manejan bien (como en el caso de ciertas series infinitas) y pueden representarse de una manera sencilla. Por ejemplo, la fracción 1/3 corresponde a un número cuya parte decimal tiene una cantidad infinita de dígitos: 0,3333333333... Pero se puede escribir de una manera muy sencilla y compacta, sin llenar el universo de libros que contengan aburridas hileras de treses. Otros tipos de expresiones decimales pueden representarse también de forma compacta.

Por ejemplo, el número decimal 0,675486754867548...se puede expresar escribiendo «tómense los dígitos 67548 y repítanse siempre».

Esta forma compacta de expresar algo que es demasiado grande para manejarlo convenientemente si se escribe completo se suele llamar algoritmo (y podría representarse tanto con palabras como mediante una sencilla expresión matemática); de una expresión que puede representarse así de manera compacta se dice que es algorítmicamente comprimible. Pero, incluso los antiguos griegos, que nunca descubrieron la idea de los números decimales, sabían que existían algunos números que no son algorítmicamente comprimibles y no pueden escribirse de forma compacta. Aún peor: *la mayoría* de los números no pueden escribirse de esta forma tan sencilla.

Los números que pueden escribirse de manera sencilla son fracciones, son cocientes entre números enteros, como 4/3 o 29847/65109 (incluso los enteros son en cierto modo cocientes y pueden escribirse como 2/2 = 1, 8/4 = 2, y así sucesivamente).

Porque se expresan mediante cocientes o razones, los griegos los llamaron números racionales. Sin embargo, hay otros números, de los cuales el más famoso es *π,* que nunca pueden escribirse como la razón de dos números. Por este motivo, se llaman números irracionales, y sólo pueden expresarse como una hilera infinita de dígitos decimales que no se repiten[31]. Los griegos, al carecer de decimales, no lo expresaron así, pero eran muy conscientes de que tales números irracionales existían. ¡Y de que *la mayoría* de los números son irracionales! Esto incide en la idea fundamental del caos y del concepto newtoniano (o laplaciano) de reversibilidad, como podemos ver intentando especificar la posición exacta de un sistema en el espacio de las fases.

Este «sistema» puede ser muy pequeño, una sola partícula desplazándose por todo el interior de una caja, bajo la influencia de la gravedad. El estado de la partícula queda determinado por su posición y su momento, y Newton demostró que se comportará como si toda su masa estuviera concentrada en un punto matemático situado en el centro de la partícula. «Todo» lo que hemos de hacer para determinar la posición del sistema en el espacio de las fases es especificar la ubicación de ese punto y el momento de la partícula.

Pero, para simplificar el problema aún más, centrémonos en la ubicación. Podemos hacer que las cosas sean aún más sencillas suponiendo que la partícula se desplaza en línea recta, quizá cayendo bajo la influencia de la gravedad. Entonces, sólo tenemos que especificar la ubicación de la partícula en dicha línea recta, y éste es el problema más sencillo que uno se puede imaginar en física. Sin embargo, esto resulta imposible, salvo en muy pocos casos. Supongamos que sabemos que la partícula se encuentra en algún lugar situado entre A y B.

Necesitamos saber qué fracción *exacta* de la distancia entre A y B ha cubierto ya. No hay problema si se trata de un tercio del recorrido sobre la línea recta, o 98/317 de todo el recorrido, o de cualquier otra fracción racional. Pero, entre cada par de puntos de la línea representados por fracciones racionales hay una cantidad infinita de puntos representados por números irracionales, y cada lino de estos puntos sólo se puede especificar mediante una hilera infinita de dígitos, sin que haya modo alguno de comprimirlos en una forma compacta.

[31] La expresión racional 22/7, utilizada a menudo como una aproximación de π, es precisamente eso, una aproximación: no es realmente el número π.

Si, por ejemplo, la partícula está a una distancia 1/π sobre la línea que une A y B, podemos representar esto con tanta precisión como queramos, calculando la expresión con tantas cifras decimales como se desee, pero no podemos representarlo *exactamente* escribiendo un número infinito de dígitos. ¡Y esto es sólo para el caso de una partícula que se desplaza en línea recta!

En un sistema que sea suficientemente sensible a las condiciones iniciales, con independencia del número de dígitos que tomemos para los cálculos, siempre es posible que, como Lorenz descubrió, todo el futuro del sistema pueda depender en gran medida del valor del dígito siguiente, el que de hecho hemos descartado.

Esto significa que se necesita un ordenador dotado de una memoria infinita para especificar el estado de *una sola* partícula. Ningún ordenador puede ser más grande que todo el universo y si definimos el universo como «todo lo que existe», esto significa que el único sistema que puede reproducir el comportamiento del universo con todo detalle es el propio universo. Incluso si, como pensó Laplace, el universo es totalmente determinista y el futuro está contenido completamente en su estado actual, no existe en absoluto modo alguno de predecir o conocer el futuro, salvo observando la evolución del universo. Tanto si existe el libre albedrío como si no, el universo se comporta como si tuviéramos esa libertad, y esto es realmente lo único que importa. El universo ignora su propio futuro y es el más rápido de todos sus simuladores.

Pero, ¿qué sucede con la reversibilidad y la flecha del tiempo? Se suele hablar de «agitar una pared mágica» para referirse a lo que es la inversión del movimiento de cada partícula del universo (o de las partículas que hay en una caja de gas) con el fin de que el tiempo corra hacia atrás. Pero actualmente esto no se considera factible. Es imposible invertir con precisión el movimiento de tan siquiera una sola partícula, incluso dejando a un lado sutilezas tales como la teoría de la relatividad, que plantea cuestiones de fondo sobre cómo se pueden invertir *simultáneamente* todos los movimientos que se dan en el universo, si las señales no pueden viajar más rápido que la velocidad de la luz, y si lo que se indica al decir «simultáneo» depende del punto desde el cual se observa el universo. Para invertir el movimiento de una partícula (salvo en casos muy raros en los que dicho movimiento se expresa para una fracción de segundo mediante números racionales) primero tendríamos

que especificar su estado presente utilizando un número infinito de cifras decimales, con el fin de hacer luego una inversión exacta.

En principio esto no es posible, y no sólo por insuficiencias de tipo humano. Lo que sí parece es un premio más que adecuado por aprender que el caos es inevitable. El universo, en principio, *no se puede* predecir con todos sus detalles; pero, del mismo modo, *tampoco se puede,* en principio, invertir el tiempo.

Las ideas que hemos presentado en este capítulo son los cimientos, la profunda simplicidad, sobre los cuales se construye la complejidad del universo. Construyendo a partir de estos cimientos, podemos empezar a ver cómo han emergido, a partir del caos, la complejidad del universo y la propia vida. Si el lector está ya familiarizado con el caos y los fractales, puede saltarse el próximo capítulo y pasar directamente al capítulo 4; pero espero que se quedará con nosotros para hacer una incursión breve, pero emocionante, en los dominios del caos auténtico, y para comentar el tema que, por su relación con el caos, ha llegado a ser obligatorio en cualquier investigación sobre el mismo: la teoría de los fractales. No hay que pensar que se trata de una digresión, ya que esto nos llevará a las fronteras de la propia vida, y hemos de tener en cuenta que el conocimiento del modo en que surgió la vida es el objetivo último de nuestra exposición.

Capítulo 3
El caos surgido del orden

El tipo de caos sobre el que vamos a hablar ahora, lo que los científicos del siglo XXI llaman «caos», no es lo mismo que la clase de caos a la que se refieren los antiguos, y tampoco lo que denominamos caos en la vida cotidiana. Este tipo de caos es completamente aleatorio e impredecible, incluso en principio. Pero el caos sobre el que vamos a hablar aquí es algo totalmente ordenado y determinista, siendo cada paso consecuencia del anterior según una cadena ininterrumpida de causa y efecto que es absolutamente predecible en cada etapa, en principio. Lo que pasa es que, en la práctica, es imposible predecir con detalle lo que va a suceder, porque no podemos hacerlo más rápido que los acontecimientos que se desarrollan en tiempo real. El ejemplo clásico de estas dificultades se toma del estudio de las turbulencias, que se producen en muchos lugares, pero podemos ilustrar su importancia con uno de los casos más sencillos: el modo en que cambia el flujo de agua de un río a medida que crece la corriente.

Para nuestro sencillo modelo, imaginemos un río que fluye en calma y en el que hay una gran roca que sobresale en la superficie de sus aguas. El flujo se divide para rodear la roca y vuelve a unirse sin que se perciba señal alguna allí donde las aguas se juntan de nuevo, de tal modo que unas pequeñas virutas de madera que floten sobre el río seguirán estas «líneas de corriente». Si llueve río arriba, el flujo aumentará y experimentará al menos tres cambios diferentes, que todos nosotros hemos visto ya alguna vez, probablemente sin pensar mucho sobre lo que significan. En primer lugar, a medida que el flujo aumenta, se forman pequeños remolinos detrás de la roca.

Estos vórtices se mantienen en el mismo lugar y una viruta de madera que vaya flotando río abajo puede quedar atrapada en uno de ellos y permanecer allí dando vueltas durante mucho tiempo. Este comportamiento es muy similar al que se observa dentro del espacio de las fases en el atractor de Lorenz, una especie de remolino en torno al cual un sistema recorre el mismo ciclo durante un largo intervalo de tiempo. En el espacio de las fases, este tipo de atractor se conoce como ciclo límite, porque, con independencia de dónde comience el sistema, en el límite

será atraído a esa pauta de comportamiento repetitiva específica[32]. Utilizando un lenguaje en cierto modo poco ortodoxo (aunque razonablemente claro), podríamos llamar «ciclo límite» a cada pequeño remolino que se forma detrás de la roca.

En la fase siguiente, cuando aumenta la velocidad del agua que fluye río abajo, se forman vórtices detrás de la roca, pero no permanecen allí. Se van (o los lleva la corriente), y se desplazan río abajo, manteniendo su existencia independiente durante un tiempo hasta que se disuelven en el flujo de agua. Mientras hacen esto, se forman nuevos vórtices detrás de la roca y se van de allí a su vez. Ahora bien, una viruta de madera podría ser atrapada por uno de estos remolinos y sería arrastrada río abajo, girando todavía dentro del vórtice mientras éste dura.

Cuando el flujo de agua aumenta aún más, la zona de la parte de detrás de la roca donde los vórtices sobreviven se queda cada vez más reducida, y dichos vórtices se forman y se rompen casi inmediatamente, produciendo una superficie picada en la que sólo parece haber fluctuaciones irregulares, turbulencias. Finalmente, cuando la corriente es lo suficientemente rápida, desaparece todo rastro de orden en la zona situada detrás de la roca. No se forman vórtices y toda la superficie del agua se rompe detrás de la roca, dando lugar a un movimiento caótico impredecible.

Dicho así, hay dos características fundamentales en el trecho que va desde el orden al caos y que se plasma en este ejemplo de las turbulencias. En primer lugar, algo está cambiando. Parece casi demasiado obvio como para que haya que mencionarlo, pero es fundamental en toda esta historia.

Lo que parece ser el mismo sistema puede describirse de una manera sencilla si se dan determinadas condiciones, o hablando de caos si las condiciones son otras, pero entre ambas posibilidades hay una zona compleja donde suceden cosas interesantes (en este caso, el «nacimiento» de vórtices). Es sólo una cosa, un parámetro, lo que está cambiando en este sistema: la velocidad a la que fluye el agua. Incrementar el valor de este parámetro único hasta más allá de un punto crítico resulta suficiente para desencadenar el comienzo del caos.

En segundo lugar, cuando se examina con detalle el modo en que los vórtices se rompen detrás de la roca durante la compleja etapa intermedia que transcurre entre el orden y el caos, se descubre algo muy interesante. Este descubrimiento requiere

[32] Hablando en sentido estricto, un ciclo límite sería un bucle cerrado en el espacio de bases, pero para lo que tratamos aquí, podemos considerarlo como un único lóbulo del atractor de Lorenz.

una cuidadosa atención a los detalles, pero no precisa un equipamiento muy técnico (al menos, no para ver la amplia pauta de lo que está sucediendo). Leonardo da Vinci había llamado la atención sobre ello medio siglo antes. Afirmó que un remolino que se aparta de la roca y desciende río abajo no se limita a desaparecer. Se rompe interiormente en vórtices más pequeños, que, a su vez, se rompen en otros aún menores, formándose remolinos dentro de los remolinos en lo que parece ser un proceso de bifurcación infinito. El camino hacia el caos incluye lo que parece ser un número infinito de opciones que operan a una escala infinitamente pequeña, al menos así es en el caso de las turbulencias. ¿Podemos ver el funcionamiento de algo similar en algún otro sitio?

Por supuesto, la respuesta es «sí». Un lugar en el que se puede ver algo similar, también relacionado con el fluir del agua, es en el goteo de un grifo. Si comenzamos con el grifo cerrado y lo abrimos muy ligeramente, nos será fácil producir un goteo continuo, que resuena al caer en la pila como el ritmo monótono de un tamborilero poco inspirado, tip, tip, tip, tip.... Es un ritmo con período uno, en la jerga de este arte. Si abrimos el grifo un poco más, seguirá siendo bastante fácil (incluso yo puedo hacerlo) ver y oír cómo el sistema cambia a un ritmo de período dos, como si el percusionista fuera un poco más habilidoso al tocar rat-at, rat- al, rat-at, rat-at... Cuando abrimos el grifo aún más, el asunto empieza a resultar más interesante, y luego completamente desordenado. Después de esta segunda fase, es bastante difícil detectar el tema que nos interesa.

En el caso de un grifo determinado que hay en la casa donde yo vivo, puedo abrirlo hasta un punto en el que más o menos me convenzo de que lo que estoy oyendo es un ritmo de cuatro tiempos: *rat-a-tat-tat, rat-a-tat-tat, rat-a-tat-tat, rat-a-tat-tat*, es decir, un ritmo de período cuatro. Sin embargo, para ser sincero, no estoy seguro de si es así sólo porque espero oírlo de esta manera, porque sé que es lo que ponen de manifiesto los experimentos realizados en laboratorios en condiciones cuidadosamente controladas[33].

Este proceso se conoce como duplicación del período, por razones obvias, pero no

[33] Curiosamente, sin embargo, después de haber escrito esta sección, me encontraba haciendo compras después de que hubiera caído un fuerte chaparrón y, mientras esperaba haciendo cola en la ferretería, observé el goteo que caía por fuera de la ventana desde un canalón agujereado. En ese momento. ¡pude ver claramente cómo caían conjuntos de cuatro gotas y oí de manera nítida cómo su impacto sobre el pavimento hacía exactamente el ritmo *rat-a-tat-tat*!

puede continuar indefinidamente. En cierto momento crítico (muy pronto, en el caso de la apertura paulatina del grifo) las duplicaciones repetidas del período (bifurcaciones repetidas) producen una pauta de goteo complicada y aparentemente aleatoria, cuando el sistema se vuelve caótico. Abriendo aún más el grifo, las gotas se fusionan para formar un flujo uniforme; si lo abrimos todavía más, el flujo se vuelve turbulento y desordenado. Pero, por ahora, no queremos ir tan lejos, aunque volveremos en breve al asunto de las turbulencias.

El mejor ejemplo de la ruta hacia el caos a través de la duplicación del período procede en su totalidad de otra área de la ciencia, e incide en lo fundamentales y amplias que son las implicaciones del caos. Hay una ecuación muy sencilla, llamada la ecuación logística, que funciona muy bien para describir cómo cambia la población de una especie de una generación a la siguiente. Para no complicar las cosas, podemos suponer que se trata de algún tipo de insecto cuya población adulta muere al completo en invierno, después de poner los huevos que se abrirán en la primavera siguiente para dar vida a una nueva generación. Partimos de una población que cuenta con *x* individuos.

Obviamente, el número de individuos de la generación siguiente (el número de los que sobreviven para reproducirse a su vez) depende del número de huevos que se abran (el índice de natalidad), que a su vez depende de cuántos huevos se habían puesto, por lo que, si como media (calculando el promedio sobre machos y hembras) cada insecto pone B huevos, la nueva población será *Bx.* Pero esto no tiene en cuenta a los individuos que, al no conseguir encontrar alimento, pasan hambre y no sobreviven para reproducirse a su vez. Esta tasa de mortandad depende del tamaño de la población inicial, cuantos más individuos haya, más difícil será que cada uno de ellos obtenga alimento suficiente.

Podemos simplificar las cosas un poco más estableciendo un límite máximo para la población (bastante razonable, si pensamos, por ejemplo, en el número de pulgones que pueden habitar en un rosal) y dividiendo la población real por este número, de tal modo que el valor que usemos para x esté siempre entre 0 y 1. Este truco se llama renormalización. A continuación, con el fin de tomar en cuenta la tasa de mortandad prematura, podemos multiplicar el factor de crecimiento *Bx* por un nuevo término (1 - x). Si las cifras de población son muy bajas (tan cerca de cero

que sean casi despreciables), casi todos los individuos sobreviven y encuentran alimento, por lo que la tasa de crecimiento es casi igual a *Bx*; si la población es muy numerosa, *x* está muy cerca del valor 1, (1 - x) está muy cerca de cero, y la mayoría de la población muere de hambre o cae presa de los depredadores. Entretanto, la población puede ascender o descender de una generación a la siguiente dependiendo del valor exacto del índice de natalidad B. Podemos ver cómo cambia la población para los distintos valores de B iterando la expresión

$$x \text{ (siguiente)} = Bx\,(1 - x)$$

donde x (siguiente) es la población de la generación siguiente. Efectuando el producto del segundo miembro de la igualdad, éste se convertirá en $Bx - Bx^2$, lo que pone de manifiesto el hecho de que este proceso no es lineal (a causa del término que lleva x^2) e incluye una retroalimentación (en el proceso de iteración).

Si B es menor que 1, esto significa que la población no consigue reproducirse a sí misma de una generación a la siguiente. Cada individuo adulto deja, por término medio, menos de un descendiente, lo cual es la receta para el desastre evolutivo, y la población finalmente debe extinguirse, con independencia del valor x con que hayamos empezado. Cuando B es mayor que 1, suceden fenómenos interesantes, y éstos pueden investigarse para distintos valores de B y x, utilizando una calculadora o un ordenador, o incluso lápiz y papel.

A partir de la década de 1950, muchos ecologistas hicieron precisamente eso, utilizando varias versiones de la ecuación logística (que se puede adaptar fácilmente para incluir los efectos de la depredación en la población, o de dos poblaciones que ejercen una interacción mutua de otras maneras, como, por ejemplo, compitiendo entre ambas para conseguir alimentos) para intentar encontrar modelos que pudieran darles una idea del comportamiento de las poblaciones reales de diversas especies.

Pero, en gran medida por las limitaciones de los aparatos de cálculo disponibles, no demostraron las posibilidades más interesantes y se centraron en las implicaciones más sencillas de la ecuación, algo equivalente a que los expertos en hidrodinámica ignorasen las turbulencias y se centrasen en el estado del río cuando los remolinos

se forman y rompen desde detrás de la roca de una manera ordenada.

Si B es mayor que 1 pero menor que 3, resulta lo que llamaríamos actualmente un atractor para esta sencilla ecuación logística. Con independencia del valor de la población que tomemos como punto de partida (cualquier valor de *x* comprendido entre 0 y 1), después de un número suficientemente grande de generaciones este valor se asienta en un nivel estable, correspondiente a una población constante. El valor exacto al que desciende crece ligeramente cuando B aumenta, y para valores de B cercanos (pero inferiores) a 3, baja a 0,66, que corresponde a 2/3 de la población máxima posible. Puede comenzar con un número pequeño y luego ascender, sobrepasando este valor y posteriormente zigzagueando en torno a 0,66 por arriba y por abajo en generaciones alternas, acercándose cada vez más al atractor (como la aproximación de π mencionada al principio del capítulo 2), o puede empezar con un número grande y descender hacia este valor, sobrepasándolo y oscilando de la misma manera.

Pero, si realizamos una cantidad suficiente de iteraciones, siempre se establece en 0,66. Si hacemos que B aumente, las oscilaciones se vuelven más extremas y pasa más tiempo hasta que logramos una estabilización, pero, mientras B sea menor que 3, estas oscilaciones convergen en el atractor. Sin embargo, cuando B llega a valer 3, sucede algo diferente.

En cuanto B es tan sólo un poquito mayor que 3, la pauta cambia. Una vez que la iteración se ha llevado a cabo bastantes veces, la población experimenta un cambio entre dos niveles diferentes y constantes en generaciones alternas. El atractor único se ha dividido en dos (se ha bifurcado) y el período se ha duplicado, pasando de 1 a 2.

Esto se puede entender en términos de poblaciones reales. En un año, hay una población muy numerosa que agota pronto todo el alimento. Como resultado, muchos individuos pasan hambre y mueren sin reproducirse. En consecuencia, la generación siguiente tiene una población pequeña, cuyos miembros encuentran todos ellos gran cantidad de alimento y ponen huevos, y así sucesivamente.

Si expresamos todo esto en una gráfica, para valores de B que inicialmente son menores que 3 y luego superan este valor, cuando el atractor único se bifurca, obtenemos un diagrama que se parece a un diapasón colocado sobre uno de sus

lados.

Todo esto es claramente suficiente para explicar la cuestión con palabras y dibujar diagramas, una vez que se han efectuado los cálculos, pero éstos son muy tediosos. El cálculo de las iteraciones para un solo valor de B resulta ya aburrido y, para ver con detalle lo que sucede cerca del valor crítico 3, necesitamos llevar a cabo muchas iteraciones para gran cantidad de valores de B ligeramente diferentes.

La primera persona que comprobó el comportamiento de la ecuación logística con el tipo de detalles necesario para mostrar cómo cambia la pauta global de comportamiento a medida que se aumenta el valor de B fue el físico nacido en Australia y convertido en ecologista Robert May, que trabajaba en Princeton a principios de la década de 1970.

Antes de cumplir los cuarenta años, May tenía justo la formación precisa para aplicar las teorías de la física y las matemáticas a la biología, y se encontraba precisamente en el momento y el lugar adecuados para aprovechar la potencia y la velocidad cada vez mayores de los ordenadores. Tras haber encontrado la bifurcación cuando B = 3, el paso siguiente era obvio y consistía en aumentar todavía más el valor de B y observar lo que sucedía. El resultado fue asombroso.

Con un valor de 3,4495, se bifurca *cada una* de las dos púas del diagrama en forma de diapasón, produciendo un sistema que oscila entre cuatro poblaciones diferentes (período 4). Con B = 3,56, cada uno de estos atractores se divide en dos y la población salta arriba y abajo entre ocho niveles diferentes; con B = 3,596, otra duplicación produce dieciséis niveles de población posibles, y en este momento cualquier biólogo que estudie una población real de seres vivos que cumpla esta sencilla ley determinista se sentirá fuertemente impulsado a ver un cierto orden entre las fluctuaciones aparentemente caóticas de una población que salta de un nivel a otro de una generación a la siguiente.

Las bifurcaciones, como ya se puede ver, se acercan unas a otras cada vez más a medida que aumenta el valor de B, y los trabajos que desarrollan la obra anterior de May muestran que con B = 3,56999 el número de atractores existentes para la población llega al infinito, de tal modo que cualquiera que esté estudiando cómo cambia la población de año en año se encontrará ante un auténtico caos determinista.

Pero, todavía hay más. Aunque en la mayoría de los casos se produce el caos para valores de B mayores que 3,56999, hay pequeñas escalas de valores de B donde el orden se restablece, una especie de ventana despejada entre la confusión del caos. Por ejemplo, si B es un poco mayor que 3,8 y un poco menor que 3,9, el sistema parece asentarse en un estado estable parecido al comportamiento que se da cuando B es menor que 3.

Sin embargo, a medida que el valor de B asciende lentamente, vemos una vez más bifurcaciones reiteradas (duplicaciones del período), como la pauta que observamos para valores de B situados por encima de 3. Pronto atravesamos las mismas etapas por las que habíamos pasado antes, y el caos reaparece. Pero la pauta que vemos es, de hecho, muy similar a la original, sólo que a una escala menor. Dentro de esta nueva versión del caos a menor escala hay una ventana, exactamente igual que la ventana en la que encontrábamos orden para valores de B comprendidos entre 3,8 y 3,9; y dentro de la ventana se repite de nuevo la pauta completa.

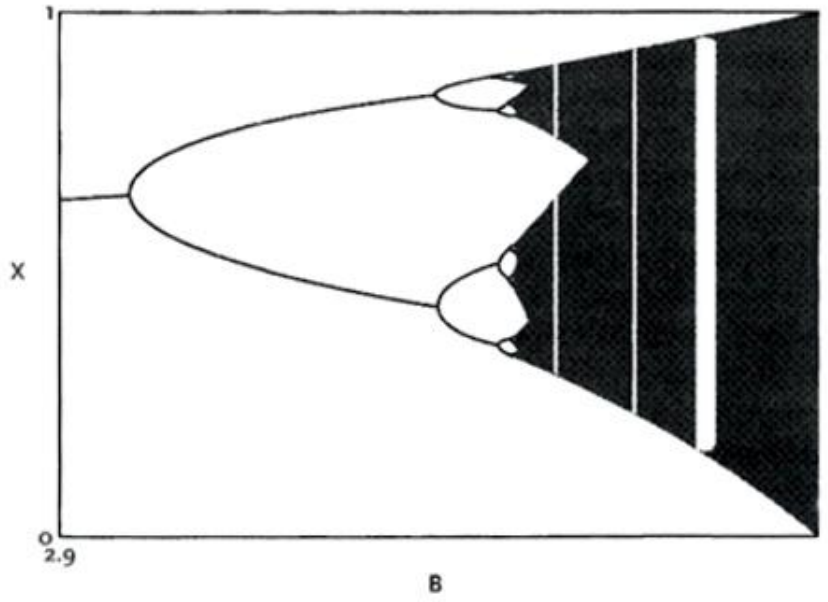

Figura 3.1a. El diagrama de Feigenbaum constituye un ejemplo de la ruta hacia el caos «con duplicación del período».

Esta repetición continúa indefinidamente, con la misma pauta de comportamiento repetida a escalas cada vez menores, como un conjunto de muñecas rusas

encajadas unas en otras que no se acaba nunca. Por razones obvias, se dice que esas pautas que se hallan dentro de las pautas son autosimilares.

En medio del orden existe el caos; pero en medio del caos hay orden. May constató que todo esto tenía implicaciones más allá de los campos de la ecología y la biología, y llamó la atención sobre sus descubrimientos en una publicación que apareció en 1976 en *Nature,* la revista científica interdisciplinar más leída.

Esto sucedía justo en la época en que numerosos estudios independientes sobre fenómenos diversos, realizados por distintos científicos, se estaban conjuntando para crear la teoría del caos.

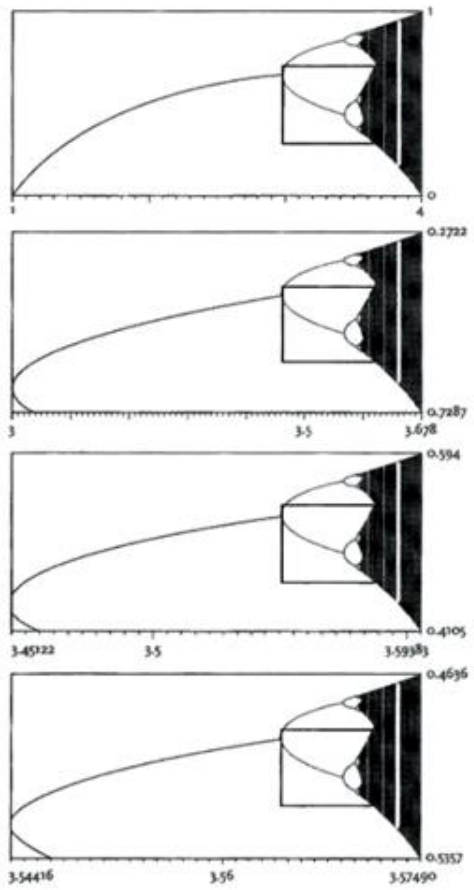

Figura 3.1b. Una característica fundamental del diagrama de Feigenbaum es la autosimilitud. Un pequeño fragmento del diagrama, ampliando la escala en la proporción debida, es exactamente igual que el diagrama original.

Hemos estado utilizando el término «caos» para describir este comportamiento,

pero hasta pasado un tiempo no se le dio este nombre, de hecho, mientras todo esto estaba sucediendo, fue cuando el caos recibió su nombre y llegó a ser ampliamente conocido. Cuando Edward Lorenz hizo su descubrimiento personal del caos a principios de la década de 1960, se encontraba trabajando en el contexto de la meteorología.

Como era natural, informó sobre su descubrimiento en congresos a los que asistían meteorólogos y publicó en revistas de meteorología, su publicación fundamental, «*Deterministic Nonperiodic Flow*», apareció en el *Journal of the Atmosferic Sciences* en 1963. Sin embargo, los matemáticos, los físicos e incluso los biólogos que podrían haber sacado provecho del descubrimiento no leían revistas de meteorología, mientras que los meteorólogos que sí las leían estaban mucho más interesados en el modo de hallar orden entre el caos (en el sentido vulgar del término) del tiempo atmosférico, que en mirar cómo podían conseguir el caos a partir del orden. Una década más tarde, entró en escena James Yorke (1941 -), un matemático que trabajaba en un instituto interdisciplinar, el Institute for Physical Science and Technology, creado en la Universidad de Maryland con el propósito específico de intentar romper algunas de las barreras que separaban a los científicos en compartimentos diferentes. Uno de los colegas de Yorke en la Universidad de Maryland, Alan Faller, trabajaba en el Departamento de Meteorología y había leído la publicación de Lorenz de 1963. En una conversación con Faller, Yorke habló de su trabajo sobre la no periodicidad y Faller le mostró la mencionada publicación, haciendo luego unas fotocopias que distribuyó en el instituto.

Yorke se dio cuenta al momento de que la publicación no sólo contenía una profunda verdad que debía tener muchas aplicaciones fuera del campo de la meteorología, sino que también se percató de que, expresando esto en términos de comportamiento de un sistema físico real, se conseguía que las matemáticas en que se basaba fueran más accesibles para los físicos. Los matemáticos habían estado jugando con los números de una manera que se reflejaba el comportamiento que habían tenido éstos durante décadas en el sencillo programa informático de Lorenz, pero antes de Lorenz nadie había establecido un vínculo entre tales abstracciones matemáticas y el mundo real.

El propio Lorenz no sabía, en aquella época, cuánto del trabajo matemático,

aparentemente abstracto, que se había realizado hasta entonces podía ser traducido en términos físicos con este planteamiento; pero Yorke conoció a un hombre que sí lo sabía. Durante una visita al campus de Berkeley de la Universidad de California, le pasó una copia de la publicación de 1963 a Stephen Smale, un topólogo que, tras haber llevado a cabo un trabajo en el campo de su especialidad, había sido premiado por ello y estaba interesado también en los sistemas dinámicos. Smale, por su parte, hizo muchas copias de la publicación y las repartió entre la comunidad matemática. El secreto había salido ya a la luz, pero aún no había recibido su nombre. Esto sucedía en 1975, cuando Yorke y su colega Tien Yien Li publicaron un trabajo titulado «Period Three Implies Chaos».

Lo que Li y Yorke demostraron es que, para ciertas familias de ecuaciones diferenciales, si hay al menos una solución de período tres, entonces debe haber también un número infinito de soluciones periódicas, con todos los períodos posibles y, además, un número infinito de soluciones no periódicas. Esto no es realmente lo que hoy en día llamamos caos; dada la presencia de todas las soluciones periódicas no caóticas que también contenían las ecuaciones, el propio Lorenz prefiere llamar «caos limitado» a lo que Li y Yorke descubrieron. Es muchísimo más probable que un sistema así esté en un estado periódico, mientras que en la situación que Lorenz llama «caos completo», y la mayoría de los científicos actuales denominan sencillamente «caos», aunque existen soluciones periódicas, es enorme la probabilidad de que el sistema caiga en un régimen caótico. Sin embargo, a pesar de que la terminología ha evolucionado desde 1975, la publicación de Li y Yorke en general está considerada como el texto que introdujo la palabra «caos», si bien inadvertidamente, como un término científico con su significado moderno.

Así pues, hacia la segunda mitad de la década de 1970 había ya una palabra para expresar el tipo de comportamiento que May había descubierto en su investigación de la ecuación logística. No obstante, con todo lo fascinante que pueda resultar, esto sería poco interesante para cualquiera, salvo para los matemáticos, si se refiriera sólo a una simple ecuación logística que, a decir verdad, ni siquiera da una representación realista de una sola especie biológica. Sin embargo, unos pocos años después de la innovación de May, a mediados de la década de 1970, Mitchell Feigenbaum, que trabajaba entonces en el Los Alamos National Laboratory, en

Nuevo Méjico, había demostrado que la novedad en cuestión tenía implicaciones mucho más amplias. Feigenbaum (1945-) demostró que la ruta hacia el caos a través de la duplicación del período no es sólo una característica especial de la ecuación logística, sino que es producto del proceso iterativo mediante el cual el sistema se retroalimenta a sí mismo, tanto si el sistema es una población animal, como si es un oscilador en un circuito eléctrico, una reacción química oscilante, o incluso (en principio) el ciclo económico. Lo importante era que los sistemas tenían que ser «autorreferenciales». En este caso, seguían todos ellos la misma ruta hacia el caos, no aproximadamente, sino *exactamente*.

Feigenbaum observó el modo en que los intervalos comprendidos entre las duplicaciones del período se vuelven cada vez más cortos a medida que B aumenta a lo largo de la rula hacia el caos estudiada por May.

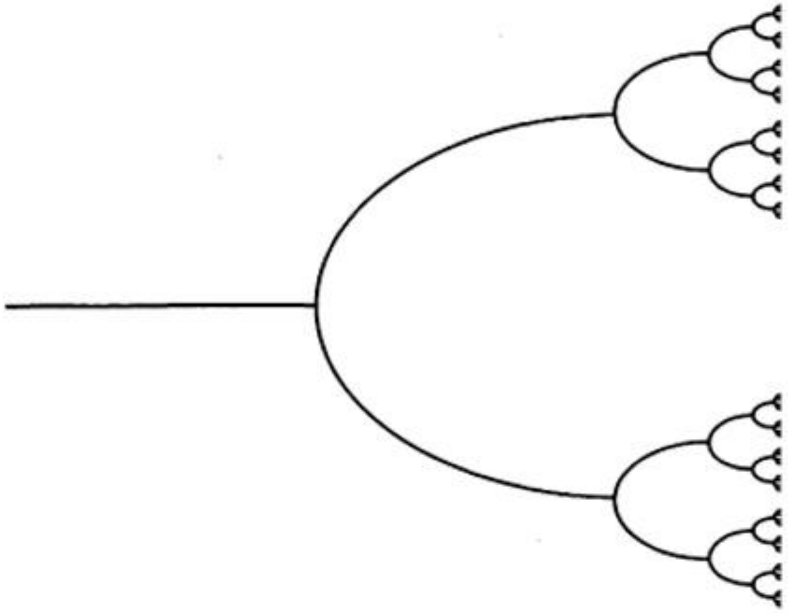

Figura 3.2a. Representación simplificada del tipo de ramificación autosimilar que se ve en un diagrama de Feigenbaum.

Descubrió que hay una proporción constante en cuanto al tamaño de un paso con respecto al siguiente, utilizando hasta tres cifras decimales, la proporción es 4,669:1; el resultado es igual si comparamos el primer paso con el segundo, el segundo con el tercero, el paso número ciento cinco con el ciento seis, y así

sucesivamente. Esta pauta de comportamiento se denomina convergencia geométrica, porque nos acerca cada vez más a cierto punto crítico, sin llegar nunca a él. Feigenbaum halló duplicación del período, bifurcación y la misma convergencia geométrica en todos los sistemas autorreferenciales que examinó. No es que encontrara algún tipo de convergencia, sino exactamente la misma convergencia geométrica con la misma proporción 4,669:1. Había hallado una verdad universal y recibió el espaldarazo definitivo cuando se le dio su nombre al misterioso número que había surgido de estos cálculos, 4,669.[34]

Hay muchos otros ejemplos de duplicación del período como ruta hacia el caos, todos con hechos interesantes que se producen justo en la transición de la sencillez al caos. Pero aquí no necesitamos explicar todos los detalles. En vez de eso, lo que podemos hacer es empezar a ver lo que sucede si volvemos a los ejemplos del grifo que gotea y los vórtices del río. Estos vórtices también muestran el efecto de bifurcación. Observando lo que sucede mucho más detalladamente que Leonardo, podemos considerar la turbulencia como el resultado de la suma de un número cada vez mayor de ciclos periódicos regulares (el primero que hizo esto fue el físico ruso Lev Landau en la década de 1940).

En un vórtice simple el movimiento se corresponde con un bucle que rodea un atractor simple, el ciclo límite. El paso siguiente sería imaginar un punto que describe un círculo en el espacio de las fases, mientras el centro del círculo describe un círculo de mayor tamaño. El atractor resultante sería un toro, una forma igual que la del tubo interior del neumático de una bicicleta, o la de un anticuado cinturón salvavidas. Al desplazarse de manera regular en torno al círculo pequeño, mientras éste se desplaza de la misma manera en torno al círculo grande, el punto que representa el estado del sistema sigue una trayectoria igual a la de un muelle curvo enrollado, o con la forma del juguete que se conoce como *slinky*[35] enrollado formando círculos, de un modo regular y predecible. En general, dos movimientos periódicos en el espacio de las fases ejercen una interacción mutua y se unen estrechamente en un ritmo repetitivo.

---

[34] El número de Feigenbaum es en realidad un número irracional, como π y muchos otros. Con unas pocas cifras decimales, puede escribirse 4,6692016090.

[35] Juguete que consiste en una larga pieza de metal o plástico enrollada en forma de espiral. Si sacudimos el *slinky* desde uno de sus extremos de manera periódica, aparece una onda transversal que se transmite por todo el juguete, haciendo que la espiral se alargue y se acorte a medida que pasa la onda. *(N. de la t.)*

Desde un punto de vista matemático es sencillo describir etapas posteriores de un comportamiento cada vez más complejo en la ruta que lleva a la turbulencia, si hablamos de «toros» con un número más elevado de dimensiones. Un ciclo límite es un atractor unidimensional que existe en dos dimensiones, la superficie de un toro es un atractor bidimensional empotrado en un espacio de las fases tridimensional, y el paso lógico siguiente sería una pauta de comportamiento descrita mediante un atractor tridimensional empotrado en un espacio de las fases tetradimensional.

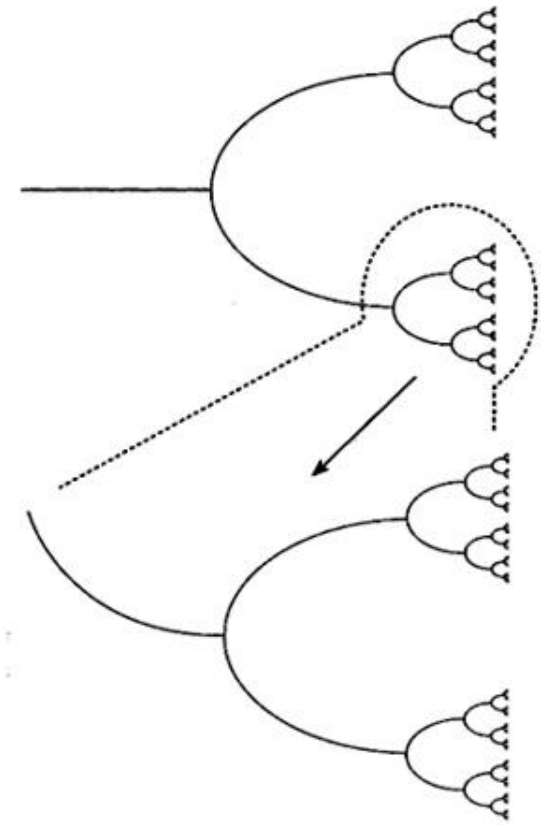

Figura 3.2b. Representación simplificada de la autosimilitud.

Pero el mundo real no va más allá por esta ruta. La turbulencia se produce en la etapa inmediatamente posterior, donde el punto del espacio de las fases representa el estado del sistema situado en la superficie bidimensional del toro, pero desplazándose de un modo mucho más complicado por dicha superficie, siguiendo una trayectoria que nunca puede pasar dos veces por el mismo punto del espacio de las fases (porque, si lo hiciera, el sistema sería periódico y su comportamiento se

repetiría) y, por consiguiente, nunca se corta consigo misma. Una característica típica es que, reproduciendo el problema original de los tres cuerpos, en el espacio de las fases nunca caen tres o más períodos en un ritmo repetitivo, y su efecto combinado no es más predecible que la órbita de la partícula pequeña en el problema restringido de los tres cuerpos (aunque *es completamente determinista,* sin que podamos poner el énfasis en que lo sea demasiado intensamente o demasiado a menudo).

La trayectoria del punto que describe el sistema a través del espacio de las fases correspondería a una línea infinitamente larga enrollada de una forma compleja, pero sin cortarse consigo misma, alrededor de una superficie finita (en realidad existe también una complicación adicional, de la cual hablaremos en breve). El belga David Ruelle y su colega holandés Floris Takens, cuando trabajaban en París, llamaron a este ente «atractor extraño», en una publicación que apareció en 1971.

Es ahora cuando los fractales entran en escena. Al igual que el caos, recibieron su nombre, como veremos luego, en 1975; pero, también como en el caso del caos, con anterioridad ya habían estado presentes en la ciencia durante mucho tiempo, sin que se apreciara plenamente su importancia. Esas entidades a las que llamaríamos actualmente fractales fueron descubiertas, con gran sorpresa (e incluso horror), por unos matemáticos a finales del siglo XIX. En aquella época, se consideraba que eran una especie de aberraciones, «monstruos» que no encajaban en el mundo ordenado de las matemáticas al uso. Para entender la razón de esta actitud, basta con examinar un ejemplo que pone de manifiesto todas sus peculiaridades.

En 1890 Giuseppe Peano (1858-1932) publicó un trabajo en el que explicaba cómo construir una curva que llenara completamente un plano[36].

Esto no le sonará demasiado horrible a alguien que no sea matemático. Pero, reflexionemos un poco. Un plano es una entidad bidimensional, posee longitud y anchura. Una línea es una entidad unidimensional, posee longitud, pero no anchura. Peano demostró cómo se podía hacer que una línea se retorciera y diera vueltas dentro de un plano, de tal modo que pasara por todos los puntos de éste sin cortarse nunca consigo misma. ¡Una línea unidimensional llenaba completamente un

[36] Tan sólo un año más tarde, David Hilbert (1862-1943) hizo el mismo descubrimiento.

plano bidimensional! Entonces, ¿cómo podía un plano ser «realmente» bidimensional, si todo punto del plano estaba en una sola línea? Aún hay más. Si nos imaginamos el plano como un cuadrado, entonces las curvas de Peano trazan los contornos de un conjunto de cuadrados menores, que serían como baldosas que llenan el plano; dentro de cada cuadrado menor, la curva dibuja un conjunto similar de «baldosas»; y así sucesivamente.

La pauta es autosimilar y el proceso nunca termina. La curva de Peano es infinitamente larga, pero está contenida dentro de un área finita. Hay aquí una clara analogía con el atractor, esa curva «que llena el espacio», enrollada en torno al toro en el espacio de las fases que describe un sistema turbulento, aunque nada de esto se conocía en la década de 1890. El lenguaje que se necesitaba para describir estas entidades lo desarrolló finalmente Benoit Mandelbrot en la década de 1970, cuando trabajaba en el Thomas J. Watson Research Center, un centro de investigación de IBM en Yorktown Heights, Nueva York.

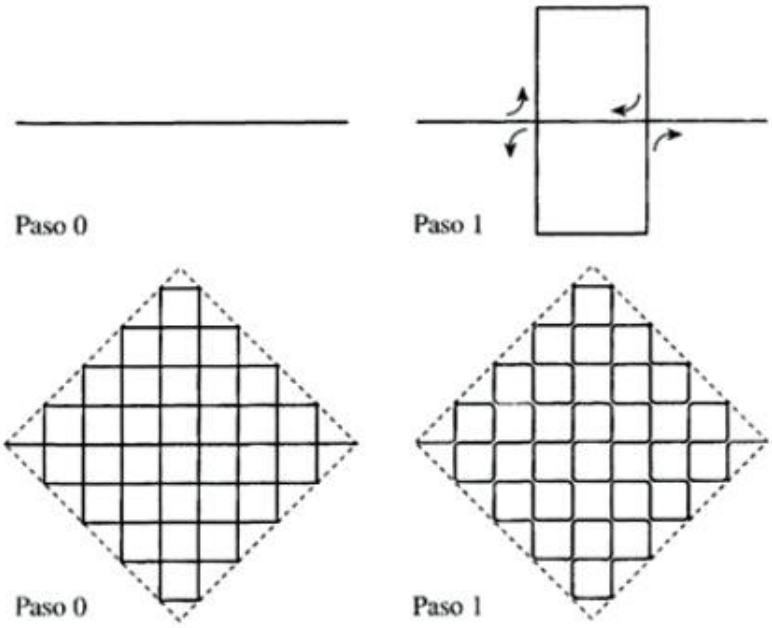

Figura 3.3. La curva de Peano, una línea que cree ser una superficie.

Mandelbrot, que había nacido en Varsovia en 1924, tuvo una formación ecléctica que le vino muy bien para ser el fundador de lo que en esencia era una nueva

disciplina científica. Su familia se trasladó a Francia en 1936, donde Benoit cursó estudios en París después de producirse la Liberación en 1944. Después de pasar unas temporadas en Caltech y en el Institute for Advanced Studies de Princeton, en 1955 regresó por un breve tiempo a Francia, antes de trasladarse definitivamente a Estados Unidos y establecerse en Yorktown Heights en 1958. Mandelbrot constató que un objeto como las curvas de Peano podía describirse como algo dotado de una dimensión intermedia, en este caso algún valor comprendido entre 1 y 2.

Una línea seguía siendo un objeto unidimensional; y un plano, un objeto bidimensional, pero, del mismo modo que las matemáticas habían asumido la idea de que hay una infinidad de números entre dos números racionales cualesquiera, también tendrían que aceptar la idea de que existen entidades con dimensiones intermedias, no enteras.

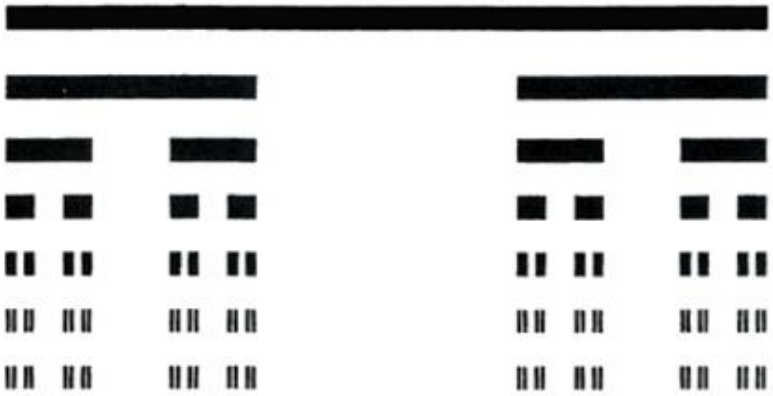

Figura 3.4. El conjunto de Cantor. Borrando siempre el tercio central de cada segmento, tal como se indica en el texto, nos queda una especie de polvillo que contiene una cantidad infinita de puntos, cuya longitud total es cero.

Si la dimensión de una de estas entidades no es un número entero, entonces debe ser fraccionario. Con el fin de disponer de una palabra para definir estas entidades, dice Mandelbrot, «*en 1975 acuñé el término fractal, derivado del latín fractus, que alude a una piedra rota, quebrada e irregular*[37]». También podríamos pensar quizá, en una especie de contracción combinada de las palabras «fraccional» y «fractura»,

[37] Véase su artículo en The New Scientist Guide to Chaos, editado por Nina Hall.

que proceden de la misma raíz latina.

Hay otros tres fractales, conocidos todos ellos durante décadas, antes de 1975, como monstruos matemáticos espantosos, de los que nos conviene saber algo antes de hablar sobre la importancia que tiene todo esto en relación con el caos. El más antiguo de estos fractales (en el sentido de que fue el primero que descubrieron los matemáticos) es el conjunto de Cantor, que lleva el nombre del matemático alemán Georg Cantor (1845-1918), quien lo descubrió en 1883[38].

La curva de Peano se puede considerar una línea que es más que una línea, que «intenta» ser un plano; el conjunto de Cantor puede verse como una línea que es menos que una línea, que «intenta» ser un punto. Es fácil explicar cómo se construye un conjunto de Cantor. Tomamos un segmento de línea recta y borramos su tercio central (teniendo cuidado de no borrar los puntos situados exactamente a un tercio y a dos tercios del recorrido). Nos quedan así dos segmentos, cada uno de los cuales tiene una longitud igual a un tercio de la que tenía el segmento inicial, y separados por un hueco de la misma longitud que cada uno de los dos segmentos.

Continuamos haciendo lo mismo, es decir, iteramos el procedimiento hasta que todos los segmentos se han borrado y nos queda una cantidad infinita de puntos colocados según un modelo regular y con huecos entre ellos. Este es el conjunto de Cantor, un fractal cuya dimensión es un valor intermedio entre 0 (un punto) y 1 (una línea), una especie de fantasma de la línea original, como la risa burlona que se desvanece en la cara del gato de Cheshire en *Alicia en el país de las maravillas.* A pesar de ser tan sencillo (y bastante soso a la vista), el conjunto de Cantor resulta tener una importancia crucial en la teoría del caos, tal como se puede deducir viendo un resumen de sus propiedades fundamentales.

En primer lugar, se produce por iteración (retroalimentación), que, como ya hemos dicho, es una de las claves del caos. En segundo lugar, hay autosimilitud. A partir de la segunda etapa de nuestra iteración (donde quedan cuatro segmentos), en cada paso el conjunto de Cantor está formado por tíos copias exactas de sí mismo, reduciéndose siempre su tamaño en un tercio. Pero hay algo más con respecto al conjunto de Cantor. Si volvemos al diagrama de bifurcación, que representa la ruta

[38] Lo que Cantor no sabía es que este conjunto ya había sido descubierto en 1875 por Henry Smith (1826-1883), un matemático dublinés ingenioso, aunque más bien modesto. Pero, como Smith había fallecido y su descubrimiento era prácticamente desconocido, fue el nombre de Cantor el que quedó asociado con este con junto.

hacia el caos por duplicación del período, justo en el punto donde se inicia el caos (que a veces se denomina punto de Feigenbaum, el límite al que tienden los siempre decrecientes intervalos entre bifurcaciones), en el último paso del proceso antes de que el caos aparezca, todos los tipos de ramas del árbol de la bifurcación forman un conjunto de Cantor, indicando (o, más bien, iluminando como con luces de neón) la profunda relación que existe entre fractales y caos.

El conjunto de Cantor tiene también un lugar importante en la historia de los fractales, porque fue una de las rutas que llevaron a Benoit Mandelbrot al trabajo que le haría famoso. Mandelbrot se interesaba por todo tipo de fenómenos que variaban en el tiempo y el espacio, así como por pautas tan diversas como el modo en que las palabras están distribuidas en un texto, la distribución de ciudades grandes y pequeñas en un país, y los altibajos del mercado bursátil. Poco después de empezar a trabajar como investigador para IBM abordó un problema que tenía gran importancia práctica para esta empresa. Entonces, como ahora, la información se pasaba de un ordenador a otro a través de líneas telefónicas, y a los ingenieros relacionados con este trabajo les atormentaba el ruido que se producía en dichas líneas, porque podía alterar datos de vital importancia.

No era un problema muy importante en las conversaciones de voz, que eran el objetivo para el cual se habían establecido inicialmente las líneas telefónicas, si se oye un ruido parásito en la línea durante una conversación, siempre cabe la posibilidad de hablar un poco más alto o de repetir lo que se ha dicho. Pero causaba estragos en la transmisión de datos en aquellos tiempos[39]. La reacción inmediata de los ingenieros fue aumentar la fuerza de la señal, lo cual era como gritar para ahogar el ruido; pero, aun así, vieron que, de vez en cuando, un ruido repentino borraba un dato de manera aleatoria.

Lo curioso con respecto a este ruido era que, aunque fuera aleatorio e impredecible, parecía llegar en ráfagas. Podía haber intervalos bastante largos en los que las transmisiones se realizaban a la perfección, seguidos de manera impredecible por intervalos en los que se producían repetidas ráfagas de ruido. Cuando Mandelbrot examinó este problema, descubrió que la pauta tenía autosimilitud. En un intervalo tranquilo, no había ruido en absoluto. Sin embargo, dentro de un intervalo ruidoso

[39] Los problemas de este tipo se han superado ampliamente con la transmisión digital de datos, pero esto sucedía en los tiempos de las señales analógicas.

siempre había intervalos más breves totalmente desprovistos de ruidos, y otros con ráfagas de ruido. Dentro de estos breves intervalos ruidosos se repetía toda la pauta, hasta donde él podía saber, indefinidamente. Mandelbrot descubrió que, a cualquier escala, el cociente entre la longitud del intervalo silencioso y la del intervalo ruidoso tenía el mismo valor.

De hecho, la distribución de las ráfagas de ruido en el sistema de transmisión era la misma que la de los puntos en el conjunto de Cantor. Esto tenía implicaciones prácticas para los ingenieros de IBM. El descubrimiento significaba que no tenía sentido malgastar dinero aumentando la intensidad de las señales, dado que el sistema siempre tendría ruidos, por lo que era mejor concentrar los esfuerzos en el desarrollo de técnicas para detectar errores y repetir las partes distorsionadas de cualquier mensaje; también significaba que se le podía evitar al personal la tarea infructuosa de intentar encontrar una causa física (como una rama de árbol que rozara los cables) de unos errores que eran esencialmente aleatorios, y hacer que trabajara en tareas más productivas (como rastrear fuentes de ruido que no encajaran con la pauta y que, por lo tanto, tendrían probablemente una causa física). Mandelbrot había descubierto uno de los primeros ejemplos reconocidos de caos en funcionamiento dentro de la tecnología humana, y lo había asociado desde el principio con los fractales, aunque nadie había dado nombre todavía al caos y a los fractales.

Mientras estábamos trabajando en este libro, nos fijamos en una noticia local que sugería que podría estar sucediendo en un pueblo cercano algo similar al proceso que inicialmente llamó la atención de Mandelbrot sobre el caos. Los habitantes de aquel pueblo se quejaban de que habían estado sufriendo repetidamente cortes en el suministro eléctrico durante el año anterior. La empresa que les proporcionaba dicho suministro eléctrico tenía una explicación perfectamente racional para cada apagón, uno había sido causado por un cisne que, mientras volaba, chocó con los cables aéreos; otro, por fuertes vientos que habían derribado un árbol, haciéndolo caer sobre el tendido; un tercero, por un relámpago; y así sucesivamente.

Hoy en día, la teoría del caos puede decirnos que tal cúmulo de desastres locales tiene que suceder de vez en cuando en algún tramo de la red de suministro de energía, pero no puede decirnos dónde, ni cuándo, por lo que esto no consuela

mucho a las personas afectadas.

También los vínculos existentes entre los procesos aleatorios, el caos y los fractales pueden explicarse examinando otro de estos monstruos matemáticos, que esta vez data de principios del siglo XX: una construcción conocida como el triángulo de Sierpinski, que fue revelada al mundo de las matemáticas en 1916 por el matemático polaco Waclaw Sierpinski (1882-1969).

Las instrucciones para hacer un triángulo de Sierpinski hablan sólo de una sencilla iteración repetitiva. Tomemos un triángulo equilátero dibujado en un papel y con su superficie pintada de color negro (actualmente lo podemos representar en la pantalla de un ordenador; en sentido estricto, no es necesario que sea un triángulo *equilátero,* pero así es más fácil visualizar lo que pasa). A continuación, tomemos los puntos medios de cada lado del triángulo y unámoslos para formar un triángulo invertido dentro del primer triángulo. Borramos el color negro del interior de este triángulo, para dejar un triángulo blanco invertido rodeado por tres triángulos negros que descansan sobre un lado y tienen un vértice en la parte superior (no invertidos), de tal modo que entre todos llenan el triángulo original.

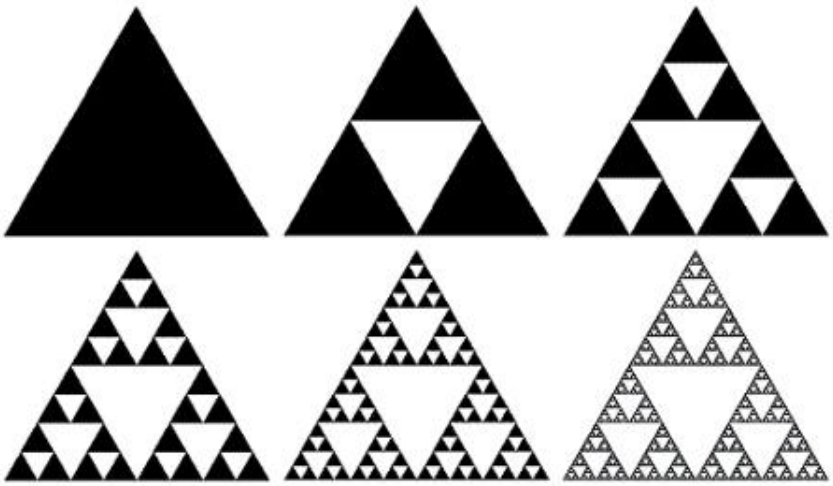

Figura 3.5. El triángulo de Sierpinski.

Se puede adivinar cuál es el paso siguiente, se repite este proceso en cada uno de los tres triángulos negros de menor tamaño que el inicial, para luego hacer lo mismo en cada uno de los nueve triángulos negros resultantes, y así sucesivamente (en principio, infinitas veces). Lo que nos queda es el triángulo de Sierpinski, una

entidad que posee claramente autosimilitud y que es también un fractal, con una dimensión comprendida entre 1 y 2 (explicaremos en breve cómo se miden las dimensiones de los fractales).

Pero existe otro modo de construir el triángulo de Sierpinski que es igualmente fácil, sólo hemos de tener paciencia. Todo lo que hay que hacer es tomar un lápiz, marcar tres puntos en un papel para representar los vértices de un triángulo equilátero, y utilizar un dado ordinario, no trucado, que utilizaremos para elegir al azar los números 1, 2 y 3. Puesto que el dado tiene seis caras, podemos hacerlo estableciendo que el 4 sea el 1, el 5 lo mismo que el 2, y el 6 lo mismo que el 3. Llamaremos a los vértices del triángulo 1, 2 y 3.

A continuación marcamos cualquier punto del papel como punto de partida y lanzamos el dado. Si sale el 1 o el 4, marcamos un punto del papel que esté a medio camino entre el punto de partida y el vértice señalado como 1. Si el dado nos da un 2 o un 5, marcamos el punto a medio camino entre el punto de partida y el vértice 2, y si sale el 3 o el 6 marcamos un punto a medio camino entre el punto de partida y el vértice 3.

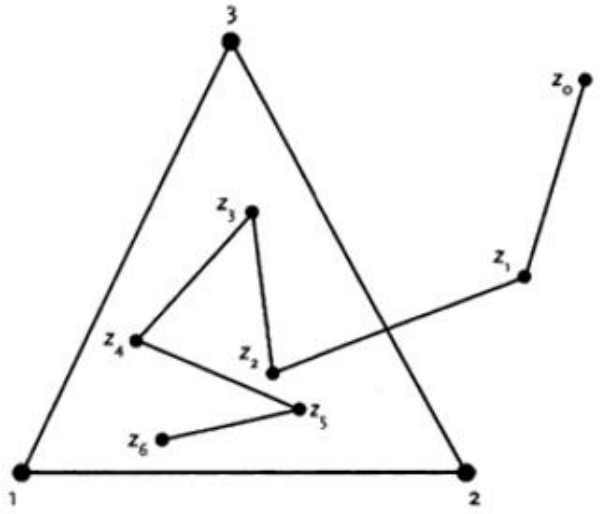

Figura 3.6a. Construcción del triángulo de Sierpinski utilizando el juego del caos. Los seis primeros pasos del juego.

Tomando el nuevo punto como punto inicial, lanzamos el dado y repetimos el proceso una y otra vez. Muy lentamente, después de marcar los primeros puntos, la

figura que éstos forman en el papel muestra la forma del triángulo de Sierpinski. Un proceso iterativo, completamente aleatorio, que obedece una regla muy sencilla, ha construido una pauta fractal.

En realidad, el triángulo de Sierpinski hace el papel de atractor en este proceso tan especial, y los primeros pasos de la iteración no se producen exactamente en el triángulo, porque comienzan en un lugar que no está dentro de éste, aunque luego se desplazan hacia él en la medida en que son atraídos. Los atrae un fractal, que así se convierte en un «atractor extraño», según el lenguaje introducido por Ruelle y Takens, siendo los fractales claramente extraños si los comparamos con objetos que encontramos en nuestra vida cotidiana, como triángulos y bandas elásticas.

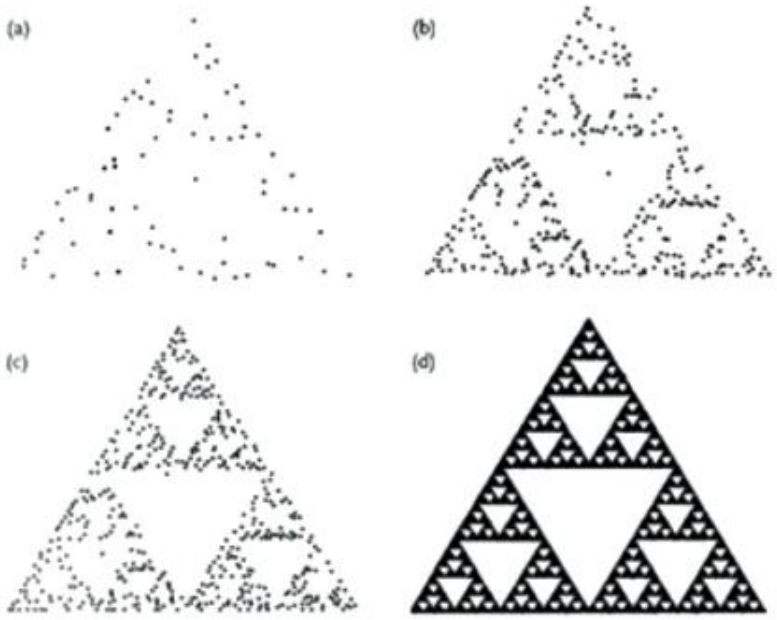

Figura 3.6b. La situación del juego después de 100, 500, 1.000 y 10.000 pasos.

Sin embargo, aparte de estos pocos puntos iniciales, con independencia de dónde comience el juego, se terminará con el triángulo de Sierpinski. Si el lector siente la tentación de entrar en este juego (en una versión sencilla suele llamarse a veces el juego del caos, donde es posible producir muchas pautas interesantes, utilizando reglas reiterativas que coinciden en ser todas ellas muy sencillas), conviene que escuche dos advertencias. La primera es la paciencia, tendrá que dar varios cientos de pasos hasta poder ver realmente algo como el triángulo de Sierpinski.

La segunda se refiere a la aleatoriedad; hay que evitar la tentación de usar un ordenador, a menos que se sea un programador lo suficientemente bueno como para estar seguro de que se utilizan números auténticamente aleatorios.

Figura 3.7. Una modalidad del juego del caos produce una imagen similar a un helecho.

Por desgracia, los generadores de números aleatorios que se usan en los ordenadores no siempre son realmente aleatorios, aunque existe la posibilidad de soslayar este problema, si se tiene la habilidad necesaria.

Algunas modalidades del juego del caos, basadas todavía en marcar puntos en una hoja de papel (o en la pantalla de un ordenador, si el jugador está seguro de que sus números son aleatorios) siguiendo una sencilla regla de este tipo, repetida de una manera aleatoria, pueden producir imágenes fractales asombrosamente parecidas a formas de seres vivos, como helechos y árboles. Para los que estén interesados, las matemáticas que se utilizan en estos juegos pueden encontrarse en *Chaos and Fractals* de Heinz-Otto Peitgen, Hartmut Jürgens y Dietmar Saupe.

Nadie dice que éste sea precisamente el modo en que los seres vivos crecen y adoptan formas complejas, pero lo importante para nosotros es que estos sistemas aparentemente complicados pueden producirse, o explicarse, mediante la aplicación

reiterada de una regla muy sencilla. Se dice a menudo que la información almacenada en el ADN de cada célula de un ser vivo contiene el «plan original» que detalla cómo está formado ese ser. Pero, en realidad, ésta es una analogía bastante pobre.

El plan original real sería muy complicado, un dibujo que representara con detalle todo lo que hay en un ser vivo y cómo está ensamblado. Una analogía mejor es la que se podría hacer con la receta de un bizcocho de pasas, que no nos indica qué aspecto tendrá éste al final (y mucho menos la ubicación exacta de cada pasa dentro de él), sino que nos dice «tome los siguientes ingredientes, mézclelos bien y póngalo todo en el horno durante tantos minutos a tantos grados de temperatura». Una receta así es como un paso del juego del caos.

Es difícil entender cómo toda la riqueza de ADN contenida en una sola célula que se desarrolla para convertirse en un ser humano, o un pino, o cualquier otra cosa, puede contener el auténtico proyecto original de todas las estructuras complejas existentes en la forma adulta final. Sin embargo, es mucho más fácil ver cómo el ADN podría contener unas pocas instrucciones sencillas, al estilo de «duplicar el tamaño en *n* pasos, dividirlo luego en dos, y repetir lo mismo en cada rama». En el juego del caos, unas instrucciones que sean sólo ligeramente más complicadas que ésta producen por iteración estructuras complejas con aspecto de helechos. Si existe un atractor para la forma de helecho, basado en ese tipo de regla sencilla, no es sorprendente que algunas plantas crezcan convirtiéndose en helechos, aunque no necesiten ser programadas para tener esa forma, del mismo modo que no se programa un dado para hacer el triángulo de Sierpinski. Es la aleatoriedad, junto con una sencilla regla iterativa (o varias), lo que hace la complejidad del mundo.

Sin embargo, antes de que pasemos a examinar más detenidamente la naturaleza de nuestro complejo mundo, quedan todavía unas pocas cosas que vale la pena mencionar a propósito de los fractales y los atractores, sin olvidar, puesto que el lector es seguramente curioso, cómo medir la dimensión de un frac- tal. Lo más sencillo para hacerse una idea de esto es utilizar el último de los monstruos matemáticos de finales del siglo XIX y principios del XX que vamos a comentar detalladamente: la curva de Koch. Además de su interés intrínseco, la curva de Koch es también importante para esta historia porque contribuyó en la década de

1960 a avivar en Mandelbrot el interés por los fractales, y porque, en sentido metafórico, afecta tangencialmente a la vida de un hombre del que ya hemos hablado, Lewis Fry Richardson. Sin embargo, uno de los aspectos curiosos que presenta la curva de Koch es que, aunque se trata realmente de una curva, el único modo en que puede tener un contacto tangencial con algo es en sentido metafórico, ya que no tiene tangente alguna y, dicho en lenguaje corriente, está toda ella formada por esquinas.

Esta curva la descubrió, o inventó, el matemático sueco Helge von Koch (1870-1924), que la presentó al mundo en un trabajo publicado en 1904. Es fácil describir el modo en que se construye la curva de Koch. Comenzamos con un segmento de línea recta y lo dividimos en tres partes iguales. Sobre el tercio central se construye un triángulo equilátero (esta vez, sí debe ser realmente equilátero), borrando a continuación la base de dicho triángulo. Ahora tenemos cuatro segmentos iguales (uno horizontal a cada lado y una V invertida en el centro, que constituyen la forma básica inicial que necesitamos para generar la curva de Koch (esta forma inicial, lógicamente, recibe el nombre de generador). Acto seguido, repetimos el mismo proceso con cada uno de los cuatro segmentos de recta, dividir en tres segmentos iguales, construir un triángulo equilátero en el centro y borrar la base de este triángulo. Seguiremos repitiendo este procedimiento en una iteración infinita para generar la curva de Koch, que está formada en su totalidad por esquinas infinitamente pequeñas en forma de V. Es una curva de longitud también infinita, aunque sus puntos extremos son también los del segmento original y, en conjunto, sólo alcanza una pequeña altura sobre dicho segmento.

Para hacer una ligera variación sobre el tema, en la etapa del generador, podemos tomar tres generadores (todos del mismo tamaño, por supuesto) y colocarlos sobre los lados de un triángulo equilátero para formar una estrella de David de seis puntas. Después de realizar unas pocas iteraciones, el perfil sumamente quebrado de la curva se parece al de un copo de nieve, por lo que esta construcción recibe a veces el nombre de copo de nieve de Koch. Sin embargo, a medida que avanzan las iteraciones, esta curva tan recortada comienza a parecerse a la estructura de una accidentada línea costera, lo cual hace que al caso límite final de este fractal se le dé otro nombre, la isla de Koch. Esta isla de Koch tiene una «línea costera»

infinitamente larga, aunque toda ella está contenida en un círculo que toca las puntas del generador original cuya forma es la estrella de David de seis puntas.

Así pues, ¿serían fractales las líneas costeras reales? El parecido existente entre la isla de Koch y una isla real, ¿es algo más que una simple analogía? Fue este tipo de preguntas lo que realmente impulsó las investigaciones de Mandelbrot sobre fractales y, según dijo él mismo[40], el detonante fue la lectura de una publicación poco conocida de Richardson en la que este meteorólogo ocasional se rompía la cabeza pensando sobre el modo en que podían calcularse las longitudes de las fronteras existentes entre España y Portugal y entre Bélgica y los Países Bajos. Richardson había constatado una diferencia del 20 por ciento en las longitudes que daban distintas fuentes para las mismas fronteras.

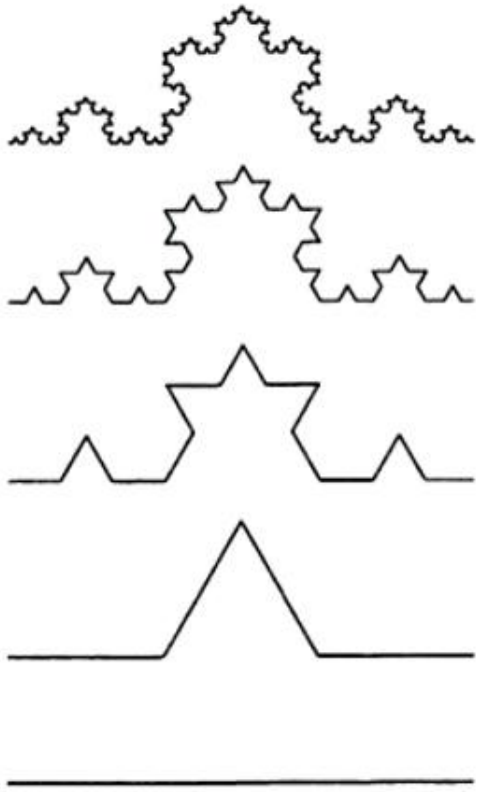

Figura 3.8. La curva de Koch (véase p. 146). Una línea costera muy accidentada, como las de Gran Bretaña o Noruega, tiene aproximadamente la misma estructura fractal que esta curva.

[40] Citado por James Gleick en Chaos.

Las diferencias surgen porque la longitud que se mide para una línea tan accidentada depende de la escala a la que se esté realizando la medición. Aunque podría haber tramos rectos en esas fronteras, en la mayor parte del recorrido éstas siguen alteraciones o características naturales del paisaje, tales como ríos y cordilleras.

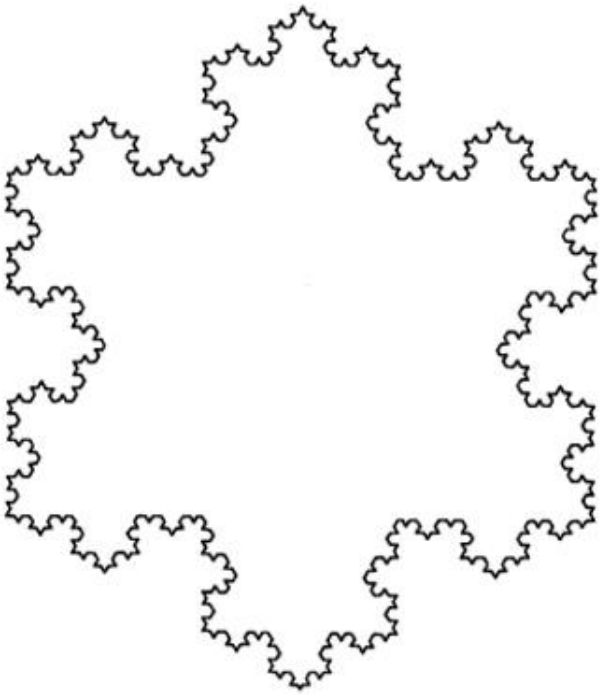

Figura 3.9. El copo de nieve de Koch

Una frontera así debe cartografiarse utilizando teodolitos y toda la parafernalia cartográfica habitual, espaciando en unos cien metros los puntos desde los que se realiza la medición, midiendo una serie de segmentos de recta de cien metros de longitud que zigzaguean por el campo.

Si recorriéramos la frontera paso a paso, seguiríamos algunos de los vericuetos que hay entre estos puntos y obtendríamos para la longitud de la misma un valor que sería mayor que el anterior; si empujáramos una pequeña rueda provista de un contador para registrar el número de rotaciones que realiza a lo largo de la frontera, obtendríamos una longitud aún mayor porque ahondaría en pequeños vericuetos en

los que no habríamos pisado al caminar; si tomáramos una cuerda muy larga y la colocáramos cuidadosamente por toda la frontera, resultaría un valor aún mayor, y así sucesivamente. La cuestión es que las fronteras, en principio, son irregulares a todas las escalas (al menos hasta el nivel de los átomos), por lo cual, cuanto menor sea la escala a la que realizamos la observación, más vericuetos veremos y más larga parecerá la frontera.

Mandelbrot trató detalladamente estas teorías en un trabajo que presentó con el título *«¿Qué longitud tiene la costa de Gran Bretaña?»*, publicado en la revista *Science* en 1967. Para lo que aquí pretendemos explicar, la línea costera de una isla tiene la ventaja de que es fundamentalmente una línea natural, sin los largos segmentos de recta que pueden aparecer en las fronteras trazadas por el ser humano, como es el caso de las que existen entre muchos estados en EE.UU.

La respuesta a la pregunta de Mandelbrot, una respuesta que está implícita en su trabajo, es que depende de la escala a la que hagamos las mediciones de la línea costera, y si se llega a escalas suficientemente detalladas, la longitud medida se aproximará a infinito (sólo «se aproximará» a infinito, porque las sustancias se descomponen cuando llegamos a la escala de los átomos). Aunque realmente no es lo que Mandelbrot pronto llamaría un fractal, resulta que la línea costera de Gran Bretaña se parece a un fractal, y su dimensión sería un valor intermedio entre 1 y *2*.

Ahora bien, ¿cómo medimos la dimensión de un fractal? La curva de Koch nos aporta una clave; posee autosimilitud. Si tomamos cualquiera de los cuatro segmentos de la curva de Koch y aumentamos su tamaño multiplicándolo por tres (ampliamos la escala al triple), obtenemos una curva que es idéntica a la curva original de Koch. Pero el hecho de poseer autosimilitud no convierte necesariamente a cualquier objeto en un fractal. Un cubo corriente es autosimilar. Si cortamos un pequeño cubo de su interior y ampliamos su escala mediante el factor adecuado, tendrá exactamente el mismo aspecto que el cubo original. Esto será cierto para *cualquier* pequeño cubo.

La autosimilitud opera a cualquier escala en muchos objetos cotidianos como el cubo, pero sólo a ciertas escalas específicas en el caso de los fractales; cuando se (rata de la curva de Koch, *debemos* tomar exactamente un cuarto del original y

ampliarlo multiplicándolo por 3 para reproducir dicho original. Para dejar clara la diferencia, aumentaremos la escala de algunos objetos cotidianos utilizando también el factor 3. Si dividimos un segmento de línea recta en tres partes iguales y multiplicamos por 3 las longitudes de estos tercios, obtendremos en cada caso un segmento idéntico al primero.

La ampliación proporcional junto con la reducción de longitud da el mismo elemento inicial, dividimos por 3 y multiplicamos por 3—, por lo que la dimensión de una línea es 1. Si tomamos un cuadrado y dividimos en tres parles iguales cada uno de los lados, obtenemos una red de nueve pequeños cuadrados. Por consiguiente, hemos de tomar una novena parte del cuadrado original y multiplicar sus lados por 3 para obtener un cuadrado igual al primero. Dividimos por nueve y ampliamos multiplicando por 3. El número 9 es 3^2, luego la dimensión del cuadrado es 2. En el caso de un cubo, si dividimos cada arista en tres partes iguales, obtenemos un conjunto de 27 pequeños cubos, cada uno de los cuales será exactamente igual al cubo original si aumentamos la escala multiplicando cada arista por 3. Dado que 27 es 3^3 esto significa que la dimensión del cubo es 3. Dividimos por 27 y multiplicamos por 3.

Volvamos a la curva de Koch. En este caso, dividimos la longitud entre 4 y luego hacemos una ampliación a escala multiplicando por 3. Sabemos que 3^1 es 3, y 3^2 es 9, pero ¿qué potencia de 3 es igual a 4? Si $3^n = 4$, ¿cuánto es *n?*

La respuesta, ajustada hasta cuatro cifras decimales) es 1,2619, como es fácil de comprobar mediante una calculadora. Así pues, la dimensión de la curva de Koch es 1,2619, es decir, un número comprendido entre 1 y 2, tal como esperábamos, pero está más cerca de la dimensión de una línea que de la dimensión de un plano. El mismo tipo de razonamiento puede aplicarse al cálculo de la dimensión de oíros fractales, e incluso de cosas que son aproximadamente fractales, la línea costera de Gran Bretaña, por ejemplo, resulta tener una dimensión de aproximadamente 1,3 que, en comparación con la de la curva de Koch, es ligeramente menos parecida a la de una línea.[41]

Los fractales son de por sí fascinantes, y el mérito principal que llevó a Mandelbrot a

[41] En realidad existen distintos métodos para medir las dimensiones de los fractales y dan «respuestas» ligeramente diferentes, por lo que es posible que en algún otro texto podamos ver que se citan valores ligeramente diferentes. Pero no es nuestra intención entrar aquí en estas sutilezas y se las dejaremos a los matemáticos. Lo importante es que mostremos aquí al menos un método realmente obvio para medir dimensiones fractales.

la fama fue su descubrimiento del conjunto que recibió su nombre, cuya estructura fractal se puso de manifiesto al realizar la iteración de una sencilla expresión matemática. La diferencia, en comparación con los conjuntos y frac- tales de los que hemos estado hablando, es que en la expresión se utiliza lo que los matemáticos llaman números complejos, que son números en los que aparece la raíz cuadrada de -1.

La cuestión clave en relación con los números complejos es que en cierto sentido son bidimensionales, mientras que los números utilizados cotidianamente, es decir, los números reales, son unidimensionales. Un solo número real es todo lo que necesitamos, por ejemplo, para especificar la posición de un punto en una línea. Un solo número complejo es lo mínimo que necesitamos para determinar la posición de un punto en lo que llamamos plano complejo (mediante su distancia a cada uno de dos de los bordes del plano).

Este número contiene dos informaciones. La iteración que da lugar al conjunto de Mandelbrot utiliza la expresión (Z^2 + C), donde Z es un número complejo variable y C es un número complejo fijo.

Según el procedimiento habitual, realizamos la iteración eligiendo un valor de Z y sustituyéndolo en la expresión, para calcular el resultado y utilizarlo como el valor de Z que se toma para la siguiente iteración. La pauta del fractal surge cuando los valores de Z se representan en el plano complejo. No necesitamos saber nada sobre números complejos para damos cuenta de que se trata de un proceso muy sencillo (pero no lineal), similar a la iteración de la ecuación logística.

Sin embargo, el fractal que resulta es posiblemente la entidad más complicada que ha investigado nunca la humanidad, y no sólo es complicado, sino hermoso, de tal modo que las porciones ampliadas del conjunto han llegado a ser iconos en el arte de ilustrar carteles y hay libros completamente llenos de ellos. Pero estas investigaciones nos llevan en la dirección opuesta a la que deseamos seguir en el resto de este libro, hacia el interior del mundo de las matemáticas, en vez de hacia el exterior constituido por el mundo de los seres vivos.

Lo único que deseamos recalcar en relación con el conjunto de Mandelbrot es que, una vez más, este objeto extremadamente complicado se obtiene, en realidad, mediante una expresión muy sencilla y por iteración. Desde este punto de vista, se

trata de una de las cosas más sencillas que ha investigado la humanidad en todos los tiempos, pero, como veremos, desde este punto de vista también es muy sencillo todo lo demás.

La cuestión clave que estamos abordando aquí es cómo surge la complejidad a partir de cosas que son sencillas. Esto, como el vínculo existente entre caos y fractales, se ve claramente si volvemos a nuestra vieja amiga, la ecuación logística, e interpretamos sus efectos en términos de topología, lo cual aporta el nivel de complejidad adicional al que hicimos alusión anteriormente. Lo que hace la ecuación logística (y el tipo de procesos iterativos similares que hemos estado comentando) es transformar un conjunto de números en otro.

Si los números originales son puntos de un plano, de la superficie de una esfera o de alguna otra superficie, es posible transformarlos uno por uno en puntos que se encuentran en otro lugar del plano, en otro plano o en alguna otra superficie. Todos conocemos procesos de este tipo, aunque son tan habituales que rara vez pensamos en lo que se está haciendo. Se trata de los procedimientos cartográficos. El tipo de mapa que utilizamos para saber desplazarnos por una ciudad, o para estudiar geografía, no reproduce todos los puntos del mundo real, pero un plano de calles es una copia del trazado de las calles de una ciudad que hace una fiel reproducción de las mismas en miniatura, sin más que reducir su tamaño para que quepan en una hoja de papel.

Un globo puede utilizarse, en principio, como un mapa terrestre sin distorsiones y fiable, ya que la Tierra es aproximadamente esférica. Sin embargo, un mapa puede ser también una copia fiable *y* presentar distorsiones. El plano del metro de Londres constituye un buen ejemplo, hay puntos que representan todas las estaciones y líneas que representan todas las líneas de trenes, de tal manera que el mapa mantiene la esencia de la relación existente entre estos puntos y estas líneas, pero se ha distorsionado para que sea más fácil de interpretar.

Hasta cierto punto, las distorsiones del plano del metro son opcionales, es posible también dibujar planos del metro de Londres que sean copias mucho más fiables del original, representando todos los vericuetos de las líneas. Pero un plano del metro que se dibuje en una hoja de papel *debe* estar distorsionado, porque lo que se pretende es dibujar la superficie de una esfera en un plano. Esta es la razón por la

cual las formas de los continentes que aparecen en la conocida proyección de Mercator y la famosa proyección de Peters (dos modos diferentes de resolver el problema) no son exactamente iguales, y también lo que explica que en ninguna de estas proyecciones aparezcan exactamente las formas reales que tienen los continentes en la superficie de la Tierra.

Pero en ambos casos es posible representar cualquier punto de la superficie terrestre mediante un punto de la superficie del mapa. De hecho, la proyección de Mercator es también un mapa de la proyección de Peters, y viceversa; ninguna de ellas es intrínsecamente «mejor» que la otra.

El proceso que describe la ecuación logística es también cartográfico. Podemos comprobarlo observando cómo afecta esta ecuación a un conjunto de números que representan los puntos de una línea recta. Recordemos que la ecuación cambia un valor de * por otro nuevo valor utilizando la transformación

$$x \text{ (siguiente)} = Bx\,(1 - x)$$

y que, gracias a la renormalización, sólo hemos que tener en cuenta valores iniciales de a; comprendidos entre 0 y 1. Estos valores iniciales pueden representar un segmento de línea recta, podemos imaginarnos, por ejemplo, una vara de medir de una unidad (1 metro) de longitud. Para hacerlo más sencillo, elegiremos un valor entero para B, por ejemplo, 3, y veremos cómo se transforma el segmento mediante la ecuación logística, cómo cambia su trazado al obtener otra disposición de los puntos.

Podemos ver lo que sucede, observando el modo en que funciona esta representación cuando se trata de puntos separados por un décimo de la longitud del segmento original. El punto x = 0, obviamente, se lleva al punto 0. Para un valor de x igual a 0,1, un décimo de la longitud del segmento, obtenemos para x (siguiente) un valor de 3 x 0,1 x 0,9, es decir, 0,27.

El punto x = 0,1 se ha llevado al punto *x* = 0,27. Del mismo modo, el punto *x* = 0,2 se lleva al punto 0,48, el 0,3 se lleva al 0,63, el 0,4 al 0,72 y el 0,5 al 0,75.

Esta mitad del segmento, desde 0 hasta 0,5, se ha alargado y cubre ahora un segmento de recta que va desde 0 hasta 0,75, lo que demuestra otra curiosa

propiedad del infinito: hay un número infinito de puntos en cada segmento y una correlación 1:1 entre los puntos de los dos segmentos, pero uno de ellos tiene una longitud superior en un cincuenta por ciento a la del otro.

Entre 0,5 y 1 sucede lo mismo, pero a la inversa, como podemos comprobar rápidamente. El punto 0,6 se lleva al 0,72, el 0,7 se lleva al 0,63, el 0,8 al 0,48 y el 0,9 al 0,27. Además, dado que (1 - 1) es 0, el punto $x = 1$ va al punto 0.

Este medio segmento no sólo se ha alargado, sino que se le ha dado la vuelta y se ha colocado sobre la representación de la primera mitad del segmento. Se ha producido una transformación topológica del segmento de recta, un alargamiento y un plegamiento, por lo que, aunque el segmento es ahora un 50 por ciento más largo, gracias al plegamiento encaja en el 75 por ciento de la longitud del segmento original.

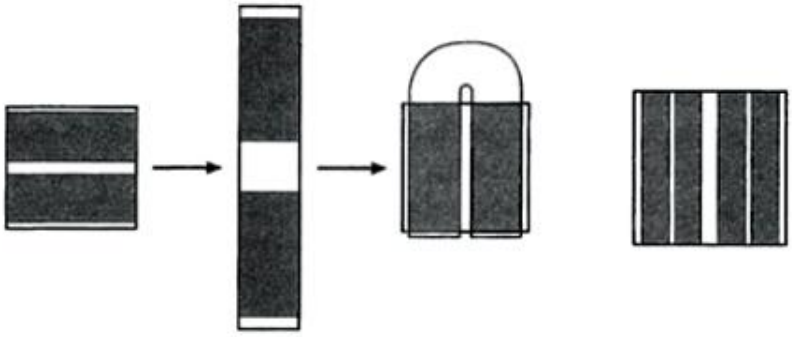

Figura 3.10. La herradura de Smale, producida por la transformación del panadero (véase en el texto). Repitiendo indefinidamente este sencillo proceso de plegamiento y alargamiento, se forma una estructura de múltiples capas espaciadas como los puntos en el conjunto de Cantor.

El vínculo entre álgebra y topología es una idea tan profunda como la constatación de que una ecuación algebraica, por ejemplo $y = x^2$, puede representarse gráficamente mediante una línea en una hoja de papel o una trayectoria en el espacio, en el ejemplo mencionado, una parábola. Pero hemos de recordar que es en el espacio de las fases, y no en el espacio real, donde tienen lugar las transformaciones topológicas que nos interesan aquí.

Si pensamos en términos físicos, el proceso de alargamiento y plegamiento que

describe la ecuación logística produce, en vez de líneas matemáticas abstractas, algo parecido a una forma de herradura a partir de un cilindro inicial, y recibe a veces el nombre de transformación de herradura (o herradura de Smale), a pesar de que la «curva» es al final una horquilla de 180 grados, y no la suave curva redondeada de una herradura real. Pero, ¿por qué íbamos a dar por finalizado el tema con una transformación de herradura?

El modo de hacer estructuras complejas a partir de reglas sencillas es, como ya sabemos, la repetición. Por consiguiente, iteramos todo el proceso, transformamos la «herradura» del mismo modo, a continuación transformamos la herradura transformada, y así sucesivamente, de manera indefinida. Este proceso se parece un poco al modo en que un panadero trabaja la masa, estirándola, doblándola sobre sí misma, aplanándola con unos golpes, estirándola y doblándola de nuevo, así una y otra vez. El proceso se conoce como transformación del panadero *(baker transformation).*

Con independencia del nombre que se prefiera, en la época en que científicos como Robert May comenzaron sus investigaciones sobre la ruta hacia el caos mediante la duplicación del período, es decir, en la década de 1970, y por suerte para el estudio del caos, las propiedades topológicas de la transformación de herradura ya habían sido investigadas, sencillamente a causa de su interés topológico intrínseco, por el brillante teórico Stephen Smale, el cual, como ya hemos visto, entró en contacto con el caos a través de Robert Yorke.

No necesitamos entrar en los aspectos matemáticos, ya lo hizo él por nosotros. Pero la sencilla imagen física de lo que sucede en el espacio de las fases cuando se realiza la transformación de herradura un número indefinido de veces está respaldada por todo el peso de las matemáticas y realmente da una idea auténtica de lo que está pasando.

Pensemos en el segmento original como en un atractor que se encuentra en el espacio de las fases e intentemos visualizar lo que le sucede al aplicarle reiteradamente la transformación de herradura. En cada etapa, la longitud del segmento aumenta, pero, al plegarse de nuevo sobre sí mismo, ocupa menos espacio «horizontal mente». Después del primer paso, tenemos dos segmentos colocados el uno sobre el otro (pero sin ocupar distancia vertical, ya que los

segmentos no tienen anchura), estando uno de ellos curvarlo; tras el paso siguiente, hay cuatro capas con tres plegamientos; después del siguiente, son ocho capas y siete plegamientos, y así sucesivamente. El número de capas se duplica cada vez y el número de plegamientos se obtiene multiplicando por 2 el anterior y sumando 1. Se termina con una curva replegada un número infinito de veces (pero con los plegamientos dispuestos de una manera muy específica: curvaturas dentro de curvaturas que a su vez están dentro de otras, formando una red autosimilar), formada por una cantidad infinita de capas, pero sin que haya distancia entre ellas ni horizontal, ni verticalmente.

Trazando una sección transversal que cortara el montón formado por esta estructura de infinitas capas, veríamos que los puntos están distribuidos como los de un conjunto de Cantor. Pero, ¿de dónde proceden los puntos contiguos de las distintas capas? A causa del proceso de alargamientos y plegamientos reiterados, dos puntos de la curva final que al principio hayan estado juntos en el segmento original pueden acabar separados en este conjunto, mientras que dos puntos que inicialmente se encontraran casi en extremos opuestos del segmento original pueden acabar estando muy juntos. Si el estado de un sistema se desplaza uniformemente a lo largo del segmento original en una dirección, pasando por todos y cada uno de los puntos, parecerá que esté dando saltos de manera aleatoria entre los puntos de un conjunto de Cantor. Esta es la topología asociada con el inicio del caos producido por duplicación del período.

La formación de infinitas capas, como un hojaldre de infinitas hojas, se produce también en el atractor de Lorenz. Como ya vimos en la figura 2.3 (y en el texto correspondiente), la línea que representa el atractor de Lorenz en el espacio de las fases parece cortarse consigo misma muchas veces, de hecho, un número infinito de veces. Pero lo que sucede realmente en el espacio de las fases es que se cruza «viajando» por una capa diferente de este espacio, es decir, por un plano diferente. Una forma de intentar visualizar esto sería pensar en un libro que tiene un número infinito de páginas infinitesimalmente delgadas y que está abierto por la mitad, de manera que queda plano. Uno de los lóbulos del atractor de Lorenz se dibuja en la página izquierda y el otro en la derecha, hay un bucle, o una serie de bucles, alrededor del lóbulo correspondiente en cada página del libro. Sin embargo, cada

vez que la trayectoria del atractor situado en el espacio de las fases cruza la parte central del libro y pasa al lado opuesto, lo que hace es desplazarse a una página diferente. Hay un número infinito de lugares de cruce, pero la línea que representa al atractor nunca se corta consigo misma.

En ambos casos, tanto en el atractor de herradura, como en el atractor de Lorenz, un número infinito de capas del espacio de las fases está contenido en una porción finita de volumen de dicho espacio. Ambos atractores son fractales, se trata de atractores extraños. Y esto es sólo el principio.

Si realizamos el mismo tipo de alargamientos y plegamientos con atractores que no comienzan como segmentos de recta en el espacio de las fases, tales como un atractor que se enrolla alrededor de un toro en dicho espacio, obtendremos al final una pauta de caos fractal aún más complicada, pero siempre basada en un conjunto de reglas muy sencillas. Finalmente, hemos conseguido toda la información que necesitamos sobre el modo en que la sencillez fundamental del mundo puede producir estructuras complejas, y es el momento de empezar a ascender a través de las capas de complejidad para aplicar estos conocimientos al nacimiento de la vida. Todavía hemos de dar varios pasos en ese camino, pero, para despertar el apetito del lector, he aquí un indicio de lo relevante que puede ser lodo esto, no sólo para la vida en general, sino para la vida humana en particular.

Recordemos el modo en que habíamos calculado la dimensión fractal de la curva de Koch utilizando las potencias, o los exponentes, que aparecen en la ley de representación a escala. Esta relación se llama una ley potencial, 3 elevado a la segunda potencia es 3^2, 3 elevado a la tercera potencia es 3^3 y así sucesivamente, y debemos elevar 3 a 1,2619 con el fin de obtener 4, el número de piezas que componen la unidad básica de la curva.

Algo tan sencillo como el volumen de un objeto sigue también una ley potencial. Si tenemos un cubo tal que la longitud de sus aristas es l, el volumen es proporcional a l^3 con independencia de cuál sea el valor de *l;* si tenemos una esfera de radio r, el volumen es proporcional a r^3 independientemente del valor de *r.*

El volumen sigue una ley potencial de exponente 3. A mediados de la década de 1980, los investigadores que estaban estudiando las tasas metabólicas de animales de distintos tamaños se quedaron intrigados al descubrir que, aunque estas tasas

también seguían una ley potencial, no se trataba de una sencilla ley de exponente 3, como se podría esperar por el modo en que se distribuyen estos tamaños.

Compararon la tasa metabólica con la masa corporal de mamíferos tan diferentes como las ratas, los perros, los seres humanos y los caballos. La masa corporal de un animal es proporcional a su volumen y, como sería de esperar, cuanta más masa tiene más alta es su tasa metabólica, porque es mayor el cuerpo que está haciendo la combustión de los alimentos y liberando energía. Sin embargo, aunque la masa aumenta según una ley potencial de exponente 3, la tasa metabólica crece siguiendo una ley potencial de exponente 2,25. En este sentido, los animales se comportan como si su tamaño no se ajustara a un volumen tridimensional, sino como algo intermedio entre un volumen y una superficie bidimensional, en concreto como una superficie frac- tal extremadamente arrugada. Un matemático (al menos un topólogo) interpretaría inmediatamente esta ley potencial como algo que sugiere que los objetos implicados son superficies fractales arrugadas dentro de volúmenes finitos.

Si observamos los cuerpos con más detalle, descubriremos que muchas de las características de los seres vivos son de tipo fractal. El modo en que las arterias y venas se ramifican, por ejemplo, tiene un carácter esencialmente fractal, lo cual hace posible que la sangre llegue a cualquier parte del cuerpo y regrese desde allí sin que las venas y las arterias ocupen tanto espacio que no quede lugar para nada más. Esto se ve claro especialmente en los riñones, donde las venas y las arterias se ramifican entrelazándose de una manera extremadamente compleja con el fin de intercambiar fluidos. El riñón es sin duda un objeto tridimensional finito; sin embargo, las venas y las arterias que están dentro de él tienden a conseguir la longitud infinita de un auténtico fractal.

Por supuesto, la analogía falla en los extremos. Los sistemas que se ubican dentro del riñón humano no se ramifican en realidad un número infinito de veces, sino sólo muchas veces; y, considerándolo en el otro sentido, no encontramos riñones encajados dentro de super-riñones, en una sucesión incesante de encajes, sino que cada sistema está contenido dentro de sí mismo. No obstante, la similitud entre muchos sistemas vivos y los fractales es más que una mera analogía y explica también, por ejemplo, el modo en que la superficie bidimensional de los pulmones

puede tener un área suficientemente amplia como para que haya un intercambio de dióxido de carbono y oxígeno a través de dicha superficie con la rapidez necesaria para mantener el cuerpo vivo, aunque los pulmones ocupen un volumen bastante pequeño.

La autosimilitud casi fractal es una característica que impregna los cuerpos de los organismos vivos. Por otra parte, recordemos que, además de hacer su trabajo con efectividad, la cantidad de ADN requerida para codificar la producción de estos sistemas es, de hecho, muy modesta, si la comparamos con la cantidad de información contenida en el plano de una estructura tan compleja como el riñón. Es la sencillez de las reglas en que se fundamenta la estructura de los fractales lo que permite que los sistemas vivos sean suficientemente complicados como para que evolucionen unas criaturas capaces de plantear preguntas sobre la naturaleza del mundo.

Lo que hemos visto es que podemos tener sistemas muy sencillos, en los que no sucede nada interesante, como un grifo goteando a una velocidad uniforme con período 1. Y también podemos tener sistemas caóticos, en los que lo esencial son las fluctuaciones caóticas, de tal forma que no hay orden y la estructura se destruye. Pero entre ambos casos, partiendo del aburrido extremo del espectro, tenemos un nivel creciente de complejidad, comenzando por el grifo que gotea con período 1 y adquiriendo cada vez un interés mayor hasta que se desencadena repentinamente el caos.

Por lo tanto, los fenómenos u objetos más complejos y más interesantes del universo se producen precisamente al borde del caos, justo antes de que el orden quede destruido. Es ahí donde encontramos grifos (o canalones con fugas) que gotean con ritmos extraños y exóticos, remolinos que se producen dentro de otros remolinos y giran describiendo pautas asombrosas, y la extraordinaria complejidad del riñón, o la superficie de la corteza cerebral humana, que se pliega y se repliega hasta resultar casi un fractal. Hasta aquí hemos observado el orden y también el caos. Ya es hora de que dirijamos la mirada al borde extremo del caos, allí donde habita la complejidad.

Capítulo 4
Al borde del caos

En cierto sentido, la termodinámica clásica pretende que el tiempo no existe. Los sistemas se describen considerando pequeños cambios infinitesimales que tardarían un tiempo infinito en hacer que esos sistemas pasaran de un estado a otro. La termodinámica clásica asume también la no existencia de flujos de energía, concretamente, el calor, a través de un sistema. Sin embargo, como ya hemos visto, en el sencillo ejemplo de la separación del gas hidrógeno del ácido sulfhídrico gaseoso por termodifusión, empiezan inmediatamente a suceder cosas interesantes cuando la energía fluye a través del sistema. Por un cadáver no fluye una cantidad significativa de energía, pero a través de nuestro cuerpo sí fluye una gran cantidad de energía, la energía química procedente de la metabolización de los alimentos que ingerimos, una energía que potencia nuestros músculos y el resto de nuestra actividad corporal, disipándose en última instancia en forma de calor. La disipación de la energía es una característica esencial en la termodinámica de la falta de equilibrio, y nos ofrece otro modo de referirnos al flujo de energía que se produce a través de un sistema, calificándolo de proceso disipador. A estos sistemas se les ha llamado también sistemas abiertos, porque (a diferencia de la hipotética caja cerrada con gas en su interior, que era el ejemplo predilecto de los pioneros de la termodinámica en el siglo XIX) no están sellados, es decir, no están aislados del mundo en sentido amplio. Es en los sistemas cerrados donde encontramos la reversibilidad en el tiempo y las recurrencias de Poincaré; en los sistemas abiertos encontramos irreversibilidad y una flecha del tiempo.

Como vimos en el capítulo 2, la termodinámica clásica se basa en una contradicción. El propio término «dinámica» nos dice que esta rama de la ciencia describe el modo en que cambian los sistemas; sin embargo, ideas tales como la entropía se derivan de cálculos en los que se utilizan sistemas en equilibrio, donde nada cambia. El equilibrio como tal carece de interés intrínseco, porque nada sucede dentro de él. Pero sí puede resultar interesante el modo en que las cosas se mueven hacia el equilibrio. El momento en que un ser vivo se encuentra más cerca del equilibrio es cuando muere. El hecho de que algo esté muerto no es ni de cerca tan interesante

como el modo en que muere, una premisa en la que se basa el éxito de todo un género de novelas de detectives, los crímenes misteriosos.

Cuando los científicos comenzaron a estudiar procesos disipadores, era natural que empezaran por investigar sistemas cercanos al equilibrio y procesos del tipo de la difusión térmica. Si los sistemas se encuentran cerca del equilibrio, generalmente responden de un modo lineal ante los cambios que se producen en su entorno, por ejemplo, si cambiamos ligeramente el gradiente de temperatura en el proceso de difusión térmica, el sistema responderá de manera lineal, con un cambio también pequeño. Los fundamentos para el estudio de la termodinámica lineal fueron establecidos por el químico noruego Lars Onsager (1903-1976), que trabajaba en la Brown University a principios de la década de 1930 (se trasladó a Yale en 1933). Su contribución decisiva, que mereció ser destacada para una mención especial en la concesión del premio Nobel (otorgado en 1968), fue el descubrimiento de lo que se conoce como relaciones recíprocas, lo cual significa que hay simetría en tales sistemas. En el caso de la difusión térmica, por ejemplo, del mismo modo que un gradiente de temperatura produce un gradiente en la concentración de la mezcla de gases, así también, según la teoría de Onsager, el establecimiento y el mantenimiento de un gradiente de concentración deben producir un gradiente de temperatura, ocasionando un flujo de calor. Más tarde se realizaron experimentos que lo confirmaron, y las relaciones recíprocas llegaron a ser conocidas también con el nombre de cuarta ley de la termodinámica.

Onsager había proporcionado a los químicos un conjunto de herramientas que podían utilizar para iniciar la investigación de fenómenos irreversibles en sistemas disipadores o, al menos, en aquellos en los que se cumplían reglas lineales (el «régimen lineal»). La persona que hizo el uso más efectivo de este instrumento fue Ilya Prigogine, nacido en Moscú en 1917, que emigró con su familia a Bélgica cuando tenía doce años de edad, y en 1945 estaba trabajando en la Universidad de Bruselas. Fue Prigogine quien demostró, como ya hemos mencionado en el capítulo 1, que un sistema disipador en régimen lineal no se estabiliza en el estado de muerte correspondiente a la máxima entropía (como si estuviera en equilibrio), sino en un estado en el que la entropía se produce lo más lentamente posible y la actividad disipadora funciona a una velocidad mínima. Ésta es la situación que se

produce en la difusión térmica cuando el gradiente de concentración de los gases se ha estabilizado para un gradiente de temperatura determinado, pero Prigogine demostró que esto se cumple en general para los sistemas que describe la termodinámica lineal.

Las cosas existen en un estado de equilibrio dentro de un régimen lineal; por ejemplo, un ser humano puede mantener su integridad durante muchos años utilizando el flujo de energía (y alimentos) que circula dentro de su cuerpo, aunque en este caso el estado de equilibrio finalmente se rompe, por razones que aún no se entienden. Pero el estado de equilibrio de un adulto humano es notablemente diferente de los cambios radicales que se producen cuando se desarrolla un nuevo ser humano a partir de la célula única del óvulo fertilizado, lo cual evidentemente es un proceso no lineal. En breve hablaremos más sobre este asunto.

Durante las dos décadas siguientes, los trabajos de Prigogine y sus colegas en Bruselas (estos científicos llegaron a ser conocidos como la «Brussels School») se centraron en el intento de proporcionar una descripción matemática de la termodinámica para sistemas que se encuentran lejos del equilibrio, dentro del régimen no lineal en que cualquier pequeño cambio que se produzca en el entorno exterior puede ser la causa de grandes cambios en el sistema. Quizá lo mejor sería considerar esto, incluso hoy en día, como un trabajo que se encuentra todavía en curso; está lejos de ser una teoría terminada. Felizmente, no necesitaremos entrar en detalles, o enredarnos en las controversias que rodean este trabajo, ya que los aspectos importantes y ampliamente aceptados pueden entenderse a partir del comportamiento de sistemas físicos y químicos sencillos. Pero hay un aspecto fundamental que debe ser tenido en cuenta: para describir estos sistemas se utilizan ecuaciones no lineales en situaciones, en las que la retroalimentación es importante.

Al mismo tiempo, Prigogine ha desarrollado también sus propias ideas sobre la naturaleza del tiempo, así como sobre la relación existente entre la termodinámica y la flecha del tiempo; son probablemente estas teorías las que han contribuido más a que haya sido conocido fuera de los círculos científicos profesionales, y ésta es la única razón por la que las mencionamos. El enigma que se pretende resolver en este trabajo es cómo pudo surgir el orden, cosas como nosotros mismos, a partir

del caos, un caos *no* determinista, sino esencialmente lo que los antiguos griegos llamaron caos, la distribución uniforme de los gases que existían cuando el universo era joven. Pero no entraremos en este aspecto de su trabajo, porque a la mayoría de los físicos les parece un callejón sin salida, ya que existe un modo natural de explicar por qué hay una flecha del tiempo en el universo, y en particular, como explicaremos a continuación, el modo en que ciertos sistemas ordenados, como nosotros mismos, pueden haber surgido del estado de desorden casi uniforme que emergió del big bang.

Lo mejor para comenzar a desarrollar un cierto conocimiento relativo a las cosas interesantes que pueden suceder en los sistemas abiertos que se encuentran lejos del equilibrio dentro del régimen no lineal es hablar de un fenómeno cuyo primer descubridor fue el francés Henri Bénard en 1900. Debido a que este fenómeno fue investigado posteriormente por el físico inglés lord Rayleigh, se conoce, por una parte, como inestabilidad de Bénard, y por otra como inestabilidad de Rayleigh-Bénard. El modo más efectivo de ver lo que sucede con esto es calentar una sartén plana que contenga una fina capa de aceite de silicona, aplicando el calor cuidadosamente desde abajo, en condiciones controladas dentro de un laboratorio. La capa de aceite debe tener sólo un milímetro de espesor, más o menos, y el calentamiento ha de ser uniforme a través de la base plana del recipiente. Se puede hacer utilizando un cazo en una cocina corriente, aunque no recomendaríamos a nadie que intentara calentar finas capas de aceite de esta manera; algo similar puede observarse a veces en una cazuela que contiene una sopa espesa cuando se lleva a ebullición lentamente. Pero el aspecto importante del experimento que vamos a explicar es que no deseamos calentar el aceite a tanta temperatura que comience a hervir de una forma caótica y desordenada. Nos interesa lo que sucede justo antes de que se desencadene el caos. Dado que el aceite de por sí es transparente, para que sea más fácil ver lo que sucede, se suele mezclar con polvo de aluminio (algo que no recomendamos para la sopa).

Lo que se produce, por supuesto, es una convección. El fluido caliente que se encuentra en el fondo de la capa de aceite se expande, se vuelve menos denso, e intenta ascender a través del fluido más denso y más frío que está encima. Si el calentamiento no es uniforme, esto no plantea problema alguno. El fluido

simplemente asciende donde está más caliente, se detiene en algún otro lugar y fluye de nuevo hacia abajo hasta la zona caliente, formando así una célula de convección circulante. Sin embargo, en un fluido que se calienta de manera uniforme, no hay un lugar especial donde este proceso pueda iniciarse. Hay una simetría uniforme en toda la base del fluido y todo él «quiere» ascender al mismo tiempo. Al principio, no sucede nada. Luego, cuando la temperatura aumenta gradualmente en la parte inferior, mientras se mantiene fría la parte superior (de tal modo que el gradiente de temperatura se agudiza), al llegar a un punto crítico la simetría se rompe y la superficie uniforme del fluido se fragmenta formando una pauta de pequeñas células de convección hexagonales. El punto en el que esto sucede depende de la diferencia de temperatura entre la parte superior y la parte inferior del fluido, el gradiente de temperatura, y las células que se forman de este modo tienen a través del fluido una altura aproximadamente igual a la profundidad del aceite.

En cada célula, el fluido caliente asciende en el centro, se mueve hacia los bordes y desciende hasta la frontera con la célula vecina. Dado que la superficie del Huido está abierta al aire, los efectos de la tensión superficial son importantes y fomentan el movimiento del fluido hacia afuera en la superficie, empujándolo desde el centro de cada célula y produciendo un efecto de hoyuelo, estando el centro de la célula por debajo de la altura a la que están los bordes, aunque el fluido esté ascendiendo. La situación que estudió Rayleigh era ligeramente distinta, ya que una fina capa de fluido se encontraba retenida entre dos placas: una placa caliente por abajo y una fría en la parte superior. Esto hace que tanto la interpretación física como la interpretación matemática del fenómeno sean mucho más sencillas, ya que la tensión superficial no ha de tenerse en cuenta, y la pauta de comportamiento que se pone de manifiesto vincula la inestabilidad de Rayleigh-Bénard con la aparición del caos determinista.

En esta versión simplificada de la convección de Bénard, se puede apreciar varias etapas de la ruta hacia el caos. Cuando la convección comienza, lo hace en forma de largas salchichas, de tal modo que la superficie del fluido adquiere un aspecto rayado. Estas salchichas tienen una sección transversal aproximadamente cuadrada, y en los giros o «rollos» que se forman en torno a ellas la convección

circula respectivamente en el sentido de las agujas del reloj o en el contrario, alternando el sentido de una a otra, ya que, si no, la convección no funcionaría de manera uniforme. Pero no hay modo de decir con antelación si un punto concreto del fluido formará parte del rollo convectivo que gira en el sentido de las agujas del reloj, o del que va en sentido contrario. La disposición exacta se determina aleatoriamente. Existe una bifurcación en la que el sistema pasa de un solo estado a un par de ellos.

Cuando el gradiente de temperatura aumenta,[42] en otro punto crítico la pauta cambia bruscamente a lo que se llama una forma bimodal, en la que hay dos conjuntos de rollos convectivos que funcionan simultáneamente formando el uno con el otro ángulos rectos. Esto da lugar a una pauta de células de convección aproximadamente cuadradas cuando se ven desde arriba, como el dibujo de una camisa a cuadros. El número de estados posibles para los componentes del fluido se ha duplicado de nuevo, de 2 a 4.

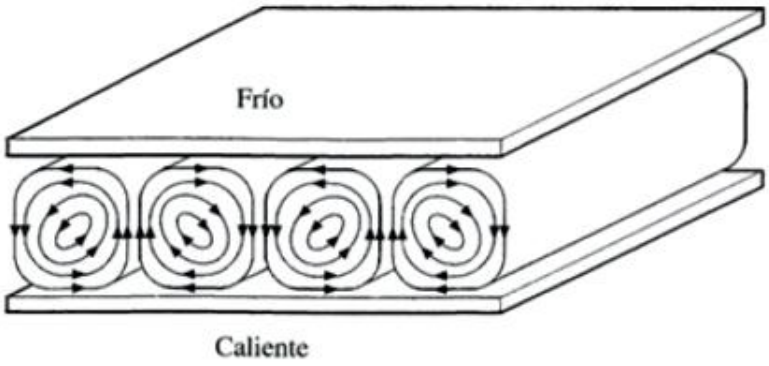

Figura 4.1. Según la versión de la convección de Bénard tal como la estudió Rayleigh, se generan rollos convectivos con una sección transversal aproximadamente cuadrada (como un rollo suizo aplastado).

Este es otro ejemplo de cómo funciona la bifurcación, una variación sobre el tema de ruta hacia el caos mediante duplicación del período (aunque, por supuesto, Rayleigh no sabía nada de esto) y, si el gradiente de temperatura (o el número de Rayleigh) aumenta lo suficiente, acabaremos con una forma turbulenta de

[42] En sentido estricto, diríamos «cuando aumenta el valor de un parámetro llamado número de Rayleigh». Este número no sólo depende del gradiente de temperatura, sino también de las propiedades del fluido.

convección en la que aparecen pautas fugaces que se deshacen y se vuelven a formar incesantemente. Pero las pautas estables más interesantes aparecen justo al borde del caos y, en el caso de la fina capa de aceite que se calienta en una cacerola abierta, a causa de la tensión superficial, la interesante pauta específica que aparece es un conjunto de hexágonos dispuestos como en un panal. Esto sucede lejos del equilibrio, gracias a la energía que fluye a través de un sistema abierto y se disipa. Aquí leñemos el secreto de la existencia del orden en el universo, y especialmente el secreto de la vida. Siempre que nos veamos confrontados con el enigma de la existencia del orden, y especialmente de la existencia de la vida, y nos preguntemos cómo es posible que el mundo haya llegado a ser así, podremos decimos a nosotros mismos: «¿Recuerdas las células hexagonales de la convección de Bénard? Pues es lo mismo».

Un sistema sólo puede mantenerse en un estado interesante fuera del equilibrio si es disipador y está abierto a su entorno, pollo que debe existir una fuente exterior de energía. En la Tierra, la energía nos llega en última instancia del Sol; si llegamos a comprender por qué éste brilla, ese conocimiento nos aclarará por qué ha surgido el orden a partir del caos y cómo el universo llegó a tener una flecha del tiempo.

Todo sucede gracias a la gravedad, que tiene una propiedad muy curiosa: la energía gravitatoria de cualquier objeto que tenga masa es negativa. Es literalmente menos que cero y, cuanto más compacto sea el objeto, más negativa es la energía gravitatoria. Como mera exposición de los hechos, esta idea puede ser difícil de aceptar. Sin embargo, es fácil entender lo que sucede si se examinan dos aspectos de la gravedad. En primer lugar, la fuerza gravitatoria que se ejerce entre dos objetos disminuye a medida que éstos se alejan uno del otro. El hecho es que la fuerza disminuye en proporción inversa al cuadrado de la distancia, tal como demostró Newton. La cantidad de energía gravitatoria almacenada en un conjunto de objetos (ya sean átomos, estrellas, bloques de hormigón o pinturas de Van Gogh) depende de la fuerza que se ejerce entre ellos.

Esta no es una característica exclusiva de la gravedad, cuando se estira un trozo de elástico de goma, es necesario tirar cada vez con mayor fuerza para estirarla más y más, y la energía almacenada en el elástico está relacionada con la fuerza que aplicamos para estirarlo. Sin embargo, para «estirar» la distancia entre dos (o más)

objetos gravitatorios se necesita tirar cada vez menos a medida que se distancian más. Ésta es la razón por la que se necesita un enorme cohete acelerador para colocar un satélite en órbita alrededor de la Tierra, pero, una vez que está en órbita, sólo se necesita un pequeño cohete para enviarlo a la Luna o a Marte. Si los objetos estuvieran a una distancia infinita, la energía total almacenada por el campo gravitatorio que los vincula mutuamente sería cero, fundamentalmente porque la fuerza es proporcional a 1 dividido por infinito elevado al cuadrado.

Imaginemos ahora un conjunto de partículas (átomos, ladrillos, o lo que sea) dispersas a distancias infinitas unas de otras, a las que se aplica un leve codazo para que empiecen a moverse tollas bajo la influencia de la gravedad. Al hacerlo, cada una de ellas adquiere energía cinética y se mueve cada vez más rápido a medida que cae. Esta energía procede del campo gravitatorio, del mismo modo que, si atamos los extremos de un elástico de goma a dos guijarros, los separamos estirando y luego los soltamos, los guijarros chocarán entre sí cuando la energía almacenada en el elástico se convierta en energía de movimiento (energía cinética).

De manera similar, la energía cinética de las partículas que caen conjuntamente procede del campo gravitatorio; pero este campo comenzó con energía cero. Lo que pasa es que ahora tiene menos que cero, energía negativa. En el momento en que las partículas se reúnen para formar una estrella, el campo gravitatorio asociado con ésta tiene mucha energía negativa. ¿Cuánta? La respuesta puede parecemos sorprendente, pero al menos estamos en muy buena compañía.

Ya he contado esta anécdota en otras ocasiones,[43] pero no hay mejor manera de ilustrar el asunto, y sería tonto suponer que lodo el que lea esto ha leído también mis libros anteriores.

Se remonta a la década de 1940, cuando los físicos ya conocían lo que podría llamarse la negatividad de la gravedad, pero la mayoría de ellos estaba ocupada principalmente haciendo trabajos relacionados con la guerra, en EE.UU., sobre todo el Proyecto Manhattan. Un físico asentado en Estados Unidos, al que no se permitía trabajar en los proyectos más secretos porque era ruso de nacimiento, fue George Gamow (1904-1968), dotado de un carácter entusiasta y con una amplia gama de intereses, el cual, entre otras cosas, desarrolló el primer modelo del origen del

[43] Véase In Search of the Big Bang. [Hay traducción al castellano: John Gribbin, En busca del big bang, Ediciones Pirámide, S.A., Barcelona, 1989, trad, de Manuel Martínez.]

universo según el big bang. Su contribución al esfuerzo de guerra fue como asesor en el Bureau of Ordnance (servicio de material de guerra) del Ministerio de Marina de EE.UU., en Washington, DC, y una de sus obligaciones era viajar a Princeton cada quince días con un maletín de documentos para Albert Einstein. Aquellos documentos no contenían grandes secretos, explicaban algunas ideas relacionadas con inventos que podían ayudar a ganar la guerra, y que enviaban a la armada algunos ciudadanos bienintencionados, pero en general bastante descaminados, y el ministerio quería la opinión de Einstein sobre si valía la pena intentar llevar a la práctica alguna de ellas. En los documentos podía haber ideas tales como la sugerencia de congelar el océano Atlántico para evitar que operaran allí los submarinos, por lo que no parece que fuera innecesario poner a prueba la habilidad de Einstein para separar el trigo de la paja, aunque, habiendo sido funcionario de patentes en Suiza, Einstein era bastante bueno para detectar los fallos en propuestas más sutilmente inviables.

Un día, mientras caminaba por Princeton con Einstein, desde su vivienda hasta el Institute for Advanced Studies, Gamow mencionó, sin darle importancia, una idea que había lanzado tiempo atrás uno de sus colegas, el especialista en física cuántica Pascual Jordán. No era más que una idea loca que se le había ocurrido a Jordán, el tipo de cosa que puede animar una conversación mientras se toma café, o un paseo por Princeton. Jordán había calculado que, en el caso de cualquier objeto material que tuviera su masa concentrada en un punto, la energía negativa del campo gravitatorio asociado sería $-mc^2$ y, por consiguiente, contrarrestaría exactamente toda la energía másica positiva de la estrella. Dicho de otro modo, tal como Gamow se lo contó a Einstein, se podía crear una estrella a partir de absolutamente nada. «Einstein se paró en seco,» dice Gamow,[44] «y, como estábamos justo cruzando la calle, varios coches tuvieron que parar para no arrollarnos».

Hay que insistir en que esto no es una mera analogía, ni el resultado del procedimiento que elegimos para medir algo como la energía. Una verdad fundamental sobre el modo en que funciona el universo es que el campo gravitatorio tiene energía negativa y que, cuando la materia está concentrada en un punto, esta energía negativa compensa exactamente la energía másica de la

[44] Véase George Gamow, *My World Line.*

materia.

La idea que hizo a Einstein pararse en seco mientras cruzaba una carretera tuvo poco impacto en aquella época, pero cuarenta años más tarde se convirtió en la piedra angular de la teoría según la cual el universo en su totalidad podía haber surgido de este modo a partir de la nada, como una burbuja de energía másica con un campo gravitatorio asociado que llevaba una cantidad de energía igual, pero de signo opuesto, de tal modo que la energía total del universo es cero, y aquella burbuja diminuta fue hinchándose, en un proceso llamado inflación, hasta llegar a ser el universo que vemos a nuestro alrededor. Todo esto lo explico en mi libro *In Search of the Big Bang,* pero lo único que nos importa aquí es que sabemos que el universo comenzó a existir a partir del big bang en un estado muy uniforme. Los detalles de lo que sucedió en el big bang son todavía tema de debate, pero las observaciones directas del universo nos dicen cómo era éste muy poco después de empezar a existir.

La característica más importante del universo actual es que está en expansión. Todas las estrellas que vemos en el cielo nocturno forman parte de un sistema en forma de disco, una galaxia llamada la Vía Láctea, que contiene cientos de miles de millones de estrellas. La Vía Láctea es una más entre cientos de miles de mi Nones de galaxias, y éstas se encuentran agrupadas en *clústeres* o cúmulos de galaxias que se mantienen juntas por efecto de la gravedad, pero estos cúmulos se van alejando unos de otros a medida que se expande el universo. Lo que causa la expansión es el propio espacio que se estira entre los cúmulos de galaxias, y no el hecho de que los cúmulos se desplacen por el espacio. Esto lo explica en todos sus detalles la teoría general de la relatividad de Einstein (de hecho, lo predijo esta teoría). Si ahora todo se está separando aún más, es obvio que en el pasado estuvieron las cosas más cercanas unas de otras y que, si nos remontamos lo suficientemente lejos hacia el pasado, estaban todas unas encima de otras, formando un solo bloque, antes del big bang. Midiendo la velocidad de expansión actual y aplicando la teoría general de la relatividad, podemos fijar el tiempo que ha transcurrido desde el big bang en unos catorce mil millones de años.

Una de las maravillas de la astronomía es que, como la luz tarda un tiempo finito en desplazarse por el espacio, vemos los objetos distantes tal como eran en el pasado.

Por ejemplo, una galaxia que está a una distancia de diez millones de años luz la vemos con la luz que emitió hace diez millones de años. Si consideramos «el comienzo» como instante cero y nos movemos hacia adelante, lo más lejano en el tiempo que podemos ver con nuestros detectores es una época en que el universo tenía unos 300.000 años de edad y era un mar de gas caliente (en sentido estricto, un plasma) casi totalmente uniforme, y con más o menos la misma temperatura que tiene ahora la superficie del Sol, aproximadamente 6.000 °C.

La radiación procedente de esa caliente bola de fuego se ha enfriado a medida que ha pasado el tiempo y el universo se ha expandido (del mismo modo que el gas caliente que hay en una caja cerrada se enfría si la caja se expande), y se detecta hoy en día como un débil silbido de interferencias radiofónicas, en la zona de microondas del espectro electromagnético, con una temperatura por debajo de *menos* 270 °C (o 2,7 grados por encima del cero absoluto).

Unas muy ligeras variaciones en la temperatura de esta «radiación cósmica de fondo en forma de microondas» que procede de distintas zonas del cielo nos dicen que el gas caliente del que se formaron las galaxias y las estrellas estaba entonces distribuido por el espacio de una manera que estaba muy cerca de ser uniforme, aunque no lo era del todo. En algunos lugares el gas era un poco más denso que en otros. Las zonas más densas habrían atraído hacia sí mismas más cantidad de materia mediante la gravedad, aumentando así la irregularidad del universo y llegando finalmente a la situación que vemos actualmente, es decir, la materia concentrada en galaxias de estrellas brillantes, con mucho espacio oscuro entre ellas.

Este es un punto de vista fundamentalmente diferente del que obtenemos estudiando una caja de gas aquí en la Tierra. Allí donde los efectos de la gravedad pueden ignorarse (es decir, los efectos de las fuerzas gravitatorias de las partículas de la caja actuando unas sobre otras), el estado de máxima entropía para una caja de gas es el que se tiene cuando éste está distribuido de manera uniforme a una temperatura también uniforme. Sin embargo, cuando no se pueden ignorar las fuerzas gravitatorias que ejercen unas partículas sobre otras para configurar el gas, como es el caso de las grandes nubes de gas y polvo del espacio, la gravedad puede atraer los objetos hasta juntarlos en grupos, creando más orden y reduciendo *al*

mismo tiempo la entropía. Como dijo Paul Davies, «la inestabilidad inducida por la gravedad es una fuente de información».[45]

Más información implica menos entropía, por lo que también se puede considerar que esto significa que cuando la información «sale» del campo gravitatorio en una nube de gas que se está deshaciendo, el campo gravitatorio está absorbiendo entropía, para compaginarla con su energía negativa. Es precisamente la energía negativa del campo gravitatorio la que hace posible que éste absorba entropía de esta manera, lo cual explica por qué el universo no se encuentra actualmente en equilibrio termodinámico.

Pero dejemos la enorme imagen del universo a gran escala y volvamos a centramos en la aparición de la vida sobre la Tierra. Lo importante es que la gravedad ha sido responsable de una especie de efecto dominó, no a pequeña escala, sino como en las gigantescas hileras de fichas que se vuelcan una tras otra para batir un récord o con el fin de recoger dinero para obras caritativas. En un proceso que se desarrolla paso a paso, se han creado en el universo cantidades cada vez mayores de organización, descendiendo (o ascendiendo) hasta el nivel de seres suficientemente inteligentes como para especular sobre el modo en que pudo suceder todo esto. Sin entrar en detalles relativos al modo en que funcionan las estrellas,[46] podemos ver que existen sólo porque la gravedad hizo que se formaran nubes de gas que llegaron a estar en su interior lo bastante calientes (gracias a la energía cinética que sus partículas constituyentes tomaron del campo gravitatorio) como para que se produjeran reacciones de fusión nuclear. Esto ha llevado a una situación en la que la estrella y su entorno no se encuentran en equilibrio termodinámico. Por el contrario, lo que tenemos es una estrella caliente situada en un espacio frío, de tal modo que la estrella desprende energía en un intento de nivelar la temperatura dentro y fuera. La flecha termodinámica del tiempo apunta a la dirección en la cual el futuro corresponde a la energía que fluye de la estrella; y no es una mera coincidencia que el futuro termodinámico sea también la dirección del tiempo a partir del big bang, a causa del papel que desempeña la gravedad en todo esto. En última instancia, es la gravedad la que dice a la flecha del tiempo hacia dónde debe apuntar.

---

[45] Véase Paul Davies, *The Fift'h Miracle.* [Hay traducción al castellano: *Quinto milagro: la búsqueda del origen* y *significado de la vida.* Editorial Crítica, Barcelona, 2000, (rad. de Javier García Sanz]

[46] Un tema que trato en mi libro *Stardust.*

Un planeta como la Tierra se encuentra bañado por el flujo de energía de una estrella, que convierte toda la superficie del planeta en un sistema disipador abierto. Toda la vida que se desarrolla en la superficie de la Tierra utiliza esta energía para mantenerse lejos del equilibrio, al borde del caos[47]. Las plantas reciben su energía directamente de la luz del Sol a través de la fotosíntesis; los herbívoros obtienen su energía a partir de las plantas; los carnívoros consiguen su energía de otros animales. Pero, en origen todo procede del Sol, y todo sucede, en origen, gracias a la gravedad. Pero el modo en que los sistemas se organizan por sí mismos, utilizando este flujo de energía, llegando a adoptar lo que parece ser formas complejas, en realidad es bastante sencillo. Podemos verlo claramente si nos dejamos guiar por el brillante matemático Alan Turing (1912-1954), que tuvo la idea loca de intentar explicar, hace más de medio siglo, lo que sucede en el proceso más complicado que conocemos: el desarrollo de un embrión a partir de una sola célula viva. Turing se adelantó a su tiempo, y la importancia de su obra en este campo no fue reconocida hasta mucho después de su muerte; utilizando nuestra percepción retrospectiva, éste es lógicamente el punto en el que hemos de iniciar la siguiente etapa de nuestra historia.

Turing, que nació en Paddington, Londres, el 23 de junio de 1912, es conocido sobre todo como criptógrafo, ya que encabezó y dirigió el equipo de Bletchley Park, Buckinghamshire, que durante la Segunda Guerra Mundial descifró los códigos de los alemanes (incluido el famoso código Enigma). Para entonces ya estaba profundamente interesado en la posibilidad de crear una inteligencia artificial, un «ordenador universal» (al que hoy en día se denomina a veces máquina de Turing) que pudiera resolver cualquier problema.

Dado su interés por el modo en que se desarrollaba la inteligencia humana, comenzó a pensar en el desarrollo embrionario. Ciertamente habría avanzado aún más si no hubiera sido por el hecho de que las autoridades le hostigaban por ser un homosexual activo (lo cual era entonces ilegal en Gran Bretaña) y se suicidó el 7 de

[47] También llega energía del interior de la Tierra, principalmente como resultado de la desintegración de elementos radioactivos que se encuentran en el núcleo del planeta. Este material radioactivo se produjo en generaciones anteriores de estrellas y se dispersó por el espacio cuando estas estrellas explotaron, llegando a formar parte de la nube interestelar a partir de la cual se formó el sistema solar. Por lo tanto, también esta fuente de energía debe su origen en última instancia a la gravedad. Las formas de vida que se alimentan de esta energía, la que escapa a través de grietas calientes situadas en el fondo oceánico, pueden hacerlo con total independencia de la energía que aporta la luz solar, pero son también un producto de la gravedad, del mismo modo que lo somos nosotros.

junio de 1954, un par de semanas antes de cumplir cuarenta y dos años, comiendo una manzana que había sumergido en cianuro. Turing se había graduado en el King's College, Cambridge, en 1934, y trabajó allí durante dos años, antes de doctorarse en Princeton, tras lo cual regresó al King's College en 1938. En 1936 escribió un trabajo titulado «*On Computable Numbers*», en el que presentó la idea de la máquina de Turing. En la época esto era meramente un artilugio imaginario, un «experimento mental» diseñado para demostrar la estructura lógica de un hipotético ordenador universal, pero los principios que Turing enunció en aquel trabajo constituyen la base de todos los ordenadores modernos, al tiempo que proporcionaba una visión del mundo de los sistemas complejos autoorganizados.

La máquina imaginaria de Turing estaba provista de una larga cinta de papel (en principio, infinitamente larga) dividida en cuadrados en los que había símbolos o números que podían leerse y, cuando conviniera, borrarse y sustituirse por otros símbolos. Turing pensaba en una cinta de papel que pudiera borrarse y reescribirse cuando fuera necesario. Hoy en día la imagen que se evoca es la de una cinta magnética, como la de un aparato de audio, o incluso el disco duro de un ordenador, o su memoria de acceso aleatorio (RAM), en estado sólido y con posibilidad de leerse o escribirse.

Lógicamente, para este propósito existen otros medios equivalentes a una u otra posibilidad. La lectura del número escrito en un cuadrado le diría a la máquina si ha de moverse hacia adelante o hacia atrás a lo largo de la cinta (o por la memoria), cómo calcular un nuevo número a partir de la información de la cinta, y cuándo reescribir el contenido de un cuadrado. Por ejemplo, el símbolo que haya en un cuadrado podría indicar a la máquina que sumara los números situados en los dos cuadrados siguientes y escribiera la respuesta en el segundo de ellos. De este modo, si las instrucciones de la cinta fueran lo suficientemente específicas, la máquina podría calcular cualquier cosa. Comenzando por un extremo de la cinta, o en algún cuadrado concreto, la máquina iría realizando su laborioso trabajo, moviéndose a veces hacia atrás y hacia adelante por un mismo tramo de la cinta, escribiendo y reescribiendo símbolos en los cuadrados, cambiando su estado interno de acuerdo con las instrucciones que leyera (por ejemplo, cambiando de «sumar» a «dividir»), y terminando finalmente en el otro extremo de la cinta, o en algún

cuadrado que le indicaran las instrucciones, tras haber escrito por sí misma un mensaje que sería la respuesta a la pregunta planteada por el mensaje original. Dependiendo de qué símbolo haya en un cierto cuadrado, siempre sabrá qué hacer a continuación.

Turing demostró que podía existir una máquina *universal* de este tipo (llamada ahora máquina universal de Turing) que sería capaz de resolver cualquier problema que pudiera expresarse en el lenguaje simbólico adecuado.[48] Sería una máquina que, según las propias palabras de Turing:

> *... puede estar hecha para realizar el trabajo de cualquier máquina específica, es decir, para llevar a cabo cualquier cálculo, si se inserta en ella una cinta que lleve las «instrucciones» adecuadas.*

Hoy en día, esto le resulta familiar incluso al más ingenuo usuario de ordenadores, porque es una descripción del modo en que funcionan estos aparatos: el soporte lógico (*software*) programado en su memoria desempeña el papel, de la cinta magnética, mientras que el soporte físico (*hardware*) representa la máquina de Turing. Resulta tan familiar que es difícil hacerse una idea de lo importante que fue este avance en 1936.

Aparte del valor práctico de estas ideas, que pronto sería evidente bajo la presión de la emergencia en tiempo de guerra, el trabajo publicado por Turing hizo surgir cuestiones de fondo que son importantes para comprender la vida y otros sistemas complejos. En muchos casos, la ventaja de una máquina de Turing consiste en que las instrucciones que le dicen cómo efectuar un cálculo (el algoritmo) son mucho más compactas que los resultados de dicho cálculo. Por ejemplo, existe un algoritmo razonablemente breve que especifica el modo de calcular π, sin tener que escribir realmente toda la serie infinita de dígitos que representan el número π. Para muchas aplicaciones, el algoritmo es *π.*

Dicho de una manera más prosaica, la técnica de «multiplicar» que todos conocemos es un algoritmo, aunque en términos cotidianos escribir, por ejemplo, 6 x 9 es casi una versión taquigráfica de 54. Turing demostró la existencia de

48 Técnicamente, Turing demostró que una máquina así podría resolver cualquier problema computable, lo cual significa cualquier problema que se pueda especificar utilizando un algoritmo adecuado; también demostró que algunos problemas no son computables, lo cual tiene un interés enorme para los matemáticos y está relacionado con la idea de que algunos planteamientos no son comprensibles desde un punto de vista algorítmico.

sistemas que no pueden «comprimirse» mediante un algoritmo y cuyas representaciones más compactas son ellos mismos, uno de los conceptos fundamentales que nos surgieron en nuestro comentario anterior sobre el caos, aunque la idea iba en otra dirección. En particular, como ya hemos visto, la descripción más breve del universo es el universo mismo.

En la época en que estalló la Segunda Guerra Mundial, Turing estaba ya intentando construir máquinas reales que demostrarían cómo funcionaban en la práctica aquellos principios, aunque, en cierto modo, se quedarían todavía lejos de ser máquinas de computación universales. Mientras se realizaban los trabajos de descifrado de códigos llevados a cabo en Bletchley Park, se desarrolló un proyecto británico secreto: el primer ordenador digital verdaderamente programable. Durante los años inmediatamente posteriores a la guerra, Turing continuó trabajando en el diseño de ordenadores, primero en el National Physical Laboratory (NPL) y luego (a partir de 1948) en la Universidad de Manchester. Fue el ímpetu que los trabajos de descifrado de códigos, y, en particular, la contribución de Turing, dieron al desarrollo de los ordenadores lo que hizo que el sueño de Lewis Fry Richardson relativo a la predicción numérica del tiempo llegara a ser una realidad cuando este científico vivía todavía, y ese mismo ímpetu fue lo que llevó a Edward Lorenz al redescubrimiento del caos.

Sin embargo, la visión de Turing iba ya mucho más lejos. El año académico 1947-1948 lo pasó trabajando en Cambridge, destinado allí por el NPL, y escribió un trabajo, que en vida de él no llegó a publicarse, sobre lo que llamaríamos actualmente redes neuronales, un intento de demostrar que cualquier sistema mecánico suficientemente complejo podía aprender de la experiencia, sin que en realidad lo programe ninguna inteligencia exterior. En 1950, cuando ya se había establecido en Manchester, estaba en condiciones de comenzar a aplicar los conocimientos que había adquirido sobre sistemas mecánicos y ordenadores electrónicos a los sistemas biológicos y al cerebro humano.

El salto de aquí a su trabajo sobre el desarrollo de los embriones no era tan grande como podría parecer, ya que Turing no sólo estaba interesado en cómo crecen los cerebros y forman conexiones; su interés por el modo en que la diversidad de sistemas vivos se desarrolla desde unas sencillas condiciones iniciales se había visto

estimulado en su juventud por la lectura del clásico libro de D'Arcy Thompson titulado *On Growth and Form*[49]. Por lo tanto, cuando Turing fue elegido miembro de la Royal Society, en 1951, por sus contribuciones a la informática, estaba ya trabajando en lo que probablemente habría sido, si él hubiera vivido, una aportación científica aún mayor.

Ni siquiera Turing podía saltar directamente desde los conocimientos de biología que existían a principios de la década de 1950 hasta un modelo del modo en que el cerebro desarrolla su red de conexiones, después de todo, la estructura de doble hélice del ADN, la molécula de la vida, no se descubrió hasta 1953, cuando Francis Crick y James Watson, que trabajaban en Cambridge, la detectaron. Lo que Turing hizo fue abordar el problema fundamental relativo al modo en que emerge la estructura en el desarrollo del embrión, desde lo que es inicialmente una gola casi esférica de células, prácticamente no diferenciadas, lo que se llama el blastocito, que se forma a partir del óvulo fecundado.

En términos matemáticos, se trataba de un problema de simetría rota, un fenómeno que en otros contextos (por ejemplo, la convección de Bénard) ya resultaba familiar para los físicos. Un buen ejemplo de ruptura de la simetría es el que se produce cuando calentamos y luego enfriamos ciertos tipos de sustancias magnéticas. Algunos materiales magnéticos, como el hierro, pueden considerarse formados por un conjunto de diminutos dipolos, como pequeños imanes de barra. Por encima de la temperatura crítica, conocida como el punto de Curie (en honor de Pierre Curie, que descubrió este efecto en 1895), hay suficiente energía calorífica para romper cualquier vínculo magnético que se forme entre los dipolos, de tal forma que los imanes puedan girar libremente y se desordenen de una manera aleatoria, apuntando a todas las direcciones, con lo que no habrá campo magnético global. En términos de magnetismo, se puede decir que el material posee una simetría esférica, porque no hay una dirección magnética preferente.

Cuando la temperatura cae por debajo del punto de Curie (760 °C para el hierro), las fuerzas magnéticas que actúan entre dipolos adyacentes superan la tendencia de su decreciente energía calorífica a desordenarlos, y los dipolos se alinean para producir un campo magnético global, con un polo norte en un extremo y un polo sur

[49] Hay versión castellana: *Sobre el crecimiento y lo forma,* Editorial Hernann Blume, Madrid, 19X0, trad. de Juan Manuel Ibeas Delgado.

en el otro. La simetría original se ha roto. Este cambio se llama transición de fases, y es similar al modo en que el agua se congela y se convierte en hielo en una transición de fases que se produce a 0 °C. El concepto de transición de fases tiene también en la física de partículas unas aplicaciones importantes que no necesitamos explicar aquí; el aspecto relevante es que, aunque tales ideas no se habían aplicado con amplitud dentro de la biología antes de 1950, en aquella época era natural que un matemático interesado por la teoría del desarrollo biológico pensara en términos de ruptura de la simetría y dispusiera de las herramientas matemáticas que describen la naturaleza general de tales transiciones.

En 1952 Turing publicó un trabajo que explicaba, en principio, el modo en que la simetría de una mezcla química inicialmente uniforme podía romperse de manera espontánea por la difusión de diversas sustancias químicas a través de dicha mezcla. El propio título de la publicación, «The Chemical basis of morphogenesis»[50], dejaba clara la importancia que iba a tener esto para la biología, y la propuesta de Turing consistía en decir que algo parecido al proceso que él describía matemáticamente podría suceder en realidad en el embrión durante su desarrollo, para producir pautas donde nada de esto existía originalmente.

A primera vista, la propuesta de Turing parece contradecir totalmente lo que podríamos intuir. Sería de esperar que la difusión mezclara las sustancias y destruyera las pautas, no que las creara allí donde no existía pauta alguna inicialmente. El ejemplo más obvio es el modo en que una gota de tinta que se echa en un vaso de agua se extiende hasta dejarlo lleno de una mezcla de agua y tinta; casi parece que Turing sugiere una inversión de los procesos termodinámicos que operan a esta escala, como si el tiempo transcurriera hacia atrás y la mezcla uniforme de agua y tinta se separara para dejar un vaso en el que una solitaria gota de tinta está rodeada por el agua limpia. Pero no era éste el caso. La clave de la teoría de Turing es que el proceso en el cual se forman las pautas necesita al menos dos sustancias químicas entre las cuales se produce una interacción mutua.

Todo ello depende del proceso conocido como catálisis, en el cual la presencia de una sustancia química concreta (el catalizador) propicia que tenga lugar una determinada reacción química. En algunos casos, la presencia de un compuesto

[50] Philosophical Transactions of the Royal Society, 1952, volumen B237, p 37; actualmente este trabajo está considerado como uno de los que más han influido en todo el campo de la biología teórica.

químico (que podemos representar mediante la letra A) dentro de una mezcla de sustancias químicas favorece reacciones en las que se produce una cantidad mayor de dicho compuesto. Esta reacción se denomina autocatalizadora y, dado que cuanto mayor sea la cantidad de A que existe, mayor será la cantidad de A que se produce, podemos ver aquí otro ejemplo de retroalimentación positiva que se produce en un proceso no lineal.

Por otra parte, hay sustancias químicas que actúan produciendo el efecto contrario, es decir, inhibiendo ciertas reacciones químicas. Lógicamente reciben el nombre de sustancias inhibidoras. Además, no se puede decir que una sola sustancia no pueda favorecer más de una reacción química al mismo tiempo. Turing dedujo que podían surgir ciertas pautas en una mezcla de sustancias químicas si el catalizador A no sólo fomentaba la producción de más A, sino también la de otro compuesto B, que era un inhibidor cuyo efecto consistía en frenar la velocidad a la que se producía A. La sugerencia decisiva que formuló Turing fue que, una vez que se hubieran formado A y B, se difundirían a través de la mezcla de sustancias químicas a velocidades diferentes, de tal modo que en algunas zonas de la mezcla habría más A que B, y en otras más B que A. Con el fin de calcular exactamente qué cantidad de A y B habría en distintos lugares, Turing tuvo que optar por las ecuaciones más sencillas que se podían utilizar, ya que los ordenadores electrónicos tenían todavía unas capacidades muy limitadas, y eran pocos los que se podían conseguir, por lo que tenía que resolver todo con lápiz y papel.

Esto significaba que debía trabajar con aproximaciones lineales que sustituyeran convenientemente a las ecuaciones no lineales que en realidad tenía que utilizar para describir la situación, y estas ecuaciones resultaron ser muy inestables, en el sentido de que un pequeño error en alguna parte de los cálculos producía más adelante un gran error. En consecuencia, Turing sólo pudo calcular lo que sucedía en los sistemas más sencillos, pero esto era suficiente para hacerse una idea de las posibilidades. El propio Turing reconoció que la investigación completa de sus teorías tendría que esperar hasta que se fabricasen ordenadores digitales más potentes, pero, al desarrollar sus ideas lo mejor que pudo, más allá de las que había esbozado en su publicación de 1952, demostró que la competición entablada entre A y B era la clave para la formación de pautas, y que resultaba esencial que B

tuviera que difundirse a través de la mezcla más rápidamente que A, de tal modo que, mientras la galopante producción de A mediante el proceso autocatalizador de retroalimentación es siempre un fenómeno local, la inhibición de A ocasionada por B es un fenómeno ampliamente difundido. También la rápida difusión de B desde el lugar donde se produce significa que no evita del todo la producción de A en su fuente.

Para hacemos una idea de lo que sucede, imaginemos una mezcla de sustancias químicas que se encuentra en reposo dentro de una jarra de cristal. A causa de las fluctuaciones aleatorias, habrá algunos lugares en el líquido en los que, existirá una concentración de A ligeramente mayor, y esto beneficiará la formación tanto de A como de B en esas zonas. La mayor parte de B se difundirá abandonando dichos lugares, evitando que se forme A en las zonas intermedias entre ellos, mientras el proceso autocatalizador garantiza que seguirá produciéndose más A (y más B) en los lugares iniciales. (También habrá lugares en la mezcla original donde las fluctuaciones aleatorias produzcan, para empezar, un exceso de B, pero, por supuesto, en ellos no sucederá nada interesante.) Supongamos ahora que la sustancia A tiene color rojo y la sustancia B tiene color verde. El resultado será que una jarra de líquido inicialmente uniforme y sin características especiales se transformará espontáneamente en un mar de color verde con manchas rojas que mantienen sus posiciones dentro del líquido (siempre que éste no se remueva, ni se chapotee en él). La pauta es estable, pero en este caso particular se trata de un proceso dinámico, en el que se producen nuevas cantidades de A y B, mientras haya una fuente de sustancias químicas a partir de las cuales puedan fabricarse, y siempre que exista un «sumidero» a través del cual puedan retirarse los productos finales. En la terminología que a estas alturas ya debe resultamos familiar, la pauta es estable y persistente, con tal de que el sistema que estamos manejando sea abierto y disipador, y se mantenga en un estado de no equilibrio. Turing describió también sistemas matemáticos en los que hay una pauta de color que se ondula por todo el líquido, donde será más evidente para cualquier observador (si tales sistemas pudieran reproducirse en experimentos reales) que se está realizando un proceso dinámico. Hoy en día, un compuesto autocatalizador, como A, se denomina accionador, mientras que B se conoce por el nombre de inhibidor; sin embargo, el

propio Turing nunca utilizó esos términos y se refirió a B llamándolo «veneno», cosa que ahora nos trae ecos espeluznantes de su propia muerte[51]. Aunque parece estar lejos de lo que es el desarrollo de un embrión (no digamos del de un cerebro), la cuestión fundamental en relación con el descubrimiento de Turing fue que aportó un modo químico natural de romper la simetría para crear pautas de manera espontánea en un sistema inicialmente uniforme, si existieran sistemas químicos reales que se comportaran de esta forma.

Por muy intrigantes que fueran las teorías de Turing, aunque su papel se considera de una importancia sumamente fecunda para la biología teórica en nuestros días, durante la década de 1950 y la mayor parte de la de 1960 suscitó poco interés entre los químicos y los biólogos, precisamente porque nadie sabía de ningún sistema químico real que se comportara del modo descrito por este modelo matemático. Es decir, nadie salvo una persona, el bioquímico ruso Boris Belousov, pero, como no era lector de *Philosophical Transactions of the Royal Society,* no supo nada sobre el trabajo de Turing, del mismo modo que Turing nunca tuvo noticias del trabajo de Belousov antes de su propia muerte prematura. Al comienzo de la década de 1950, Belousov estaba trabajando en el Ministerio de Salud soviético y su interés se centraba en el modo en que la glucosa se descompone en el cuerpo para liberar energía.

Pasaba ya de los cincuenta años, una edad inusualmente avanzada para que un científico realice un importante descubrimiento, y tenía la experiencia de haber trabajado en un laboratorio militar, cuestión sobre la cual disponemos de pocos datos, salvo que, antes de retirarse de esta actividad después de la Segunda Guerra Mundial, llegó a alcanzar el grado de *combrig,* equivalente más o menos al de coronel del ejército, lo cual constituía una distinción extraordinariamente alta para un químico. Como muchos otros procesos metabólicos, la descomposición de la glucosa que interesaba a Belousov se ve favorecida por la acción de las enzimas, diferentes tipos de moléculas proteínicas que actúan como catalizadores en distintos pasos de la serie correspondiente de reacciones químicas. Belousov ideó una mezcla

[51] Parece ser que Turing estuvo obsesionado con el veneno. Su biógrafo Andrew Hodges relata cómo Turing acudió a ver la película *Snow White and The Seven Dwarfs* (Blancanieves y los siete enanitos) en Cambridge, el año 1938, y quedó muy impresionado «por la escena en que la malvada bruja colgaba una manzana de un hilo y la sumergía en una hirviente cocción de veneno, murmurando: «Sumerge la manzana en la pócima hirviente. Que rezume la muerte durmiente». Según parece, a Turing le gustaba cantar este pareado «una y otra vez» mucho antes de que adaptara la acción a los versos.

de sustancias químicas que, según creía él, podía imitar al menos algunas características de este proceso, y se quedó atónito cuando vio que la solución que tenía delante iba cambiando de ser transparente e incolora a adquirir color amarillo y viceversa, con un ritmo regular y repetitivo. Fue como si se hubiera puesto a mirar un vaso de vino tinto, para acabar viendo cómo desaparecía el color del vino y luego volvía a aparecer, no una sola vez, sino muchas veces, como por arte de magia.

Esto parecía escapar a lo establecido en la segunda ley de la termodinámica, tal como ésta se entendía en la época. Sería del todo razonable que el líquido cambiara de incoloro a amarillo, si el amarillo indicara un estado más estable con mayor entropía. Y también sería del todo razonable que el líquido cambiara de amarillo a incoloro, si el hecho de ser incoloro indicara un estado más estable con mayor entropía. ¡Pero era imposible que *ambos* estados tuvieran cada uno de ellos una entropía mayor que el otro! Era como si, utilizando las teorías originales del siglo XIX sobre la relación entre la termodinámica y el tiempo, la propia flecha del tiempo estuviera invirtiéndose continuamente, haciendo virajes hacia atrás y hacia adelante en el interior del fluido.

Belousov se habría quedado menos atónito si hubiera conocido algún trabajo anterior que en cierto modo presagiara tanto sus propios experimentos, como los modelos matemáticos de Turing. Remontándonos a 1910, otro creador de modelos matemáticos, el austríaco Alfred Lotka (1880-1949), había desarrollado una descripción matemática de un hipotético sistema químico que oscilaba de este modo, produciendo primero un exceso de un compuesto, invirtiéndose luego para producir un exceso de otro compuesto, volviendo a invertirse de nuevo, y así sucesivamente[52]. Dando un claro ejemplo del modo en que ciertos procesos sencillos con retroalimentación pueden a menudo, describirse mediante las mismas leyes en circunstancias que, a primera vista, parecerían muy diferentes, el italiano Vito Volterra (1860-1940) demostró en la década de 1930 que las ecuaciones de Lotka funcionaban bastante bien como descripción del modo en que las poblaciones de peces cambian cuando existe una interacción entre una especie que es la presa y otra que actúa como depredadora, siguiendo un ciclo de auge y precariedad de la

[52] Estos sistemas oscilantes se conocen actualmente como «relojes químicos» por la regularidad de sus ritmos; pero esta regularidad es sólo relativa, y no son suficientemente precisos para ser utilizados como verdaderos relojes.

población cuando prolifera primero una de las especies y luego la otra. Ya en 1921 el químico de origen canadiense William Bray (1879-1946), que entonces trabajaba en la Universidad de California, en Berkeley, había descubierto que una reacción química en la que participaban agua oxigenada y yodato producía una mezcla de yodo y oxígeno en la cual las proporciones de los dos productos oscilaban más o menos del modo que Lotka había descrito. Aunque Bray hizo referencia al modelo de Lotka cuando anunció su descubrimiento, la respuesta de sus colegas fue básicamente que, puesto que los resultados contradecían la segunda ley de la termodinámica, debía haber algo erróneo en su experimento, de tal modo que el «descubrimiento» sería un efecto extraño causado por algún descuido en la mezcla y el seguimiento de los ingredientes. Lotka, Volterra y Bray habían fallecido ya todos ellos cuando Belousov halló casi la misma respuesta cuando intentaba publicar un informe en el que explicaba sus descubrimientos de 1951, un año antes de que Turing publicara su trabajo fundamental. Los hallazgos de Belousov, según dijo el editor de la revista a la que los presentó," contravenían la segunda ley de la termodinámica, por lo que su procedimiento experimental debía tener fallos.

Para hacernos una idea de cómo en aquellos tiempos la consideración que tenía dicha segunda ley de la termodinámica se aproximaba a la que se concedía a las Sagradas Escrituras, basta leer una famosa cita extraída de los escritos del astrofísico británico Arthur Eddington. En su libro *The Nature of the Physical World,* publicado por Cambridge University Press en 1928, escribía lo siguiente:

> *La ley según la cual la entropía siempre aumenta, la segunda ley de la termodinámica, tiene, a mi parecer, el rango más elevado entre todas las leyes de la naturaleza. Si alguien nos indica que nuestra teoría favorita del universo está en desacuerdo con las ecuaciones de Maxwell, pues peor para ellas. Si se descubre que la observación la desmiente, bueno, podemos decir que esos experimentalistas a veces son unos chapuceros. Pero, si resulta que nuestra teoría contradice la segunda ley de la termodinámica, no hay esperanza para nosotros; lo único que podemos hacer es cargar con la mayor de las humillaciones*.[53]

[53] Siguiendo el asesoramiento de un experto ajeno a la revista, en la mayoría de las publicaciones culturales, el

Sustituyendo la palabra «teoría» por «experimento» en la frase final, tendremos en esencia la respuesta que halló Belousov, era un «experimentalista chapucero» y tuvo que «cargar con la mayor de las humillaciones». Confieso que he citado previamente esta observación de Eddington con aprobación, pero sin detenerme a explicar que, aunque es verídica en lo que dice, la formulación original de la segunda ley de la termodinámica no es como él pensaba, la verdad última, y debe ser reconsiderada en situaciones de falta de equilibrio y allí donde intervenga la gravedad. Pero en 1951 aún se estaba lejos de entender esto.

La reacción de Belousov ante el rechazo que sufrió su publicación fue, quizá, lo que era de esperar en un hombre de su edad y su historial. Se sintió insultado personalmente y lo tomó como una afrenta, dada su pericia profesional como experimentador, por lo que decidió no trabajar más en aquel campo, si sus logros iban a ser despreciados de esa manera. Uno de sus jóvenes colegas, S. E. Shnoll, intentó animarle a perseverar, pero no lo consiguió. Después de pretender durante años, sin éxito, que le publicaran su obra, en 1959 Belousov logró meter en la imprenta, de tapadillo, un resumen de sus descubrimientos en dos páginas añadiéndolas a un informe que le iban a publicar sobre un tema completamente diferente, que había escrito para un simposio médico sobre las radiaciones, celebrado en Moscú el año anterior. Luego abandonó totalmente este trabajo.[54]

Las actas de este congreso (que no estuvieron sometidas a proceso alguno de control o aprobación por parte de un editor) se publicaron sólo en ruso; casi nadie las leyó fuera de la URSS, y tampoco allí tuvieron muchos lectores. Pero Shnoll siguió interesándose por la obra de Belousov y, en la década de 1960, la puso en conocimiento de uno de sus estudiantes de postgrado, Anatoly Zhabotinsky, animándole a seguir la trayectoria iniciada por Belousov. Sólo una persona de la siguiente generación de químicos, Zhabotinsky, se sintió incentivada para aprovechar el resumen de dos páginas realizado por Belousov, pero sólo una era necesaria para hacerse cargo del descubrimiento y hacer que el mundo científico le prestara atención.

«árbitro» solía revisar los trabajos presentados, para decidir si merecían ser publicados.

[54] Una traducción inglesa de aquel escrito original que había sido rechazado se publicó finalmente en 1985, en *Oscillations and Travelling Waves in Chemical Systems*, editado por R. J. Field y M. Burger, Wiley, Nueva York.

Zhabotinsky era un estudiante graduado de la Universidad del Estado de Moscú cuando se le dio a conocer el descubrimiento de Belousov y se sintió lo suficientemente intrigado como para intentar llevar a cabo la reacción por sí mismo (es difícil que alguien no se quede intrigado cuando su director de tesis le «sugiere» que examine un determinado tema de investigación), confirmando que esta reacción funcionaba del modo que había dicho Belousov, y haciendo a continuación algunos cambios en los ingredientes hasta que obtuvo una mezcla que mostraba un cambio de color mucho más drástico, ya que pasaba del rojo al azul y viceversa. No tendría que sorprendernos el hecho de que fuera un estudiante quien retomara la idea, ya que los jóvenes investigadores están, en general, menos constreñidos por la tradición que sus mayores y más dispuestos a considerar la posibilidad de romper con las leyes sacrosantas (aunque en la mayoría de los casos éstas resisten bien a sus presiones). Zhabotinsky expuso sus resultados en un congreso internacional que se celebró en Praga en 1968, donde los científicos occidentales oyeron hablar por primera vez sobre el intrigante comportamiento de lo que se dio en llamar la reacción de Belousov-Zhabotinsky, o reacción BZ. El tema causó aún más impacto porque, aunque algunos de ellos conocían ya los trabajos de Turing, como ya hemos visto, no habían pensado que pudiera ser relevante para los sistemas químicos reales. A diferencia de Turing, Belousov vivió lo suficiente para ver cómo su descubrimiento se aceptaba de esta manera, pero falleció en 1970, antes de que se llegara a valorar plenamente la importancia de estas reacciones químicas.

No es de extrañar que uno de los primeros que retomó el trabajo de Zhabotinsky y desarrolló un modelo teórico para describir el tipo de oscilaciones observadas en la reacción BZ fuera Ilya Prigogine, que había conocido a Turing en Inglaterra en 1952, poco después de que éste hubiera escrito su informe sobre la química de la creación de pautas, y había comentado con él ciertos aspectos del trabajo mencionado. Cuando ya estaba trabajando en Bruselas con su colega René Lefever, y partiendo del trabajo de Turing, antes de finales de 1968, Prigogine había desarrollado un modelo en el que intervenían dos sustancias químicas que se convertían en otras dos diferentes en un proceso de múltiples pasos en el que aparecían también dos productos intermedios durante un breve espacio de tiempo. El modelo se conoció como el *brusselator* (o bruselador); no necesitamos entrar en

detalles relativos a cómo funciona este modelo (aunque sólo son ligeramente más complicados que el modelo de Turing relativo a la forma de hacer manchas), pero la cuestión verdaderamente importante es que las reacciones se caracterizan por la retroalimentación y la no linealidad. Si nos imaginamos que los productos de la serie de reacciones son respectivamente rojo y azul, el *brusselator* nos dice que, mientras la mezcla se mantenga en un estado disipador, suficientemente lejos del equilibrio, con una aportación constante de materias primas que se le añaden y con retirada de productos finales, cambiará regularmente de rojo a azul, y viceversa, sin estabilizarse en el color púrpura uniforme que se esperaría desde el punto de vista ingenuo de la fe en la segunda ley de la termodinámica. De hecho, todo el proceso es coherente, si se tiene en cuenta que esta segunda ley debe ser modificada en la medida en que ha de adaptarse a las condiciones de equilibrio desarrolladas ya por Prigogine y sus colegas.

Durante la década de 1970 se avanzó tanto en la creación de modelos, como en la investigación de los sistemas químicos reales en los que la estructura se crea de manera espontánea en un proceso de autoorganización. En el campo experimental, los químicos no tardaron en encontrar procedimientos para hacer que unas ondas de colores se desplazaran a través de las mezclas de sustancias químicas, en la propia reacción BZ, dentro de una bandeja plana que contenga el potaje químico adecuado es posible producir círculos concéntricos y espirales en rojo y azul que se desplazan hacia el exterior desde la fuente donde se originan. Entre la enorme variedad de pautas que se desarrollaron mediante experimentos similares durante las décadas siguientes, los químicos hallaron por fin en la década de 1990 un modo de producir las pautas *estacionarias* de puntos exactamente iguales a las que había descrito Turing originalmente.

A principios de la década de 1970, los detalles químicos de la reacción BZ fueron objeto de investigaciones llevadas a cabo por un equipo de la Universidad de Oregón, que identificó al menos treinta clases de sustancias químicas distintas que participaban en el engranaje de las reacciones químicas que producen en general los cambios de color, incluidas las sustancias intermedias de corta vida, como las que aparecen en el *brusselator.* Esto les llevó, en 1974, a descubrir un modelo que describe los pasos clave del proceso BZ utilizando sólo seis tipos de sustancias

químicas que participan en interacciones mutuas en cinco pasos diferentes, incluida la importantísima influencia de la autocatálisis. La diferencia entre, por una parte, este modelo y, por otra, el modelo de Turing y el *brusselator* estaba en que, mientras los últimos se refieren a sustancias hipotéticas llamadas A, B, etc., el modelo del equipo de Oregón utiliza compuestos químicos reales que participan en reacciones químicas también reales. El modelo llegó a ser conocido con el nombre de *oregonator* (u oregonador). Una vez más, no es preciso que entremos en detalles; la cuestión es que lo que parece ser una complicada pauta de autoorganización se puede explicar hablando de unas pocas interacciones sencillas.

Pero, aún hay más. Realmente es posible dejar la mezcla BZ en un estado invariable y uniforme, si esperamos el tiempo suficiente sin añadir nuevos ingredientes, finalmente se estabilizará por sí misma.

Ahora bien, si añadimos algunos reactivos más, comenzará el comportamiento oscilatorio que hemos descrito. Se ha producido una bifurcación en la cual el sistema cambia de un estado de período 1 a un estado de período 2. Siempre podemos adivinar cuál viene a continuación. Si aumentamos gradualmente la velocidad a la que los nuevos ingredientes fluyen entrando en el sistema y se eliminan los «productos residuales», al llegar a un umbral crítico la pauta oscilatoria se complica más, presentando un ritmo doble cuando el sistema se bifurca una vez más para entrar en un estado de período 4.

Si seguimos aumentando el flujo de reactivos, la cascada de duplicaciones del período, que ya conocemos por el caso del grifo que gotea y otros ejemplos anteriormente mencionados, aparecerá en toda su gloria, las pautas periódicas se volverán menos obvias y el sistema llegará al borde del caos (en este caso lo hará realmente de una manera bastante rápida, precipitándose todo ello después de alcanzar el período 4).

Todas las cosas interesantes que hemos estado explicando, especialmente la autoorganización y la aparición espontánea de pautas a partir de sistemas uniformes, se producen al borde del caos. Todo ello se puede explicar en el lenguaje del espacio de las fases, los ciclos límite y los atractores, exactamente igual que en los ejemplos que hemos comentado antes. Incluso hay indicios de la existencia de atractores extraños asociados a las trayectorias que en el espacio de las fases

describen la evolución de una reacción BZ[55]. Sin embargo, aunque está bien ver cómo todo esto se relaciona con la historia que hemos contado hasta ahora, parece que nos hemos desviado un largo trecho, dejando a un lado la esperanza que tenía Turing de proporcionar una explicación del proceso de morfogénesis, el aspecto de la embriología relacionado con el desarrollo de pautas y formas en el embrión que está creciendo. No es así; aunque se ha de demostrar todavía que sus teorías son una aportación fundamental al desarrollo global del embrión (en gran medida los trabajos están aún en fase de realización), hay un área en particular donde han tenido un éxito espectacular y claramente visible.

Este triunfo del mecanismo de Turing se refiere al modo en que se desarrollan marcas tales como rayas y manchas en la piel y el pelaje de los mamíferos y, más en general, a la manera en que se forman ciertas pautas en la superficie de otros animales. Una importante contribución a esta investigación fue la realizada por James Murray, inicialmente en la Universidad de Oxford y luego en la Universidad de Washington, en Seattle, que reunió muchos de sus descubrimientos en un artículo muy interesante publicado en *Scientific American* en 1988 con el título «How the leopard gets its spots»[56]; un informe completo, más técnico pero menos entretenido, relativo a este trabajo puede encontrarse en su libro *Mathematical Biology.* Murray descubrió que no sólo las manchas del leopardo, sino también las rayas de una cebra, las manchas de una jirafa, e incluso la carencia de pauta alguna en la piel de un ratón o de un elefante, pueden explicarse mediante el mismo y sencillo proceso, en el que interviene la difusión de las sustancias químicas accionadoras e inhibidoras en toda la superficie del embrión en un momento crucial de su crecimiento. Nadie ha demostrado hasta hora que éste sea definitivamente el modo en que funciona el proceso de creación de pautas, pero ciertamente las pautas que se forman son del tipo de las que se formarían si este proceso de creación estuviera funcionando. La teoría tiene mucho atractivo, en gran parte a causa de su sencillez. A un cierto nivel, en lo relativo al hecho de que el código del ADN especifica la construcción de cada uno de los cuerpos, el almacenamiento de la información que dice «libérense estas dos sustancias químicas en esta fase del

[55] Esto lo descubrió en 1983 un equipo de la Universidad de Texas, en Austin, en el que participaba Harry Swinney, que más tarde sería uno de los primeros que hicieron el modelo del «punto de Turing».
[56] Volumen 258, número 3, p. 80.

desarrollo» ocupa mucho menos espacio (menos memoria, si utilizamos la analogía del ordenador) que un plan original en el que se detalle con precisión la ubicación exacta de cada mancha y cada raya del cuerpo adulto. A otro nivel, disponer de un mecanismo sencillo que explique cómo y por qué aparecen pautas de distintos tipos en los cuerpos de animales diferentes, y también por qué algunos no tienen ninguna pauta, ahorra mucho más esfuerzo que tener un proyecto diferente para describir cada tipo de pauta en cada uno de los distintos animales.

Finalmente, como veremos, el sencillo mecanismo que propuso inicialmente Turing y discutieron con todo detalle Murray y sus contemporáneos ofrece unas perspectivas interesantes a la hora de examinar los mecanismos de la evolución. No es *siempre* cierto en las ciencias que la solución más sencilla de un problema tenga que ser necesariamente la correcta, pero este planteamiento, conocido como la navaja de Ockham[57], ha demostrado ser un método empírico sumamente fiable en la mayoría de los casos y, ciertamente, siempre es aconsejable elegirla solución más sencilla, salvo que haya razones de peso para no hacerlo. En este caso, el proceso de Turing es la solución más sencilla para este problema.

Las pautas que observamos en la superficie corporal de los mamíferos son colores de la piel o colores que han adquirido los pelos mientras crecían en una zona particular de la piel. En cualquiera de estos casos, es la presencia de algo existente en la piel lo que determina el color. Es sorprendentemente escasa la cantidad de colores diferentes con los que se obtiene la gama completa que aparece en las marcas de un gato carey, negro, blanco, marrón y toda una gama de colores situados entre el naranja y el amarillo. Los colores dependen de la presencia o la ausencia de dos pigmentos producidos por las células de la piel, dependiendo la intensidad del color de la cantidad de cada pigmento que esté presente, la eumelanina, que da un color negro o marrón, y la feomelanina, que da un color amarillo o naranja (la ausencia de ambas melaninas deja el pelo o la piel de color blanco).

Pero ¿qué decide si ciertas células se «accionan» para producir uno u otro tipo de melanina y cuánta van a producir? El triunfo de Murray fue demostrar que, aunque no sepamos todavía con precisión qué sustancias químicas intervienen, las pautas

[57] Según el nombre del lógico y filósofo inglés William of Ockham (c. 1285-1349), que decía lo siguiente: «Los entes no deberían multiplicarse más allá de lo estrictamente necesario».

que observamos en animales reales vivos son exactamente (y *únicamente)* las que se generarían mediante reacciones de Turing en las que se produjera la difusión de un accionador y un inhibidor a través de la superficie del embrión al principio de su desarrollo, a pocas semanas de su concepción (en la cebra, por ejemplo, hay pruebas de que las pautas de la piel se establecen aproximadamente entre 21 y 35 días después de la concepción, en un período total de gestación de 360 días).

Si la presencia de una de estas sustancias químicas pusiera en marcha también la capacidad de una célula para producir melanina, el resultado sería que una pauta equivalente a la observada en las versiones del recipiente plano de la reacción BZ quedaría impresa de manera invisible en estas células, pero sólo se manifestaría más tarde, cuando alguna otra sustancia desencadenante (o el crecimiento del pelo) enviara un mensaje con la orden de comenzar a producir melanina, un mensaje que recibirían todas las células de la piel, pero sólo sería ejecutado por aquellas células que hubieran sido preparadas para ello durante la reacción de Turing.

Por lo tanto, sin preocuparse demasiado por los procesos bioquímicos que están en juego, «todo» lo que Murray tuvo que hacer fue desarrollar un modelo matemático que se pudiera utilizar para predecir el modo en que las pautas se formarían como resultado de la reacción de Turing en superficies configuradas como las de los embriones de los mamíferos en distintas etapas de su desarrollo. Dado que este proceso incluye ondas que se desplazan por (o a través de) las superficies, tanto el tamaño como la forma de éstas afectan a la pauta producida por la reacción.

Como señala Murray, la situación es superficialmente bastante similar al modo en que los sonidos producidos por la piel tensada de un tambor dependen del tamaño y de la forma de ésta, porque las ondas sonoras de distintos tamaños (es decir, de diferentes longitudes de onda, correspondientes a las diversas notas musicales) encajan claramente en pieles de distintas medidas (en realidad existe una analogía matemática bastante estrecha entre ambos sistemas, aunque los procesos físicos son muy diferentes). Murray descubrió que si la superficie es muy pequeña en el momento en que se pone en marcha el proceso de Turing, no se puede formar pauta alguna. No hay espacio suficiente para que el mecanismo se ponga a funcionar o, si preferimos pensarlo en estos términos, la «longitud de onda» asociada con la reacción es mayor que el tamaño de la piel, por lo que la pauta no

puede verse (sería como intentar pintar finos detalles en un pequeño lienzo utilizando un rodillo de pintar paredes). En el otro extremo, si las superficies son relativamente grandes, las interacciones llegan a ser demasiado complicadas para permitir que surjan pautas globales de cualquier tipo.

Es como si se estuvieran desarrollando distintas conversaciones al mismo tiempo en una habitación, produciendo un efecto global que no es sino una algarabía uniforme. De hecho, en grandes superficies es posible una «creación de pautas» dentro de una escala de finos matices; si se examina a una distancia suficientemente corta, veremos que no todos los pelos que hay en la superficie corporal de, por ejemplo, un elefante tienen exactamente el mismo color, pero a mayor distancia el color de un elefante parece uniforme, del mismo modo que en una pintura puntillista existe una estructura muy fina de puntos que se funden en un color uniforme cuando se ven desde una posición algo alejada (y del mismo modo que, en una habitación llena de personas que charlan, podemos captar lo que dice quien está a nuestro lado, a pesar del ruido de fondo). Por lo tanto, según el modelo, tanto los mamíferos pequeños como los muy grandes deberían carecer de pautas en sus superficies corporales, y esto es exactamente lo que vemos en la naturaleza. ¿Qué sucede entre estos extremos?

Comenzando por los pequeños y aumentando gradualmente el tamaño, resulta que el primer tipo de pauta que puede formarse es una que presenta franjas bastante anchas, luego otra de rayas, seguida de la de pequeñas manchas, pasando luego a la de grandes manchas separadas por estrechas franjas que forman las fronteras entre las primeras, mientras que en superficies mayores las grandes manchas se fusionan para dar un color uniforme. Las pautas generales que se producen se parecen claramente a la gama de pautas presentes en la naturaleza, desde las manchas del leopardo hasta las rayas de un tigre o una cebra, para ir después a las grandes manchas de una jirafa. Murray lo resume en su libro de la siguiente manera:

Vemos que existe una sorprendente similitud entre las pautas que genera el modelo y las que se observan en una amplia variedad de animales. Incluso con las restricciones que impusimos a los parámetros para realizar nuestras simulaciones, la riqueza de las pautas posibles es notable. Las pautas dependen estrechamente de

la geometría y la escala del dominio de reacción, aunque el posterior crecimiento puede distorsionar la pauta inicial... Resulta atractiva la idea de que un solo mecanismo pueda generar todas las pautas observadas en la piel o el pelo de los mamíferos.

Pero, por supuesto, un animal no ha de tener *necesariamente* una pauta, aunque exista un mecanismo bioquímico que le permita tenerla. Siempre es posible desconectar el mecanismo y hay razones evolutivas muy obvias por las cuales, por poner un caso, un oso polar ha de tener un color blanco uniforme. Sin embargo, entre los casos en que las pautas han evolucionado, con respecto a la correlación entre el tipo de pauta y el tamaño de la superficie, puede verse un ejemplo muy claro en las colas de muchos miembros de la familia de los felinos.

En el caso de las colas que son más o menos cilíndricas, las pautas observadas pueden ser, o bien manchas, o bien franjas o rayas circulares que rodean la cola. Pero cuando las colas se afilan en el extremo final, como la del jaguar, aun en el caso de que la base de la cola esté cubierta de manchas, la punta está marcada por franjas de rayas, de acuerdo con la predicción del modelo según el cual las franjas siempre se forman en las superficies pequeñas, mientras las manchas aparecen en superficies mayores.

Sin embargo, una de las características fundamentales del modelo es que el tipo de pauta que se forma sobre la superficie de un animal no depende del tamaño y de la forma del adulto, sino del tamaño y de la forma del embrión en el momento en que está funcionando el proceso de Turing. Está claro que existe alguna correlación con el tamaño del adulto, porque desde muy poco tiempo después de la concepción los embriones de elefante tienden a ser más grandes que los de ratón en la misma etapa de su desarrollo; sin embargo, la importancia del tamaño del embrión queda muy bien ilustrada mediante las diferencias que presentan las rayas de dos tipos de cebras, *Equus burchelli* y *Equus grevyi.*

La primera tiene menos rayas, pero más anchas que las de la segunda, lo que hace que sean claramente diferentes cuando se ven una al lado de otra, a pesar de que los adultos son más o menos del mismo tamaño. Contando el número de rayas que hay en cada caso, y teniendo en cuenta el modo en que la pauta se ha distorsionado

durante el crecimiento del animal, en la década de 1970 J.B.L. Bard demostró que la pauta observada en la *burchelli* debe quedar establecida en el embrión cuando éste tiene veintiún días, mientras que la pauta que presenta la *grevyi* quedaría formada en el embrión cuando éste tiene ya cinco semanas.

Figura 4.2. Las pautas que aparecen en el cuerpo de un animal como consecuencia de los procesos químicos autoorganizados que tienen lugar durante el desarrollo del embrión (véase el texto) están relacionadas con el tamaño del animal. Los procesos de difusión química producen rayas en superficies de menor tamaño y manchas en otras mayores.

Esto se sabía antes de que Murray encontrara su modelo matemático del efecto de Turing, pero las diferencias se ajustan exactamente a las predicciones del modelo, correspondiendo las rayas más anchas a una difusión más temprana del accionador y del inhibidora través de la superficie del embrión. El modo drástico en que se combinan la genética y el entorno (la «naturaleza» y la «nutrición») quedó claro a principios del año 2002, cuando se difundió la noticia del nacimiento del primer gato clonado. Dado que la pauta según la cual se disponen los colores en los animales multicolores resulta de una combinación de su herencia genética con lo que sucede

en el útero (incluida, por ejemplo, la cantidad de alimento que recibe el embrión durante su desarrollo), el gatito no presenta en su piel exactamente la misma pauta de marcas que su madre, ni siquiera cuando ambos animales tienen el ADN idéntico.

Esto nos lleva también a considerar la importancia que tiene todo ello para nuestro conocimiento de la evolución. Las diferencias visibles entre las pautas de las dos especies de cebras se producen simplemente cambiando el momento en que el efecto de Turing actúa sobre el embrión. Hasta donde llegan nuestros conocimientos, en este caso concreto no hay ventajas evolutivas en ninguna de las pautas (no todas las características anatómicas han de estar relacionadas con la adaptación). Pero si hubiera alguna ventaja por tener rayas más estrechas (o más anchas), quizá porque proporcionaran un camuflaje mejor, es fácil ver cómo las variaciones naturales de un individuo a otro podrían proporcionar la materia prima para responder a la presión de la selección y desplazar toda la población de una especie de cebra en esa dirección, sin que cambiara nada, salvo la *programación temporal* de un suceso determinado durante el desarrollo del embrión, una de las «mutaciones» mínimas imaginables. Tendremos más que decir sobre la evolución, mucho más, en el resto de este libro.

Sin embargo, para aclarar la situación, vale la pena explicar unos pocos ejemplos más relacionados con el modo en que la forma externa parece estar controlada por unos sencillos procesos químicos. Dimos algunos detalles sobre el modelo de Murray relativo al modo en que los mamíferos adquieren sus marcas, en parte porque fue uno de los primeros modelos de este tipo, y en parte porque se deduce de forma natural a partir de nuestro comentario sobre el mecanismo de Turing y la reacción BZ; pero no será necesario entrar en tantos detalles al hablar de estos otros modelos, ya que funcionan basándose en principios muy similares.

Sin dar preferencia a nuestro último comentario sobre la evolución, deberíamos quizá resumir primero brevemente el proceso darwiniano de selección natural, a causa de la gran confusión que existe en cuanto a la relación entre evolución y selección natural. La evolución es un hecho, un proceso que vemos funcionar lauto en los seres que actualmente viven como en el registro fósil[58].

[58] Mis ejemplos favoritos sobre la evolución tal como está actuando hoy en día es el modo en que los

Asimismo, son también hechos el modo en que las manzanas caen de los árboles y la Luna describe su órbita alrededor de la Tierra. El movimiento de las manzanas y de la Luna (y de otras cosas) se explica mediante la teoría de la gravedad. La teoría de Newton es adecuada para la mayoría de los fenómenos a nivel humano, pero la teoría de Einstein es la apropiada si estudiamos fenómenos extremos, tales como la destrucción de las estrellas. La evolución se explica mediante una teoría, la teoría darwiniana de la selección natural, que es adecuada para la mayoría de los fenómenos que se dan actualmente en la Tierra, pero que, como la teoría de la gravedad de Newton, se han modificado para tener en cuenta otras cosas que eran desconocidas para su creador.

La teoría de Darwin es a la evolución lo que la teoría de Newton es a la gravedad, pero todavía no tenemos un equivalente evolucionista de la teoría de la gravedad de Einstein, es decir, una teoría completa de la evolución que vaya más allá de la teoría de Darwin para incorporar en ella todos los hechos que conocemos actualmente, aunque se han presentado varias sugerencias sobre el modo en que debe ser modificada la teoría darwiniana. La parte esencial de la teoría de Darwin sigue siendo el núcleo de nuestra forma de entender la evolución. Afirma que los descendientes se parecen a los progenitores, pero de una manera imperfecta, de tal modo que existen variaciones de una generación a la siguiente y también entre individuos de una misma generación. Los que están mejor adaptados a su entorno (los más «apios», en el sentido del modo en que una pieza encaja en un rompecabezas) se las arreglan mejor para encontrar alimento, aparearse y (lo que es decisivo) reproducirse, por lo que dejan más descendencia para la generación siguiente.

Dado que estos descendientes se parecen a sus progenitores, heredarán los atributos particulares que dieron el éxito a sus padres, pero quizá con ligeras variaciones. Y así va sucediendo en posteriores generaciones.

Este es un proceso de selección natural, que en cada generación selecciona al más apto (en el sentido del rompecabezas, no en el sentido atlético, aunque ambos aspectos pueden ir juntos). En el ejemplo clásico, si es una ventaja llegar a las hojas jugosas que hay en las cimas de los árboles, los herbívoros dotados de cuellos

---

«supermicrobios» han evolucionado, desarrollando una resistencia a los antibióticos, y cómo las plagas de insectos han desarrollado una resistencia a los insecticidas.

largos lo harán mejor en la lucha por la supervivencia y criarán más descendientes, mientras que los individuos que tienen cuellos cortos conseguirán menos alimento y tendrán más dificultades para criar a sus hijos. Siempre que la tendencia a tener cuellos largos o cortos sea hereditaria, aunque haya otras variaciones entre los individuos (algunos cuellos son más largos o más cortos que la media, sea ésta cual sea), durante muchas generaciones, la presión de la selección a favor de cuellos más largos llevará a la evolución de una criatura como la jirafa.

La variedad no es solamente la sal de la vida, sino que está en el núcleo fundamental del modo en que ésta funciona. La teoría darwiniana explica a las mil maravillas por qué las especies deben estar extraordinariamente bien adaptadas para vivir en sus nichos ecológicos, encajando en ellos como una llave en la cerradura. Sin embargo, es importante darse cuenta de que también permite la existencia de características que son neutras desde un punto de vista evolutivo y en cuanto a la selección, ya que no suponen ventajas ni inconvenientes para los individuos que las poseen. Entre estas características están incluidas variaciones tales como las pautas de rayas que se ven en las conchas de algunos caracoles de huerta, donde miembros de la misma especie pueden mostrar aspectos mucho más diferentes entre sí que los que se observan en una cebra *burchelli* y una *grevyi,* aunque ambas pertenecen a especies diferentes. Con esto ya tenemos bastante sobre fundamentos de la evolución, para continuar ahora con lo nuestro.

A partir del trabajo de Murray sobre las rayas y manchas que forma el mecanismo de Turing, tanto el propio Murray como otros científicos han investigado con el mismo método muchas más pautas de la naturaleza. Entre estas investigaciones, una de las más importantes para la historia que estábamos contando fue la realizada por Hans Meinhardt, que trabaja en el Instituto Max Planck de Biología Experimental, en Tubinga, y su colega André Koch. Utilizando un planteamiento similar al de Murray, pero basado en el mecanismo de la reacción BZ, en vez de la reacción de Turing, descubrieron que en su modelo matemático podían producirse pautas muy parecidas a las reales (incluidas las correspondientes a las manchas del leopardo) cuando se estimulaba la producción de un accionador en lugares aleatorios de la piel del embrión en el momento adecuado durante su desarrollo. La ventaja de este modelo reside en que es capaz de producir pautas más

complicadas, aunque la química en que se basa sigue siendo muy sencilla. Las pautas que se observan en las conchas de las criaturas marinas corresponden también a las que esperaríamos que se produjesen mediante ciertos procesos químicos en los que participarían compuestos accionadores e inhibidores, y muchos biólogos creen haber descubierto una especie en la que pueden ver cómo funciona un proceso de este tipo.

En el caso del pez ángel *Pomacanthus imperator,* el animal adulto tiene rayas paralelas que van de la cabeza a la cola, a lo largo de todo el pez. Cuando este animal crece, se forman nuevas rayas de tal forma que (odas siguen teniendo el mismo tamaño, manteniéndose también iguales los espacios que las separan. Las nuevas rayas se desarrollan a partir de unas bifurcaciones existentes en algunas de las rayas anteriores, que se ramifican como cuando la vía del ferrocarril se bifurca en una serie de puntos para convertirse en dos vías paralelas.

En la década de 1990 Shigeru Kondo y Rihito Asahi, que trabajaban en la Universidad de Kyoto, desarrollaron un modelo matemático que reproduce exactamente esta pauta de comportamiento, utilizando el mecanismo de Turing. Esto sugiere que el proceso de Turing sigue funcionando en estos peces adultos, es decir, no se trata de un suceso puntual que se produce durante el desarrollo embrionario, por lo que surge la esperanza de que pronto puedan ser identificadas las sustancias químicas que realmente intervienen en el proceso.

Se han utilizado modelos similares para imitar las pautas producidas en las alas de las mariposas, que es sólo uno de los muchos casos posibles, pero lo vamos a elegir por la importancia que tiene para lo que vendrá más tarde. Murray estudió los mecanismos que actúan para producir muchas de las características de estas pautas de las alas, como es la formación de grandes manchas que tienen un parecido superficial con unos ojos (se cree que estas manchas han sido seleccionadas por la evolución, ya que un depredador que eche un vistazo rápido a la mariposa podría ver, no un jugoso pedacito de comida, sino los ojos de una criatura de mayor tamaño que le está mirando). Un aspecto de este trabajo de Murray mostraba lo fácil que es formar estas pautas a partir de unos procesos químicos básicos, sin necesidad de tener una planificación complicada almacenada en los genes de la mariposa y, por consiguiente, la gran probabilidad de que una característica de este

tipo pueda ser seleccionada por la evolución.

Pero, en este caso particular, el modelo mostraba también que una característica fundamental de la pauta de la mancha en forma de ojo, su radio, cambia gradualmente si las condiciones en que se producen las reacciones químicas correspondientes cambian también de una forma gradual. Por ejemplo, las mariposas que se desarrollan a una temperatura determinada tendrán manchas de ojos de un determinado tamaño, mientras que las que se desarrollan a otra temperatura tendrán las manchas de un tamaño diferente, y existirá una correlación continua y uniforme entre el tamaño de las manchas y la temperatura.

Es interesante mencionar esto porque no es la única pauta de comportamiento que puede darse como resultado del tipo de reacciones químicas básicas que hemos estado comentando. En algunos procesos, los pequeños cambios que se dan en el entorno en que tienen lugar las reacciones producen poco efecto hasta que se alcanza un punto crítico, y es entonces cuando el proceso entra en una nueva fase de operación. Un ejemplo citado por Murray en relación con un proceso que presenta esta sensibilidad es el desarrollo de las extremidades de los vertebrados. Si los procesos bioquímicos que tienen lugar en la etapa de desarrollo en que crecen los dedos sufre una leve interrupción, el resultado no es una mano de tamaño ligeramente mayor o menor, sino una mano con un sexto dedo, listo puede surgir de forma natural a causa de una mutación (un pequeño cambio en el código del ADN de un gen, quizá como resultado de un error de copia) que produce un ligero cambio en el desarrollo; entonces el ADN alterado se transmitirá a las generaciones siguientes, salvo que la mutación sea perjudicial.

Esta es la razón por la que tales rasgos suelen transmitirse en las familias, como pasaba en la de Ana Bolena, una de las esposas de Enrique VIII, que nació con seis dedos en una mano, aunque le corlaron rápidamente uno de ellos. En otro ejemplo famoso entre los biólogos experimentales, un hombre de Boston tenía una especie de doble mano, sin pulgar y con siete dedos dispuestos en dos grupos, uno de tres y otro de cuatro, cada uno de ellos al lado del lugar donde «debería» haber estado el pulgar. Estructuras similares pueden producirse también injertando células de una extremidad en otra[59]. Además, pueden describirse de manera matemática

[59] Lewis Wolpert y sus colegas llevaron a cabo en Londres experimentos importantes en esta misma línea, utilizando

mediante los modelos adecuados, modelos que describen procesos tales como el mecanismo de Turing, que actúa en sistemas disipadores al borde del caos. La cuestión que hay que tener presente en todo esto es que, a veces, unos pequeños cambios en el entorno, o unas pequeñas mutaciones, pueden tener grandes efectos en la forma corporal que se está desarrollando. Este es uno de los aspectos relativamente nuevos que fueron desconocidos para Darwin y que nos ofrece una visión más clara de cómo funciona la evolución. Pero recordemos que en otras circunstancias lo que se produce es un cambio gradual. Como dice Murray:

Dependiendo del mecanismo y de la característica específica de la formación de pautas que estemos considerando, podemos tener un cambio gradual o discontinuo en la forma... está claro que, para comprender cómo se produce la evolución, debemos conocer los procesos morfogenéticos que intervienen en ella.

De hecho, esto es lo máximo que queremos profundizar en la morfogénesis. Todo lo que deseamos extraer de esta discusión es el conocimiento de que en la morfogénesis y en la biología evolutiva, a veces, unos pequeños efectos producen grandes cambios, y que esto se puede entender, en principio, en términos de modelos de formación de pautas en los que interviene una química muy sencilla. Además, resulta que unos pequeños cambios aleatorios pueden también tener efectos grandes o pequeños en el mundo en general, concretamente cuando se trata de sistemas disipadores al borde del caos, y que tener conocimiento de cómo y por qué sucede esto puede ayudarnos a comprender mejor la aparición de la vida y de la inteligencia.

---

polluelos.

Capítulo 5
Terremotos, extinciones y emergencias

Cuando oímos a los científicos referirse a los «sistemas complejos», se crea a veces una barrera, ya que para muchas personas «complejo» significa «complicado», y se supone automáticamente que, si un sistema es complicado, será difícil de comprender. Ninguna de estas suposiciones es necesariamente correcta. En realidad, un sistema complejo es tan sólo un sistema que está formado por varios componentes más sencillos que ejercen entre sí una interacción mutua.

Como ya hemos visto, desde los tiempos de Galileo y Newton los grandes triunfos de la ciencia se han logrado, en gran medida, descomponiendo los sistemas complejos en sus componentes simples y estudiando el modo en que se comportan esos componentes (en caso necesario, como primera aproximación, dando el paso suplementario de pretender que los componentes son aún más sencillos de lo que son en realidad). En el ejemplo clásico del éxito que ha logrado este planteamiento para conocer el mundo que nos rodea, buena parte de la química puede entenderse mediante un modelo en el que los componentes simples son átomos, y para esto importa poco de qué están formados los núcleos.

Ascendiendo un nivel, las leyes que describen el comportamiento del dióxido de carbono encerrado en una caja pueden entenderse pensando en unas moléculas más o menos esféricas que rebotan unas contra otras y contra las pare des de su contenedor, y poco importa que cada una de estas moléculas esté formada por un átomo de carbono y dos de oxígeno unidos entre sí. Ambos sistemas son complejos, en sentido científico, pero fáciles de entender. La otra clave para llegar al conocimiento, como ilustran estos ejemplos, consiste en elegir los componentes adecuados más sencillos para realizar el análisis; una buena elección nos proporcionará un modelo que tendrá amplias aplicaciones, del mismo modo que el modelo atómico se aplica a toda la química, no sólo a la del carbono y del oxígeno, y el modelo de las «bolas que rebotan» se aplica a todos los gases, no sólo al dióxido de carbono.

A un nivel más abstracto, el mismo principio básico se aplica a lo que los matemáticos optan por llamar números complejos. Este nombre ha asustado a

muchos estudiantes, pero los números complejos son en realidad muy sencillos, ya que sólo tienen dos componentes, lo cual no justifica en absoluto el uso del término «complejo». Los dos componentes de un número complejo son en sí mismos números corrientes, que se diferencian entre sí en que uno de ellos está multiplicado por una constante universal denominada *i.*

Así, mientras un número ordinario se puede representar mediante una sola letra (por ejemplo, X), un número complejo se representa mediante un par de letras (por ejemplo, $A + iB$). Resulta que i es la raíz cuadrada de -1, de tal modo que $i \times i = -1$, pero esto en realidad no importa. Lo que sí importa es que hay un conjunto de reglas extraordinariamente sencillas que nos dicen cómo manejar los números complejos, lo que sucede cuando se multiplica un número complejo por otro, o se suman dos números de este tipo, etc...

Verdaderamente estas reglas son sencillas, mucho más que, por ejemplo, las del ajedrez. Sin embargo, su utilización abre dentro de las matemáticas todo un mundo nuevo que tiene amplias aplicaciones en física, por ejemplo para describir el comportamiento de la corriente eléctrica alterna, así como en las ecuaciones de onda de la me canica cuántica.

Pero tenemos otro ejemplo más casero para ilustrar la sencillez de la complejidad. Las dos «máquinas» más sencillas que existen son la rueda y la palanca. Una rueda dentada, como el piñón de una bicicleta de carreras, es en efecto una combinación de la palanca y la rueda[60].

Una rueda, incluso un piñón, no es un objeto complejo. Pero una bicicleta de carreras, que no es fundamentalmente más que un conjunto de ruedas y palancas, es un objeto complejo, en el sentido científico del término, aunque sus componentes individuales, y el modo en que ejercen interacciones unos con otros, resultan fáciles de entender. Esto ilustra la otra característica importante de la complejidad, tal como se utiliza este término actualmente en la ciencia, la importancia del modo en que las cosas ejercen interacciones mutuas. Un montón de ruedas y palancas no sería por sí mismo un sistema complejo, aunque en ese montón estuvieran todas las piezas necesarias para fabricar una bicicleta de carreras. Las piezas simples han de ser conectadas unas con otras adecuadamente, de tal modo que ejerzan

[60] Como una rueda dentada de un cronómetro de cuerda, que sería un ejemplo igualmente válido para lo que se explica a continuación.

interacciones para producir un conjunto que es algo más que la suma de sus partes. Esto es precisamente la complejidad, basada en una profunda sencillez.

Cuando los científicos se enfrentan a la complejidad, su reacción instintiva es intentar comprenderla examinando componentes más sencillos y el modo en que se producen sus interacciones. Así esperan encontrar una ley sencilla (o varias) que puedan aplicarse al sistema que están estudiando. Si todo va bien, resultará que esta ley es aplicable también a un conjunto más amplio de sistemas complejos (como sucede con el modelo atómico de la química, o con el modo en que las leyes de las ruedas dentadas son aplicables tanto a las bicicletas como a los cronómetros), y que han descubierto una verdad profunda relativa al funcionamiento del mundo.

Este método ha dado buenos resultados durante más de trescientos años como guía para el comportamiento de sistemas próximos al equilibrio. En la actualidad se está aplicando a sistemas disipadores que se encuentran al borde del caos, y, ¿qué mejor ejemplo terrestre podría existir, en cuanto a un sistema en el que se disipan grandes cantidades de energía, que un terremoto?

Una de las preguntas más naturales que pueden plantearse sobre los terremotos es con qué frecuencia se producen terremotos de diferentes magnitudes. Aparte de su interés intrínseco, esto tiene una gran importancia práctica si vivimos en una zona propensa a sufrir terremotos, o si representamos a una compañía de seguros que intenta determinar qué primas se han de cargar en seguros de terremotos. Son muchas las maneras en que los terremotos se pueden distribuir a través del tiempo. Los terremotos, en su mayoría, pueden ser muy grandes, liberando enormes cantidades de energía que luego se tarda mucho tiempo en acumular de nuevo.

Otros pueden ser pequeños y liberar energía de una forma casi continua, de tal modo que nunca es suficiente para desencadenar un gran seísmo. Podría existir algún tamaño característico de los terremotos y que fueran relativamente raros los fenómenos más intensos o más débiles (del mismo modo en que se distribuyen las alturas de las personas, en torno a algún valor medio).

O podrían ser completamente aleatorios. No tiene sentido intentar adivinarlo; el único modo de averiguarlo consiste en examinar todos los registros relativos a terremotos y calcular cuántos se han producido en cada nivel de intensidad. La

primera persona que realizó esta tarea de una manera adecuada fue Charles Richter (1900-1985), el creador de la escala de su mismo nombre que se utiliza hoy en día para medir la intensidad de los seísmos[61].

La escala de Richter es logarítmica, de tal modo que un aumento de una unidad en la escala corresponde a un aumento de la cantidad de energía liberada del orden de 30 veces; un terremoto de magnitud 2 es treinta veces más potente que uno de magnitud 1; un terremoto de magnitud 3 es 30 veces más potente que uno de magnitud 2 (y, por consiguiente, 900 veces más potente que un terremoto de magnitud 1), y así sucesivamente.

Aunque esta escala lleva únicamente el nombre de Richter, éste la elaboró a principios de la década de 1930 junto con su colega Beno Gutenberg (18891960) y, a mediados de la década de 1950, el mismo equipo se dedicó a investigar la frecuencia de terremotos de distintas magnitudes. Examinaron registros de terremotos que se habían producido en todo el mundo y los agruparon en «compartimentos» correspondientes a pasos de media unidad en la escala de Richter—es decir, todos los terremotos cuya magnitud estaba comprendida entre 5 y 5,5 se incluyeron en un compartimento, todos los de magnitud comprendida entre 5,5 y 6 estaban en el siguiente compartimento, y así sucesivamente.

Teniendo en cuenta que la escala de Richter es logarítmica, para comparar valores tomaron el logaritmo de cada uno de los números obtenidos en este recuento[62]. Cuando trazaron una representación gráfica correspondiente al logaritmo del número de terremotos incluidos en cada compartimento en relación con los valores de la magnitud (lo que se llama un «gráfico log-log»[63]), descubrieron que era una línea recta. Hay una cantidad enorme de terremotos de baja intensidad, muy pocos de gran intensidad, y el número intermedio queda, para cualquier magnitud que tomemos, sobre la línea recta que une estos dos puntos extremos.

---

[61] La denominación «escala de Richter» se aplica en la actualidad a las versiones modernas de esta escala, que difieren ligeramente de la original de Richter, pero estas diferencias no son importantes para lo que estamos comentando aquí.

[62] «Logaritmo» es otra palabra que asusta a todos los tienen aversión a las matemáticas (causada generalmente por una mala enseñanza). Es simplemente un modo de expresar números en términos de potencias de diez, o potencias de algún otro número elegido como base. Por ejemplo, 100 es 10^2, por lo que el logaritmo decimal de 100 es 2. Éste es todo el misterio de los logaritmos, y lo único que hay que recordar es que los logaritmos son un modo de expresar números mediante el exponente de una potencia.

[63] Hacer un gráfico del logaritmo, en vez de hacerlo directamente con los números, representándolo en el eje vertical del gráfico, significa sencillamente que los intervalos iguales marcados en el eje como 0, 1,2, 3, etc., representan I, 10, 100, 1.000, etc.

Esto significa que el número de terremotos correspondiente a cada magnitud sigue una ley potencial, por cada 1.000 terremotos de magnitud 5 hay aproximadamente 100 terremotos de magnitud 6, 10 de magnitud 7, y así sucesivamente.

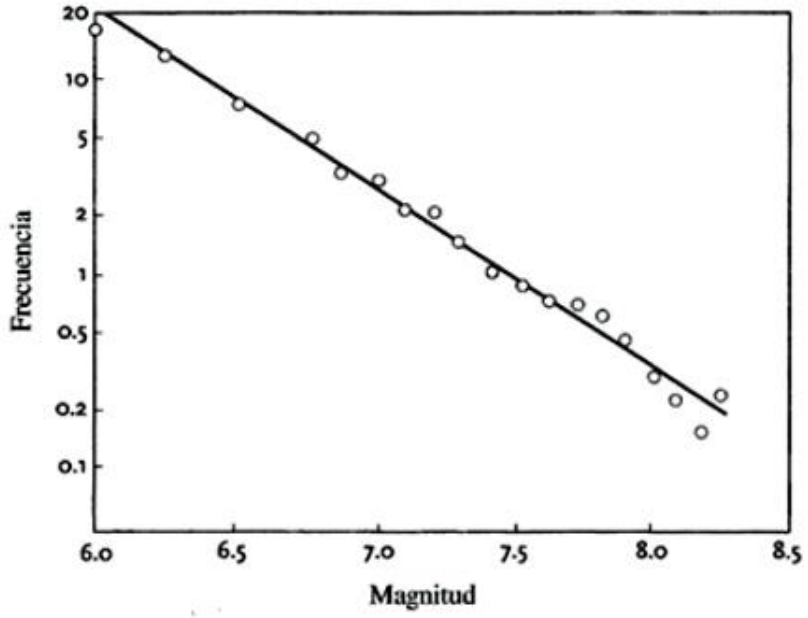

Figura 5.1. La ley de Gutenberg-Richter relaciona las magnitudes de los terremotos con la frecuencia con la que se producen. Es un ejemplo de ley potencial que pone de manifiesto un comportamiento de las magnitudes inverso a la frecuencia.

En la actualidad esto se conoce como ley de Gutenberg-Richter; es un ejemplo clásico de una ley sencilla que subyace a lo que a primera vista parece ser un sistema complejo. Pero, ¿qué significa exactamente? ¿Tiene esto aplicaciones amplias?

En primer lugar, hay que insistir en lo poderosa que es esta ley de la naturaleza. Un terremoto de magnitud 8, un poco menos intenso que el famoso terremoto de San Francisco de 1906, libera 20.000 millones de veces la energía de un terremoto de magnitud 1, que corresponde al tipo de vibración que percibimos en nuestra casa cuando pasa un camión pesado por la calle. Sin embargo, en toda esta amplia gama de energías es aplicable la misma ley, que además es sencilla. Está claro que esto nos dice algo fundamental sobre el modo en que funciona el universo. Parte de ese algo está relacionado con lo que aprendimos sobre el caos y los fractales en el

capítulo 3, donde también se perfilaban leyes potenciales. De hecho, si nos parece bien, podemos tomar uno de los ejemplos arquetípicos de un fractal que se observa en la naturaleza, la longitud de una línea costera, y lo podemos expresar exactamente del mismo modo. Un ejemplo especialmente adecuado es el que nos ofrece la costa de Noruega[64], donde grandes fiordos se ramifican para formar fiordos de menor tamaño, que a su vez hacen lo mismo, y así sucesivamente.

Para describir esto en términos de dimensión fractal de la línea costera, vamos a imaginar que disponemos de un mapa muy detallado de la costa (o incluso que lo hacemos sobre la costa real) y lo cubrimos con una red de cuadrados idénticos. Si el cuadrado es lo suficientemente grande, con uno bastará para cubrir toda Noruega. Sin embargo, a medida que tomemos cuadrados de menor tamaño, necesitaremos cada vez más cuadrados para cubrir cualquier entrante sinuoso de la línea costera. Es obvio que si hacemos cuadrados que tengan la mitad del tamaño, nos hará falta más del doble para esta tarea.

La velocidad a la que crece el número de cuadrados que necesitamos, comparada con la velocidad a la que se reduce el tamaño de cada cuadrado, nos da la dimensión fractal de la línea costera. En este caso la dimensión fractal resulta ser 1,52.

Esto se calcula trazando el gráfico del logaritmo de la longitud de costa medida en cada nivel (para cada tamaño de cuadrado) frente al logaritmo del tamaño de cada cuadrado. Se obtiene una línea recta, y la dimensión fractal es la pendiente que tiene esta línea del gráfico.

Aparte del valor de la pendiente (el grado de inclinación de esta línea en el gráfico), se trata exactamente del mismo tipo de ley potencial que la de Gutenberg-Richter. Parece existir una relación entre la geometría fractal de la costa de Noruega (y otras líneas costeras) y la frecuencia con que se producen terremotos de distintas intensidades en todo el mundo (y, de hecho, en zonas sísmicas de todo el globo).

Parte de esta relación se basa en una propiedad de los fractales con la que ya estamos familiarizados, son independientes de la escala. Ésta es la razón por la que el gráfico de la línea costera obedece una ley potencial, y lo que podemos deducir es que el hecho de que se produzcan terremotos también es invariable con respecto

[64] Véase J. Feder, Fractales.

a la escala.

Figura 5.2a. La longitud de la línea costera de Noruega se puede calcular cubriéndola con una red de cuadrados, tal como se explica en el texto. Cuanto más pequeños son los cuadrados, más larga es la línea costera que se mide. Cuando esta relación se representa en un gráfico log-log, se obtiene una línea recta, según una ley potencial. En este caso la pendiente de la línea (su potencia) es una medida de la dimensión fractal de la costa de Noruega, que resulta ser 1,52, casi exactamente a medio camino entre una línea (dimensión 1) y una superficie (dimensión 2).

Aunque esto resulte interesante para satisfacer nuestra curiosidad, lo que significa en la práctica es que no hay diferencias esenciales entre un gran terremoto y otro pequeño, salvo por su intensidad. *No* es necesario apelar a ningún efecto físico especial, raro o peculiar para explicar por qué se producen grandes terremotos, simplemente suceden—; se presentan con una frecuencia mucho menor que la de

los pequeños, pero se producen fundamentalmente por el mismo proceso físico que los de menor intensidad, y la ley potencial nos dice esto, aunque no tengamos ni idea de cuáles son las causas físicas de un terremoto.

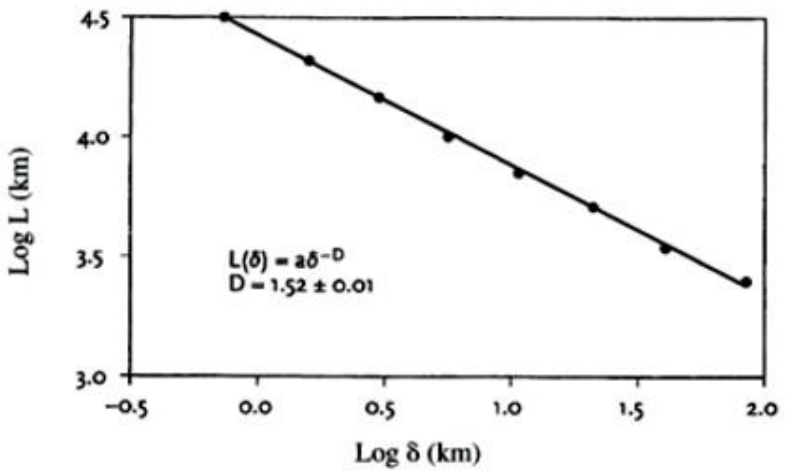

Figura 5.2b. (Véase el pie de la ilustración anterior.)

Esto difiere bastante de esa idea, ampliamente difundida entre los legos e incluso entre algunos geólogos, según la cual los grandes terremotos requieren grandes causas desencadenantes. Esta idea sugeriría que los grandes terremotos (como el de San Francisco en 1906) sobrevienen después de haberse creado en la corteza terrestre una gran tensión, que luego «ocasiona» la aparición de un punto débil. Una reconfortante consecuencia de esto sería que después de un gran terremoto habría que esperar mucho tiempo hasta que se produjera el siguiente en el mismo lugar. La estadística, y la ley potencial en particular, dice otra cosa. 'Panto los terremotos grandes como los pequeños se producen aleatoriamente, pero (y esto es muy importante) con frecuencias diferentes.

En cualquier momento podría desencadenarse un terremoto de cualquier magnitud en una zona sísmica y, del mismo modo que la probabilidad de obtener cara al lanzar una moneda sigue siendo 1 entre 2 aunque los tres últimos lanzamientos hayan dado cruz (suponiendo que la moneda no esté trucada), asimismo las probabilidades de otro terremoto como el de 1906 no eran más (ni menos) en 1907 o en 1905. Las leyes potenciales siempre significan que aquello que describe la ley

es invariable con respecto a la escala, por lo que los terremotos de cualquier intensidad están gobernados en todos los casos por las mismas leyes.

I lay un bonito ejemplo de invariabilidad de la escala en un sistema físico sencillo, que, a pesar de ser conocido a través de muchos relatos, no por eso deja de ser adecuado, Mark Buchanan lo expresa de una manera especialmente clara en su libro *Ubiquity,* donde intenta audazmente aplicar estas ideas a lo que él llama «la verdadera ciencia de la historia». Imaginemos un experimento en el que se lanzan patatas congeladas contra un muro, de lo cual resulta que quedan hechas añicos.

Si se hace esto es porque el modo en que se rompen las patatas congeladas es muy similar al modo en que se rompen trozos de roca cuando chocan entre sí, y las observaciones relativas a este proceso de fragmentación nos pueden dar una idea de, por ejemplo, la manera en que unos trozos de roca chocan entre sí en el espacio y se fragmentan para formar los asteroides que dan vueltas alrededor del Sol en un cinturón que se encuentra entre las órbitas de Marte y Júpiter. Los fragmentos de las patatas tienen tamaños muy diferentes, observándose muchos trozos diminutos, unos pocos pedazos grandes y otros de tamaño intermedio. Resulta sencillo clasificar estos trozos en compartimentos según su peso, del mismo modo que Gutenberg y Richter clasificaron los terremotos según su intensidad. Al principio ignoraremos todos los fragmentos realmente diminutos, dejándolos a un lado, y clasificaremos todos los demás.

Luego podremos hacer un gráfico del número de fragmentos que hay en cada compartimento, confrontándolos con el peso correspondiente a cada uno de estos intervalos. Se obtiene una ley potencial. A continuación, examinamos los diminutos fragmentos de patata que antes habíamos ignorado y utilizamos algún tipo de tecnología más precisa para medir sus pesos y repartirlos en compartimentos de la misma manera, pero a una escala menor. Se obtiene la misma ley potencial. En experimentos llevados a cabo en la Universidad del Sur de Dinamarca a principios de la década de 1990, los investigadores descubrieron que es aplicable la misma ley potencial para analizar de este modo trozos de patata cuyo peso oscila desde 10 gramos hasta una milésima de gramo.

Esto significa que, si tuviéramos el tamaño de una hormiga y nos arrastráramos entre los fragmentos de las patatas, el tipo de paisaje que veríamos sería (desde un

punto de vista estadístico) exactamente el mismo que si tuviéramos el tamaño de una mariquita que se mueve entre los mismos restos. No sería, por ejemplo, como comparar las suaves colinas de Sussex con las escarpadas cimas de la cordillera del Himalaya. El «paisaje» que ofrecen los trozos de patata tiene el mismo aspecto a cualquier escala.

Resulta que el paisaje de la Luna también es invariable con respecto a la escala, por lo que se refiere al número de cráteres de diferentes tamaños que dan relieve a su superficie. Dado que se produjeron por el impacto de rocas de distintos tamaños procedentes del espacio, y que estos fragmentos de roca a su vez eran, casi con toda certeza, el resultado de colisiones entre asteroides de mayor tamaño, que se rompieron según las mismas leyes que rigen la fragmentación de las patatas congeladas, todo esto no sorprende; pero extiende la escala de esta ley potencial concreta desde las patatas hasta los planetoides.

Hay otro modo de describir todos estos tipos de variaciones (en realidad, como ya hemos visto, diferentes versiones del *mismo* tipo de variación). Los grandes «acontecimientos» (terremotos, alargamientos de líneas costeras, fragmentos de patata) son más raros, y esto se puede expresar diciendo que la frecuencia de un suceso es igual a 1 dividido por alguna potencia de su magnitud. A la inversa, se puede decir que la magnitud de un acontecimiento es proporcional a 1 partido por alguna potencia de su frecuencia, ya que la potencia exacta no es especialmente importante; en general se llama a esto «ruido 1 partido por *f*», y se escribe ruido 1/*f* Puede parecer una extraña manera de referirse a los terremotos, pero la razón por la que se hace así es que el ruido 1/*f* aparece en muchos tipos de sistemas que varían de forma natural, y ya se había estudiado matemáticamente antes de que entraran en escena los expertos en terremotos. Para lo que estamos tratando aquí, los términos «comportamiento según una ley potencial» y «ruido 1/*f*» son sinónimos.

Este tipo de ruido contiene variaciones en todas las escalas, que van desde un rápido parpadeo hasta una pulsación lenta, todo superpuesto lo uno sobre lo otro. Como conocedor de la astronomía, para mí el clásico ejemplo de ruido 1/*f* es la variación en la luz emitida por un quásar, que cubre toda la gama de fluctuaciones desde el parpadeo a una escala de minutos hasta pulsaciones en períodos de años y

décadas;[65] este mismo tipo de pauta puede observarse en la luz procedente de algunas estrellas.

Cuando el parpadeo de un objeto de este tipo se representa en un gráfico, en cuyos ejes figuran los valores del brillo y del tiempo (los astrónomos lo llaman la «curva de la luz»), obtenemos una línea irregular de dientes de sierra que parece el perfil de una cordillera, con ascensos y descensos que se producen en todas las escalas. Esto es el ruido 1/*f*. En un extremo, el contraste directo con el ruido 1/*f* sería lo que llamamos ruido blanco, que es completamente aleatorio, y, en el otro extremo, el contraste se produciría con una señal pura que contiene una sola frecuencia, como una nota musical aislada.

Como se puede deducir a partir de los nombres utilizados, estos fenómenos se estudiaron por primera vez en acústica, rama de la física en la cual las señales que se investigan son ondas sonoras; el ruido blanco es el monótono silbido de interferencias que se oye en un receptor de radio de modulación de amplitud cuando movemos el selector y dejamos de sintonizar una emisora; un «ruido» de una sola frecuencia sería un tono puro (pero igualmente monótono) con una sola nota musical; y el ruido *1/f(que* a veces se llama también ruido rosa) resulta un sonido interesante para el oído humano. El sonido de la música tiene una estructura *1/f,* como el de las palabras habladas (incluida la algarabía que se oye en una fiesta). El ruido *1/f* contiene información, como en una cáscara de nuez.

Pero es necesaria una advertencia. Siempre que se realizan nuevos descubrimientos científicos, se produce un afán por subirse al carro del triunfador, lo cual hace que muchas personas intenten explicar todo mediante esa novedad que les entusiasma. El ruido 1/*f* se ha convertido en un tema candente dentro de la ciencia, y sus entusiastas lo utilizan para demostrar todo lo que tienen entre manos, cosa que puede abrirnos nuevas perspectivas, o puede poner de manifiesto las limitaciones de este campo (o ambas cosas a la vez). Pero está claro que el ruido *1/f no* es el único efecto que funciona en todos los fenómenos vinculados a la variación en el tiempo que se encuentran entre el ruido blanco y la frecuencia única. Por ejemplo, consideremos el ascenso de las temperaturas medias globales que se ha producido desde mediados del siglo XIX. Se parece mucho al tipo de curva de la luz que

[65] Y probablemente más largos, pero los quásares no se descubrieron hasta principios de la década de 1960, y los registros fotográficos de objetos de los cuales ahora se sabe que son quásares se remontan a la década de 1880.

interesa en la astronomía de los quásares, donde hay una tendencia global al crecimiento, pero con muchos picos y valles en un perfil de dientes de sierra. El análisis muestra que estas fluctuaciones son en realidad el ruido 1/*f*.

Podría ser que la tendencia creciente a largo plazo fuera sólo una parte de esta pauta global, una fluctuación del ruido a muy largo plazo. Pero en este caso, la tendencia a largo plazo se ha producido exactamente en la época en que las actividades humanas han estado liberando gases que retienen el calor en la atmósfera y, cuando las fluctuaciones a corto plazo alcanzan un promedio, la amplitud de este ascenso durante siglo y medio se corresponde exactamente con el calentamiento que sería de esperar por la cantidad de estos gases de invernadero que han sido añadidos a la atmósfera.

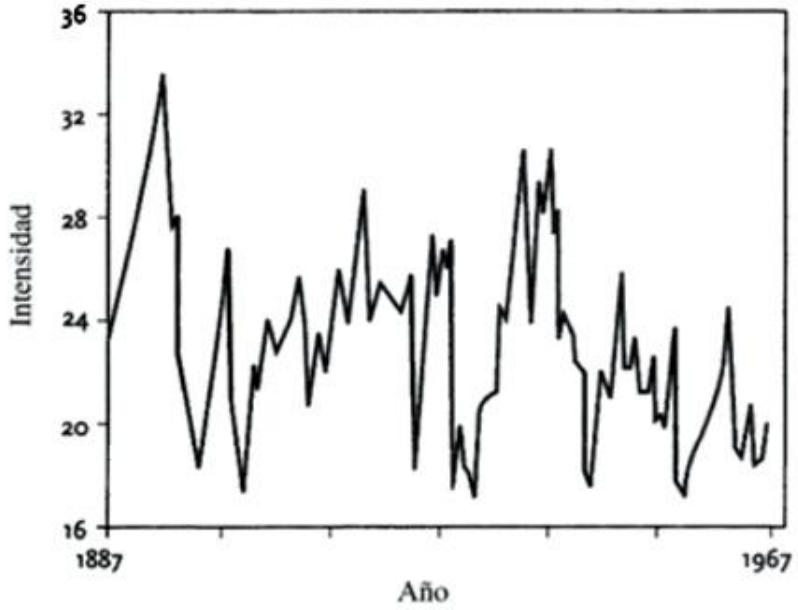

Figura 5.3. Los cambios en el brillo de la luz emitida por un quásar durante un intervalo comprendido entre 1887 y 1967 constituyen un ejemplo de ruido 1/f.

Negar aquí la relación existente entre causa y efecto tendría más o menos tanto sentido como negar que un terremoto que se produjo en el desierto de Nevada justo después de que una bomba nuclear explotara bajo la superficie estuviera causado por dicha bomba, y afirmar que fue sólo parte del ruido de fondo 1/*f*.

La conclusión es que, aunque hay mucho ruido en el sistema climático, y es en gran

medida ruido 1/*f* (lo cual es en sí mismo un descubrimiento interesante e importante), también existe en cuanto a la temperatura una auténtica tendencia ascendente vinculada a las actividades humanas. Lo que resulta especialmente interesante en relación con el hecho de que el ruido existente en el sistema climático sea ruido 1/*f* es, por supuesto, lo mismo que resulta interesante en relación con los terremotos que siguen una ley potencial. Significa que los sucesos de cualquier intensidad pueden ocurrir en cualquier momento, pero que los grandes sucesos (en este caso, grandes desviaciones con respecto a la temperatura media) se producen raras veces. Tradicionalmente, las compañías de seguros y los planificadores estatales han calculado las probabilidades de fenómenos extremos basándose en la historia.

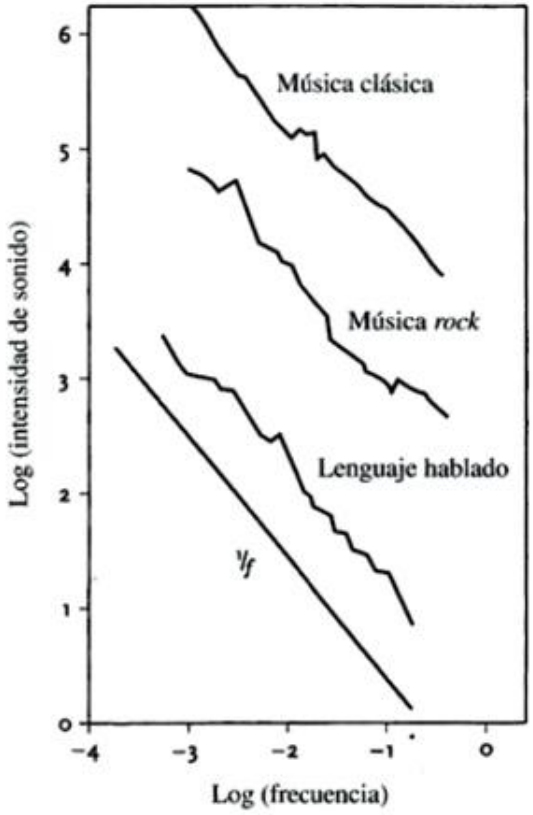

Figura 5.4. Utilizando un típico gráfico log-log, se observa que la música clásica, la música rock y las palabras habladas muestran todas ellas las características del ruido 1/f, porque todas contienen información.

Examinan con qué frecuencia se ha producido en el pasado un suceso de un tipo y una magnitud concreta (por ejemplo, una grave sequía), y pronostican que sucederá con la misma frecuencia en el futuro. Suelen hablar de cosas tales como una sequía «de las que suceden una vez cada cien años». Pero esto es sólo una media verdad. La señal 1/*f* nos dice que tales sucesos siguen realmente una ley potencial, y esto significa que, si durante un verano se produce una sequía «de las que suceden una vez cada cien años», la probabilidad de sufrir una segunda de igual gravedad el año siguiente es exactamente la misma que si no hubiera habido una sequía ese año.

No hay seguridad alguna para suponer que, después de una sequía «de las que suceden una vez cada cien años», vayan a pasar cien años hasta que se produzca la siguiente. Sin embargo, del mismo modo, el ruido 1/*f* implica que a una sequía de las que baten todos los récords le podría seguir una helada que batiera todos los récords. El año siguiente podría ser el más frío que se hubiera registrado jamás, sin que esto invalidara de ningún modo la hipótesis del efecto invernadero, gracias a una fluctuación aleatoria del sistema climático natural. Esta es la razón por la que es tan importante examinar las tendencias a largo plazo, y no sólo los sucesos aislados.

En la naturaleza hay muchos otros ejemplos de leyes potenciales y de ruido 1/*f,* pero no es necesario hacer aquí un catálogo de todos ellos. Sin embargo, hay dos que nos gustaría mencionar antes de volver a nuestro tema principal, que es la importancia de todo esto en la aparición de la vida, ya que estos ejemplos están mucho más cerca de la vida cotidiana que cosas como las curvas de luz de los quásares y las estadísticas de tendencias de la temperatura a largo plazo. La mayoría de los lectores de este libro habrá vivido en una gran ciudad en algún momento de su existencia, aunque ahora mismo viva en otro tipo de lugar.

Volviendo a la década de 1940, George Zipf, de la Universidad de Harvard, estudió las poblaciones de ciudades de todo el mundo. Obviamente, hay unas pocas ciudades que tienen un número muy grande de habitantes, y muchas otras con cifras de población menores. Lo que resultaba menos obvio en aquella época (aunque el lector probablemente ya habrá adivinado lo que viene ahora) es que, cuando en un gráfico log-log el número de ciudades que tienen una población

determinada se acota en intervalos confrontados con otro eje en el que se marcan las cifras de población, se obtiene una línea recta. Tanto si se trata del mundo en su totalidad, como si consideramos distintas regiones del globo, el modo en que las personas se congregan en las ciudades sigue una ley potencial, y la misma pauta es tan cierta hoy en día como lo fue en la década de 1940.

Todos tomamos una decisión sobre dónde vivir de manera independiente, ejerciendo nuestra libre voluntad; sin embargo, en cierto modo estamos sometidos a las mismas leyes que gobiernan la luz parpadeante de los quásares y la aparición de terremotos. Está claro que estamos tras la pista de una profunda verdad, y podemos empezar a comprender lo que sucede observando sistemas más sencillos, comenzando por uno que también incluye seres humanos.

Aunque mucha gente vive en ciudades, no todas ellas se ven afectadas por terremotos, que siguen siendo un fenómeno abstracto para la mayoría de nosotros; pero todo el que vive en una ciudad ha sufrido alguna vez un embotellamiento del tráfico, y resulta que los embotellamientos siguen también una ley potencial. El tipo de retenciones del tráfico a que nos referimos son las que se producen en una autopista abierta, en una carretera de dos carriles sin obstrucciones, aparentemente sin razón alguna.

Cualquier conductor (de hecho, cualquier pasajero que no esté dormido) sabe que cuando hay poco tráfico se circula de manera fluida y uniforme, pero, si el tráfico se hace demasiado denso en la carretera, se producen retenciones aunque parezca que no haya razón alguna por la que no se pueda seguir circulando a una velocidad uniforme, por ejemplo a ochenta kilómetros por hora.

El problema es que, si un vehículo frena por cualquier motivo (quizá sencillamente porque el conductor se ha acercado demasiado al vehículo que circula por delante), se produce un efecto en cadena cuando todos los que van detrás reducen la velocidad, y es más fácil frenar que acelerar (afortunadamente para la seguridad en carretera). Este tipo de comportamiento lo han estudiado unos investigadores de la Universidad de Duisberg que, a principios de la década de 1990, desarrollaron un modelo mediante ordenador para expresarlo sobre una base matemática[66].

El modelo no pretende describir el mundo real de las autopistas de varios carriles,

[66] K. Nagel y M. Paczuski, «Emergent Traffíc Jams», Physical Review, 1995, E51, p. 2909

pero (como todo buen modelo) es suficiente para dar una idea de lo que sucede. En este modelo sólo hay una fila de vehículos que se desplazan por una carretera de carril único, sin adelantamientos. Todos los vehículos («coches») son del mismo tamaño y la velocidad de cada uno de ellos se expresa como el número de longitudes de coche que puede recorrer entre un paso de la simulación por ordenador y el siguiente.

Si un vehículo se desplaza tan rápido que puede llegar a chocar con el de delante, ha de reducir su velocidad; si hay un gran trecho libre, aumentará su velocidad (hasta un determinado límite) para acercarse al coche que circula por delante. Finalmente, se cuantifica la capacidad de reducir y aumentar la velocidad, siendo más fácil frenar que acelerar. Todo lo que tiene el modelo es una fila de coches que se desplazan en fila y en línea recta, obedeciendo estas reglas. Sin embargo muestra todo el comportamiento de los atascos de tráfico reales, desplazándose los coches libremente cuando el tráfico es fluido, con algunos embotellamientos cuando el tráfico es denso, y con retenciones más o menos considerables cuando se alcanza una cierta densidad de vehículos.

Lo que es más, tanto en los modelos como en los estudios del tráfico real nos encontramos con una ley potencial, el número de embotellamientos de diferentes calibres (medidos según el número de vehículos implicados) sigue el mismo tipo de ley que el número de terremotos de diversas magnitudes. Al igual que los embotellamientos reales, los del modelo se propagan hacia atrás a través del tráfico, a medida que hay más coches que frenan tras la obstrucción, mientras los que van delante se alejan a toda velocidad, y, también como en el tráfico real, los embotellamientos se «disuelven» a medida que el propio atasco reduce la velocidad media del tráfico. Así como hay grandes y pequeños embotellamientos, también existen embotellamientos dentro de otros embotellamientos, siguiendo una pauta fractal. Todo esto constituye un ejemplo clarísimo del ruido $1/f$ en pleno funcionamiento.

Se puede extraer de todo esto dos lecciones prácticas. En primer lugar, no necesitamos un «desencadenante» extraordinariamente fuerte (como un accidente) para producir un gran embotellamiento[67]; un atasco de cualquier intensidad puede

[67] Aunque, por supuesto, algunos embotellamientos se producen a causa de accidentes o choques, al igual que

ser desencadenado incluso por una perturbación mínima, como un vehículo que se aproxima demasiado al de delante y frena bruscamente. Ningún vehículo choca con otro en ningún momento, no hay causa obvia que pueda ser visible a los ojos de las personas implicadas, pero se produce un trastorno que puede afectar a cientos de personas. En segundo lugar, si la densidad del tráfico aumenta, se puede mantener la circulación más fluida limitando la velocidad máxima, porque esto reduce el impacto de la diferencia entre el tiempo que se tarda en acelerar y el que se tarda en frenar. Realmente es cierto que si todo el mundo respeta los límites de velocidad cuando la autopista está abarrotada, todos llegarán a su destino antes que si intentan conducir más rápido.

Hay otra área de actividades humanas en la que las leyes potenciales y el ruido *í/f* pueden tener relevancia y que ha sido objeto de gran atención recientemente. Se trata de la economía, y, en particular, de la economía bursátil. No deseamos entrar aquí en detalles, porque eso nos llevaría demasiado lejos de nuestro tema principal, que es la aparición de la vida y su lugar en el universo[68]. Pero, ciertamente, vale la pena mencionar esta relación. En la década de 1960, poco antes de centrar su atención en los fractales, Benoit Mandelbrot examinó el modo en que los precios de productos básicos tales como el acero y el algodón habían experimentado variaciones en la Bolsa de Nueva York, y descubrió que las fluctuaciones de precios seguían una ley potencial, como el ruido *1/f* La deducción inmediata es que la economía constituye un sistema que sigue las mismas leyes que la aparición de terremotos (o de embotellamientos de tráfico), y los grandes acontecimientos (en este caso, sucesos tales como el hundimiento de los mercados bursátiles, por ejemplo, el que se produjo en octubre de 1987) pueden surgir de repente, como consecuencia de pequeños factores desencadenantes.

Este descubrimiento no fue del gusto de los economistas, a los que les gusta pensar que las economías pueden ser controladas por actuaciones gubernamentales, por ejemplo ajustando los tipos de interés. Sin embargo, si las fluctuaciones de la Bolsa siguen realmente una ley potencial, un pequeño ajuste de los tipos de interés podría (aunque sólo sea raras veces) ser el desencadenante de una gran fluctuación del

---

algunos terremotos se originan como consecuencia de pruebas nucleares y parte del calentamiento global de la atmósfera tiene por causa las actividades humanas.

[68] Si el lector desea saber más sobre este aspecto de la aplicación de las leyes potenciales, recomendamos los libros *Complexity,* de Mitchell Waldrop, o *Ubiquity,* de Mark Buchanan.

mercado, o podría no producir efecto alguno.

Sin embargo, el aspecto que debemos recordar es el mismo que se da en los casos de terremotos, embotellamientos de tráfico y cosas por el estilo. Globalmente la pauta puede ser el ruido 1/*f*, pero a veces la intervención del ser humano a escala suficientemente amplia producirá algún efecto, pocos dudan, por ejemplo, de que la Segunda Guerra Mundial dio a la economía de EE.UU. un impulso que la sacó de la depresión. La pregunta es la siguiente: ¿en qué punto entre un 0,25 por ciento de ajuste de los tipos de interés y una guerra mundial llega a ser importante la intervención gubernamental?

Parte de la respuesta a esta pregunta se encuentra probablemente en un trabajo impulsado en gran medida por el irlandés de nacimiento Brian Arthur, que trabajó primero (a finales de la década de 1970) en el International Institute for Applied Systems Analysis (IIASA), una fábrica de pensamiento situada en Austria, y posteriormente en la Universidad de Stanford (en la década de 1980). La economía clásica está basada en principios muy similares a los de la termodinámica clásica, y trabaja con sistemas (economías) que están cerca del equilibrio.

Esto incluye un concepto conocido como «rendimiento decreciente», que es una forma de retroalimentación negativa. Dicho en lenguaje vulgar, significa que, aunque podamos haber hecho una fortuna inventando algo nuevo y vendiéndolo, cuando ya lo tiene todo el mundo se hace más difícil venderlo y los beneficios disminuyen. Arthur, que tenía conocimientos de ingeniería, constató que lo contrario podía ser cierto, que podemos tener una retroalimentación positiva, o un rendimiento creciente. Una vez que hemos acaparado el mercado, todo el mundo tiene que comprar nuestro producto y ganamos enormes cantidades de dinero, actualmente el ejemplo más obvio es el de Bill Gates y Microsoft. Además, Arthur constató que ni siquiera es preciso tener el mejor producto. Todo el mundo sabe que los ordenadores Apple son técnicamente superiores a los que se basan en los programas de Microsoft (esto estaba muy claro en la década de 1980).

Pero su introducción en el mercado no fue tan efectiva al principio y, cuando algo se convierte en un estándar industrial, el efecto «yo también» garantiza a que genere enormes ventas y, en consecuencia, enormes beneficios.

Todo esto es muy lógico para un ingeniero o un físico, pero los economistas que han

tenido una formación clásica siguen peleando desesperadamente en la retaguardia contra estas ideas heréticas. En este momento no nos interesa ese debate, salvo para desarrollar la analogía existente entre el modo en que la termodinámica clásica difiere de la de los sistemas que no están en equilibrio, y el modo en que la economía clásica difiere de lo que también podríamos llamar economía de los sistemas que no están en equilibrio. Los economistas modernos como Arthur trabajan con sistemas dinámicos cambiantes, en los que aparecen retroalimentaciones positivas y a través de los cuales fluye una forma de energía (en este caso, el dinero).

Visto así, la historia resulta familiar. Gracias especialmente a la retroalimentación, las economías son en realidad sistemas autoorganizados que se encuentran al borde del caos, con todo lo que esto implica. La situación es bastante más complicada, y resulta más difícil que los árboles dejen ver el bosque, porque estamos mirando hacia afuera desde el interior de la economía, y los seres humanos somos una parte integral del sistema que tratamos de analizar. No obstante, incluso al nivel más simple no hay lugar a dudas en cuanto a que las fluctuaciones de la Bolsa se comportan como el ruido 1/*f*, y cualquiera que crea que los gobiernos controlan la economía vive en las nubes.

Pero volvamos a cosas más sencillas, cosas como la muerte de los dinosaurios y el significado de la vida. Los dinosaurios se extinguieron[69] hace unos 65 millones de años, como parte de una extinción masiva de la vida en la Tierra, algo tan claramente reflejado en el registro fósil que se utiliza para marcar el final de un período del tiempo geológico, el cretáceo, y el comienzo de otro, el terciario. Puesto que la «C» ya se ha utilizado como inicial en un contexto similar en relación con el período cámbrico, este marcador se suele denominar frontera K-T, con una «K» de *Kreide,* que es el nombre del cretáceo en alemán.

No fueron sólo los dinosaurios los que resultaron afectados, aunque son los que aparecen con mayor protagonismo en los relatos populares cuando se habla de este desastre. Alrededor del 70 por ciento de todas las especies que vivían en la Tierra a finales del cretáceo habían desaparecido a principios del terciario, lo cual indica que se trató realmente de una extinción «en masa» y explica por qué los geólogos y los

[69] No todos; se cree que los pájaros descienden de una o más líneas de dinosaurios.

paleontólogos utilizan la frontera K-T como un marcador importante en el registro fósil. Dadas las dificultades que plantea la interpretación de unas pruebas de tiempos tan remotos, y la lentitud con la que se acumulan los estratos geológicos, todo lo que podemos decir realmente sobre la velocidad a la que se produjo aquella extinción es que sucedió en menos de unas pocas decenas de miles de años o en unos cien mil años, pero en ningún caso durante muchos millones de años; sin embargo, esto se considera un cambio brusco en relación con la escala de tiempo utilizada en geología.

Las preguntas obvias que esto plantea son las mismas que surgen tras un gran terremoto, por qué sucedió, si podría suceder de nuevo y, en este caso, cuándo. En el caso del suceso K-T, hay un candidato muy adecuado para ser el desencadenante que hizo que la extinción se produjera hace 65 millones de años, en vez de suceder, por ejemplo, hace 60 o 55 millones de años. Los restos de un enorme cráter que data justo de entonces han sido descubiertos bajo lo que es actualmente la península de Yucatán, en México, y por lodo el mundo se han hallado estratos de hace 65 millones de años que contienen restos de iridio, un metal pesado que es raro en la corteza terrestre, pero del que se sabe que es un componente de algunos tipos de meteoritos.

Lo que se deduce de esto es que una gran roca procedente del espacio, quizá de unos 10 kilómetros de diámetro, chocó con la Tierra y causó una gran devastación por todo el planeta. La capa de iridio es tan delgada que tuvo que depositarse en menos de 10.000 años (quizá mucho menos), lo cual es coherente con la teoría de que el suceso K-T fue desencadenado en su totalidad, de manera más o menos instantánea, por un gran golpe que llegó del espacio.

No es difícil explicar por qué pudo suceder todo esto. La energía cinética contenida en tal impacto sería equivalente a la explosión de unos mil millones de megatoneladas de TNT y arrojaría al espacio unos detritos en forma de grandes bloques que se desplazarían siguiendo trayectorias balísticas (como las de los misiles balísticos intercontinentales) y volverían a entrar en la atmósfera por todo el globo terráqueo, difundiendo calor y aumentando la temperatura en todas las zonas. Se produciría un electo de calentamiento de 10 kilovatios por cada metro cuadrado de la superficie terrestre durante varias horas, un fenómeno descrito

gráficamente por Jay Melosh, de la Universidad de Arizona, como algo «comparable a un horno doméstico encendido para “asar a la parrilla”»[70]. A continuación, unas diminutas partículas de polvo lanzadas al interior de la parte superior de la atmósfera se extenderían alrededor de todo el planeta y, combinadas con el humo de todos los incendios desencadenados por el «asado a la parrilla», bloquearían el paso de la luz del Sol, causando la muerte de todas las plantas que dependían de la fotosíntesis y congelando temporalmente el planeta. No es necesario aquí entrar en más detalles, aunque vale la pena recalcar que toda esta situación se ha estudiado meticulosamente y se ha situado en su contexto geológico, por lo que no se trata de una teoría disparatada o mal construida; dicho esto, lo que queremos hacer es precisar que los impactos procedentes del espacio pueden ser negativos para la vida en la Tierra[71]. Pero no se acaba aquí la historia. En primer lugar, hay pruebas de que en otras ocasiones, especialmente hace unos 35 millones de años, la Tierra sufrió unos impactos parecidos procedentes del espacio sin que se produjera una extinción del calibre del suceso K-T. Aunque los factores desencadenantes tengan la misma magnitud, no siempre ocasionan sucesos de la misma magnitud. En segundo lugar, hay también algunas pruebas de que los dinosaurios y otras especies estaban ya en decadencia en los últimos dos millones de años del cretáceo. Los paleontólogos todavía discuten sobre la interpretación de esta evidencia y no está nada claro que esto fuera una decadencia terminal (los dinosaurios habían experimentado varios altibajos durante los 150 millones de años que pasaron vagando por la Tierra), mientras algunos dicen que no se encontraban en absoluto en decadencia.

Pero, ciertamente, es posible que el impacto de Yucatán (que definitivamente sucedió hace 65 millones de años, en la frontera K-T) hubiera sido «la gota que colmó el vaso» y desencadenara amplias extinciones en la Tierra, en un tejido vital que ya estaba en una situación dramática por otras razones (tales como, quizá, el cambio climático causado por la deriva continental). La cuestión es que cualquier suceso aislado podría ser un caso especial que por sí mismo no nos dice gran cosa sobre la causa fundamental de sucesos similares, o la probabilidad de su recurrencia, al igual que el estudio de un solo terremoto tampoco nos dice mucho

[70] *Nature,* 1990, vol. 343, p. 251.
[71] Para más detalles, véase nuestro libro *Fire on Earth.*

sobre los terremotos en general y la frecuencia con que se producen. Necesitamos examinar las pautas de varias extinciones a lo largo de la historia geológica con el fin de hacernos una idea clara de lo que sucede, y para averiguar si la muerte de los dinosaurios fue un acontecimiento especial o solo uno de tantos.

El suceso K-T es en realidad sólo una entre cinco catástrofes similares (en la medida en que afectó en aquella época a la vida en la Tierra) a las que los geólogos denominan en conjunto las «cinco grandes», y no es en absoluto la mayor. Cada una de ellas se utiliza como marcador entre períodos geológicos y todas han sucedido durante los últimos 600 millones de años. La razón por la que nos centramos en este pasado geológico relativamente reciente es que fue en esa época cuando los seres vivos desarrollaron por primera vez algunas características, tales como las conchas, que podían fosilizarse fácilmente, dejando rastros que pueden reconocerse en los estratos que se estudian en la actualidad.

Antes de aquella época (durante la larga era geológica conocida como precámbrico) había florecido la vida en los océanos durante casi cuatro mil millones de años, en forma de criaturas de una sola célula y de cuerpo blando que nos han dejado poco que estudiar. Sin embargo, hace alrededor de 600-590 millones de años, al comienzo del período geológico conocido como cámbrico (por lo tanto, en el precámbrico; el cámbrico es el primer período de la era paleozoica), hubo una explosión de vida que dio lugar a diferentes variedades de formas multicelulares, y luego a criaturas vivas poco más evolucionadas. Obviamente, cuanto más nos acercamos a la actualidad, conocemos cada vez más sobre las pautas cambiantes de la vida en la Tierra, y el tipo de conocimiento que nos interesa aquí, que abarca grandes extinciones de vida, no empieza a verse claro hasta después de concluir el precámbrico.

Tomándolas cronológicamente, las cinco grandes extinciones se produjeron hace unos 440 millones de años (que marcaron la frontera entre los períodos ordovícico y silúrico), hace 360 millones de años (entre el devónico y el carbonífero), 250 millones de años (entre el pérmico y el triásico), 215 millones de años (en la frontera entre el triásico y el jurásico) y 65 millones de años (en la frontera K-T). Hay muchas otras extinciones en el registro fósil, y también se utilizan a menudo como marcadores en el «calendario» geológico, por ejemplo, la frontera entre el

jurásico y el cretácico, hace unos 145 millones de años, se define también mediante una extinción.

Pero las cinco extinciones que hemos mencionado son las más importantes. La más espectacular de todas ellas es el suceso que tuvo lugar hace unos 250 millones de años, al final del pérmico. Barrió al menos el 80 por ciento, y posiblemente hasta el 95 por ciento, de todas las especies que vivían en nuestro planeta en aquellos tiempos, tanto en tierra como en los océanos, y lo hizo durante un intervalo de menos de 10.000 años. En conjunto se calcula que más de un tercio de lo das las especies que han vivido siempre en la Tierra han desaparecido en extinciones masivas. Sin embargo, dado que también se calcula que el 99 por ciento de todas las especies que han vivido en la Tierra se han extinguido, esto significa que son el doble las que han desaparecido en sucesos de menor importancia.

La cuestión que nos intriga es si las extinciones en masa son realmente acontecimientos especiales, de carácter diferente al de las extinciones de menor importancia, o si son el mismo tipo de suceso, pero a gran escala—¿son las extinciones de vida en la Tierra unos hechos cuya naturaleza es independiente de su magnitud, como los terremotos y todos los demás fenómenos que hemos descrito con anterioridad? La respuesta sincera es «no lo sabemos», pero hay bastantes evidencias como para intuir que ésta es una posibilidad muy real.

Disponemos de la mayor parte de estas evidencias gracias al meticuloso trabajo de Jack Sepkoski, de la Universidad de Chicago. Sepkoski ha reunido una enorme base de datos relativos a extinciones, combinando la información relativa a miles y miles de especies diferentes a partir de los registros que han sido publicados por otros investigadores y se encuentran disponibles en cualquier buena librería científica (si se dispone de la paciencia necesaria para llevar a cabo la consulta). Por supuesto, la cosa no es tan sencilla. Se ha de conocer bien la especie que se está estudiando (Sepkoski se centró en los invertebrados marinos) y hay que tener el sentido común necesario para distinguir las extinciones reales de los vacíos producidos por la dificultad de obtener restos fósiles, teniendo en cuenta también el hecho de que resulta más fácil encontrar restos en estratos más recientes.

Después de realizar toda esta tarea, Sepkoski pudo trazar un gráfico en el que mostraba cómo ha fluctuado durante los últimos 600 millones de años el nivel de

extinciones que se produjo en cada intervalo de cuatro millones de años. Las fluctuaciones globales del gráfico parecen claramente aleatorias, con largos intervalos en los que las variaciones son relativamente escasas y parece que sólo unas pocas especies se extinguían de vez en cuando, y otros que están salpicados de dramáticos sucesos en los que se produce un gran número de extinciones, las extinciones en masa que ya hemos mencionado.

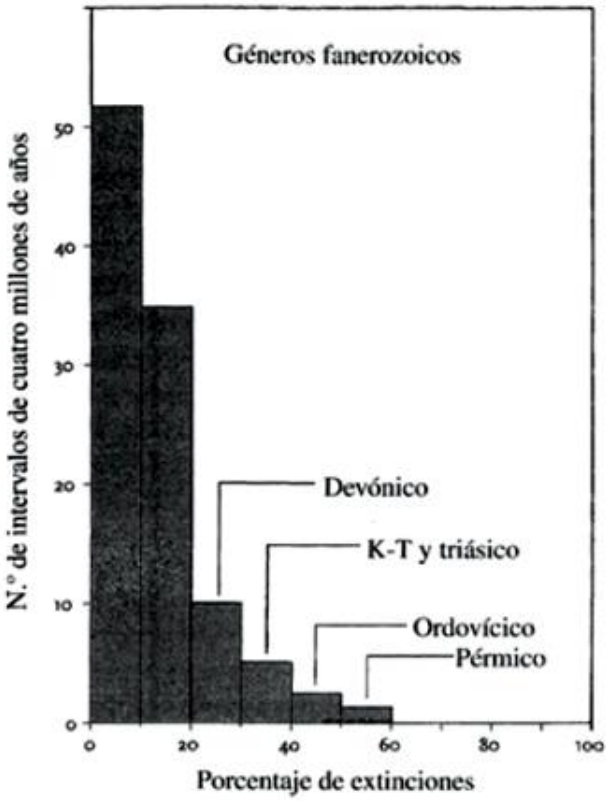

Figura 5.5. Los datos de Jack Sepkoski representados en un histograma para mostrar el número de intervalos de cuatro millones de años durante los cuales se produjo la cantidad de extinciones indicada. Más del 90 por ciento de las extinciones se producen en menos del 50 por ciento de los intervalos. No es una distribución uniforme.

Los porcentajes que aparecen en el gráfico, tomados de un trabajo de Sepkoski publicado en 1993, no coinciden exactamente con los antes mencionados porque se refieren a familias, no a especies, y sólo a algunas familias (los invertebrados

marinos), no a todos los seres vivos de la Tierra; pero la correspondencia es, sin embargo, bastante buena, y muestra, por ejemplo, que la «muerte de los dinosaurios» fue también la «muerte de los invertebrados marinos». La pregunta que se plantea es qué tipo de aleatoriedad es ésta, si realmente son sucesos aleatorios. Resulta que es una ley potencial, nuestro viejo amigo, el ruido 1/*f*. Una de las demostraciones de esto es la que hizo David Raup, que también trabajaba en la Universidad de Chicago, y «almacenó» los datos de Sepkoski de la manera habitual, añadiendo el número de intervalos de 4 millones de años en los que se extinguió hasta el 10 por ciento de los géneros importantes, el número de intervalos de 4 millones de años en los que se extinguió entre el 10 y el 20 por ciento de los géneros, y así sucesivamente.

El histograma resultante parece exactamente el equivalente del histograma realizado para los terremotos. Raup también ha analizado las pruebas mediante un histograma en el que el tiempo de vida de un género, tal como se pone de manifiesto en el registro fósil, se compara con el número de géneros que tienen ese tiempo de vida. Una vez más, la pauta sigue una ley potencial[72]. Exactamente la misma pauta se ha encontrado desde entonces en bases de datos fósiles mucho mayores analizadas por Michael Boulter, de la Universidad de East London, y sus colegas.

Ahora bien, no parece probable que todas las extinciones de vida que han sucedido en la Tierra hayan tenido como causa impactos procedentes del espacio. Lo que parece estar diciéndonos el registro fósil es que las extinciones se producen en todas las escalas, todos los tiempos, y que (como en el caso de los terremotos) puede producirse una extinción de cualquier magnitud en cualquier época. Algunas extinciones podrían ser desencadenadas por impactos de meteoritos; otras, por períodos glaciares. Sin embargo, la otra lección que podemos extraer de lo que conocemos sobre las leyes potenciales y el ruido 1/*f* es que no necesitamos un gran desencadenante para poner en marcha un gran suceso. Una extinción de cualquier magnitud podría iniciarse mediante un desencadenante de cualquier magnitud. Lo que importa es que estamos ante un sistema complejo, la vida en la Tierra, que es

[72] No es preciso que nos preocupemos demasiado por las sutilezas de la nomenclatura biológica, pero, para el registro, una especie es el nivel inferior del esquema de clasificación; una especie es una de las categorías que forman parte de un género, y un género es una de las categorías que forman parte de una familia.

autoorganizador, se alimenta a partir de un flujo de energía, y existe al borde del caos.

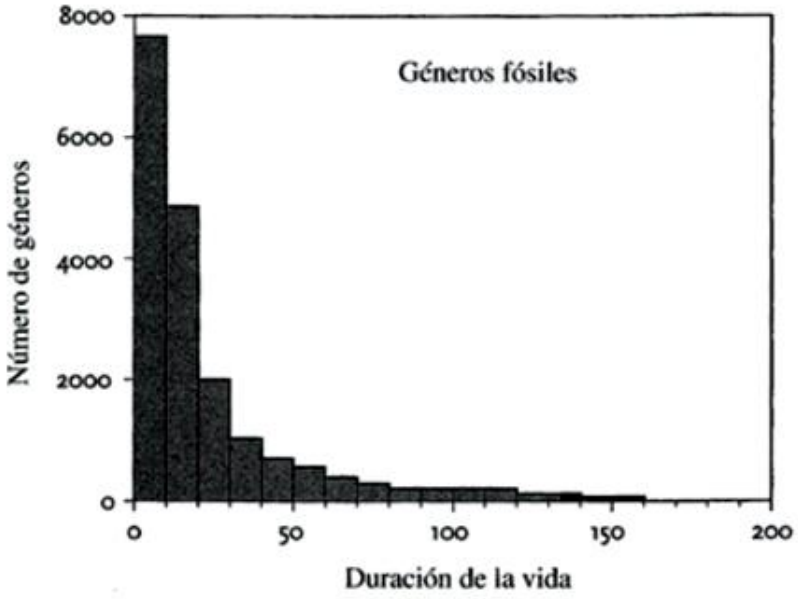

Figura 5.6a. Gráfico de David Raup en el que se representa, mediante un histograma y como gráfico log-log, el número de géneros que sobrevive durante un tiempo concreto. Las duraciones de las vidas de los distintos géneros siguen una ley potencial, con un exponente cercano a 2.

Ya estamos casi preparados para centrarnos en lo que todo esto significa para la vida en sí misma. Pero queda un último paso que demostrará ser de gran ayuda para comprender el surgimiento y la evolución de la vida. Dentro de la mejor tradición científica, tras haber encontrado una ley universal, la ley potencial, que es aplicable a una amplia categoría de situaciones complicadas (y complejas), queremos prescindir de todo, menos de las cosas esenciales, y examinar la cuestión con profundidad para encontrar un modelo sencillo que vaya al fondo del asunto y se pueda utilizar para penetrar en el misterio de los sistemas complejos que se construyen sobre esta profunda sencillez.

Del mismo modo que una mirada profunda en las complejidades de la gravedad comienza con la compresión del sistema gravitatorio más sencillo, un único objeto que esté en órbita alrededor de otro, antes de desplazarse hacia niveles superiores

para aportar una idea sobre el funcionamiento del sistema solar y del movimiento de las galaxias en el universo a gran escala—, así también resulta que una mirada profunda a las complejidades del tejido de la vida puede empezar con el estudio de un sistema mucho más sencillo: un montón de arena colocado sobre una mesa.

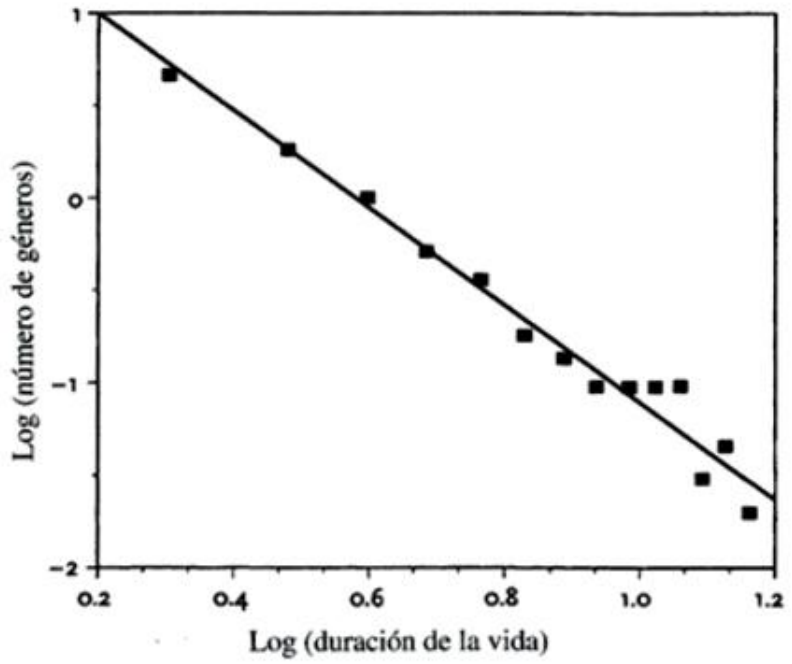

Figura 5.6b. (Véase el pie de la ilustración anterior.)

En la segunda mitad de la década de 1980, el físico de origen danés Per Bak, que trabajaba entonces en el Brookhaven National Laboratory, en Long Island, Nueva York, se interesó por el comportamiento complejo de sistemas que estaban al borde del caos.[73] Lo que le suscitó este interés fueron sus trabajos de física fundamental, cuyos detalles no nos conciernen en este libro, pero pronto percibió que estaba ante un fenómeno de muy amplia repercusión, si no de importancia universal. Bak constató que las leyes potenciales y el ruido 1/*f* siempre están asociados con grandes sistemas que se componen de muchas partes[74], lo que hemos estado llamando sistemas complejos. También se dio cuenta, por supuesto, de la importancia que tenía el hecho de que los sistemas fueran abiertos, es decir, que

[73] Bak (1947-2002) falleció justo cuando este libro se estaba terminando de escribir.
[74] Verdaderamente, esto es casi una tautología. En este contexto, «grande» significa en realidad «formado por muchas partes», y no algo que tenga que ver necesariamente con el tamaño físico del sistema.

tuvieran un suministro de energía procedente del exterior, si iban a llegar al borde del caos, un estado que se llegó a denominar criticalidad autoorganizada. A medida que crecía su interés, Bak comenzó a buscar sistemas de este tipo fuera del ámbito de la física fundamental, y conoció cosas tales como la ley de Gutenberg-Richter para los terremotos y el trabajo de Zipf sobre el tamaño de las ciudades. La respuesta instintiva de un físico ante la constatación de que está funcionando una ley general es encontrar un modelo sencillo que describa lo que está sucediendo, y Bak, trabajando conjuntamente con dos colegas más jóvenes (Chao Tang y Kurt Wiesenfeld), descubrió el ahora famoso modelo del montón de arena.

Al principio, se trataba solamente de un modelo conceptual, que podía expresarse mediante palabras y podía reproducirse en un ordenador, pero no había montones de arena reales, eso llegó más tarde. La clave para este modelo, y para la visión del comportamiento de sistemas complejos que proporcionaba, estaba en la diferencia existente entre un montón de arena colocado sin modificaciones en una mesa en equilibrio (algo tan monótono como los sistemas en equilibrio que se estudiaban en la termodinámica clásica) y el mismo montón de arena, colocado en la misma mesa, al que se van agregando más granos de arena, de uno en uno.[75]

Nuestra experiencia cotidiana, y los recuerdos de haber estado jugando en la playa, nos dicen esquemáticamente lo que va a suceder. La arena se apila en un montón, hasta que la pendiente de dicho montón alcanza un valor crítico. Entonces, si se añade más arena, se producirá un corrimiento de tierras, o una serie de corrimientos de tierras, con lo cual el montón de arena se desplomará. Añadir más arena hace que el proceso se repita, hasta que la mesa queda totalmente cubierta y la arena se derrama por el borde cuando se producen corrimientos de tierras.

En este estado, la cantidad de arena del montón sigue siendo por término medio siempre la misma, ya que son iguales la cantidad de arena que se derrama por el borde de la mesa y la que se añade desde arriba. El sistema se encuentra en un estado de criticalidad autoorganizada, alimentándose del flujo de energía que aportan los nuevos granos de arena que se deja caer sobre el montón. Y así, del

[75] De uno en uno, porque nos interesan los sistemas que se encuentran en algún tipo de estado estacionario, como el cuerpo de un ser humano vivo, que, al recibir un flujo continuo de energía, sólo se ve perturbado ligeramente. La situación cambia de manera radical si volcamos sobre el montón un cubo lleno de arena, del mismo modo que sería diferente lo que sucedería con la vida en la Tierra si, en vez de un flujo continuo de energía procedente del Sol, se liberase toda la energía solar en una enorme explosión.

mismo modo que los terremotos reales de cualquier magnitud pueden ser desencadenados por estímulos de la misma magnitud, la adición de un solo grano de arena puede causar una gran avalancha, o una serie de pequeñas avalanchas, o dejar el nuevo grano balanceándose delicadamente sobre el montón. Pero éste siempre permanece cerca del punto crítico.

*Bak y Tang se basaron en esta idea para dar la clave del modo en que funcionan los terremotos y de la importancia de la ley potencial. Idearon un modelo para ordenador que simulaba el com*portamiento de una zona sísmica como la Falla de San Andrés, donde dos bloques de la corteza terrestre se deslizan uno junto al otro, con movimientos intermitentes. Como ya hemos mencionado, el modelo tradicional de terremotos decía que, cuando la falla se detiene, se acumula la tensión hasta que se libera de una vez una gran cantidad de energía, lo que hace que la tensión vuelva a cero. Sin embargo, en el modelo desarrollado por Bak y Tang, como en el montón de arena, la tensión se acumula y las rocas se deforman hasta que están a punto de resbalar, en un estado crítico autoorganizado. Entonces, la zona sísmica sufre repetidos deslizamientos a todas las escalas, unos deslizamientos que por sí mismos no «liberan toda la energía» de una vez, sino sólo la energía suficiente para mantener la falla en el estado crítico. A diferencia del modelo tradicional, este modelo explica por qué los terremotos siguen una ley potencial.

Los modelos del «montón de arena» desarrollados por Bak y sus colegas mediante ordenadores especificaban cosas tales como la redondez de los granos que se añadían al montón, su adherencia, etc. Examinaron lo que sucede si los granos se añaden aleatoriamente sobre la superficie del montón, y si se dejan caer siempre desde el mismo lugar. Pero, no necesitamos un ordenador (ni un montón de arena) para jugar a este juego y hacemos una idea de una de las leyes más fundamentales que afectan a los sistemas que no están en equilibrio, incluidos los sistemas vivos. En su libro *How Nature Works,* Bak explica cómo puede cualquiera hacerse una idea del modelo del montón de arena utilizando los cubos de construcción con los que juegan normalmente los niños.[76]

Sobre los cuadrados de una red (un tablero de ajedrez sería lo ideal) podemos colocar bloques aleatoriamente hasta una altura máxima de, por ejemplo, tres

[76] Sugiere los de Lego, pero, al fin y al cabo, Bak era danés. Los cubos de madera funcionan mejor.

bloques, de tal forma que en algunos cuadrados no haya bloque alguno, en otros haya uno, en otros dos, y en algunos tres (sirven igualmente unas fichas de jugar a las damas, si podemos conseguir suficientes, o también monedas). Definimos una regla arbitraria que diga que, cuando un montón alcanza la altura de cuatro bloques, hemos llegado a la situación crítica, y entonces retiramos los cuatro bloques de ese montón (que representan granos de arena) y añadimos uno de ellos a cada uno de los cuatro cuadrados contiguos a los lados del cuadrado crítico, o, en caso necesario, dejamos que «desaparezca» el bloque, si el cuadrado crítico está junto al borde del tablero.

Si, como resultado de esto, uno de los cuadrados adyacentes se convierte en un cuadrado crítico, repetimos el proceso. A continuación, añadimos a la red más bloques, de uno en uno y aleatoriamente (podemos elegir el cuadrado al que vamos a añadir un bloque lanzando un dado o utilizando un generador aleatorio de números en un ordenador[77]).

Cada vez que demos este paso, tendremos que mover luego los bloques según las reglas establecidas. Observaremos que el sistema evoluciona hacia el estado crítico, produciéndose «avalanchas» de diferentes magnitudes y cayendo «arena» por el borde de la red, empujada por el nuevo material que se añade desde arriba. En el astado crítico autoorganizado, tanto en los modelos realizados mediante ordenador, como en este «modelo de un modelo», las avalanchas siguen una ley potencial.

Por supuesto, también podríamos intentar algo similar mediante un montón real de arena. La única pega es que la arena de verdad no se comporta del todo igual que un modelo sencillo, porque la arena no es muy adherente y cada grano tiene mucha inercia, por lo que es difícil detener las avalanchas. Este sistema seguiría también una ley potencial, pero la pequeña magnitud de avalanchas menores es más difícil de detectar. Resulta que los montones de arroz de grano largo constituyen una aproximación mucho mejor al montón «ideal» de arena de los modelos realizados por ordenador, ya que los granos largos del arroz producen más fricción, lo cual hace posible que se peguen formando interesantes configuraciones que duran más tiempo y no se desploman de repente en una avalancha de todo o nada.

Los primeros que llevaron a cabo experimentos realizados mediante arroz fueron

[77] Con este modelo de juguete no importa si los números que genera el ordenador no son realmente aleatorios.

Vidar Frette y algunos colegas suyos en la Universidad de Oslo, en Noruega. Para hacer las cosas aún más sencillas, el arroz de grano largo utilizado en los experimentos se introdujo en un hueco situado entre dos paneles verticales de cristal (como una ventana de cristal doble), lo que hacía que el experimento fuera fundamentalmente bidimensional, con un espesor de sólo un grano de arroz (otro ejemplo del afán de los físicos por hacer las cosas sencillas), y posibilitaba que el comportamiento del arroz fuera grabado en vídeo a medida que se introducían los granos de uno en uno dentro del hueco.

Las magnitudes de las avalanchas de arroz que se producen cuando el sistema llega al estado crítico, analizadas de la forma habitual, siguen la ley potencial que se ha predicho. Aún mejor, el análisis de una instantánea de la pauta que forma el borde de la pila de arroz cuando se detiene entre dos avalanchas muestra que, como en el caso de la línea costera de Noruega, esta pauta es fractal. Pero, una de las ideas más importantes que aportó el equipo noruego surgió cuando colorearon algunos de los granos de arroz que echaban por la rendija de la parte superior de la «ventana», de tal modo que podían seguir el movimiento de cada uno de esos granos. Resulta que los granos introducidos, como ya habrá adivinado el lector, no siempre se deslizaban hacia abajo por la superficie para formar parte de la siguiente oleada de avalanchas.

En vez de eso, cualquiera de los granos podía quedar profundamente enterrado dentro del montón, antes de emerger, mucho más tarde, en una avalancha. Cualquier grano puede permanecer en el montón durante cualquier intervalo de tiempo. Por el contrario, ningún grano se queda permanentemente en el montón. Incluso los granos enterrados a más profundidad acaban por abrirse camino hacia la superficie y desplomarse con las avalanchas. Todavía no se entiende del todo por qué y cómo sucede esto, pero el hecho nos indica con claridad que, incluso en este ejemplo extremadamente sencillo de criticalidad autoorganizada, simplificado hasta quedar en lo básico, la totalidad del sistema ejerce una importante influencia sobre cada una de las partes que lo componen, y cada uno de sus componentes ejerce una influencia importante sobre la totalidad del sistema. No hay componentes que se queden inmóviles y permanezcan básicamente inactivos.

La idea de colorear algunos «granos» también demostró ser fructífera en la versión

del juego del montón de arena realizada para el ordenador, que ha conservado este nombre, aunque, dados sus antecedentes, podría ser más exacto denominarla juego de los granos de arroz.

Mientras que en el juego del arroz las conclusiones se obtenían observando el montón de arroz lateralmente, la nueva conclusión obtenida a partir del juego del montón de arena se obtenía observándolo desde arriba—una de las ventajas que ofrece la simulación mediante ordenador es, por supuesto, que podemos «observar» el juego desde cualquier perspectiva, e incluso, si lo deseamos, podemos tomar una sección transversal. Desde este punto de vista, el equipo de Brookhaven dio instrucciones al ordenador para que codificara mediante colores los granos de arena visibles en la superficie del montón, utilizando el verde para los granos situados en zonas llanas o de poca pendiente, donde la probabilidad de avalanchas es pequeña o nula, y el color rojo para las zonas situadas en pendientes pronunciadas, cercanas al punto crítico en que se desencadena una avalancha (en nuestra versión del juego sobre un tablero de ajedrez, esto sería equivalente a pintar de rojo todos los montones que contienen (res bloques, y de verde el resto).

El experimento (que por supuesto, se desarrolla totalmente dentro del ordenador) comienza a diseminar granos de arena, pintados de verde, por toda la superficie del tablero. Cuando el montón va creciendo, aparecen linealmente unas pocas manchas rojas.

A medida que crece más, tendiendo hacia el estado crítico, aumenta la cantidad de manchas rojas y éstas se fusionan en hebras que recorren como una red la superficie del montón de arena. Cuando sólo hay unas pocas manchas rojas aisladas, aunque caiga un grano más sobre el montón en una de estas manchas, el efecto que puede producir es sólo local, ocasionando en los granos de arena del montón una reorganización local a pequeña escala.

Pero cuando la red de hebras rojas es suficientemente densa, un grano más que caiga sobre una mancha roja produce una perturbación que afecta a otras manchas rojas próximas, en un efecto dominó que puede desencadenar una pequeña reorganización de la arena o una serie de avalanchas que afectan a la mayor parte de la superficie del montón de arena. Lo que importa es la densidad de la red de

puntos críticos que hay en el sistema[78].

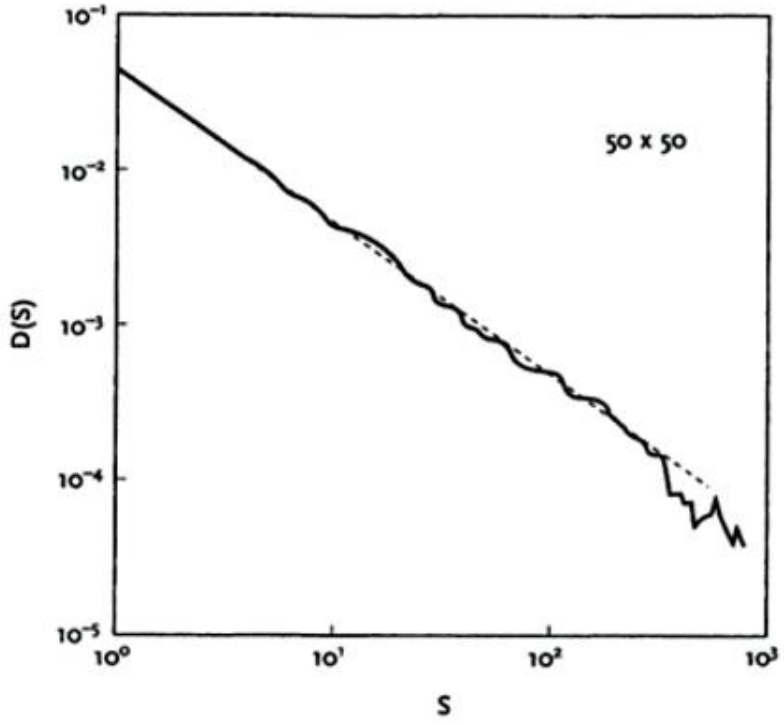

Figura 5.7. El número de avalanchas de magnitudes diferentes que se producen en el modelo del montón de arena sigue también una ley potencial, dando lugar a una línea recta en un gráfico log-log.

Pero hay otra cosa que tiene la misma importancia. Al igual que las zonas sísmicas, el sistema no vuelve a ponerse a cero, independientemente de lo complicada que sea la red. En el estado crítico autoorganizado, la caída de un solo grano más sobre la red puede desencadenar una reorganización de la arena a gran escala; sin embargo, después de haberse producido la reorganización, seguirá habiendo una red compleja de estados críticos, aunque esté organizada según una pauta diferente. *Plus ça change, plus c'est la méme chose.*

Estamos preparados para dar el paso siguiente, y para hacemos una primera idea real de lo que pudo ser la aparición de la vida. Tras haber desbrozado el estudio de la complejidad hasta dejarla en sus fundamentos esenciales, y haber constatado

[78] Para un astrónomo también resulta intrigante que la pauta de esta red se parezca a la pauta en forma de red que presentan las galaxias que están diseminadas por el universo; pero no tengo pruebas de que se trate de una similitud significativa.

que la misma verdad profunda está en la base de cosas tan diferentes como los terremotos, el mercado bursátil y el movimiento de las poblaciones humanas, descubrimos que todo está formado por redes, interconexiones entre las partes simples que componen un sistema complejo. La importancia de tales redes en general, y su importancia específica para la aparición de la vida, han sido investigadas por Stuart Kauffman, del Instituto de Santa Fe, en Nuevo Méjico.

Kauffman ha planteado una curiosa analogía que pone de manifiesto tanto lo que queremos decir al hablar de complejidad, como la importancia de las redes en los sistemas complejos emergentes. Nos pide que imaginemos un gran número de botones corrientes, quizá unos 10.000, repartidos por todo el suelo. Los botones pueden unirse mediante un hilo, pero al comienzo de este experimento imaginario no hay tales uniones.

Desde luego, no se trata de un sistema complicado, según el significado habitual de esta palabra. Pero el modo en que se unen los botones lo convierte en un sistema complejo. Comenzamos sólo con los botones que están en el suelo y con unas buenas reservas de hilo, pero sin que haya ningún botón conectado con otro. Elegimos dos botones al azar y los atamos con un solo hilo, dejándolos en sus posiciones originales sobre el suelo. Repetimos el proceso unas cuantas veces y, si resulta que elegimos un botón que ya está conectado con otro, no hemos de preocuparnos; basta con que utilicemos el hilo para conectarlo también a otro botón. Tras haber hecho esto unas pocas veces, habrá una pequeña cantidad de estructura en la colección de botones. Cada vez con más frecuencia observaremos que, en algunas ocasiones, elegimos un botón que ya está conectado a otro; a veces, incluso escogeremos un botón que tiene ya dos conexiones y lo uniremos a un tercer componente de lo que se ha convertido en una red creciente de conexiones.

Para ver el grado de interés que ha llegado a suscitar esta red, basta con tomar unos pocos botones, de uno en uno, y contar el número de conexiones que cada uno tiene con otros. Cada racimo de botones que ya se han conectado constituye un ejemplo de lo que se conoce como un componente en una red; los botones son ejemplos de nodos, es decir, puntos en los que se realizan las conexiones. El número de botones del racimo más grande (el tamaño máximo de los componentes)

constituye una medida del grado de complejidad a que ha llegado el sistema. El tamaño del racimo más grande crece lentamente al principio, más o menos de manera lineal, a medida que aumenta el número de hilos que unen pares de botones; dado que la mayoría de ellos no tienen muchas conexiones, sólo hay una pequeña probabilidad de que cada nueva conexión añada uno o dos botones más al racimo de mayor tamaño.

Sin embargo, cuando el número de hilos se acerca y luego sobrepasa la mitad del número total de botones, el tamaño del racimo más grande aumenta con una rapidez extrema (fundamentalmente, de manera exponencial) siempre que se añade un nuevo hilo, porque al estar ahora la mayoría de los botones formando racimos hay una buena probabilidad de que cada nueva conexión una sólo racimos menores (no sólo un botón aislado) al racimo de tamaño máximo ya existente. Muy pronto se forma un único racimo de tamaño extraordinario, una red en la que la gran mayoría de los botones está unida en un solo componente. Entonces la tasa de crecimiento se va reduciendo, ya que el hecho de añadir más hilos lo único que hace es aumentar las conexiones entre botones que ya están conectados y sólo ocasionalmente conecta uno de los pocos botones marginales que quedan fuera del súper-racimo.

Aunque la red haya cesado de cambiar de manera notable a medida que se realizan nuevas conexiones, es ya, sin lugar a dudas, un sistema complejo.

Podemos ver nosotros mismos cómo funciona este tipo de efecto si nos ponemos a jugar con botones reales (sería más sensato elegir más o menos unos cincuenta, en vez de 10.000), pero la pauta de comportamiento surge con mucha claridad utilizando para sistemas de este tipo un modelo sencillo en el ordenador. El aspecto importante es que, cuando el número de conexiones de la red excede la mitad del número de nodos, cambia con gran rapidez de encontrarse en un estado bastante monótono (muchos bolones con pocas conexiones entre ellos) a otro estado estable mucho más estructurado, pero en el que hay pocas perspectivas de que se produzcan más cambios. Este es un sencillo ejemplo del fenómeno conocido como transición de fases, que Kauffman compara con el modo en que el agua experimenta una transición de fases cuando se congela y se convierte en hielo.

En la red, la complejidad ha surgido de forma natural a partir de un sistema muy

sencillo, sin más que añadir conexiones, con una interesante actividad asociada al cambio radical que se produce de repente, justo en la transición de fases, cuando el sistema pasa de un estado a otro como resultado de realizar una cantidad muy pequeña de conexiones finales.

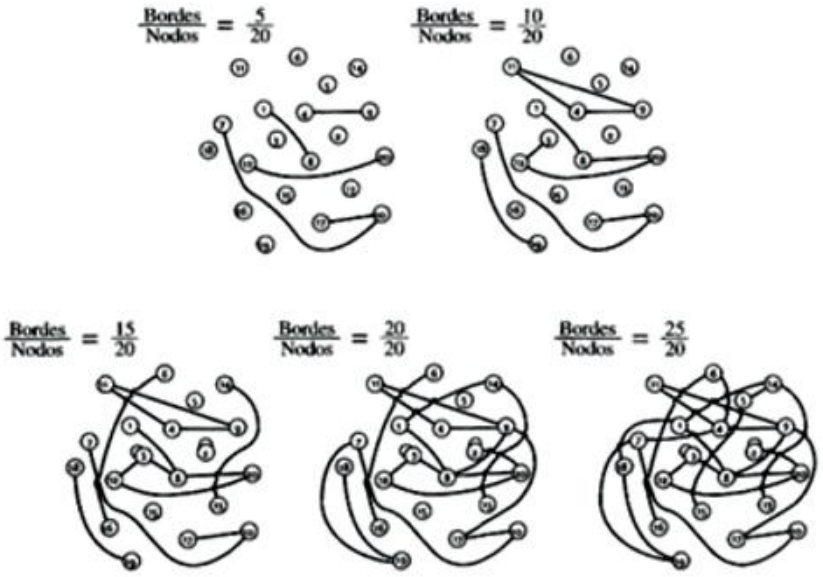

Figura 5.8a. Modelo de los «botones» con el que Stuart Kauffman explica cómo se forman las redes.

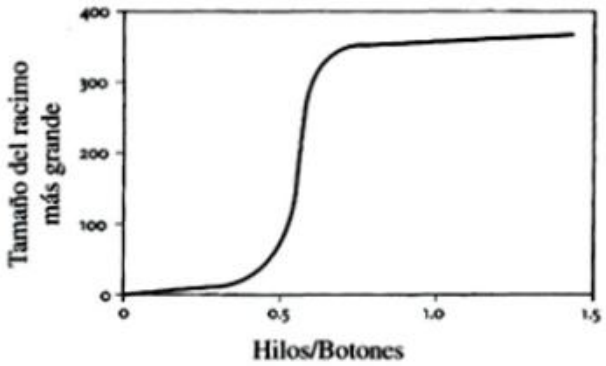

Figura 5.8b. Cuando el número de conexiones aumenta, se produce una brusca transición desde un estado en el que hay muchos botones «aislados» y unas pocas conexiones, hasta el estado en el cual casi todos los botones forman parte de la red.

Por supuesto, no es necesario que las conexiones sean trozos reales de hilo. En el

modelo del montón de arena, las conexiones vienen determinadas por las leyes de la física relativas al caso, la ley de la gravedad, que lleva los granos de arena hacia la parte de abajo del montón, las leyes del rozamiento, que favorecen la permanencia de los granos de arena en su posición, etc... Lo que importa es que, a medida que el montón se vuelve más empinado, se establecen más conexiones entre los granos, porque el movimiento de uno de ellos afecta de manera creciente la estabilidad de sus vecinos.

En un montón de arena plano, el hecho de quitar (o añadir) un grano no afecta mucho a sus vecinos. Cuando se ha llegado al ángulo crítico, la retirada de un grano puede significar que se quita la pieza del rompecabezas que, como la clave del arco de un puente, evitaba que otras piezas se fueran abajo y, cuando éstas se deslizan, afectan a muchos otros granos, y así sucesivamente. Hacer que el ángulo del montón de arena sea más empinado tiene el mismo efecto, en cuanto a incrementar la complejidad de la red, que añadir más conexiones entre botones. De manera similar, como ya hemos visto, un asteroide que describe su órbita entre Marte y Júpiter no está sometido sólo a la influencia aislada del Sol, sino que es parte de una compleja red de conexiones, creada por la gravedad, que implica (en principio) a todos los demás objetos gravitatorios del universo, e incluso, a un nivel sencillo, a todos los planetas importantes del sistema solar, así como al propio Sol. Esto es lo que hace que el comportamiento del asteroide sea impredecible.

La razón por la que Kauffman estaba interesado en el tema de las redes era que le llamaba la atención el fenómeno emergente más importante de todos, cómo surgía la vida a partir de la materia no viva. ¿Qué fue lo que sucedió en la «sopa» química primordial, ya fuera en la Tierra, o en algún otro lugar, para que algunas de aquellas sustancias químicas se unieran incorporándose a sistemas que formaban réplicas de sí mismos? Kauffman (1939-) estudió en el Darmouth College, en New Hampshire, y luego, desde 1961 hasta 1963, en la Universidad de Oxford con una beca Marshall. Aunque cuando empezó este curso se veía a sí mismo como filósofo, en Oxford la filosofía se enseñaba junto con un extenso programa de biología, con el fin de formar a los estudiantes de tal manera que pudieran plantear preguntas tales como la relación entre las imágenes del mundo que tenemos en nuestras mentes y el modo físico en que nuestros ojos y sistemas nerviosos interpretan la

información procedente del mundo exterior. Al desilusionarse cada vez más por la tendencia de la filosofía a la contemplación del propio ombligo, y sintiéndose cada vez más intrigado por la biología, Kauffman decidió pasarse a la medicina, y en la Universidad de California, en San Francisco, estudió para ser médico.

Fue allí donde desarrolló su interés por el modo en que actúan los genes dentro de las células y por la aparición de la vida, llegando a ser oficialmente un biólogo teórico cuando se trasladó a la Universidad de Chicago, en 1969, después de terminar sus estudios de medicina. Pero este interés por los aspectos fundamentales de la aparición de la vida quedó en barbecho hasta la década de 1980, sencillamente porque estaba demasiado ocupado con otro trabajo (primero en Chicago, luego en los National Institutes of Health y, más tarde, en la Universidad de Pensilvania) como para tomar el hilo.

El trabajo fundamental se publicó finalmente en 1986, y su idea central puede resumirse de manera muy sencilla utilizando el tipo de redes de las que ya hemos hablado; aunque todavía formaba parte del personal de la Universidad de Pensilvania, donde impartía docencia, fue el Santa Fe Institute el centro donde llevó a cabo básicamente sus investigaciones, y allí desarrolló sus teorías en colaboración con otros colegas de esta institución, en gran medida en un entorno interdisciplinar donde se oían murmullos incesantes sobre ideas tales como el caos, la complejidad y la criticalidad autoorganizada.

Todo lo que tenemos que hacer es imaginarnos que en el caldo químico inicial había algunas sustancias que actuaban como catalizadores para la formación de otras sustancias, como en los procesos catalíticos que tiene lugar en la reacción BZ. La sustancia química A cataliza la formación de la sustancia catalítica B. Dada la gran variedad de sustancias químicas presentes como materias primas[79], es difícil imaginarse que no sucediera algo así, e incluso si el impulso que A daba a la formación de B era sólo ligero, aumentaría en todo caso la concentración de B dentro de la mezcla. Supongamos ahora que la presencia de B fomenta la formación de C, que a su vez, actúa como catalizador para la formación de D, y así

[79] Algunas investigaciones modernas relativas al universo a gran escala han revelado que el tipo de nubes existentes en el espacio, a partir de las cuales se han formado sistemas planetarios como el nuestro, contiene una gran variedad de moléculas, incluso aminoácidos, que habrían sido traídas por cometas a la superficie de la Tierra cuando el sistema solar era joven, sembrando nuestro planeta, si no de vida, sí ciertamente de las materias primas que necesita ésta. Volveremos sobre ello más adelante.

sucesivamente.

En algún lugar de esta cadena, uno de los compuestos que aparecen cataliza la formación de A, con lo que tenemos un bucle automantenido de interacciones, que, de hecho, se alimenta de la materia prima disponible y hace que entre en el bucle más cantidad de los compuestos mencionados, con ayuda de energía procedente del Sol o del calor de chimeneas volcánicas. Puede haber también otras interacciones, como vimos en nuestros anteriores ejemplos relativos a interacciones químicas automantenidas, D podría actuar como catalizador para la formación de A, así como para la de E, y asimismo podrían participar en esto sustancias inhibidoras, de nuevo como en la reacción BZ. Es bastante fácil ver cómo puede surgir una red de conexiones entre las sustancias químicas contenidas en el caldo, una red autocatalítica que se sustenta a sí misma. Según dice Kauffman, así es como surge la vida, como una transición de fases en un sistema químico que posee un número suficiente de conexiones entre los nodos (compuestos químicos) de la red. Una característica crucial y convincente de esta argumentación es que, al igual que la aparición de un racimo de tamaño extraordinario (un súper-racimo) en la red de botones e hilos, éste es un suceso de todo o nada. Si la red está conectada de forma insuficiente, no hay vida; pero basta con añadir una o dos conexiones más para que la vida se convierta en algo no sólo posible, sino inevitable. No es necesario formar una larga cadena de sucesos químicos improbables para que emerja la vida y no existen estados intermedios en los cuales no podamos estar totalmente seguros de si el sistema está vivo o muerto.

Sobre esto hay mucho más que decir, incluyendo detalles tales como interacciones químicas reales, y la necesidad de pensar cómo fue posible que las materias primas se hubieran acercado unas a otras lo suficiente para que se produjeran estas series de interacciones, una posibilidad es que estuvieran concentradas en una fina capa de materia, ya fuera como una mancha de líquido extendida sobre una superficie, o una capa adherida a una superficie de barro. Hemos de recalcar que estas teorías son al mismo tiempo especulativas y discutibles; no es un argumento que convenza a todo el mundo, y existen otras teorías sobre el modo en que la vida podría haber comenzado. Pero tampoco estas otras teorías convencen a todo el mundo, y nadie sabe con seguridad qué ocurrió cuando la vida surgió de la materia inerte. Incluso

en el modelo de Kauffman quedan muchos detalles sin explorar. Pero nos parece que, de forma global, el paquete completo que forman estas teorías resulta convincente, en cierta medida por el hecho de que sitúa la aparición de la vida en el mismo conjunto de sistemas complejos basados en leyes sencillas que encontramos tan a menudo en otros contextos. Como dice Kauffman, «la vida cristaliza a partir de un nivel crítico de diversidad molecular, debido a que la propia clausura catalítica cristaliza»[80].

Sean cuales sean los *orígenes* de la vida, las teorías que incluyen redes, conexiones y criticalidad autoorganizada proporcionan unas ideas nuevas y poderosas sobre el modo en que funciona la vida una vez que ha surgido, un tema que comenzaremos a comentar en el próximo capítulo. Sin embargo, antes de iniciar ese comentario, es conveniente explicar otra de las especulaciones de Kauffman. Aunque no es menos discutible que las teorías relativas a la aparición de la vida, dicha especulación ofrece un ejemplo sorprendente de la medida en que la complejidad de los seres vivos (lo más complejo que hay en todo el universo) puede estar basada en una profunda sencillez.

Kauffman estuvo, y está, interesado en el modo de funcionar de las células, al nivel de los genes que aportan las instrucciones que gobiernan lo que a veces se llama de una manera imprecisa la maquinaria de la célula. Estas instrucciones se encuentran en última instancia codificadas en el ADN, las grandes moléculas de las que están constituidos los genes; pero tanto la maquinaria como la estructura del cuerpo están hechas de proteínas. Elementos tales como el pelo y las uñas de los dedos, así como los músculos, son tipos de proteínas, y también lo son sustancias como la hemoglobina, que transporta el oxígeno en la sangre, y las enzimas, que son los catalizadores biológicos esenciales que favorecen las reacciones químicas importantes para la vida. Las propias proteínas son grandes moléculas formadas por subunidades llamadas aminoácidos, y ésta es la razón por la que resulta tan intrigante el descubrimiento de que los aminoácidos existen en el tipo de nubes interestelares a partir de las cuales se formaron el Sol y su familia de planetas.

El código genético que está en el ADN contiene instrucciones para fabricar proteínas y, luego, estas proteínas realizan las tareas de que se compone la vida. Pero, en

[80] Véase Stuart Kauffman, *Al Home in the Universe.*

este proceso hay otro paso que resulta sorprendente. Cuando un gen se activa (cómo y por qué sucede esto es algo que va más allá de los objetivos de esta explicación), la información que interesa en ese momento se copia primero en una molécula muy similar llamada ARN.

Posteriormente, la maquinaria de la célula lee el ARN y actúa según sus instrucciones para fabricar la proteína adecuada. Este proceso de dos pasos probablemente nos esté diciendo algo sobre el modo en que se originó la vida, y existe alguna posibilidad de que el ARN se «inventara» antes que el ADN. En la situación que describe Kauffman, la «cristalización» de la vida tiene lugar en el nivel de las proteínas, en una sopa química rica en aminoácidos, donde surgieron las primeras redes autocatalíticas de la vida; en este modelo encaja fácilmente la posibilidad de que el ARN participara en una fase temprana y que, posteriormente, las presiones evolutivas asociadas con la competencia entre distintas redes autocatalíticas pudieran haber conducido al sistema a la situación que vemos en la actualidad. Pero, nos estamos desviando.

Los puntos relevantes en los que se apoya la investigación de Kauffman sobre el modo en que funcionan las células son, por una parte, el hecho de que los genes actúan para controlar la maquinaria celular y, por otra (siendo éste el aspecto crucial) que los genes pueden afectarse mutuamente, cuando un gen activa o desactiva a otro[81]. La investigación llevada a cabo por Kauffman sobre la maquinaria de las células en términos de redes se remonta a su interés extraoficial por este concepto cuando era un estudiante de medicina, pero dio fruto en el trabajo que realizó en el Santa Fe Institute a finales de la década de 1980.

En la época en que Kauffman desarrolló este trabajo, se pensaba que había unos 100.000 genes diferentes en el ADN humano, es decir, en el genoma humano. Desde entonces, el proyecto del genoma humano ha demostrado que esta estimación era excesiva, y que sólo hay alrededor de un tercio de dicho número de genes para especificar lo que debe ser una criatura humana. Todos ellos están presentes en el ADN de cualquier célula del cuerpo humano, pero no todos están

[81] Ésta es una razón de peso para pensar que la ingeniería genética nunca será tan sencilla como algunos informes hayan podido hacernos creer. Incluso en el caso de que sea posible modificar un solo gen, por ejemplo para aumentar la capacidad de los pulmones (algo que en sí mismo es sumamente improbable), al cambiar este gen cambiaríamos asimismo, de un modo que hoy por hoy es impredecible, el comportamiento de todos los genes que forman parle de la misma red.

activos en todas las células del cuerpo. Los distintos tipos de células están especializados para realizar diferentes tareas, así, por ejemplo, las células del hígado no hacen lo mismo que las células de los músculos.

Esta diferenciación de las células en formas especializadas se produce durante el desarrollo del embrión, siendo el conocimiento del proceso de desarrollo y la especialización de las células una de las áreas más importantes de la investigación biológica. Sin embargo, sea cual sea el modo en que se realiza el proceso, en un ser humano plenamente desarrollado hay 256 tipos diferentes de células especializadas. En cada caso, sólo los fragmentos adecuados de ADN (los genes adecuados) se «activan» siempre en el transcurso normal de la vida, de tal modo que una célula del hígado es siempre sólo una célula del hígado. Sin embargo, sigue estando ahí todo el resto de la información genética, como se ha demostrado drásticamente en la clonación, donde el ADN de una célula especializada se transfiere a un óvulo, que luego se desarrolla para formar un nuevo ser adulto, que es la réplica de aquel de quien procede el ADN.

La red que participa en la organización del funcionamiento de una célula se puede describir como un sistema que posee un nodo para cada gen y también conexiones que vinculan los genes entre sí, como los hilos que unían los botones en el modelo anterior. Con un número de genes implicados que oscila entre 30.000 y 100.000, es fácil adivinar que el problema de describir el comportamiento de una red como ésta sería muy complicado de resolver, incluso utilizando los ordenadores más avanzados. Kauffman utiliza una analogía en la que los nodos de la red son bombillas enlazadas al azar con un cableado.

En este sistema, puede ser que todas las bombillas estén encendidas, o que todas estén apagadas, y entre estos dos extremos hay un gran número de estados diferentes en los que las bombillas que están encendidas se presentan en combinaciones diferentes. De una manera más precisa, si en total hay *N* bombillas, dado que cada una puede estar en cualquiera de los dos estados (encendida o apagada), el número de estados posibles para todo el sistema es 2^N. Cuando tenemos una bombilla, hay dos estados; con dos bombillas, hay cuatro; con tres, hay ocho; y así sucesivamente. Si tenemos 100 bombillas conectadas aleatoriamente mediante cables, hay 2^{100} estados posibles que son diferentes, y 2^{100}

es aproximadamente 10^{30}. Es imposible captar un número como éste desde cualquier tipo de perspectiva humana, pero, en la medida en que una comparación puede ayudar, diremos que sólo han transcurrido aproximadamente 4 x 10^{17} segundos desde el big bang en que nació el universo, y recordemos que 10^{30} no es 13 veces mayor que 10^{17}, sino 10^{13} veces mayor (es decir, diez billones de veces mayor).

Sin embargo, Kauffman y sus colegas continuaron impávidos sus investigaciones sobre el comportamiento de estas enormes redes en modelos realizados mediante ordenador, con lo que parece ser la complicación añadida de la retroalimentación, en la que el hecho de que una bombilla esté encendida o apagada depende del estado de las bombillas a las que está conectada. Aplicaron las reglas de la lógica booleana (o álgebra booleana)[82], de tal modo que, por ejemplo, si una bombilla está conectada mediante cable a otras dos, podría estar encendida sólo si las otras dos lo están (una operación de ADN), o podría estar encendida si una u otra de las otras dos lo está (una operación OR, con el nombre en inglés de la conjunción disyuntiva), y así sucesivamente. En estas investigaciones examinaron las pautas globales de comportamiento de redes formadas por distintos números de bombillas, con diferentes números de conexiones en cada nodo, y con distintos operadores booleanos asignados aleatoriamente a cada nodo. No es de extrañar que esta investigación les llevara años.

El equipo estaba buscando pautas estables que fueran el resultado de aplicar reglas sencillas. Cualquiera de estos sistemas-tiene un número finito de estados, aunque este número pueda ser absurdamente grande. Si iniciamos el sistema en algún estado elegido de forma aleatoria, con algunas bombillas encendidas y otras apagadas, las luces parpadearán y cambiarán de acuerdo con el conjunto concreto de reglas booleanas que se haya aplicado y, si tenemos suerte, el sistema se estabilizará en una pauta repetitiva, siguiendo su propia cola indefinidamente, recorriendo siempre el mismo ciclo de un estado a otro (llamado, como es lógico en este contexto, ciclo de estados, que es, en esencia, lo mismo que los ciclos límite que hemos comentado anteriormente).

El ciclo de estados puede ser lo suficientemente corto como para que podamos

[82] Llamada así porque fue el matemático inglés George Boole (18151864) quien desarrolló este concepto. El álgebra booleana tiene gran importancia en muchos lenguajes de programación.

percibir la pauta, o puede contener tantos cambios que nos obligue a esperar hasta la eternidad y nunca lleguemos a notar repetición alguna. También puede ser que el sistema llegue a congelarse en un estado en el que algunas bombillas estén encendidas y otras no, de tal modo que nunca vuelva a producirse un parpadeo. Los ciclos de estados actúan como atractores del sistema, y en algunos casos se convierten en atractores muy poderosos, de tal modo que, con independencia del estado en que iniciemos el sistema, iremos muy rápidamente a uno de estos ciclos de estados. Ésta puede ser la clave de la vida de las células.

Kauffman y sus colegas afirman que, si hay sólo una entrada por cada bombilla, nunca sucederá algo interesante, con independencia del número de nodos que tengamos y de cuáles sean las reglas de la lógica booleana que apliquemos. El sistema, o bien entrará en un ciclo de estados muy corto, o se quedará congelado. Si el número de conexiones en cada nodo es mayor que 2, reinará el caos. Hay atractores, pero éstos tienen un número enorme de estados, y son también muy sensibles, en el sentido de que un pequeño cambio (por ejemplo, apagar o encender tan sólo una de las bombillas) desplazará el sistema a favor de otro atractor, algo muy parecido al efecto mariposa.

Los *únicos* sistemas que se comportan de un modo lo suficientemente complicado como para resultar interesantes, y lo suficientemente estable como para poder comprenderlos, son aquellos en los que cada nodo está conectado exactamente a otros dos. En el borde del caos suceden cosas interesantes, siendo la retroalimentación un ingrediente esencial de lo que les da esa cualidad. Sólo en esos sistemas, cada ciclo de estados tiene una longitud igual a la raíz cuadrada del número de nodos. El sistema que hemos comentado anteriormente, a partir de sus 2^{100} posibilidades, se estabilizará en estas condiciones en un ciclo de estados que tiene sólo 10 pasos (ya que 10 es la raíz cuadrada de 100), y su pauta será evidente para cualquiera que lo esté observando. En un sistema que tenga 100.000 nodos, un típico ciclo de estados tendrá una longitud de sólo 317 pasos.

Lo que aún es más interesante es que estos ciclos de estados son potentes atractores, y el sistema, tras iniciarse en un estado aleatorio cualquiera, se desplaza rápidamente hacia uno de estos ciclos y se estabiliza, no siendo sensible a pequeñas perturbaciones. Podemos establecer un sistema que tenga 2^{100000} estados diferentes

(unos $10^{30\ 000}$ estados), iniciarlo en cualquier pauta elegida de forma aleatoria, y casi en un abrir y cerrar de ojos se estabilizará en un ciclo estable, visitando repetidamente 317 de esos estados, siempre en el mismo orden[83].

Aún hay más. Se puede aplicar el mismo tipo de estadística a los distintos ciclos de estados que pueden existir en un sistema de este tipo, aunque sólo de una manera más bien aproximada. En la clase de sistemas que acabamos de describir debe haber un número de atractores diferentes aproximadamente igual a la raíz cuadrada del número de nodos.

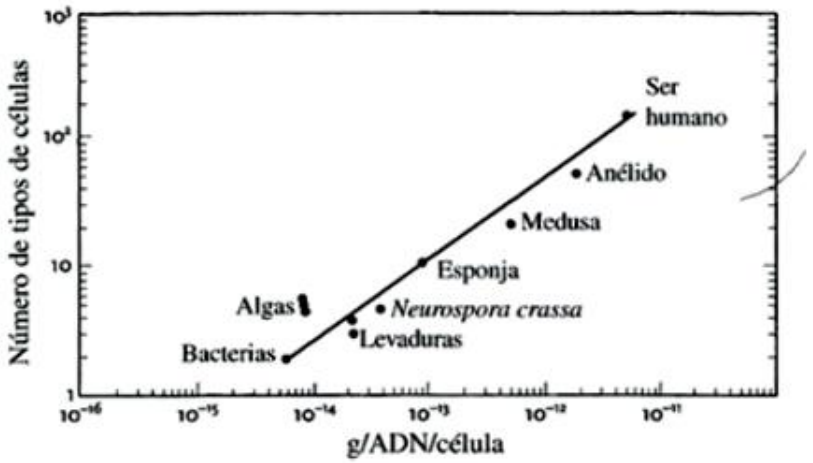

Figura 5.9. Un gráfico log-log en el que se compara la cantidad de ADN de una célula con el número de tipos diferentes de células de diversos organismos. El trazado sigue rigurosamente una línea recta cuya pendiente es 0,5 (una ley de la raíz cuadrada, o una potencia de exponente 1/2), como pronosticó Stuart Kauffman basándose en la suposición de que el número de genes estructurales y reguladores es proporcional a la cantidad de ADN que contiene la célula.

Si hay 100.000 nodos diferentes, habrá unos 317 atractores diferentes. En el caso de ser 30.000 los nodos, habrá alrededor de 173 atractores distintos. Existen entre 30.000 y 100.000 genes en el genoma humano, y existen 256 tipos diferentes de células en el cuerpo humano. ¿Podría ser que cada tipo de célula represente un ciclo de estados particular del genoma humano, en el que unos genes determinados estén activados y otros desactivados? Para comprobar esta posibilidad, Kauffman ha

[83] Hay otro parámetro, relacionado con la elección de las reglas booleanas, que afecta a este sistema, pero se trata de una especie de sintonización fina y no incide en el aspecto fundamental que estamos señalando.

comparado el número de genes y el número de tipos de células en distintos organismos. Las bacterias tienen uno o dos tipos diferentes de células, las levaduras pueden tener tres, la mosca de la fruta sesenta, y así sucesivamente. El número de tipos de células crece en proporción a la raíz cuadrada de la cantidad de ADN presente en el organismo y, aunque todavía no se ha obtenido el mapa genético de todas las especies, parece razonable, desde un punto de vista empírico, suponer que el número de genes es proporcional a la cantidad de ADN. El número de tipos de células aumenta a medida que crece la raíz cuadrada del número de genes.

De esto se puede deducir que las mismas reglas que rigen para los montones de arena y las redes de bombillas operan también en las células de nuestros cuerpos. En números redondos, los tipos de células humanas ascienden a un par de cientos, porque hay un par de cientos de ciclos de estados que son atractores en una red de decenas de miles de genes que ejercen interacciones entre ellos mismos, siguiendo las reglas de la lógica booleana. En cada tipo de células, los genes están en su mayoría congelados en el estado de «desactivación», pero una minoría de ellos parpadea, activándose y desactivándose unos a otros siguiendo una pauta que pasa por un ciclo cuya longitud es de un par de cientos de pasos, durante el cual opera la maquinaria química de la célula de acuerdo con el tipo de célula que es, para luego repetirse el mismo ciclo. Y aquí está todo, el secreto de la vida a nivel celular.

Éste es un trabajo que todavía está en fase de realización, y se ha de hacer mucho más para convertir estas teorías provocativas y persuasivas en una teoría completa de la diferenciación y el desarrollo celulares. Pero, hasta ahora, la historia se sostiene, planteando con contundencia que incluso las criaturas como nosotros, que somos los sistemas más complicados del universo, estamos construidas a partir de unas reglas muy sencillas.

Toda la complejidad aparente de un sistema interactivo formado por decenas de miles de genes se reduce a unos pocos cientos de estados posibles, gracias a la profunda sencillez del mundo. Como dice Kauffman, «somos la expresión natural de un orden más profundo». Ya es hora de que pasemos a ver cómo afecta este orden más profundo a las interacciones, no sólo entre genes y otras sustancias químicas, sino entre los distintos tipos de organismos vivos que hay en todo el mundo.

Capítulo 6
Las cosas de la vida

La evolución es un hecho, al igual que lo es la forma elíptica de la órbita que describe un planeta alrededor del Sol. Tanto en el registro fósil como en los estudios realizados sobre la vida actual en la Tierra, se puede encontrar un número considerable de pruebas relativas al modo en que actúa la evolución, transformando una especie en otra[84]. La teoría de la selección natural, a la que llegaron de manera independiente Charles Darwin y Alfred Russel Wallace en la segunda mitad del siglo XIX, es un modelo que ofrece una explicación de por qué se produce la evolución, del mismo modo que la teoría de la gravedad, desarrollada por Isaac Newton durante la segunda mitad del siglo XVII, es un modelo que explica por qué los planetas describen órbitas elípticas. Ni la teoría, ni el modelo, constituyen la última palabra sobre la cuestión que abordan. La teoría de Newton fue mejorada a principios del siglo XX por Albert Einstein, que descubrió un modelo más completo para explicar cómo actúa la gravedad, la teoría general de la relatividad.

La teoría de Einstein es más completa que la de Newton porque se cumple, con gran precisión, en situaciones en las que la teoría de Newton no resulta adecuada para describir el universo en que vivimos, sobre todo en condiciones que determinan un intenso campo gravitatorio. Pero el hecho de que la teoría de Einstein sea correcta no significa que la de Newton esté equivocada. La teoría newtoniana de la gravedad sigue siendo perfectamente adecuada para explicar (o calculará por ejemplo, lo que sucede cuando una manzana cae a tierra desde un árbol, o la trayectoria de una nave espacial que viaja desde la Tierra a la Luna. Lo que pasa es que, en determinadas circunstancias, necesitamos algo más que la teoría de la gravedad de Newton con el fin de reflejar plenamente una situación.

A finales del siglo XX, también estaba claro que en determinadas circunstancias se necesitaba algo más que la teoría de Darwin-Wallace de la selección natural para ver de forma completa cómo actúa la evolución, sobre todo cuando hay muchas especies diferentes ejerciendo interacciones mutuas en un ecosistema. Pero esto no significa que el modelo de Darwin-Wallace sea erróneo, ya que sigue siendo

[84] Véase, por ejemplo, *Evolution,* de Cari Zimmer, o *The Beak of the Finch*, de Jonalhan Weiner.

perfectamente adecuado para describir algunas situaciones evolutivas, especialmente aquellas en las que hay unas pocas especies ejerciendo interacciones mutuas y con su entorno físico. A esto lo llamaremos «evolución darvinista», siguiendo una tradición que se remonta al propio Wallace, que fue el primero en reconocer a Darwin como descubridor fundamental de la teoría, en gran medida porque había dado el paso decisivo de presentarla ante una amplia audiencia (en su libro *Sobre el origen de las especies,* publicado por primera vez en 1859). Existen trabajos recientes que han proporcionado una visión más completa del modo en que funciona la evolución cuando se incluye el complejo entramado de interacciones existente en una red ecológica formada por muchas especies que actúan influyéndose recíprocamente. Para valorar la importancia de estos trabajos, necesitamos en primer lugar hacemos una idea clara de lo que es la evolución darvinista.

La esencia de la evolución darvinista puede resumirse en tres pasos, que forman una cadena de una lógica tan directa que es casi una tautología. Esto llevó a Thomas Henry Huxley, un contemporáneo de Darwin y defensor de la teoría de la selección natural, a comentar cuando se encontró por primera vez con dicha teoría, «¡Qué estúpido he sido por no haberlo pensado yo antes!»[85]. En primer lugar, los descendientes se parecen a sus progenitores. Las características se heredan, transmitiéndose de una generación a la siguiente. En segundo lugar, el proceso (sea cual sea) que copia las características de una generación para transmitirlas a la siguiente no es perfecto, por lo que hay ligeras variaciones (a veces más que ligeras) entre los individuos de cada generación, incluso entre hermanos. Por ejemplo, todo el mundo tiene nariz, pero algunos la tienen más larga que otros. Una nariz larga puede ser característica de una familia, pero dos hermanos de esa familia no tendrán exactamente el mismo tamaño de nariz, aunque cada uno de ellos tenga una nariz más larga que la de los miembros de otra familia.

Finalmente, y ésta es la cuestión crucial a la que se aferraron tanto Darwin como Wallace, son más los individuos nacidos en cada generación que los que sobreviven para convertirse en adultos y reproducirse a su vez[86]. Pero, *¿por qué* algunos

[85] Véase Cari Zimmer, *Evolution,* p. xii.

[86] Por supuesto, el mismo razonamiento puede aplicarse a las plantas. Lo único que hacemos es referirnos a la vida animal para explicarlo, ya que esto refuerza la idea de que también puede aplicarse a nosotros mismos.

individuos sobreviven para reproducirse, mientras otros perecen antes de tiempo? La gran idea, que llevó a Huxley a realizar su famoso comentario, es que los individuos que sobreviven son los que se adaptan mejor a su entorno. Están los que son hábiles para encontrar alimento, para encontrar (o atraer) a una pareja, y para evitar ser devorados por algún otro animal. Esto es lo que significa la «supervivencia del más apto». No nos referimos necesariamente a los individuos que tienen un físico más saludable y atlético (aunque en algunos casos esto podría formar parte de la historia), sino a aquellos que encajan mejor en su entorno (incluido el «entorno» que crea la presencia de otras especies), como el encaje de una pieza en un *puzle*, o el de una llave en la cerradura correspondiente. No los que *están* en forma, sino los que *se adecúan* a las formas.

Ahora, retrocedamos al comienzo de la cadena lógica. Los individuos que encajan mejor sobreviven y se reproducen. Así transmiten sus características a la generación siguiente, incluidas, en primer lugar, aquellas que les permiten tener éxito. Mientras el entorno permanece invariable (una condición muy exigente, como veremos más adelante), este proceso irá adecuando a los miembros de la especie para que encajen aún mejor en sus nichos ecológicos. De hecho, parte de las pruebas derivadas de la observación que llevaron a Darwin a concebir la idea de selección natural (aunque pasó bastante tiempo hasta que dio en el clavo) procedía de sus estudios sobre los pinzones de las diferentes islas que forman el Archipiélago de Colón, también llamadas Islas Galápagos. Aunque los pájaros de las distintas islas son todos miembros de la misma familia, en una isla donde las semillas de que se alimentan los pinzones son duras y difíciles de abrir, los pinzones tienen unos picos cortos y fuertes. En otra isla, donde el alimento existente ha de ser extraído de rincones y grietas, los mismos de los anteriores tienen picos más largos y delgados. Cada grupo está bien adaptado a la vida de su isla, pero si lleváramos a los pájaros de una isla a vivir en otra y los soltáramos allí, realmente tendrían que luchar para sobrevivir.

Memos de insistir en que esa lucha no se produce entre los pinzones y otras especies, ni siquiera contra las plantas de cuyas semillas se alimentan. La competencia por obtener alimento se produce *entre miembros de la misma especie.* En este ejemplo, para un pájaro que viva en una isla donde sólo hay semillas duras

para comer, tener un pico más fuerte le proporciona una ventaja sobre otras aves. Su característica especial significa que puede conseguir más alimento, lo cual le ayudará a criar una nidada más numerosa, que heredará su fuerte pico. Los pájaros que tengan un pico algo menos fuerte conseguirán una cantidad correspondientemente menor de alimento, por lo que tendrán menos éxito a la hora de criar a su descendencia, y en la generación siguiente, por término medio, los picos serán más fuertes que en la generación anterior[87].

De manera similar, un conejo que escapa de un zorro no compite con éste en términos evolutivos. Compite con otros conejos. Los que corran más rápido y zigzagueen de la manera más imprevisible sobrevivirán y se reproducirán; los que no, serán devorados. Así, en generaciones sucesivas, los conejos serán cada vez mejores en eso de correr con rapidez, pero erráticamente. El zorro, por supuesto, compite con otros zorros, el zorro que corra más rápido, o sea más astuto para capturar su presa, sobrevivirá y tendrá más descendencia (heredera de sus características) que un zorro lento y bobo. Volviendo a nuestro ejemplo de las variaciones en el tamaño de la nariz, podemos imaginar que un cambio en el entorno podría situar en desventaja a la gente que tenga un determinado tipo de nariz, algún tipo de polución podría hacer que fuera ventajoso tener una gran nariz.

En ese caso, los bebés que tuvieran narices pequeñas, tendrían más probabilidades de morir en la infancia, y recordemos que, a causa de ciertas imperfecciones en el proceso de copiado de las características que pasan de una generación a otra, incluso los hermanos pueden tener narices de diferentes tamaños. Las familias cuyos miembros tienen narices grandes encajarán en el nuevo entorno mejor que aquellas cuyos miembros tienen narices pequeñas, pero, incluso dentro de estas familias, los hermanos que tengan narices de mayor tamaño se las apañarán mejor que sus hermanos y hermanas[88].

Después de muchas generaciones, el tamaño medio de la nariz aumentará en toda

---

[87] Confiamos en que todo el que lea esto conozca al menos la teoría de que las características físicas de un cuerpo (el fenotipo) están determinadas por el conjunto de genes contenido en sus células (el genotipo), y que los cambios observados en el fenotipo se producen porque hay cambios en el genotipo, cosa que puede suceder, por ejemplo, como resultado de copiar errores cuando se hace una réplica del ADN, y a veces a esto se le llama mutación. Si deseamos recordar cómo funciona todo esto, nos será de gran ayuda cualquier libro de Richard Dawkins, o podríamos intentarlo leyendo *In Search of the Double Helix,* de John Gribbin. [Hay traducción al castellano: *En busca de la doble hélice: la evolución de la biología molecular*, Salvat Editores, S.A., Barcelona, 1995, (rad. de Carlos Oppenheimer.]

[88] «Apañárselas mejor» significa «dejar más descendencia»; éste es el único criterio de éxito en cuanto a la evolución.

la población en general.

Pero no necesitamos limitamos exclusivamente a ejemplos hipotéticos, porque este mismo tipo de efecto se ha observado y medido en las propias poblaciones de pinzones que le ayudaron a Darwin a abrir los ojos y ver lo que sucedía. A principios de la década de 1970, un equipo formado por un matrimonio que trabajaba en la Universidad de Princeton, Peter y Rosemary Grant, iniciaron un largo programa de estudio de los pájaros de la isla de Dafne, en las Islas Galápagos. En esta isla vive una especie de pinzón de tierra, el *Geospizafortis*, que utiliza su fuerte pico para cascar semillas.

En 1977 hubo una gran sequía en la isla, durante la cual murieron más de 1.000 de los aproximadamente 1.200 pinzones a los que se hacía el seguimiento dentro del programa de estudio. Murieron porque desaparecieron muchas de las plantas de las que dependían para alimentarse y, a causa de esto, hubo una feroz competencia entre las aves para conseguir alimento. Los pinzones supervivientes eran todos ellos pájaros de gran tamaño dotados de grandes picos que pudieron comer semillas que los pájaros de menor tamaño no podían romper con el pico. Cuando la población se recuperó después de la sequía, con todos los nuevos pájaros descendientes de los que habían sobrevivido, el equipo de los Grant midió los picos de la nueva población de pinzones y descubrió que eran, por término medio, un 4 por ciento más largos que la media anterior a la sequía.

Ésta es la esencia de la evolución darvinista, y constituye un ejemplo muy bueno para explicar situaciones relativamente estables (como mucho, de evolución lenta) en las que están implicadas sólo unas pocas especies, como en el ejemplo clásico de la vida en las Islas Galápagos. Podemos llevar esta línea argumental hasta su límite lógico, considerando cómo este tipo de selección natural puede producir un empate completo, o equilibrio entre dos modos posibles de vida, siendo dicho equilibrio el resultado de lo que se llama una estrategia evolutivamente estable[89].

La idea resulta de la aplicación de las reglas matemáticas de la teoría de juegos a la biología evolutiva. Aunque su nombre suena a cosa agradable y puede aplicarse a juegos como el ajedrez o las damas, la «teoría de juegos» ha sido objeto de mucho interés, especialmente durante la guerra fría, por su importancia para los «juegos»

[89] *En el original, en inglés, ESS* = Evolutionarily Stable Strategy. (N. de la T.)

de guerra, situaciones hipotéticas en las que se comprueban las implicaciones de distintas estrategias militares y políticas. Resulta que estas técnicas son importantes para las competiciones evolutivas, al igual que lo son para la carrera de armamentos; uno de los más destacados defensores de este planteamiento para el estudio de la evolución es John Maynard Smith, de la Universidad de Sussex, del que procede el ejemplo que veremos a continuación.

Nos pide que imaginemos una población donde sólo hay una especie animal en la que cada miembro actúa de una o dos maneras características. Algunos son «halcones», lo cual significa que se comportan con mucha agresividad en sus disputas con otros miembros de su especie a la hora de conseguir alimento, mientras que otros son «palomas», porque se comportan de manera más sumisa en las mismas situaciones. Las auténticas palomas son, según reconoce Maynard Smith, unas aves bastante agresivas, pero, dado que a menudo se utilizan como símbolo de la paz, la elección del nombre es correcta en este contexto.

Cuando un halcón encuentra alimento y está presente otro miembro de la especie, siempre luchará por conseguirlo. Si el oponente resiste, uno de ellos finalmente se rendirá, salvo que sea una lucha a muerte, dejando el alimento al vencedor. Cuando una paloma encuentra algo que comer y está presente otro miembro de la especie, nunca iniciará una lucha. Si la atacan, emprenderá la huida inmediatamente; si no la atacan, hará un gesto de amenaza, pero cuando dos aves se hacen mutuamente estos gestos, una acabará rindiéndose y se retirará de la confrontación con más o menos donaire. Sin embargo, todo lo que hace la teoría de juegos es introducir números en tales situaciones, de tal forma que las diferentes estrategias puedan compararse de una manera cuantitativa y, como ejemplo, podemos decir que el valor del alimento para este animal es 50 unidades. Si lo ingiere, gana 50 puntos.

Si huye, no obtiene punto alguno, pero no tiene penalización por escapar. Si lucha por conseguir el alimento, puede ser herida, con lo que perderá 100 puntos, o puede vencer, en cuyo caso gana el premio de 50 puntos; si hace un gesto de amenaza antes de huir, o antes de ingerir el alimento, esto le cuesta 10 puntos.

Podemos estudiar las implicaciones empezando con una población totalmente formada por palomas. En cualquier confrontación, cada individuo pierde 10 puntos, pero uno de ellos gana 50 puntos (quedando un saldo de 40 puntos) si consume el

alimento. Por término medio, todo el mundo consigue 15 puntos en cada confrontación (40 menos 10, dividido por 2). Nadie resulta herido y todo el mundo consigue alimento. Se encuentran tan cerca de la utopía como lo permita este modelo.

Sin embargo, imaginemos ahora que, a causa de una mutación, surge un halcón en esa población de palomas. En todas las confrontaciones es él quien consigue el alimento, 50 puntos cada vez, sin coste alguno. Mientras haya sólo unos pocos halcones, las confrontaciones se realizan en la mayoría de los casos entre dos palomas, con una recompensa media de unos mismos 15 puntos, lista claro que los halcones prosperarán y tendrán muchos descendientes que compartirán con ellos las tendencias propias de un halcón.

Pero ¿qué sucede si todos son halcones? Un desastre total. Cualquier confrontación acaba en una pelea, en la que el ganador obtiene 50 puntos, mientras que el perdedor pierde 100 puntos, y la media es un desastroso -25. Una población así no tardaría en extinguirse, a menos que hubiera mucho alimento y pudiera conseguirse sin peleas. Pero, imaginemos que en esta situación surge una paloma mutante. Nunca llega a pelearse, por lo que en cualquier «conflicto» su puntuación es cero, aunque también ella consigue el alimento que está a libre disposición de todos. La paloma se las arreglará mejor que los halcones, y sus características de paloma se propagarán entre la población.

Los extremos nunca son estables, y desde cada uno de ellos hay una evolución hacia el término medio. Para esta elección particular de los números, el punto de equilibrio se alcanza realmente cuando hay cinco palomas por cada siete halcones. En esta situación, resulta que cada individuo consigue una media de 6,25 puntos por cada conflicto. El primer aspecto importante que observamos es que hay una situación estable, una estrategia evolutivamente estable, y que el proceso evolutivo lleva a la población a este punto. Sin embargo, es igualmente importante que éste no es el mejor de todos los mundos posibles.

Una puntuación media de 6,25 es notablemente menor que la de 15 que se obtiene en la sociedad formada íntegramente por palomas. La situación *estable* no es necesariamente la *mejor* situación, es decir, la que elegiría un colectivo razonable de individuos que actuaran para preservar lo mejor posible sus intereses mutuos.

Además, hay que reseñar otro aspecto más sutil, aunque intrigante, la situación es la misma cuando en la población hay cinco palomas por cada siete halcones y cuando cada miembro de la población actúa como una paloma durante cinco doceavas partes del tiempo y como un halcón durante las siete doceavas restantes, eligiendo aleatoriamente la estrategia que debe emplear en cada enfrentamiento, sin planificarla con antelación.

El planteamiento de la estrategia evolutivamente estable ha demostrado ser útil en la biología evolutiva, ya que proporciona información sobre las pautas reales de comportamiento observadas en algunas poblaciones. Pero, como el mismo nombre indica, se refiere a situaciones estables, en las que no se produce esencialmente cambio alguno, el equivalente evolutivo de la termodinámica clásica. Estas situaciones contienen un pequeño número de componentes que ejercen interacciones mutuas (en este caso, la estrategia del halcón es un componente, la estrategia de la paloma es el otro componente, y están conectadas mediante un vínculo que incluye una retroalimentación).

Hay ejemplos de estrategia evolutivamente estable algo más complejos, pero todos son como las sencillas redes formadas por unas pocas bombillas, vinculadas según las reglas de la lógica booleana, que se congelan en un solo estado, o recorren ciclos que atraviesan un número muy limitado y monótono de estados. No obstante, hay una posibilidad mucho más interesante, en la que un sistema que parece encontrarse en un estado estable, está en realidad cambiando constantemente, manteniéndose en equilibrio sólo porque ambos (o todos los) componentes evolucionan todo lo rápido que pueden con el fin de mantenerse a la par con los demás.

Esto guarda una analogía muy clara con la carrera de armamento nuclear de la guerra fría, asunto que no intentamos incluir aquí; en biología evolutiva, se conoce habitualmente como el «efecto de la Reina Roja», según el personaje que aparece en *Alicia en el país de las maravillas,* de Lewis Carroll, que debe correr tan rápido como pueda, con el fin de permanecer en el mismo lugar[90].

El mejor modo de ver de qué trata la hipótesis de la Reina Roja, de manera global, es tomar un ejemplo, Kauffman es especialmente aficionado a explicar esto

[90] La terminología de la «Reina Roja» fue obra del experto en biología evolutiva Leigh Van Valen, de la Universidad de Chicago, a principios de la década de los 70.

hablando de una especie imaginaria de rana que se alimenta de una especie imaginaria de mosca, y nosotros vamos a seguir este ejemplo. Son muchos los modos en que ejercen interacciones mutuas las ranas, que quieren comer moscas, y las moscas, que quieren evitar ser devoradas.

Las ranas pueden evolucionar para conseguir lenguas más largas que les permitan capturar más moscas; las moscas pueden evolucionar para volar más rápido y así poder escapar. Las moscas podrían evolucionar hasta adquirir un sabor desagradable, o incluso segregar venenos que dañen a las ranas, y así sucesivamente. Tomaremos una de estas posibilidades (hipotéticas).

Si una rana posee una lengua especialmente pegajosa, le resultará más fácil capturar moscas. Pero, si éstas tienen unos cuerpos especialmente escurridizos, les será más fácil escapar, incluso cuando las lenguas de las ranas lleguen a tocarlas. Imaginemos una situación estable en la que un cierto número de ranas vive en una charca y come cada año una cierta proporción de las moscas que hay por allí. En este caso, hay una especie de equilibrio, pero es inestable, no es una estrategia evolutivamente estable, como podemos ver si nos imaginamos que, a causa de una mutación (o tan sólo por las variaciones naturales de los individuos), una rana desarrolla una lengua excepcionalmente pegajosa.

Se las arreglará muy bien, en comparación con otras ranas, y los genes de las lenguas excepcionalmente pegajosas se propagarán por toda la población de ranas. Al principio, será devorada una gran proporción de moscas. Pero las que no resulten devoradas serán las más escurridizas, por lo que los genes que transmiten esta capacidad se propagarán por (oda la población de moscas. Después de un cierto tiempo, habrá en la charca el mismo número de ranas que antes, y cada año será devorada la misma proporción de moscas. Parece como si nada hubiera cambiado, pero las ranas tienen ahora lenguas más pegajosas y las moscas han conseguido unos cuerpos más escurridizos.

En el mundo real, este tipo de cosas sucede continuamente, en cada generación, no solo como resultado de mutaciones. En cada generación hay algunas ranas que tienen lenguas más pegajosas que otras, y algunas moscas cuyos cuerpos son más escurridizos que los de otras, de tal modo que, pasito a pasito, a medida que pasan las generaciones, la tendencia a tener lenguas más pegajosas y cuerpos más

escurridizos avanza sin cesar. A este proceso no se le ve un final obvio, las ranas se las arreglan cada vez mejor para atrapar moscas (de un modo u otro), mientras que las moscas son cada vez más hábiles escapando de las ranas. Cada especie ha de correr todo lo que pueden para seguir en el mismo lugar.

Por supuesto, en el mundo real hay más de dos especies en interacción mutua. Hasta cierto punto, esto quiere decir que hay muchos más procesos consistentes en correr sin moverse un ápice del lugar en que se está. Pero, al igual que sucedía con las conexiones establecidas dentro de las redes de bombillas, existen importantes diferencias entre unas redes y otras, según el modo en que las especies (que son en este caso como los nodos de la red) estén conectadas.

Si lo que le sucede a una especie afecta a todas las demás directamente, es imposible que se desarrolle ni siquiera una situación pseudoestable de Reina Roja. En el sistema hay un caos completo (pero determinista), porque incluso los pequeños cambios producen grandes efectos que, en la práctica, son impredecibles. Sin embargo, el mundo real no es así. En la vida real, lo que le sucede a una especie sólo afecta de forma directa a unas pocas especies más, aunque lo que les sucede a estas otras especies puede afectar a otras más, y así sucesivamente en una ola cada vez mayor de interacciones. Esto se parece mucho a lo que ocurre en una red de bombillas con escasas conexiones, o en el juego del montón de arena en el que el movimiento de un grano sólo afecta directamente a sus vecinos inmediatos, pero el modo en que éstos se mueven, afecta a los vecinos de cada uno, *y* así sucesivamente.

Tenemos otro ejemplo de una situación que se aproxima bastante a la vida real y que puede ayudarnos a aclarar las cosas. Los zorros comen conejos, por lo tanto, lo que les ocurre a los zorros (si su población crece o disminuye por algún motivo) afecta a los conejos, y lo que les sucede a éstos afecta directamente a los zorros. Hay nodos conectados en la red. Pero los zorros no comen hierba, por lo que a este nivel lo que le ocurra a los zorros no tiene incidencia alguna en la hierba. Los nodos de los zorros y los de la hierba no están conectados entre sí de manera directa.

Pero los conejos comen hierba, esos dos nodos están conectados—, por consiguiente lo que les sucede a los zorros, dado que afecta a los conejos, también produce algún efecto sobre la hierba. La existencia de más zorros (si lo demás

permanece igual) significa menos conejos, lo cual da a la hierba la oportunidad de crecer más. Esto último, a su vez, afectará a otras especies herbívoras, y así va pasando el efecto de unos a otros. Volviendo a las ranas de Kauffman, dado que éstas (o sus renacuajos) son devoradas por los peces que viven en la charca, y esos peces son devorados por los osos que viven en los bosques cercanos, si la población de moscas sufre un exterminio súbito (quizá porque se ha empezado a fumigar con insecticida en la zona), los osos pasarán hambre.

Las redes ecológicas reales, por supuesto, son más complejas que todo esto, y tienen algunas interconexiones (aunque no demasiadas), por ejemplo, los peces pueden comer moscas, además de ranas. Sin embargo, es interesante el hecho de que, coloquialmente, se aludiera a estos sistemas denominándolos «el entramado de la vida», incluso antes de que científicos como Kauffman aplicaran en este contexto la teoría de las redes.

Lo emocionante en todo esto es que, aparentemente, un proceso de coevolución, en el que todas las especies implicadas en una red sufren cambios cuando una de ellas cambia, impulsará de forma natural a los ecosistemas complejos desde los extremos hacia la interesante zona de la criticalidad autoorganizada, en la transición de las fases que se produce al borde del caos. Si un grupo de organismos está bloqueado en una estrategia estable, es probable que una mutación que afecte a una de las especies desbloquee la red, permitiendo su evolución. La evolución por selección natural garantizará que un cambio perjudicial para las especies implicadas vaya desapareciendo a lo largo de varias generaciones; pero todo cambio beneficioso se propagará, y al hacerlo, desbloqueará otras redes, impulsando el sistema hacia el borde del caos. En el otro lado de la transición de las fases, en el régimen caótico, sucederá lo mismo, pero a la inversa.

Dado que las reglas del juego de la Vida cambian con cada generación, cualquier grupo de individuos que consiga hasta cierto punto aislarse del caos, reduciendo el número de sus conexiones con el mundo exterior, tendrá una oportunidad de evolucionar por selección natural, hasta llegar a un estado que se beneficia de las oportunidades que hayan surgido. La lógica de este sencillo razonamiento está respaldada por las simulaciones numéricas y por los experimentos con racimos de bombillas conectadas mediante distintos cableados que han llevado a cabo

Kauffman y sus colegas. Sencillamente por el hecho de que cada especie actúe para maximizar su propia aptitud evolutiva, el ecosistema completo evoluciona hacia el borde del caos. Esto es, con toda exactitud, análogo al modelo del montón de arena, en el sentido de que el hecho de añadir a la red (o retirar de ella) una sola especie se corresponde con el hecho de añadir al montón de arena (o retirar de él) un solo grano; también es en esencia lo mismo que el modelo de las bombillas, y añadir o retirar una sola especie se corresponde con encender o apagar una sola bombilla. Los sistemas complejos evolucionan de forma natural hacia la transición de las fases al borde del caos.

Todo esto se traduce estupendamente a un lenguaje con el que ya estaban familiarizados los estudiosos de la biología evolutiva, y extiende ese modelo de una manera que los pioneros de esta teoría ni siquiera se hubieran atrevido a soñar, pero que debería resultar familiar a partir de lo que hemos aprendido sobre el modo en que la termodinámica clásica se amplió mediante los trabajos de científicos como Ilya Prigogine. Este modelo abarca lo que se conoce como paisajes adaptativos, y procede de los trabajos realizados por R. A. Fisher (1890-1962; se le concedió el título de «Sir» en 1952) en el Rothamsted Experimental Station, un centro de investigación británico, durante la década de 1920.

Fisher fue el primero que dotó de una base matemática a la biología evolutiva, haciendo los cálculos relativos al modo en que cambia la distribución de los genes en una población determinada, como resultado de la selección natural que actúa sobre los miembros de dicha población. Pero Fisher, al igual que los clásicos expertos en termodinámica, sólo estuvo interesado en las situaciones de equilibrio (un primer paso bastante razonable) y se dedicó a buscar la mejor situación matemática posible para cada problema evolutivo.

Este planteamiento tuvo un desarrollo posterior a manos del experto en genética Sewall Wright (1889-1988), que trabajaba en la Universidad de Yale a principios de la década de 1950. La analogía que presentó es la de un paisaje, donde las colinas representan las estrategias evolutivas acertadas («genes buenos» o, más exactamente, paquetes buenos de genes que actúan conjuntamente) y los valles indican estrategias evolutivas fracasadas («genes malos»). Cualquier individuo concreto puede estar representado mediante un solo punto del paisaje, que es

realmente un tipo de espacio de las fases, siendo las colinas los atractores. Sin embargo, dado que los miembros de una misma especie no son idénticos entre sí (recordemos que éste es uno de los puntos clave de la evolución darvinista), una especie estará representada en el paisaje mediante un pequeño conglomerado de puntos, como un pequeño rebaño de ovejas.

La oveja que se encuentra un poco más arriba en la pendiente tendrá más éxito en cada generación, dejando una progenie más numerosa, mientras que las que aparecen más abajo en la pendiente se las arreglarán peor, dejando una descendencia menos numerosa. Por lo tanto, a medida que avanzan las generaciones, el pequeño rebaño va ascendiendo por la pendiente hasta alcanzar el pico, y luego permanece en él.

Este planteamiento es ideal para entender la evolución cuando consideramos lo que le sucede a una sola especie que ejerce interacciones contándolo unas pocas especies más en un entorno estable, precisamente las condiciones que habían producido la variedad de pinzones que vio Darwin en las Islas Galápagos. Se cree que éstos son todos ellos descendientes de unos pocos pájaros (quizá una sola pareja) que llegó a las islas en algún momento del pasado.

En origen todos estos pájaros fueron miembros de una sola especie, que estaba escasamente adaptada para alimentarse con lo que podían encontrar en las islas, un pequeño rebaño situado en un valle entre las colinas del paisaje adaptativo. Algunos miembros del rebaño evolucionaron desarrollando unos picos más largos, subiendo así a una colina especial del paisaje, mientras otros evolucionaban adquiriendo picos más fuertes, y ascendían a una colina distinta, hasta que se convirtieron en especies diferentes. Este proceso finaliza cuando, en un caso, el pico llega a ser tan largo y delgado que ya no se puede utilizar para sondear en busca de alimento sin que se parta; o, por otro lado, cuando el pico se vuelve tan pesado y poco manejable que ya no se puede usar para manipular las semillas.

El paisaje adaptativo, apuntalado por las matemáticas de Fisher, presenta una imagen maravillosamente sencilla; incluso las fórmulas matemáticas son bastante sencillas. Pero, hay un problema. Supongamos que el pico al que ha llegado la especie en su evolución no es más que una colina al pie de una montaña y que hay picos mucho más altos (estados mucho más avanzados desde un punto de vista

evolutivo) que todo lo que pueden tener en las proximidades. No hay modo de que la especie cruce el valle que se interpone entre su colina y esos picos, porque para hacerlo tendría que retroceder en su evolución y volver a un estado peor (menos adaptado) que el que ya ha alcanzado.

En efecto, la especie ha llegado en su evolución a un callejón sin salida del que no hay modo de escapar. El paisaje de Fisher no es sólo estático, sino que es un estado en el que las especies sólo pueden lograr un equilibrio local, y han de permanecer en él.

El efecto de la Reina Roja es un ejemplo del modo en que las interacciones entre especies cambian esa simple imagen. Cuando una especie cambia, ya sea por mutación o, en el caso más extremo, por extinción, todas las especies con las que está en interacción resultan afectadas. El efecto es como si las alturas de los picos correspondientes del paisaje hubieran cambiado.

Cada una de las especies sigue evolucionando hacia los picos que tiene más cerca, pero, a causa de la coevolución, la posición y el tamaño de esos picos está cambiando continuamente. En el régimen estático, nada cambia; en el caótico, el paisaje cambia tan rápidamente que no hay tiempo para que suceda nada interesante. En el borde del caos, el paisaje está siempre cambiando, pero, generalmente, lo hace con lentitud, y esto abre continuamente nuevas posibilidades evolutivas para cada una de las especies y para los grupos de especies que hay en el ecosistema. Se trata de un paisaje infinitamente cambiante, una especie de paisaje de goma que se deforma como una cama elástica mientras las especies «se pasean» por su superficie[91], que proporciona la posibilidad de que se produzca una evolución infinita. En nuestro ejemplo de la rana y la mosca, la llegada de los seres humanos con bidones de insecticida echa abajo de golpe el pico en que se encontraban las moscas y las deja en la profundidad de un valle; pero, si no se extinguen todas las moscas, pueden evolucionar desarrollando una resistencia al insecticida y desplazándose a otro pico a medida que la población se recupera del retroceso.

Obsérvese que no se trata de un pico nuevo, estuvo ahí desde siempre, correspondiendo a la resistencia al insecticida. Sin embargo, el pico aumentó en

[91] En realidad, esta atractiva imagen está trastocada si la comparamos con los paisajes de Sewall Right, donde los *picos,* y no las pendientes, son los lugares en los que se espera encontrar a las especies; pero se capta la idea.

altura una vez que se hubo propagado el insecticida, ya que éste hizo que fuera más útil la capacidad de resistir a sus efectos. Aun así, las moscas no pudieron desplazarse a ese pico hasta que se hubo hundido el anterior. A diferencia de la situación esbozada por Fisher, aquí no hay un único «lugar óptimo» para que se asiente una especie en el paisaje evolutivo; todo lo que ésta puede hacer es desplazarse hacia un lugar mejor cuando surja la oportunidad de hacerlo, aunque lo que es mejor para una generación puede ser peor pasadas unas cuantas generaciones.

Obsérvese, sin embargo, que aquí no se trata ya únicamente de especies individuales que se desplazan hacia un estado mejor. John Holland, de la Universidad de Michigan, en Ann Arbor, ha llamado la atención sobre la importancia de las subunidades del ecosistema, formadas por especies distintas que funcionan conjuntamente en grupos. De hecho, esta imagen de grupos de individuos actuando conjuntamente es aplicable a todos los niveles, desde el de los genes, donde un grupo de ellos puede actuar como una unidad para controlar un aspecto del funcionamiento de la célula, pasando por el de las células, donde un grupo de ellas funciona conjuntamente formando un órgano, como en el caso del hígado, hasta llegar al de las especies, donde varias de ellas pueden asociarse en una red interactiva, siendo esta situación algo intermedio entre una estrategia evolutivamente estable y la independencia total.

Como dijo Holland, «El objetivo que persigue la evolución no es sólo formar un buen animal, sino encontrar unos buenos bloques de construcción que puedan ensamblarse para hacer muchos buenos animales.[92]»

Su analogía favorita es la del modo en que las imágenes del «Identikit» solían unirse, antes de la era de los ordenadores, para hacer un esbozo del rostro de un criminal sospechoso, utilizando descripciones de testigos. A partir de una gama de diez características faciales diferentes (forma del nacimiento del pelo, tamaño de la nariz, etc.), presentada cada una de ellas con diez variedades, resulta sorprendentemente rápido y fácil realizar un «retrato» que se parezca a cualquier individuo determinado, no sorprende tanto si recordamos que estamos manejando 10^{10} combinaciones, es decir, diez mil millones de caras diferentes, bastante más

[92] Citado en *Complexity*, de Milchell Waldmp.

que el número de personas que viven en la Tierra actualmente. Holland podría haber ido aún más lejos, al señalar que la evolución, en su pugna continua, actúa también para encontrar «buenos bloques de construcción» en forma de redes interactivas constituidas por especies que operan en algunos casos como unidades evolutivas diferenciadas. Se trata de la coevolución, no sólo de la evolución.

Aunque no lo expresó de este modo, era esto lo que Leigh Van Valen había descubierto, a través de sus estudios de fósiles marinos, que le llevaron en 1973 a desarrollar la teoría del efecto de la Reina Roja. Descubrió que cualquier género perteneciente a un grupo (por ejemplo, los peces óseos) tiene una probabilidad constante de extinguirse y desaparecer del registro fósil.

Con independencia de que el género haya existido durante mucho o poco tiempo, las probabilidades de extinguirse en cualquier intervalo dado de tiempo geológico son idénticas. Lo mismo se puede decir a otros niveles del entramado de la vida. En este sentido, a las especies no les va mejor en cuanto a supervivencia (a pesar de la evolución), ni les va peor a medida que pasa el tiempo; se extinguen de manera aleatoria. La interpretación que hizo Van Valen al respecto fue que la lucha por la supervivencia sigue siendo tan dura como lo era antes.

Dado que en esta lucha inciden las interacciones mutuas entre individuos, esto ha de entenderse como que todo el mundo se vuelve cada vez más eficiente, hasta que sucede algo que elimina una especie (o un género) del escenario. La consecuencia de ello es una reorganización de las especies en los nichos ecológicos disponibles, seguida de una nueva carrera de armamentos que produce al instante mucho más movimiento. Todas las especies deben evolucionar tan rápido como puedan, para no quedar rezagadas con respecto a las otras especies, tanto por defecto de «no ser menos que el vecino», como por el efecto de la Reina Roja.

Uno de los resultados de todas estas interacciones, que ha producido una viva emoción en algunos especialistas en biología evolutiva, es que las especies afortunadas consiguen evolucionar cada vez mejor a medida que pasa el tiempo, una mutación que le permita a una especie adaptarse con mayor rapidez a un entorno cambiante tiene en sí misma más probabilidades de éxito y de extenderse a toda una población. Un ejemplo de esto ha sido la evolución de la reproducción sexual, que es un modo de reproducción que acelera la evolución (porque mezcla

material genético de los dos progenitores en su descendencia dentro de una sola generación), dando a unas criaturas multicelulares, de gran tamaño y crecimiento lento, como nosotros mismos, una oportunidad de mantener posiciones en la guerra de armamentos en competición con organismos pequeños y de reproducción veloz, como las bacterias y los parásitos[93].

Ahora es preciso que hagamos un alto para explicar cómo utilizamos algunas palabras de la terminología evolutiva, con lo que es probable que aburramos a los expertos en biología evolutiva, pero los términos han de quedar claros para todo el mundo y ser aceptables fuera del ámbito académico. En gran parte de nuestra exposición hablamos de la extinción y la evolución de las especies, considerando que ambas cosas forman un todo. Por supuesto, somos muy conscientes de que la evolución opera al nivel del individuo y de que en la lucha por la supervivencia son los individuos peor adaptados los que fracasan (porque no dejan descendientes). Pero, si esto le sucede a un número suficiente de miembros de una especie, entonces ésta se extingue.

Parece razonable considerar que las especies evolucionan de este modo en conjunto, aunque sabemos que el mecanismo subyacente es la selección individual, e incluso hablar de que una especie «muta» de tal manera que se produce un cambio en un nodo de la red ecológica, generándose unas ondulaciones que afectan a otras especies. Después de todo, solemos hablar de un equipo de fútbol que compite por un trofeo como si habláramos de una sola entidad, «el equipo local quedó eliminado de la copa»—, aunque sabemos que el equipo está formado por once jugadores en un momento dado, y un número mayor durante toda la temporada, sin olvidar que depende también de la habilidad del entrenador y de otros individuos que no juegan en el campo.

Ahora se trata de entender por qué es razonable, cuando se habla de los cambios evolutivos observados en el registro geológico, describirlos abreviadamente aludiendo a «mutaciones» o extinciones de especies completas, en vez de centramos en la selección de los individuos. Para que esto quede claro, tomaremos un ejemplo muy sencillo que refleja el modo en que los cambios evolutivos afectan a los distintos miembros de una especie. Esto sucede de tal manera que los cambios

[93] Véase The Mating Game, de John Gribbin y Jeremy Cherfas (Penguin, 2001).

son demasiado pequeños para poder ser percibidos en el lapso de tiempo que dura una vida humana, pero pueden producir efectos que se manifiestan como cambios repentinos y drásticos en el registro fósil, cuando una especie es reemplazada por otra en lo que la geología considera «de la noche a la mañana». Nuestro ejemplo favorito es el modo en que este posible proceso afecta a un ratón hipotético que evoluciona lentamente para convertirse en algo así como un elefante.

Imaginemos que una cría de ratón crece 1/20.000 de la diferencia entre su tamaño y el de un elefante, y que el mismo aumento de tamaño se produce en cada una de las generaciones siguientes. Por consiguiente, tendrían que pasar 20.000 generaciones para que el ratón «evolucionara» hasta convertirse en un superratón, del tamaño de los elefantes actuales. Los ratones se crían muy rápidamente, y los elefantes muy lentamente, con lo cual, por lo que respecta a este hipotético ejemplo, podemos decir que cada generación tarda cinco años en llegar a la madurez, tomando un valor en cierto modo intermedio entre el tiempo de crianza de los ratones y el de los elefantes. Según esto, el cambio de ser un ratón a convertirse en un superratón duraría 100.000 años.

En relación con la escala de tiempo de los fósiles, este cambio sería instantáneo. Datar de forma diferenciada los fósiles que han sido depositados desde hace 100.000 años es casi imposible y, si tuviera lugar en una especie este tipo de cambio evolutivo, lo que veríamos en el registro fósil sería una capa de estratos con restos del tamaño de un ratón y, en la capa siguiente, restos del tamaño de un elefante, sin nada entre ambas capas de estratos.

En el mundo real, un cambio tan drástico no se produciría de manera aislada; también otras especies se verían afectadas por la evolución del superratón (suponiendo que tal cambio fuera factible) y veríamos en el registro fósil toda una serie de cambios que se habrían producido de forma «instantánea». Pero lodo ello sería el resultado de lentos cambios que se producirían de una generación a la siguiente, según las reglas de la evolución darwinista.

Es importante recalcar esto, porque ha habido un vivo debate entre los expertos en torno a lo que los teóricos de la evolución llaman «equilibrio discontinuo», una pauta observada en el registro fósil cuando hay largos intervalos en los que el cambio evolutivo es escaso o nulo (salvo para las especies que lleguen a estar

mejor adaptadas a sus nichos ecológicos a través del efecto de la Reina Roja), con lo que el proceso es «discontinuo» en el sentido de «interrumpido» por breves intervalos (en ocasione^ aparentemente instantáneos) de cambios drásticos en los que algunas especies se extinguen y otras evolucionan hacia nuevas formas. A veces, incluso los que deberían conocer mejor el tema, presentan esto como una teoría que entra en conflicto con las teorías darvinistas de la evolución, a las que aluden entonces calificándolas de «gradualismo» para marcar el contraste.

Pero no existe conflicto alguno. Incluso durante los intervalos de cambio aparentemente rápido (los signos de puntuación del proceso evolutivo), lo que está sucediendo a escala individual sigue siendo la evolución darvinista, una evolución por selección natural y supervivencia de los más aptos. Lo que sucede es simplemente que han cambiado las reglas relativas a qué es lo que hace apta a una especie. Y el cambio se produce siempre lentamente de una generación a la siguiente, un ratón no puede parir un elefante, ni tampoco a la inversa. En realidad, lo que les sucedió a los pinzones de las Islas Galápagos que estudió Darwin, y que le dieron la idea clave sobre cómo funciona la evolución, es un ejemplo de equilibrio discontinuo.

Los pinzones que llegaron a las islas estaban magníficamente adaptados a un estilo de vida que era el adecuado en el continente. Evolucionaron rápidamente (para lo que son los baremos geológicos, aunque sólo experimentaron unos cambios muy pequeños de una generación a la siguiente) para encajar en los nichos ecológicos que existían en las Islas Galápagos, y, una vez que se adaptaron, cambiaron sólo ligeramente para adaptarse todavía mejor a su entorno. Darwin se hubiera sentido perfectamente cómodo con la teoría del equilibrio discontinuo y hubiera opinado que sus teorías no estaban en absoluto amenazadas por dicha idea[94].

La analogía del montón de arena contribuye a aclarar esta cuestión. Cuando el montón se encuerara en un estado crítico autoorganizado, experimenta repentinos corrimientos de tierra, intercalados con largos intervalos de calma relativa, mientras los granos de arena se van acumulando. Sin embargo, cada grano de arena está sometido continuamente a las mismas leyes de la física, las leyes del movimiento, del rozamiento y de la gravedad de Newton, entre otras. No existen diferencias

[94] En una escala temporal adecuada, pueden tenerse en cuenta también los acontecimientos asociados con la sequía de 1977 en las Islas Galápagos como un ejemplo de equilibrio interrumpido.

fundamentales entre las dos situaciones, excepto una relacionada con la escala temporal. Si echamos granos sobre el montón cada segundo, pero sólo lo comprobamos cada media hora, nunca tendrá el mismo aspecto y parecerá que se encuentra en un estado de cambio constante.

Si lo miramos cada microsegundo, casi cada vez que lo miremos veremos los granos de arena casi exactamente en las mismas posiciones en que estaban la última vez que los miramos, y pensaremos que el montón sólo experimenta cambios graduales. Si la situación de las especies vivas de una red ecológica permanece aproximadamente igual durante cien millones de años, y luego se producen los cambios al cabo de más o menos un millón de años, lo que tenemos es un equilibrio discontinuo, incluso cuando las especies implicadas pudieran no ser conscientes de ninguno de los cambios drásticos que se producen en su entorno.

Si los intervalos de cambio abarcan cien millones de años, y los de calma duran un millón de años, podríamos hablar de «caos discontinuo», pero esta situación cumplirá las mismas reglas evolutivas que se aplican a las redes escasamente conectadas. La cuestión es que las redes en las que hay pocas conexiones evolucionan de manera natural hacia un estado de este tipo, en el que hay intervalos de equilibrio interrumpidos por otros de cambio, aunque nuestro conocimiento de estas redes es todavía demasiado inadecuado para poder decir si hay algo especial en relación con el balance concreto entre discontinuidad y equilibrio que parece haber sido alterado por la red de la vida en la Tierra. Sin embargo, tenemos una idea clara de cuál es el último ingrediente que se necesita para llevar esa red de la vida a la transición de las fases que se produce al borde del caos, un aspecto de menor calibre, pero importante, que a Kauffman le pasó desapercibido, pero que Per Bak y sus colegas descubrieron a mediados de la década de 1990.

En todos los primeros trabajos de este tipo, realizados utilizando modelos mediante ordenador para simular la evolución de redes de especies que ejercen interacciones mutuas, los investigadores habían elegido aleatoriamente una especie para que «mutara», y observaban el modo en que las ondulaciones resultantes se propagaban por la red como una avalancha de granos en el modelo del montón de arena. Esto produjo simulaciones que, vistas superficialmente, se parecían mucho al

modelo del montón de arena, pero que en realidad, no evolucionaban hacia el borde del caos. El avance decisivo llegó cuando Kim Sneppen, del Instituto Niels Bohr, de Dinamarca, visitó a Bak en Brookhaven, Estados Unidos, en 1993. Sneppen estaba interesado por las interacciones entre superficies, un ejemplo que Bak utiliza es el modo en que el café se extiende empapando una servilleta de papel, siendo la superficie de contacto, en este caso, la cambiante frontera entre las zonas mojada y seca de la servilleta.

Sin embargo, los cálculos de Sneppen son aplicables a una amplia gama de fenómenos físicos; este científico descubrió que los cambios en la frontera, o superficie de contacto, tienen más probabilidades de producirse en aquellos lugares en que el valor de algún parámetro adecuado alcanza su máximo o su mínimo, una situación denominada «dinámica extrema». En el ejemplo concreto (casi trivial) del café derramado, el líquido se extiende con mayor facilidad a través de la servilleta en aquellos lugares en que se encuentran los poros más anchos, porque éstos permiten que el líquido fluya más fácilmente.

En las conversaciones mantenidas entre Bak y Sneppen sobre el fracaso en los intentos de conseguir que las simulaciones de redes ecológicas evolucionasen hacia la transición de las fases, seleccionando aleatoriamente las especies para su mutación, ambos constataron que sería más apropiado aplicar también aquí alguna forma de dinámica extrema. Después de todo, incluso en el modelo arquetípico del montón de arena, no se desencadenan nuevas avalanchas en lugares aleatorios, sino en el punto en que la pendiente de la arena es más pronunciada; por otra parte, los terremotos siempre comienzan en el punto en que la tensión que se ha ido generando supera primero un cierto umbral crítico, aunque el alcance del terremoto resultante (como la amplitud de la avalancha resultante en el modelo del montón de arena) esté luego determinado por otros factores.

Los casos extremos dentro de las redes ecológicas que han estado estudiando científicos como Kauffman y Bak son las especies menos adaptadas a su entorno. ¿Qué sucedería si el equipo de Bak volviera a poner en marcha las simulaciones, pero seleccionara sólo la especie menos adaptada para hacerla cambiar a cada paso de la simulación? Sería un caso de pura evolución darvinista, la extinción del menos apto. Pero no tiene que ser necesariamente una extinción; puede ser sólo un

cambio. En cualquier ámbito ecológico real, las especies peor adaptadas no se quedan impasibles con su mala adaptación; o bien se extinguen, o evolucionan para encajar mejor en sus entornos (como los pinzones de Darwin).

Las especies mejor adaptadas cambiarán poco (salvo por el efecto de la Reina Roja), a menos que suceda algo en sus entornos, una rana que esté extraordinariamente bien adaptada para devorar moscas se verá en un serio problema si, por ejemplo, las moscas se extinguen. En el caso de las especies menos aptas, cualquier cambio que no las lleve a la extinción será casi con toda seguridad una mejora, que las enviará a una posición más elevada en las estribaciones de las colinas que pueblan el paisaje evolutivo. Si una de estas especies parte de una posición situada en una colina de poca altura, es relativamente fácil para ella que evolucione para salir de esa colina, cruzar el val le y ascender a una colina cercana más alta. Todo esto es lógico; y resultó aún más lógico cuando se pusieron en marcha las simulaciones por ordenador.

El modelo que idearon Bak y Sneppen es asombrosamente semejante. Sólo toma un millar de «especies» y representa cada una de ellas mediante un único número elegido al azar entre 0 y 1, que indica su aptitud darvinista. Si el valor es alto, corresponde a una especie que se adapta bien al conjunto de la red y está situada en un pico elevado dentro del paisaje adaptativo. Un número bajo corresponde a una especie que está realmente luchando por la supervivencia y se encuentra en una posición fluctuante sobre una colina baja del paisaje adaptativo. En este sentido, la cantidad de «adaptación» se corresponde con el tamaño del pico del paisaje adaptativo.

Con el fin de tener en cuenta las escasas conexiones existentes entre los nodos de la red, se asoció aleatoriamente cada uno de los números resultantes con otros dos, de tal modo que cada «especie» actúa de manera recíproca sólo con otras dos especies (como si las ranas únicamente comieran moscas, y ellas mismas sólo fueran devoradas por unos peces). Después, el modelo se programó para eliminar al miembro menos adaptado de la población, *y a los dos elementos relacionados con él, y* reemplazarlos por tres «nuevas» especies cuyas adaptaciones se asignaban al azar. No es necesario que estas especies sean nuevas en sentido literal; la cuestión es que la adaptación del miembro peor adaptado de la red cambie de manera

aleatoria, y que esto haga que la adaptación de cada una de las especies asociadas, podríamos pensar en ellas como si fueran sus vecinas en la red ecológica, cambie también aleatoriamente, con independencia de lo bien o mal adaptadas que estén al principio. Lo que entra en acción es el paisaje de goma. No hay más que repetir esta sencilla acción reiteradamente en el ordenador y observar qué sucede.

Al principio, dado que se ha constituido la red al azar, hay nodos con valores de adaptación que cubren todo el intervalo de 0 a 1. Por consiguiente, la «especie» peor adaptada tendrá un valor de adaptación muy bajo. Pero, cuando la especie peor adaptada se elimina y se sustituye por otra que tiene una adaptación asignada al azar, es altamente probable que la nueva especie tenga un valor de adaptación más elevado, porque es probable que un número aleatorio situado entre 0 y 1 sea mayor que el número muy bajo que estaba asociado con la especie eliminada.

Sólo con eliminar la especie «menos apta» y sus dos vecinas en cada iteración, y sustituirlas por nuevas especies cuyo valor de adaptación se ha establecido al azar, estas simulaciones realizadas mediante ordenador elevarán muy rápidamente la adaptación global de toda la red, hasta que todos los nodos estén asociados con picos relativamente altos. De forma específica, según las reglas que hemos especificado aquí, el sistema se asienta en un estado en el que todas las especies tengan una adaptación de, al menos, dos tercios. En la vida real, esto corresponde a una red ecológica estable, en la que cada especie ocupa un alto pico en el paisaje adaptativo. Tal situación debería durar mucho tiempo, pero, ¿qué sucede cuando una de las especies se extingue, o, a causa de una mutación, sufre una pérdida de adaptación, que afecta al paisaje de sus dos vecinas? En este caso, el único camino posible es descendente.

Reemplazar tres especies con valores de adaptación relativamente altos por especies a las que se ha asignado valores de adaptación aleatorios es más apropiado que producir un descenso en la adaptación global de este pequeño racimo de nodos. Esto significa que las tres nuevas especies (o especies que han sufrido una mutación), y sus vecinas, son las que tienen mayor probabilidad de verse afectadas por las iteraciones subsiguientes, generándose una oleada de extinciones (o, al menos, de reducciones de la adaptación) que se propagan desde el lugar donde se ha producido originalmente el cambio. Recordemos que una reducción de

la adaptación no significa necesariamente que una especie ha cambiado; es el paisaje adaptativo el que cambia, de tal forma que una especie muy bien adaptada a una situación pasa a encontrarse en otra situación diferente.

Una especie de avalancha, o terremoto, atraviesa el sistema, antes de que éste se ajuste y se reestructure en el estado en el que todos los niveles de adaptación son superiores a dos tercios, y el sistema pueda estabilizarse para permanecer en calma durante otro intervalo de tiempo. Este modelo produce, de manera natural, una situación en la que los intervalos de estabilidad están separados por extinciones masivas, aunque sean aplicables las mismas reglas durante los intervalos de calma y las extinciones. No deberíamos sorprendemos al saber que el resultado viene dado por una ley potencial, lo cual significa que las extinciones son autosimilares, aunque el valor exacto de la ley potencial (la inclinación de la pendiente en el gráfico de la ley potencial) no encaja del todo con el de la ley potencial de las extinciones reales que se observan en el registro fósil.

En las críticas realizadas con respecto a este trabajo se ha argumentado que los modelos son simplemente demasiado sencillos para resultar realistas, o para permitir que nos hagamos una idea de las complejidades de los ecosistemas reales y del modo en que evolucionan. Esto no es apenas sorprendente, dado lo sencillos que son los planteamientos utilizados por Kauffman, Bak y sus colegas. Pero estas críticas olvidan el hecho de que los modelos sencillos ofrecen una buena descripción del mundo real. Si estos modelos no lograran producir las pautas de extinciones que describen las leyes potenciales, nadie se los tomaría en serio. La prueba que debe superar un buen modelo no es la relativa a lo sencillo que es, sino hasta qué punto sirve para hacemos una idea de lo que sucede en los sistemas reales. En la física atómica, por ejemplo, resulta a primera vista casi ridícula la simplificación consistente en tratar a los átomos como si fueran sistemas solares en miniatura, con los electrones girando «en órbita» alrededor de un núcleo central.

Sabemos que los átomos son más complicados que todo eso. Pero este sencillo modelo, desarrollado por el danés Niels Bohr durante la segunda década del siglo XIX, funciona en realidad muy bien para predecir las longitudes de onda exactas de las líneas que aparecen en los espectros correspondientes a los distintos elementos. Por lo tanto, es un buen modelo, aunque sabemos que hay átomos que no son

«realmente» así. Un modelo que pudiera determinar, no sólo que las extinciones obedecen a una ley potencial, sino también el valor exacto de la potencia según esa ley, sería tan valioso en su propio contexto como el modelo atómico de Bohr, y nos diría asimismo una profunda verdad sobre la sencillez fundamental del mundo. Para describir otras características más complicadas del comportamiento de los átomos basta con añadir más detalles al sencillo modelo de Bohr, sin descartar la información que proporciona.

El hecho de que modelos sencillos como los de Kauffman y Bak reproduzcan de manera inmediata las características esenciales de las extinciones que tienen lugar en el mundo real, así como su invariabilidad con respecto a la escala de los fenómenos, sugiere también que los investigadores van por el buen camino y que deberían ser capaces de encontrar una correspondencia más ajustada entre los modelos y el mundo real sin más que añadir unos pocos detalles para hacer que los modelos sean más realistas. De hecho, resulta que añadir tan sólo un detalle más es suficiente para conseguir con exactitud la ley potencial correcta.

Todavía queda mucho trabajo en fase de realización y, probablemente, lo más honrado sea decir que en la primera década del siglo XXI sabemos tanto (o tan poco) sobre las redes evolutivas reales como Bohr sobre los átomos reales en la segunda década del siglo XX. Pero uno de los refinamientos más intrigantes de este tipo de modelos llegó justo a finales del siglo XX, en 1999, cuando Luis Amaral, que trabajaba en el MIT, y Martin Meyer, de la Universidad de Boston, introdujeron en el modelo un detalle crucial (y, desde luego, importante).

Añadieron una sencilla representación de otra profunda verdad relativa al mundo real, algunas especies devoran a otras. Amaral y Meyer añadieron depredadores y presas al juego inventado por Bak y Sneppen. Una vez más, estamos ante una versión simplificada del modo en que funciona el mundo real, pero esta vez se ha hecho una modificación que mejora radicalmente la correspondencia entre el modelo y la realidad.

En esta variación sobre el mismo tema, hay seis niveles en la cadena alimenticia. Las especies de cada nivel pueden (de hecho, deben) alimentarse de varias especies diferentes pertenecientes al nivel inferior, y en cada nivel hay un millar de nichos ecológicos equivalentes a los nodos de la versión del juego ideada por Bak y

Sneppen. Cuando el juego comienza, estos nichos ecológicos están en su mayoría vacíos, pero hay unas pocas especies asignadas aleatoriamente a los del nivel más bajo de la cadena. Las reglas, aplicadas como es habitual en cada iteración, funcionan de la siguiente manera: en primer lugar, hay una pequeña probabilidad de que cada una de las especies presentes al principio de cada paso iterativo se divida en dos (como la división de los primeros pinzones de las Islas Galápagos, que dio lugar a distintas especies), produciendo una nueva especie que es asignada al azar a uno de los nichos vacíos del mismo nivel, o de un nivel por encima o por debajo del correspondiente a la especie progenitora.

A cada una de las especies que surgen de esta manera se le asignan unas pocas especies del nivel inferior que son sus presas. En segundo lugar, en cada paso de la iteración se extinguen unas pocas especies del nivel más bajo elegidas al azar. La diferencia con el modelo de Bak y Sneppen es que aquí, cuando una especie desaparece de esta manera, no afecta a sus vecinas, que, por otra parte, son sus iguales, sino a los depredadores que se alimentan de ella. Los depredadores que pierden una parte mayor de su provisión de alimento se las arreglan peor que aquellos que pierden una parte menor y, en casos extremos, pueden desaparecer todas las presas de un depredador, con lo que éste se extingue, afectando su desaparición a los depredadores que se encuentran situados en el nivel siguiente de la cadena alimenticia.

Según una pauta que ahora ya nos resulta familiar, en la que funcionan sólo estas sencillas reglas y unas condiciones iniciales aleatorias, el juego se organiza por sí mismo en un estado que corresponde al borde del caos, un estado crítico en el que la desaparición de tan sólo una de las especies que actúan como presas y se encuentran en la parte más baja de la cadena (quizá sencillamente porque han sido devoradas todas ellas por los depredadores de los niveles situados por encima) puede a veces desencadenar una oleada de extinciones que se desplaza como una onda hacia arriba, atravesando las capas de depredadores y presas. No es sólo que volvamos a encontramos de nuevo con que la pauta de estas oleadas es invariable con respecto a la escala y sigue una ley potencial; esta vez, el valor exacto de la potencia según dicha ley es el mismo que el de la potencia de la ley que rige las extinciones reales de vida en la Tierra. Las especies que ejercen interacciones

mutuas en una red ecológica tienden hacia un estado crítico autoorganizado en el que las extinciones masivas que siguen una ley potencial se producirán todas ellas sin ayuda de intervención exterior alguna.

Sin embargo, sabemos que hay intervenciones externas, el ejemplo clásico es el del meteorito que chocó con la Tierra en la ó poca de la extinción de los dinosaurios. A pesar de que dijimos, como hipótesis, que ésta no fue la causa fundamental de las extinciones que se produjeron en la frontera K-T, ciertamente divo que tener algún efecto sobre la vida en la Tierra. Está claro que es importante averiguar cómo podría tal influencia externa afectar a la pauta de las extinciones, pero esto añade otro nivel de complejidad a los cálculos.

Como es habitual, el primer impulso de los creadores de modelos es descartar todo lo que no necesitan incluir en éstos y partir del modelo nuevo más sencillo posible para comenzar a comprender el «nuevo» fenómeno. En este sentido, Mark Newman, de la Comell University, decidió a mediados de la década de 1990 investigar un modelo en el que las influencias externas fueran la *única* causa de las extinciones, ignorando los efectos de las interacciones que se producían entre las distintas especies en el paisaje adaptativo de goma. Lo que se espera es que, examinando tanto las extinciones causadas por interacciones entre especies en ausencia de impactos exteriores, como las ocasionadas por impactos exteriores sin interacciones entre las especies, podamos ver qué versión está más cerca de la realidad y cómo podrían combinarse los dos efectos.

La estructura básica del modelo de Newman es muy similar a la del modelo de Bak y Sneppen. La primera diferencia es que, en vez de eliminar únicamente la especie que está «menos adaptada que todas las demás», el modelo elige al azar un nivel de adaptación (entre 0 y 1) y elimina de golpe *todas* las especies que tienen una adaptación menor que este valor, equivalente a un acontecimiento externo, como un meteorito que choca con la Tierra, que hace desaparecer todas las especies, salvo las mejor adaptadas—. Entonces, los nichos ecológicos vacíos se llenan con unas nuevas especies cuya adaptación se elige aleatoriamente (como siempre entre 0 y 1).

A continuación, para tener en cuenta el hecho de que las reglas del juego han cambiado (el paisaje evolutivo ha quedado modificado), unos pocos supervivientes,

las especies más aptas de la red ecológica original, se eliminan y se sustituyen por otros cuya adaptación se ha elegido de manera aleatoria. Esto es equivalente a alterar la adaptación de las especies que ocupaban originalmente estos nichos.

Como una opción final para establecer las reglas de cada reproducción del modelo que hacía el ordenador, era posible seleccionar por adelantado el tipo de impactos externos que los investigadores querían aplicar, podría haber una proporción relativamente elevada de pequeños impactos que sólo eliminaran las especies menos aptas, o unos pocos impactos de gran magnitud que barrieran casi todo lo que había en la red; o bien el modelo podía estructurarse justo al contrario, o con cualquier mezcla estadística de impactos externos grandes y pequeños, aunque en cada iteración el modelo elige al azar el tamaño del impacto siguiente de acuerdo con las reglas estadísticas especificadas. El sorprendente y drástico resultado fue que en todos los casos, salvo en unas pocas versiones extremas de este juego, el sistema se instala en un estado crítico en el cual la pauta de extinciones cumple una ley potencial similar a la que observamos en la pauta de extinciones del registro fósil real[95].

Un modo de imaginarse esto es pensar que, al igual que las tensiones se acumulan en una zona sísmica hasta que algo cede a la presión (sin que podamos decir con antelación qué dimensiones tendrá eso que «cede»), aquí es como si la evolución darvinista gradual generara cada vez más tensión en la red ecológica, por el efecto de la Reina Roja, hasta que algo se rompe y toda la red (o quizá sólo una parte de ella) se colapsa. No hay manera de predecir con antelación si la oleada siguiente de extinciones que va a barrer la red ecológica será grande o pequeña, del mismo modo que no hay manera de predecir si el próximo terremoto que sacuda San Francisco será intenso o leve. Tanto los sistemas vivos como los de la materia inerte siguen el mismo imperativo profundo.

Una vez más, el modelo es muy sencillo y no podemos pretender que describa perfectamente lo que le sucede a una red ecológica real cuando el desastre golpea desde fuera. Sin embargo, lo que nos dice es que la pauta de extinciones según una ley potencial es una característica muy marcada del modo en que los impactos

[95] Nuestros amigos matemáticos insisten en que señalemos que aquí no se trata de un «estado crítico» acorde con su estricta definición del término; sin embargo, se corresponde estrechamente con el tipo de estados críticos que hemos descrito hasta ahora, y esta sutileza no afecta a lo que estamos explicando.

afectan a estas redes. Nos indica, en primer lugar, que es absolutamente irrelevante qué tipo de influencia externa se ejerza, siempre que ésta ocasione muertes. Los pequeños impactos frecuentes, los grandes impactos ocasionales, o cualquier combinación de ambos (incluidos los pequeños impactos ocasionales) producen la misma pauta. Por lo tanto, una mezcla de distintos sucesos naturales, meteoritos que chocan contra la superficie terrestre, estallidos de intensa actividad volcánica, glaciaciones o cualquier otra cosa—, cada uno con su propia pauta, se combinará para producir exactamente la misma imagen de extinciones que siguen una ley potencial.

Aún más interesante es que, como ya hemos visto, los meros cambios evolutivos que se producen en la red ecológica, *sin* impactos externos de ningún tipo, dan lugar a la misma pauta de ley potencial. Todo esto explica por qué es tan sencilla la pauta real que observamos en el registro de restos fósiles distribuidos a través del tiempo. Todas las causas, aun siendo diferentes, producen la misma pauta. Es posible que nunca seamos capaces de decir si los sucesos acaecidos en la frontera K-T fueron originalmente el resultado del impacto de un meteorito, o de un cambio evolutivo, o de una actividad volcánica importante, pero sólo algún tipo de combinación de estas tres causas (y posiblemente de otras) proporcionó el detonante de aquella oleada de extinciones. Como en la novela de Agatha Christie *Asesinato en el Orient Express,* cualquiera puede ser culpable.

Desde la perspectiva de los biólogos, o de los geólogos, o de los astrónomos, que esperan todos ellos encontrar la prueba que demuestre la validez de sus hipótesis favoritas, esto podría ser frustrante; pero, desde el punto de vista de alguien que intenta comprender cómo evolucionan las complejas redes de formas de vida que ejercen interacciones mutuas, esto resulta muy interesante, e incluso estimulante. Sugiere que, a pesar de la extrema sencillez de todos estos modelos, hemos descubierto una verdad fundamental con respecto al modo en que trabaja la naturaleza, una verdad aplicable a cualquier sistema que ha sido sacado de su estado de equilibrio, incluidos los sistemas concretos que nos interesan aquí, es decir, los sistemas vivos. Con una gama de posibilidades extremadamente amplia, cualesquiera que sean las condiciones de las que partamos y los impactos que apliquemos a los sistemas vivos (externos, internos o ambos), llegamos al estado

crítico autoorganizado que se produce al borde del caos, donde incluso un pequeño detonante puede, en ocasiones, producir un cambio muy amplio dentro del sistema en su conjunto.

Como dice Newman, poniendo todo el énfasis en ello, cuando cambia el entorno físico, cambia asimismo el paisaje adaptativo. Pero también la vida en sí misma modifica este paisaje, y el destino de cada uno de los componentes de la vida en la Tierra en cualquier momento de su historia depende del modo en que las dos influencias ejerzan interacciones mutuas. Un ejemplo excelente es el que nos proporciona la mayor de todas las extinciones masivas, la que ocurrió a finales del pérmico. Hemos visto que incluso un pequeño detonante podría haber sido responsable de aquella gran oleada de extinciones; sin embargo, esto sólo habría sido posible si las especies de la época en cuestión hubieran estado en su mayoría conectadas dentro de un entramado ecológico.

Hoy en día, si algún tipo de desastre afectara a América del Norte (y recordemos que en este contexto la extinción de una sola especie podría causar un desastre y desencadenar una oleada de extinciones), los efectos podrían extenderse en una serie de ondas tanto por el norte como por el sur de América. Sin embargo, es difícil imaginarse cómo podrían afectar seriamente a las especies que viven en Australia. África o Eurasia. Esta es una de las razones pollas cuales se suele recurrir a mencionar acontecimientos verdaderamente globales, tales como las consecuencias del impacto de un meteorito gigante, para intentar explicar las extinciones globales masivas. Sin embargo, al final del pérmico, casi toda la masa terrestre de nuestro planeta estaba unida en un supercontinente que se ha bautizado con el nombre de Pangea. Esta situación favoreció la posibilidad de que prácticamente todo tipo de vida de la superficie terrestre y de los mares poco profundos que bordeaban aquel supercontinente formara parte de una única red. Un impacto que afectara a una parte de la red (un pequeño impacto procedente del exterior, o una especie que se extinguiera únicamente a causa de las presiones evolutivas) podía haberse propagado como una ola que tragaría a casi todos los seres que estuvieran vivos en aquella época.

No hay pruebas de que fuera esto lo que sucedió. Nunca podría haberlas. Pero, no es ésta la cuestión. Lo interesante es que hace 250 millones de años, al final del

pérmico, un pequeño desencadenante *podría* haber causado la extinción de la mayoría de las especies que existieran en aquella época, mientras que un desencadenante igualmente pequeño que afectara en la actualidad a un solo continente no podría producir una oleada global de extinciones igual de grande. El entorno físico y el entorno biológico están interconectados mucho más sutilmente que lo que parece a primera vista.

Pero ¿podría esta conexión actuar al revés? ¿Podría el entorno biológico de la Tierra afectar asimismo al entorno físico, de tal modo que los dos juntos formaran parte de una única red sometida al mismo conjunto de reglas profundamente sencillas? Esta especulación es el núcleo de la «hipótesis de Gaia», que presentó el científico británico independiente Jim Lovelock.

La teoría de Lovelock, que condujo finalmente a la hipótesis de Gaia, se remonta a 1965, y antecede a gran parte de los trabajos sobre la complejidad y el surgimiento de la vida que hemos estado comentando. Sin embargo, al tener en cuenta también las teorías de Lovelock, el estudio del caos, de la complejidad y del surgimiento de la vida se amplía hasta abarcar todo nuestro planeta y aporta ideas sobre el modo en que puede haber evolucionado la vida en mundos que se encuentren más allá de nuestro sistema solar.

Capítulo 7
Vida más allá

La gran diferencia entre el modo en que hemos observado la vida hasta ahora y el modo en que lo vamos a hacer de ahora en adelante, es que anteriormente la hemos mirado desde dentro hacia afuera, mientras que ahora vamos a mirarla desde fuera hacia el interior. Podemos estudiar el modo en que ciertas moléculas, como los aminoácidos y el ADN, ejercen interacciones mutuas en el interior de las células, y cómo las células trabajan conjuntamente para formar un cuerpo, con lo que aprenderemos mucho sobre la manera de actuar de la vida. O también podemos observar a un ser humano (o a un perro, o a una medusa) mientras hace su vida cotidiana, y obtener así una perspectiva diferente de lo que es la vida.

El cambio de perspectiva, que nos proporciona nuevos puntos de vista sobre la naturaleza global de la vida en la Tierra, llegó como resultado de dos cosas: una imagen espectacular y el trabajo de un hombre, dos aspectos que están relacionados con la exploración del espacio. La imagen en cuestión fue una fotografía tomada por los astronautas del Apolo, donde se podía ver la Tierra como nuestro hogar en el espacio, un único oasis de vida azul y blanco rodeado por un desierto negro.

El hombre fue Jim Lovelock, al que se le ocurrió la idea de que los componentes vivos y no vivos del entorno terrestre ejercen interacciones mutuas en una red que mantiene condiciones estables y adecuadas para la vida en nuestro planeta. Su teoría estaba basada directamente en los principios de la termodinámica y en las diferencias entre los sistemas que están en equilibrio y los que no lo están, cuestión esta última que ya hemos comentado, aunque, dada la formación de Lovelock, esta idea surgió de una manera bastante independiente a partir del trabajo que ya hemos descrito, relativo a la aplicación de la termodinámica de los estados no equilibrados a situaciones en las que participan sistemas que se encuentran al borde del caos.

Lovelock nació en 1919 y, después de abandonar la Universidad de Manchester en 1941 con una licenciatura en química, pasó dos décadas dedicado a la investigación médica, consolidando su formación como un experimentador «práctico» que

diseñaba y construía gran parte de sus instrumentos. Fue a principios de la década de 1960 cuando empezó a dar sus primeros pasos hacia una independencia por la que más tarde sería famoso, consiguiendo financiación al cabo de un tiempo a partir de los instrumentos científicos que él mismo había inventado y siendo libre para investigar cualquier tema que le gustara, aunque siempre había tenido una mentalidad independiente e inconformista, incluso en comparación con los niveles de excentricidad que se dan a veces entre los investigadores científicos.[96]

Entre los primeros pasos hacia esa independencia cabe mencionar sus trabajos para la NASA y el Jet Propulsión Laboratory (JPL) como asesor en el diseño de instrumentos para las naves espaciales que iban a aterrizar en la Luna (alunizar) y en Marte (amartizar), y para analizar las muestras de materiales que dichas naves encontraran allí. Después de contribuir en todo lo que pudo a los aspectos químicos del programa, Lovelock se acercó a los ingenieros que diseñaban los aparatos informáticos, convirtiéndose en un vínculo entre ellos y los biólogos que planificaban la búsqueda de pruebas de la existencia de vida en Marte, ya que podía hablar los lenguajes de ambos equipos.

La mayoría de los lectores de este libro probablemente no recordará el estado primitivo que presentaban los aparatos electrónicos domésticos en la década de 1960, cuando un técnico podía ganarse bien la vida reparando televisores. La idea de fletar este tipo de aparatos hacia el espacio, hacer que se posaran sobre la superficie de Marte, y esperar que funcionaran y remitieran información útil a la Tierra mediante un transmisor cuya potencia no pasaba de 100 vatios (la de una bombilla corriente), le parecía a la mayoría de la gente una idea fantástica.

Como dijo Lovelock, la ingeniería necesaria estaba tan lejos de la que se utilizaba en un televisor doméstico de la época, como éste podía estar de la ingeniería que se conocía en tiempos de los romanos. Sin embargo, funcionó; pero la cantidad de información que podía enviar aquel débil transmisor tenía un límite infranqueable, por lo que fue una cuestión de prioridad urgente decidir qué era exactamente lo que se iba a medir, si la sonda amartizadora (o *lander)* conseguía llegar a la superficie de Marte.

La idea clave que le hizo llegar a definir el concepto de Gaia surgió, según recuerda

[96] La deliciosa autobiografía que Lovelock escribió, Homage to Guia, explica los avatares de su formación con una honestidad a veces penosa, pero siempre total.

Lovelock, en septiembre de 1965. Por aquel entonces, estaba ya instalado de nuevo en Inglaterra, pero seguía visitando regularmente el JPL como asesor[97].

Durante una de aquellas visitas, en 1964, asistió a una reunión en la que el equipo discutía el tipo de experimentos que habría que incluir en la sonda amartizadora con el fin de buscar vestigios de vida en Marte. Le chocó el hecho de que todas las ideas que allí se discutían estuvieran basadas en el supuesto de encontrar la típica forma de «vida tal como la conocemos», como dijo Lovelock, los experimentos habrían sido útiles para hallar vida en el desierto de Mohave, situado justo al este de las instalaciones del JPL, cerca de Los Angeles, pero parecía que nadie consideraba la posibilidad de que la vida en Marte pudiera ser completamente diferente de la que existe en la Tierra en condiciones extremas. Lovelock sugirió que lo que se necesitaba era un experimento mediante el cual se pudiera buscar los atributos generales de la vida, no tipos específicos de vida, y, cuando le pidieron que explicara con qué clase de experimento podía hacerse aquello, replicó que lo que se necesitaba era un experimento que buscase la disminución de la entropía.

Tal como Lovelock lo entendía, y como ya hemos visto anteriormente en este libro, algo característico de los seres vivos es que aportan orden local a los sistemas, haciendo que la entropía «retroceda» mientras ellos tengan una fuente externa de energía para alimentarse. Siguiendo la mejor tradición del «puedes hacerlo», típico de aquellos primeros tiempos de la exploración espacial, le dieron a Lovelock un par de días para que encontrara una idea factible con la que llevar a cabo un experimento para la medición de la disminución de la entropía; «exponlo o cállate» que es realidad lo que aquellos científicos le dijeron.

Una vez que su mente se concentró en el problema, éste resultó sorprendentemente fácil de resolver. El mejor modo de buscar procesos de disminución de la entropía que estuvieran funcionando en Marte sería medir la composición química de su atmósfera. Si no hubiera vida en Marte, los gases de su atmósfera se encontrarían eh un estado de equilibrio termodinámico y químico dominado por componentes estables tales como el dióxido de carbono. Si hubiera

[97] Aunque Lovelock se convirtió en un investigador independiente que trabajaba desde su residencia de la campiña inglesa, en la década de 1970 también fue profesor invitado honorario en la Universidad de Reading. Este cargo no remunerado le proporcionó lo que él llamaba «una cobertura respetable», que le ayudó a conseguir que sus trabajos se publicaran en revistas cuyos editores no se fiaban de las misivas enviadas desde domicilios particulares; también le permitió establecer contactos con otros investigadores y estudiantes con los que pudo contrastar sus teorías.

vida, entonces los productos de desecho de los procesos vitales se verterían a la atmósfera marciana, aportando a ésta gases reactivos tales como el metano y el oxígeno, que disminuirían la entropía de la atmósfera.

Había otras posibilidades, algunas de las cuales explicó Lovelock posteriormente en un trabajo publicado en la revista *Nature,* incluida la posibilidad de detectar y analizar los sonidos de la atmósfera de Marte. Como ya hemos visto, los sonidos emitidos por seres vivos contienen información (entropía negativa) y se caracterizan como ruido 1/*f*, que es bastante diferente del ruido blanco de las fluctuaciones aleatorias. Ya fuera el equivalente marciano del canto de un pájaro, o del chirrido de los grillos, o de la música de Mozart, un sencillo análisis de las pautas sonoras podría revelar la presencia de vida. Estas y otras ideas dejaron impresionados a los planificadores del JPL (aunque muchos de los biólogos no se impresionaron lo más mínimo).

Los planificadores de la NASA se quedaron igualmente impresionados y, ante el asombro de Lovelock, le nombraron director científico en funciones, por lo que pasó a ser el encargado de desarrollar aquellos experimentos físicos de detección de vida para una misión en Marte que ya estaba propuesta. Era demasiado bueno para ser verdad. En septiembre de 1965 los fondos para dicha misión no obtuvieron la aprobación del Congreso de Estados Unidos, y el proyecto Viking, que era menos ambicioso, tras haber dejado instrumentos sobre la superficie de Marte en 1975, terminó con el conjunto estándar de experimentos biológicos, algunos de ellos basados en los diseños de Lovelock, los biólogos deseaban utilizar su pericia técnica, aunque desconfiaran de sus puntos de vista relativos a la biología.

Cuando le preguntaron cómo se sentía después de ver que se dedicaban tantos esfuerzos a un proyecto cuyo resultado era una conclusión prevista de antemano, respondió que se sentía como si estuviera «diseñando un robot para buscar señales de vida en el desierto del Sahara y lo hubiera equipado con una caña de pescar». Sabemos que los peces existen y constituyen ejemplos de sistemas vivos, por lo que encontrar un pez en el Sahara sería la prueba de que hay vida en el desierto. Sin embargo, los resultados negativos de los experimentos biológicos realizados mediante las sondas Viking revelaron realmente poco o nada sobre la presencia de vida en Marte, salvo que no es el tipo de vida que puede encontrarse en el desierto

de Mohave, del mismo modo que el fracaso de una expedición de pesca al Sahara lo único que demuestra es que no hay peces en el desierto, no que no haya vida. Hasta ahora, la mejor prueba de que no hay vida en Marte no la han aportado las sondas espaciales, sino unos estudios espectroscópicos de la composición de la atmósfera marciana, se da la coincidencia de que dichos estudios espectroscópicos se dieron a conocer por primera vez en septiembre de 1965, justo diez años antes de que' las sondas Viking llevaran a la superficie de Marte sus «cañas de pescar» maravillosamente diseñadas, pero inútiles.

Aquel mismo mes, Lovelock se encontraba en el JPL cuando I legaron noticias de que unos astrónomos del Observatorio del Pie du Midi, en Francia, habían obtenido una información espectroscópica detallada sobre la atmósfera de Marte a partir del análisis de la luz emitida por el planeta en la parte del espectro correspondiente a los rayos infrarrojos. La espectroscopia es un proceso en el cual se analiza la luz procedente de un objeto, descomponiéndola en el espectro del arco iris mediante un prisma o algún otro instrumento, y examinando luego las líneas del espectro producidas por los diferentes átomos o moléculas del objeto que se está estudiando. Estas líneas se parecen a la pauta de rayas que hay en un código de barras y son igual de específicas; una pauta determinada revelará sin ambigüedades la presencia de hierro, por ejemplo, mientras que otras (más importantes en este contexto) están asociadas con el oxígeno, o con el metano, o con el dióxido de carbono. Los resultados obtenidos en el observatorio francés mostraron por primera vez que la atmósfera de Marte está constituida, casi en su totalidad, por dióxido de carbono, con tan sólo unos vestigios de la presencia de otros gases. Es una atmósfera estable, no reactiva, que se encuentra en un estado de alta entropía correspondiente a un equilibrio termodinámico, y, como ya hemos visto, en un equilibrio termodinámico no sucede cosa alguna que pueda resultar interesante.

El mismo mes en que se esfumaron las esperanzas de Lovelock de enviar un experimento a Marte para medir la entropía, estos resultados del observatorio francés pusieron de manifiesto que dicho experimento no era necesario. Las mediciones que podían ser interesantes ya las habían hecho otros por él, y demostraban sin ambigüedades que Marte es hoy en día un planeta muerto (por supuesto, no decían nada sobre las condiciones que podrían haber existido en Marte

hace millones, o miles de millones, de años).

Fue la llegada de aquellas noticias procedentes de Francia lo que hizo que Lovelock se pusiera a reflexionar sobre el contraste existente entre la atmósfera de Marte, copada por el dióxido de carbono, y la atmósfera terrestre, en la que predominaba el nitrógeno, pero había alrededor del 21 por ciento de oxígeno, uno de los gases más reactivos, y sólo vestigios de dióxido de carbono, que no es reactivo. Lovelock sabía que el nitrógeno de la atmósfera participa constantemente en reacciones con el oxígeno (una especie de combustión lenta), para formar en última instancia ácido nítrico, que se disuelve en el mar produciendo nitratos estables, que luego las bacterias descomponen (utilizando la energía de la luz solar) para devolver nitrógeno al aire. «De repente, como una ráfaga de inspiración», le llegó a Lovelock la idea de que, para que la atmósfera de la Tierra permaneciera en un estado aparentemente estable durante cientos de millones de años, «algo debía estar regulando la atmósfera y manteniéndola con una composición constante. Además, si la mayoría de los gases procedían de organismos vivos, debía ser la vida existente en la superficie la que se encargaba de hacer esa regulación»[98].

Sin pararse a reflexionar más sobre ello, planteó de buenas a primeras este punto de vista ante sus compañeros de aquel entonces, un colega de la NASA llamado Dian Hitchcock, y otro visitante del JPL, el astrónomo Cari Sagan. Ésta fue la semilla de la que creció la idea de Gaia, es decir, de la Tierra como un sistema autorregulador (el nombre lo sugirió, dicho sea de paso, uno de los vecinos que tenía Lovelock en Inglaterra, el escritor William Golding).

Por enfocarlo de una manera ligeramente diferente, digamos que nos podríamos preguntar qué le sucedería a todo el oxígeno altamente reactivo de la atmósfera terrestre si no fuera renovado constantemente por la acción de los seres vivos. Si eliminámos toda la vida que hay en el planeta, en muy poco tiempo la totalidad del oxígeno quedaría bloqueada dentro de compuestos químicos estables, tales como los nitratos, el dióxido de carbono, el agua, los óxidos de hierro y las rocas silíceas. Dicho de una forma más precisa, sin la intervención de la vida, todo el oxígeno de la atmósfera quedaría bloqueado en menos de diez millones de años.

Esto indica lo sensible que es el entorno físico aparentemente estable de nuestro

[98] Véase *Homage to Gaia* de Lovelock.

planeta a la presencia (o ausencia) de vida. El asunto no resulta demasiado preocupante a una escala de tiempo humana, el mito popular según el cual, si mañana desapareciera la selva amazónica, nos asfixiaríamos todos, está lejos de ser verdad, pero diez millones de años representa sólo alrededor del 0,2 por ciento de la antigüedad de la Tierra hasta el momento presente. Si un astrónomo que está observando un planeta como la Tierra constata que dicho planeta posee una atmósfera rica en oxígeno, esto significa que, o bien está siendo testigo de un suceso raro y transitorio que tiene lugar en ese planeta, o que la atmósfera se mantiene en un estado que se encuentra lejos del equilibrio.

La idea de que la vida puede formar parte de un sistema autorregulador que determina la naturaleza física de la superficie actual de la Tierra (al menos en la «zona de la vida», una fina capa que va desde el fondo del océano hasta la parte más alta de la troposfera, es decir, hasta unos 15 kilómetros por encima de nuestras cabezas) fue recibida inicialmente de manera hostil por los biólogos, y aún tiene oponentes, muchos de ellos desanimados por lo que consideran (equivocadamente) insinuaciones místicas, cuasi religiosas.

Existe también un movimiento místico, cuasi religioso, a favor de Gaia (más o menos tan irritante para Lovelock como lo fue la Tolkien Society para J. R. R. Tolkien), que se fundamenta en una mala interpretación de lo que Lovelock y sus colegas decían. Además, están aquellos que entienden de manera totalmente equivocada de qué trata el asunto. Mi copia de la versión para CD de la *Encyclopaedia Britannica,* que en realidad debería estar mejor informada que otros, dice qué «la hipótesis de Gaia es muy discutible porque da a entender que cualesquiera especies (por ejemplo, las antiguas bacterias anaerobias) podrían sacrificarse a sí mismas en beneficio de todos los seres vivientes». ¡Desde luego no es así! Esta afirmación tiene más o menos la misma lógica que decir que la teoría de Darwin es muy discutible porque sugiere que los conejos se sacrifican a sí mismos en beneficio de los zorros.

Quizá tengamos que explicar que Lovelock *no* dice que Gaia sea Dios, ni que la Madre Tierra cuide de nosotros, ni que una especie haga sacrificios por el bien de todos. La verdad es sencillamente que Lovelock encontró una manera de describir todos los procesos relativos a la vida que tienen lugar en la Tierra, incluidos muchos

que tradicionalmente se han considerado procesos físicos no relacionados con la vida, como parte de una compleja red de interacciones, un sistema autorregulador (o autoorganizador), que ha evolucionado hasta llegar a un estado interesante, pero crítico, en el cual se puede mantener el equilibrio durante períodos de tiempo que resultan muy largos con respecto a los estándares humanos, pero en el que pueden ocurrir unas fluctuaciones repentinas que lo aparten del equilibrio (análogo al equilibrio discontinuo de la evolución biológica).

En el lenguaje de lo tratado en el capítulo anterior, lo que Lovelock dice es que el comportamiento de la vida en la Tierra altera el paisaje físico (en el que el término «físico» incluye cuestiones tales como la composición de la atmósfera) y también el paisaje biológico, y que ambos cambios afectan de manera global al paisaje adaptativo, siendo la retroalimentación un componente clave de las interacciones. En el contexto general de las teorías que comentamos en este libro, el concepto de Gaia es una consecuencia lógica y directa de lo que hemos expuesto con anterioridad; la sorpresa, que indica el poder de su intuición, es que Lovelock comenzó a desarrollar el concepto de Gaia antes de que muchas de las teorías que hemos mencionado aquí hubieran llegado a formar parte de la ciencia respetable, por lo que sólo de manera retrospectiva podemos ver ahora que todo ello encaja conjuntamente, y podemos pensar (haciéndonos eco del comentario de Thomas Henry Huxley cuando conoció por primera vez a través de Darwin la teoría de la evolución por selección natural) «qué estupidez más grande que no se me haya ocurrido a mí todo esto».

No hay necesidad de contar con todo detalle la historia completa de cómo Gaia llegó a ser ciencia respetable, pero, aprovechando esa mirada retrospectiva, podemos tomar dos ejemplos del funcionamiento de la teoría de Gaia, uno de ellos un modelo teórico y el otro tomado del mundo real, que muestran de qué modo se produce la autorregulación a partir de la interacción entre los componentes biológicos y físicos de un planeta vivo. El primero, un modelo llamado «Daisyworld» («Un mundo de margaritas»), es especialmente apropiado ya que se construye directamente a partir de un enigma que Sagan le planteó a Lovelock poco después de que éste tuviera su ráfaga de inspiración en el JPL, y además el modelo resuelve este enigma; es también un claro ejemplo de la idea principal de nuestro libro, la del

surgimiento de la vida, considerando que el total es mayor que la suma de las partes. Y Lovelock dice que es «el invento del que me siento más orgulloso».

El enigma que Daisyworld resuelve se conoce entre los astrónomos como «la paradoja del joven Sol que palidece», aunque en realidad sólo era un enigma, no una paradoja, y, gracias a Lovelock, ahora ya ni siquiera es un enigma. El enigma procede del hecho de que los astrónomos pueden decir que el Sol emitía mucho menos calor cuando era joven que en el momento actual.

Han llegado a saber esto combinando informaciones relativas a interacciones nucleares obtenidas en experimentos realizados en la Tierra, simulando mediante el ordenador las condiciones existentes en el interior de las estrellas, y comparando los resultados de sus cálculos con informaciones sobre la emisión de energía y la composición de estrellas de diferentes tamaños y edades, obtenidas mediante espectroscopia. Este es uno de los grandes logros de la física del siglo XX (en gran medida no difundido en público), pero aquí no tenemos espacio para entrar en detalles, y el lector sencillamente tendrá que aceptar por ahora que los astrónomos no saben cómo funcionan las estrellas[99].

Lo importante es que podemos decir con seguridad que, cuando el sistema solar era joven, el Sol estaba entre un 25 y un 30 por ciento más frío que en la actualidad, o, por decirlo de otra manera, desde que se asentó como una estrella estable, la emisión de energía procedente del Sol ha crecido entre un 33 y un 43 por ciento. El sistema solar se estabilizó en lo que es más o menos su configuración actual hace aproximadamente 4.500 millones de años y sabemos, por las pruebas que aportan los fósiles hallados en las rocas más antiguas que se encuentran en la superficie terrestre, que el agua en estado líquido y la vida existían ambas en la superficie de nuestro planeta hace 4.000 millones de años.

El enigma es por qué el aumento de la emisión de calor procedente del Sol, aproximadamente un 40 por ciento durante 4.000 millones de años, no hizo hervir el agua de la superficie terrestre, secándola y dejándola sin rastro de vida.

No hay problema alguno para explicar por qué la Tierra no era una bola de hielo cuando el Sol estaba más bien frío. Sabemos ahora que en la atmósfera de Venus, como en la de Marte, predomina el dióxido de carbono, y este compuesto, junto con

[99] La historia completa se explica en mi libro *Stardust*

el vapor de agua, es una parte importante de los gases liberados por la actividad volcánica.

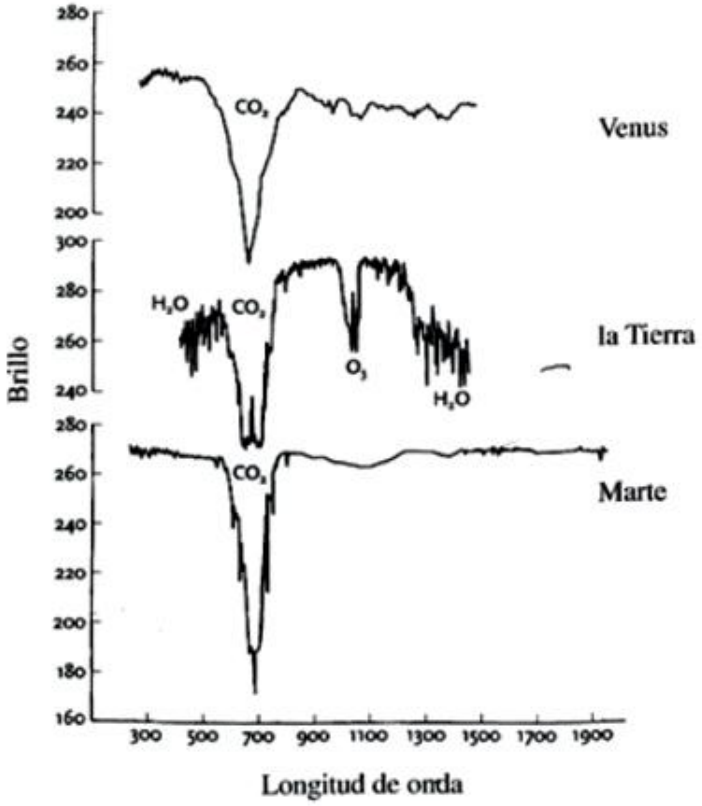

Figura 7.1. Comparando los espectros de Venus, la Tierra y Marte, se pone de manifiesto inmediatamente cuál de los tres planetas se encuentra en un estado no equilibrado y tiene probabilidades de constituir un hogar adecuado para la vida.

No hay razón alguna para pensar que la atmósfera de la Tierra en los primeros tiempos fuera, de algún modo, diferente de las atmósferas de sus dos vecinos planetarios más próximos, y una atmósfera rica en dióxido de carbono sería buena para captar el calor procedente del Sol en las proximidades de la superficie del planeta, manteniéndola caliente por el llamado efecto invernadero. Aunque ésta no es la razón principal por la que el aire de un invernadero se mantiene caliente (es en gran medida porque el tejado de cristal impide que el aire caliente se eleve por convección y escape), el efecto invernadero es fácil de entender. La luz visible que llega del Sol pasa a través de gases tales como el dióxido de carbono (y, de hecho,

a través del oxígeno y del nitrógeno) sin ser absorbida, y calienta la superficie de la Tierra.

Esa superficie caliente irradia la energía devolviéndola al espacio, pero con longitudes de onda mayores, situadas en la zona de infrarrojos del espectro. El dióxido de carbono (y, de hecho, también el vapor de agua, pero no el nitrógeno o el oxígeno) absorbe parte de esta radiación infrarroja, y luego irradia la energía que ha absorbido, enviándola en todas las direcciones, volviendo parte de ella a la superficie para calentar el planeta, mientras otra parte se libera hacia el espacio.

El resultado global es que, visto desde el exterior, hay un descenso en el espectro de la luz procedente de la superficie del planeta en las longitudes de onda donde la radiación está siendo absorbida en la atmósfera. Un astrónomo que se encontrara en Marte, provisto de un telescopio y un espectrómetro de sensibilidades adecuadas, podría asegurar, midiendo esta radiación infrarroja característica, que había un rastro de dióxido de carbono en la atmósfera de la Tierra, del mismo modo que el equipo de científicos del Pie du Midi detectó dióxido de carbono en la atmósfera de Marte. Pero la proporción de dióxido de carbono es en la atmósfera de la Tierra mucho menor que en la de Marte.

La potencia del efecto invernadero se puede ver contrastando la temperatura media que se da en la actualidad en la superficie terrestre con la de la Luna, que no tiene aire, aunque está prácticamente a la misma distancia del Sol que nosotros. En la superficie lunar la temperatura media es de -18 °C; en la Tierra, la temperatura media de la superficie es de 15 °C. La diferencia, un incremento de 33 °C, se debe a la presencia de tan sólo un 0,035 por ciento de dióxido de carbono en la atmósfera, además del vapor de agua y vestigios de gases como el metano, que son también gases de efecto invernadero. Más del 99 por ciento de nuestra atmósfera, el nitrógeno y el oxígeno, no contribuye en absoluto al efecto invernadero. Si, hoy en día, una proporción tan pequeña de gases de efecto invernadero puede producir un efecto tan grande, es fácil entender por qué la temperatura de la joven Tierra nunca bajó del punto de congelación, aunque el Sol fuera tenue.

Pero, la Tierra no sólo no se ha congelado cuando era joven, tampoco lo ha hecho a medida que su edad ha ido haciéndose más avanzada, y realmente ha mantenido una temperatura de una uniformidad notable durante miles de millones de años,

aunque el Sol se ha vuelto cada vez más caliente.

Es bastante sencillo imaginarse diversos modos en los que la temperatura del planeta ha podido mantenerse constante gracias a cambios en la composición de la atmósfera; científicos como Carl Sagan habían formulado ya varios débiles razonamientos siguiendo estas líneas, antes de que Lovelock presentara su concepto de Gaia.

Sin embargo, lo que les faltaba a estos argumentos era cualquier tipo de justificación que no fuera el deseo de los teóricos de mantener una temperatura constante en la Tierra. ¿Qué proceso natural podía conducir a esa estabilidad? Nadie lo sabía. Entonces, ¿era sólo cuestión de suerte? Cualesquiera que sean las razones, lo que hemos de hacer, por supuesto, es reducir de manera continuada, a medida que la Tierra envejece, la cantidad de gases de efecto invernadero que hay en la atmósfera para contrarrestar el calor cada vez mayor que irradia el Sol. Es fácil hacerse una idea general del modo en que podría haber sucedido esto.

Las primeras formas de vida terrestre basadas en la fotosíntesis (aquellas «antiguas bacterias anaerobias» a las que se refiere la *Encyclopædia Britannica)* habrían tomado dióxido de carbono del aire y lo habrían utilizado para formar sus cuerpos, pero habrían emitido metano al aire, con lo que el dióxido quedaría sustituido por otro gas, también de efecto invernadero, pero con unas propiedades de absorción de infrarrojos distintas de las del dióxido de carbono. Cuando estas bacterias son más activas, el equilibrio se descompensa a favor del metano; cuando son menos activas, el equilibrio se descompensa a favor del dióxido de carbono.

La clave para empezar a comprender cómo podría funcionar esto en la naturaleza fue la introducción de una percepción retrospectiva en los cálculos. Con un sencillo modelo que tenía en cuenta la creciente producción de calor del Sol, Lovelock pudo mostrar que, si se permite que las bacterias aumenten a una velocidad máxima cuando la temperatura es de 25 °C, pero con menos rapidez a temperaturas superiores o inferiores, y en ningún caso cuando la temperatura baja de los 0 °C o supera los 50 °C, se podría mantener una temperatura constante durante más o menos los primeros mil millones de años de la historia de la Tierra. Entonces, se ponían en marcha otros procesos, especialmente el surgimiento de formas de vida que emitían oxígeno al aire, donde este elemento reaccionaría con el metano para

eliminar de la red este componente, y también la disminución gradual de las concentraciones de dióxido de carbono a través de los tiempos.

Se puede hacer funcionar todo ello de una manera plausible, pero las críticas a este planteamiento señalaban que dependía en gran medida de la percepción retrospectiva y que quedaba lejos de estar claro dónde operaban los vínculos entre las especies vivas y el entorno físico, a muchos les parecía como si el carácter constante de la temperatura de la Tierra pudiera haber sido cuestión de suerte, y no de que la vida estuviera manipulando el entorno físico para su propio provecho (a través de retroalimentaciones naturales, no de una manera consciente). Aquí es donde entra en escena Daisyworld.

Inicialmente Daisyworld fue un modelo desarrollado por Lovelock y sus colegas a principios de la década de 1980 y, desde entonces, ha cobrado vida por sí mismo (quizá adecuadamente), con variaciones sobre el tema que han sido desarrolladas por varios científicos, e incluso, en la década de 1990, se ha integrado en un juego de ordenador llamado SimEarth. Fue en diciembre de 1981 cuando Lovelock tuvo por primera vez esta idea, como respuesta a las críticas que recibieron las primeras versiones de la teoría de Gaia. Presentó el modelo en una conferencia pronunciada en los Países Bajos en 1982; posteriormente tomó como ayudante a Andrew Watson, uno de sus antiguos estudiantes de doctorado de Reading, para que colaborara en una versión más matemática del modelo, que se publicó en la revista *Tettus* en 1983.

Daisyworld comienza como un planeta igual que la Tierra, pero sin vida, que recorre una órbita alrededor del Sol, a la misma distancia que lo hace la Tierra. En las versiones más sencillas del modelo, la superficie del planeta es principalmente tierra firme, con el fin de ofrecer un lugar donde puedan crecer las margaritas, y la composición de la atmósfera se mantiene constante, por lo que hay un efecto invernadero constante. Las margaritas se presentan en dos colores, blancas o negras, y crecen cuando la temperatura es de 20 °C. Les va proporcionalmente peor cuando la temperatura desciende por debajo de este valor óptimo, y no pueden crecer por debajo de 5 °C; también les va peor en proporción cuando la temperatura asciende por encima del valor óptimo, y no consiguen crecer por encima de 40 °C.

El modelo se pone a funcionar cuando la temperatura del Sol virtual aumenta lentamente del mismo modo que lo hacía el Sol real en su juventud. Una vez que la temperatura en el ecuador de la Tierra del modelo alcanza los 5 °C, se diseminan semillas de margarita de ambas variedades por su superficie y se deja que actúen por su cuenta, con la condición de que se reproduzcan de verdad, de tal modo que las margaritas blancas tengan siempre descendencia blanca y las margaritas negras produzcan siempre otras también negras.

Como ya sabe cualquiera que haya subido a un coche negro que ha estado aparcado al sol un día de verano, los objetos de colores oscuros absorben el calor del Sol con mayor eficacia que los objetos de colores claros. Por lo tanto, un macizo de margaritas negras absorberá calor y calentará la pequeña superficie en la que se encuentre, mientras que un macizo de margaritas blancas reflejará el calor y refrescará la tierra sobre la que está plantado.

Mientras Daisyworld está fresco, las margaritas negras tienen una ventaja, ya que calientan su entorno, llevando la temperatura a un valor cercano al óptimo, y crecen. En las generaciones siguientes, las margaritas negras se propagan por la superficie del planeta a expensas de las blancas, de tal modo que todo el planeta se vuelve más eficaz para absorber el calor procedente del Sol, y su temperatura asciende aún más rápidamente que si lo hiciera sólo como resultado del aumento de la temperatura del Sol. Sin embargo, una vez que la temperatura supera los 20 °C en cualquier lugar de la superficie terrestre del modelo, son las margaritas blancas las que tienen ventaja, porque al refrescar la superficie hacen que la situación vuelva a tender a la temperatura óptima.

Aunque la temperatura del Sol continúe aumentando, dado que ahora las margaritas blancas se propagan a expensas de las negras, la temperatura del planeta se queda rondando los 20 °C hasta que toda la superficie planetaria queda cubierta de margaritas blancas. Entonces, como la temperatura del Sol sigue aumentando, las margaritas lo tienen cada vez más difícil, hasta que la temperatura alcanza los 40 °C y mueren todas. La gama total de producción de energía solar que cubre esta versión del modelo va desde el 60 por ciento hasta el 140 por ciento de la producción actual de energía de nuestro Sol.

El efecto global es que durante un largo período de tiempo, aunque la producción de

calor del Sol del modelo aumenta, la temperatura de la Tierra del modelo no sólo permanece constante, sino que se mantiene en la temperatura óptima para la vida, sin que las margaritas hagan ninguna planificación consciente, y sin indicios de que otra clase de margaritas «se esté sacrificando a sí misma» en beneficio de todos los seres vivos.

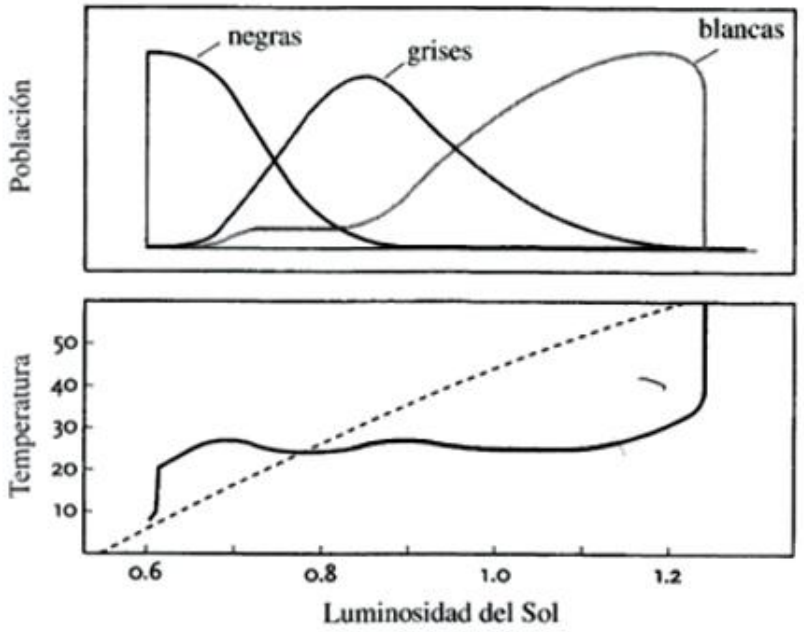

Figura 7.2. En un modelo de Daisyworld en el que aparecen tres especies de margaritas y un Sol que aumenta constantemente su brillo, las poblaciones de las distintas margaritas cambian a medida que pasa el tiempo (gráfico superior), pero la temperatura que hay en la superficie del planeta permanece casi constante (gráfico inferior), incluso cuando la luminosidad del Sol aumenta drásticamente. La línea discontinua indica cómo aumentaría la temperatura sin la influencia de las margaritas.

Ambas variedades actúan sólo en su propio interés. Pero, ¿puede un sistema muy sencillo, como es éste, ser realmente representativo del modo en que la naturaleza actúa en realidad? Una de las críticas que se le hicieron a este modelo fue que podría ser susceptible de «manipulación tramposa». Tanto las margaritas blancas como las negras han de invertir parte de su energía en la fabricación de un

pigmento que afecta a la temperatura de su entorno, y sería lógico aventurar que unas margaritas incoloras podrían crecer invadiendo Daisyworld y llevándose los beneficios del «esfuerzo» que habían invertido las otras dos variedades. Para contemplar esta posibilidad, Lovelock construyó un modelo en el que se añadía a la mezcla unas margaritas incoloras, y tanto a las blancas como a las negras se les cargaba un «impuesto por el color» del 1 por ciento por permitirles fabricar el pigmento, reduciendo así su eficiencia en el crecimiento.

Como en el caso anterior, las margaritas negras florecen cuando la producción de calor solar es baja, y las blancas florecen cuando la producción de calor solar es alta, pero ahora las margaritas incoloras tienen su mejor momento cuando la producción de energía solar es justo la correcta para que la temperatura en la «Tierra» esté en su valor óptimo sin ayuda de ninguna clase, es decir, cuando se dan las condiciones en las que el coste de la pigmentación no da a ninguna de las margaritas de color ventaja selectiva alguna. Pero el efecto global sobre la temperatura sigue siendo el mismo; la temperatura de la Tierra del modelo está cerca del valor óptimo, mientras que la del Sol del modelo aumenta drásticamente.

Otra crítica que recibe el modelo, incluso después del ajuste anterior, es que no permite que las margaritas evolucionen. En consecuencia, Lovelock hizo otra variación sobre el mismo tema, comenzando con unas margaritas grises, que no producen efecto alguno sobre la temperatura, pero permitiéndoles que mutaran de manera aleatoria a tonos ligeramente más oscuros o ligeramente más claros en cada generación. La selección natural garantizaba que, una vez más, las margaritas oscuras florecieran cuando la producción de energía solar era baja, que las margaritas claras crecieran cuando la energía solar era alta, y que la temperatura de la superficie terrestre se mantuviera casi igual durante mucho tiempo.

El modelo Daisyworld puede perfeccionarse para tener en cuenta muchos otros factores, incluidas muchas más variedades de margaritas, o para añadir animales que comen margaritas, e incluso depredadores que devoran a los herbívoros. Este planteamiento relaciona el entorno físico del planeta con el tipo de relaciones depredador-presa que describieron hace mucho tiempo Lotka y Volterra (véase el capítulo 4), lo cual es perfectamente lógico, ya que, de hecho, Alfred Lotka fue uno de los primeros científicos (probablemente *el* primero) en constatar que el hecho de

tener en cuenta la cambiante relación entre las especies vivas y su entorno físico haría realmente que entendiéramos con mayor claridad la evolución. En 1925, casi al principio de su clásico libro *Elements of Physical Biology,* escribía lo siguiente:

Es costumbre discutir la «evolución de una especie de organismos». Al hacerlo, veremos muchas razones por las que deberíamos tener presente de manera constante la evolución del sistema como un todo (organismos y entorno). Puede parecer a primera vista que así el problema resultaría más complicado que la consideración de una sola parte del sistema. Pero, a medida que avancemos en el planteamiento, se pondrá de manifiesto que *las leyes físicas que gobiernan la evolución adoptan con toda probabilidad una forma más sencilla cuando se refieren al sistema como un todo, que cuando se aplican a cualquier parte de él.*[100]

Si esto no es una piedra angular de la teoría de Gaia, no sé qué es y data de hace más de tres cuartos de siglo. Después de combinar el modelo de Lotka-Volterra con el de Daisyworld, para crear un escenario en el que las margaritas sirven de alimento a los conejos y éstos a los zorros, Lovelock comprobó la consistencia de su mundo virtual introduciendo en él cuatro catástrofes -unas plagas que barrían cada una de ellas un 30 por ciento de la población de margaritas. A pesar de estar sucediendo todo esto, como muestra la ilustración de la página siguiente, las margaritas seguían arreglándoselas para mantener la temperatura del planeta en la zona de la vida, con tan sólo unas breves interrupciones producidas por las plagas, una forma de equilibrio discontinuo.

Como todos los modelos, el Daisyworld sólo tiene por objeto mostrar lo que es posible, no decir cómo «es en realidad» el mundo. Por supuesto que la temperatura de la Tierra no está controlada por las margaritas, sino por las cambiantes propiedades de absorción de infrarrojos que tiene la atmósfera. Sin embargo, el Daisyworld muestra lo poca que es la retroalimentación que permite (o exige) a la vida llevar a cabo la regulación de la temperatura de un planeta.

Si la Tierra se quedara demasiado fría como para que la vida pudiera prosperar, se formaría más dióxido de carbono en la atmósfera, y el efecto invernadero aumentaría; cuando la vida prospera, el dióxido de carbono es extraído de la

[100] La cursiva es nuestra.

atmósfera y se evita un efecto invernadero galopante.

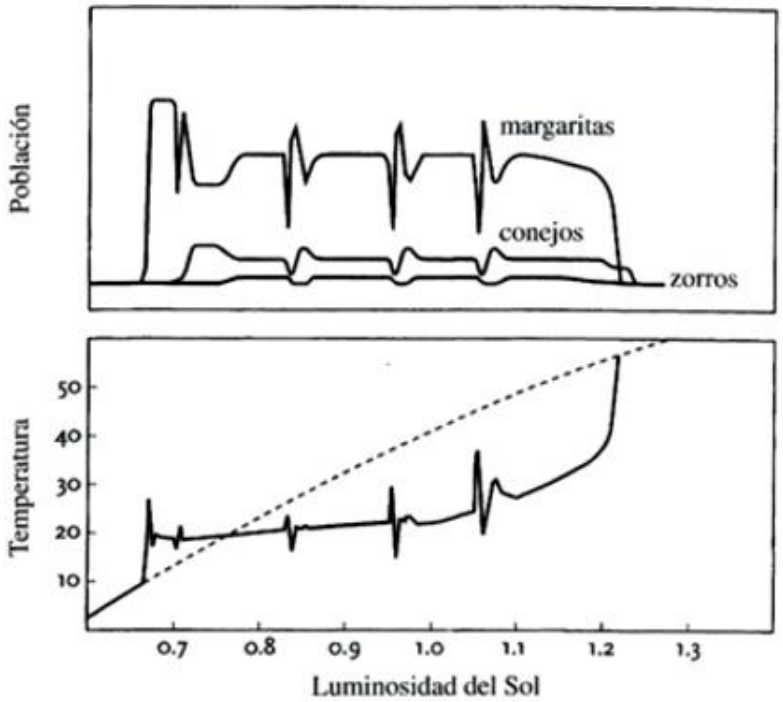

Figura 7.3. Una variación sobre el tema de Daisyworld en la cual los conejos se comen las margaritas, los zorros devoran los conejos, y se producen cuatro plagas que matan cada una de ellas un 30 por ciento de las margaritas. Mientras tanto, el brillo del Sol aumenta; sin embargo, el sistema demuestra ser lo bastante fuerte como para mantener la temperatura de la superficie del planeta dentro de unos límites confortables.

Queda pendiente saber qué sucede cuando todo el dióxido de carbono ha desaparecido y el Sol continúa volviéndose cada vez más caliente, quizá se produzca una catástrofe final para la vida, como el desbocamiento termal que se produce finalmente en los sencillos modelos Daisyworld, o quizá actúe algún otro mecanismo para mantener a raya el calentamiento galopante, posiblemente mediante un incremento de la cobertura de nubes, para rechazar por reflexión una mayor cantidad de la energía solar que llega. Esta posibilidad es la que sugiere nuestro otro ejemplo del modo en que funciona Gaia, un ejemplo extraído del mundo real, que mareó un punto de inflexión cuando Gaia dejó de ser una mera

hipótesis para convertirse en una teoría. Por esta razón, aunque Daisyworld es el logro del que Lovelock pudo estar más orgulloso, de hecho él consideró que su descubrimiento más importante fue el de un mecanismo que vinculaba la cobertura de nubes y la actividad biológica en los océanos.

La diferencia entre una hipótesis y una teoría es que, aunque una hipótesis es una idea sobre el modo en que algo podría funcionar, sin embargo no se ha comprobado mediante experimentos u observaciones. Si una hipótesis formula una predicción (o mejor, una serie de predicciones) sobre cómo serán los resultados de un nuevo experimento, o sobre lo que se descubrirá realizando nuevas observaciones, y si esa predicción queda corroborada por los acontecimientos, entonces la hipótesis se convierte en teoría. La teoría de la gravedad de Newton, por ejemplo, hace predicciones sobre las órbitas de los cometas, y Edmond Halley utilizó esta teoría para predecir el regreso del cometa que ahora lleva su nombre.

Precisamente gracias al éxito de estas predicciones, hablamos ahora de *teoría* de la gravedad de Newton, en vez de llamarla hipótesis. A menudo esta diferencia se difumina; pero lo que sucedió con Gaia durante la segunda mitad de la década de 1980 es un claro ejemplo del modo en que una hipótesis se hace mayor de edad y se convierte en una teoría.

En una fase muy temprana del desarrollo de sus ideas sobre Gaia (de hecho, incluso antes de que William Golding le diera este nombre), a principios de la década de 1970, Lovelock estaba interesado por el modo en que el azufre se transmitió desde los océanos a tierra firme. El azufre es un ingrediente esencial de la vida, pero los continentes lo están perdiendo constantemente en forma de compuestos sulfurados que son arrastrados por el flujo de los ríos. Sin la existencia de algún mecanismo que hiciera que el azufre regresara, la vida en tierra firme pronto tendría dificultades para mantenerse. En aquella época, según los conocimientos convencionales, se producían emisiones de azufre desde los océanos a la atmósfera en forma de ácido sulfhídrico, el nocivo gas que da el olor a «huevos podridos» de las bombas fétidas que suelen lanzar los escolares.

Esto no tiene nada que ver con el olor que habitualmente asociamos con el mar y, en todo caso, Lovelock era muy consciente de que el ácido sulfhídrico se descompone fácilmente mediante reacciones en las que intervenga el oxígeno que

llevan disuelto las aguas marinas. Pero, también sabía que dos décadas antes unos investigadores de la Universidad de Leeds habían descubierto que muchos organismos marinos liberan azufre en forma de un compuesto conocido como dimetil sulfuro; además, como químico de la vieja escuela, también sabía que el dimetil sulfuro, en pequeñas concentraciones dentro de la atmósfera, tiene un olor fresco, como el del «pescado fresco directamente pescado del mar, pero no el olor del pescado fresco de agua dulce».

Con la curiosidad que sentía por averiguar si el dimetil sulfuro (abreviadamente DMS) podía ser realmente el principal transportador del azufre desde los océanos a tierra firme, Lovelock diseñó y construyó un instrumento sensible para medir vestigios de DMS en el aire, y consiguió que este aparato fuera en el barco de investigaciones marinas *Shackelton* durante uno de sus viajes rutinarios de ida y vuelta entre Gran Bretaña y la Antártida. Este viaje, que finalizó en 1972, forma parte ahora de la historia de la ciencia, porque otro de los instrumentos de Lovelock que lite transportado en el mismo viaje reveló por primera vez el modo en que se habían difundido por la atmósfera de nuestro planeta los clorofluorocarbonados, unos gases producidos por las actividades humanas y de los que se sabe actualmente que son responsables de la destrucción del ozono en la estratosfera. Pero ésta es otra historia. Lovelock localizó el DMS, tal como había pronosticado, pero tuvieron que pasar diez años más, hasta principios de la década de 1980, para que posteriores observaciones y mediciones demostraran que, de hecho, hay una cantidad suficiente de DMS que se está produciendo en los océanos y entra en la atmósfera cada año para caer luego en forma de lluvia sobre la tierra y reemplazar las pérdidas del azufre que se ha ido con las aguas fluviales.

Sin embargo, los organismos marinos microscópicos que fabrican DMS no lo hacen porque «quieran» ayudar a las formas de vida que se desarrollan en tierra firme. Al igual que a todos los seres vivos, sólo les interesa maximizar sus propias posibilidades de supervivencia. Los organismos que liberan DMS, las algas marinas, han de luchar en una batalla continua para evitar que el cloruro sódico (sal común) del agua del mar penetre por las membranas de sus células y altere la química de la vida que tiene lugar en su interior. Un modo de mantener la sal fuera es desarrollar una presión adecuada dentro de la célula, utilizando un compuesto no tóxico que no

produzca un efecto adverso en los procesos vitales en los que participan el ADN, el ARN y los aminoácidos. Muchas algas marinas utilizan para ello un compuesto llamado propionato de dimelilsulfonio, que se forma en torno al azufre. Esta sustancia tiene todas las propiedades químicas requeridas y es conveniente para las algas utilizar azufre como ingrediente fundamental porque en el mar hay una gran cantidad de este elemento (gracias a todos los sulfatos que las aguas arrastran desde la tierra firme). Cuando las algas mueren o son ingeridas por un animal, el propionato de dimetilsulfonio se descompone, liberando DMS en la atmósfera. Pero ¿qué tiene que ver todo esto con la autorregulación del clima en la Tierra según la teoría de Gaia?

En 1986, durante una visita a la Universidad de Washington, Seattle, Lovelock se sintió sorprendido al enterarse, por el científico de la atmósfera Robert Charlson, de que nadie sabía cómo se formaban las nubes sobre los océanos. Es bastante fácil provocar la lluvia. Cuando el aire húmedo y caliente se eleva por convección y luego se enfría, a una altura adecuada la humedad se condensa formando grandes gotas de agua, que vuelven a caer directamente hacia abajo. Pero, para fabricar nubes, necesitamos hacer que, a partir de moléculas de agua que se encuentran en la atmósfera, se formen pequeñas gotas que flotan en el aire.

Sólo se puede lograr esto si hay unas «semillas» aún más diminutas (llamadas a veces núcleos de condensación de nubes, o NCN) alrededor de las cuales pueden reunirse las moléculas de agua. Sobre la tierra firme siempre hay gran cantidad de estas semillas, algunas producidas por el polvo que lleva el viento, otras por alguna actividad orgánica, e incluso las hay que proceden de la contaminación atmosférica causada por actividades humanas. Charlson le dijo a Lovelock que ciertas muestras del aire situado sobre el océano Pacífico mostraban gran abundancia de semillas adecuadas, en forma de gotitas de ácido sulfúrico y sulfato de amonio. Pero él y sus colegas no tenía idea alguna sobre la procedencia de estos compuestos, hasta que oyeron a Lovelock describir el proceso de reciclaje del azufre en que participa el DMS. El DMS puede oxidarse en el aire para producir los núcleos de condensación de nubes.

Está clara la importancia de esta contribución a la red global. 1 .a nubes reflejan una cantidad tan grande de la energía solar entrante, que, sin una cobertura de

nubes, la temperatura media en la superficie de la Tierra sería de 35 °C, es decir, 20 °C más alta que en la actualidad. Dado que los océanos cubren alrededor del 70 por ciento de la superficie de nuestro planeta, y que las oscuras aguas oceánicas son especialmente buenas para absorber el calor procedente del Sol, si no hubiera nubes sobre los océanos, la Tierra sería desagradablemente caliente (una *media* de 35 °C para todo el planeta supondría temperaturas mucho más altas en los trópicos).

Lo que se deduce de todo esto es que las formas de vida microscópicas de los océanos desempeñan un papel crucial en el control del clima de la Tierra. En un proceso de retroalimentación natural, si las algas se volvieran más activas, la cobertura de nubes sobre los océanos aumentaría, habría menos luz solar para la fotosíntesis, y la actividad biológica se reduciría; pero, si la actividad biológica se reduce, las algas liberan menos cantidad de DMS, y se forman menos nubes, con lo cual hay más luz solar para la fotosíntesis y la vida prospera.

Es exactamente como el tipo de retroalimentación autoorganizada que vimos funcionar en Daisyworld, y los zarcillos de la red se extienden hasta alcanzar muchos aspectos de la vida en la Tierra.

Podemos ofrecer aquí alguna idea sobre la complejidad de estos vínculos, y el lector puede encontrar muchos más detalles en los diversos libros que Jim Lovelock escribió sobre Gaia. Un aspecto clave es que el océano abierto, lejos de tierra firme, es esencialmente un desierto, escasamente poblado por seres vivos en comparación con las ricas aguas de la plataforma continental.

La razón es que lejos de tierra son escasos los nutrientes de los que puede alimentarse la vida. En cambio, cerca de los continentes siempre hay gran cantidad de materia para nutrir a los seres vivos, porque la transportan los ríos hacia el mar. Actuando como, núcleos de condensación de nubes, las moléculas de DMS producidas por las algas marinas pueden contribuir a que los desiertos oceánicos florezcan (al menos un poco) de dos maneras diferentes. En primer lugar, una mayor cobertura de nubes tiene un electo directo sobre el clima local, ya que hace que los vientos tengan más fuerza. Éstos agitan las capas superficiales del mar, llevando a la parte superior nutrientes que se encuentran en capas más profundas, bajo la zona en que se produce la fotosíntesis. En segundo lugar (y esto es

probablemente lo más importante), las nubes y la lluvia afectan al polvo que los vientos llevan desde los continentes, y que es transportado por las capas altas de la atmósfera hasta los océanos, incluso hasta los más remotos del planeta.

Se ha encontrado polvo del Sahara, por ejemplo, en las Indias Occidentales, y polvo del centro del continente asiático es transportado habitualmente a través del Pacífico hasta Hawai. Este polvo es rico en nutrientes minerales que son esenciales para la vida, pero las partículas de polvo no poseen las propiedades físicas adecuadas para actuar como núcleos de condensación de nubes.

Sin nubes, lo único que hacen es permanecer en las capas altas de la atmósfera, porque no se verían afectadas por el tipo de lluvia sin nubes que produce meramente la evaporación. Sin embargo, en las nubes que se forman como resultado de la presencia de DMS en la atmósfera, las precipitaciones de lluvia limpian el polvo del aire y lo llevan a los océanos, donde pueden utilizarlo las propias algas para fabricar DMS. Por lo tanto, la relación entre los océanos y la tierra firme que resulta de la producción de DMS por las algas marinas funciona como una carretera de doble dirección, con la que se proporciona un beneficio para la vida en ambos lugares.

Gran parte del azufre contenido en el DMS cae con la lluvia sobre la tierra, donde actúa como fertilizante para la vida terrestre; pero parte de este azufre es responsable de las nubes y la lluvia que depositan en los océanos nutrientes procedentes de tierra firme y actúan como fertilizante para la vida oceánica. Pero no hay nada que sugiera que la vida existente en un lugar actúe sacrificándose en beneficio de la vida que pueda existir en otro lugar. Cada una hace lo que es mejor para sí misma. Éste es precisamente el tipo de interacción que, como hemos visto anteriormente, es importante en las redes autoorganizadas.

Pero, incluso con esto, lo único que se explica es una situación ya existente. Lo realmente importante en relación con el funcionamiento del DMS es que también realiza una predicción sobre cómo sería el mundo cuando se produjera un cambio en la situación global, y esta predicción encaja con lo que sabemos sobre cómo era el mundo durante la última serie de glaciaciones.

En los últimos tiempos, durante unos pocos millones de años, las glaciaciones han seguido un ritmo muy regular, de tal modo que glaciaciones de unos 100.000 años

de duración están separadas por intervalos relativamente cálidos, llamados períodos interglaciales, que duran más o menos entre 10.000 y 15:000 años. Vivimos ahora en un período interglacial que comenzó hace unos 10.000 años. La pauta detallada de estos cambios de glaciación a período interglacial se corresponde exactamente con la pauta detalladla del modo en que se producen los cambios en el equilibrio de calor entre las estaciones como resultado del modo en que la Tierra se inclina y oscila cuando recorre su órbita alrededor del Sol.

La cantidad total de calor que la Tierra recibe del Sol cada año es constante, pero, a veces, hay un contraste más fuerte entre las estaciones (veranos calientes e inviernos fríos), y otras veces el contraste es menor (veranos frescos e inviernos suaves). Según la pauta, parece ser que los períodos interglaciales sólo comienzan cuando los veranos son calientes al máximo en el hemisferio norte (donde está actualmente la mayor parte de la tierra firme)[101].

Pero estos cambios no son en sí mismos lo suficientemente grandes como para explicar la transición del período glacial al interglacial y viceversa. Debe haber algún otro proceso desencadenado por los ritmos astronómicos, que actúa luego como una retroalimentación positiva para aumentar la influencia astronómica. La explicación obvia es la cantidad de dióxido de carbono que haya en el aire.

Sabemos que, cuando hay menos dióxido de carbono en el aire, el efecto invernadero se reduce y la Tierra se enfría, esto figura en el centro mismo de la teoría de Gaia. También sabemos que durante la glaciación más reciente, no sólo fue la concentración global del dióxido de carbono del aire inferior a la actual, sino que dicha concentración fluctuaba a lo largo de los milenios exactamente de acuerdo con las variaciones de la temperatura media del globo terráqueo. Lo sabemos porque existen registros exactos, tanto de las variaciones de temperatura, como de las fluctuaciones del dióxido de carbono, obtenidos a partir de largos testigos de hielo que se han extraído perforando las capas glaciales de Groenlandia y de la Antártida.

Hay varios registros de este tipo, pero el ejemplo arquetípico procede de la base de Vostok, en la Antártida rusa, donde los investigadores han obtenido un testigo continuo de hielo que cubre más de 160.000 años, lo cual es suficiente para revelar

[101] Los detalles de la teoría astronómica de las glaciaciones pueden verse en nuestro libro *Ice Age* (Penguin, 2001).

detalles de la totalidad de la última glaciación. Las perforaciones comenzaron en 1980 y dieron como resultado un testigo de 2,2 kilómetros de longitud, que contiene el hielo depositado, año tras año, en capas de nieve que luego se convirtieron en hielo por el peso de la nieve nueva caída sobre ellas. El hielo de la parte inferior del testigo se produjo a partir de nieve caída hace más de 160.000 años y se ha comprimido de tal manera que, en este nivel, una acumulación de nieve caída durante un año se ha reducido a una capa de hielo de un centímetro de espesor.

Las capas sucesivas de hielo que hay en el testigo pueden datarse mediante técnicas geológicas estándar. Además de hielo, el testigo contiene burbujas de aire que quedó atrapado cuando los copos de nieve se comprimieron para formar hielo, y el contenido de estas burbujas de aire puede analizarse para averiguar cómo cambió la composición de la atmósfera durante el último ciclo de glaciaciones y períodos interglaciales. Al mismo tiempo, el agua del hielo puede también analizarse para saber cómo fue cambiando la temperatura durante el mismo intervalo de tiempo. Hay varios modos de hacer esto, pero todos utilizan variaciones de la misma técnica.

Por ejemplo, las moléculas de agua contienen hidrógeno, y este elemento se presenta en dos variedades (dos isótopos), que son la forma común del elemento y el llamado hidrógeno pesado, o deuterio. El hidrógeno pesado es literalmente más pesado que el hidrógeno común, pero las dos formas son químicamente lo mismo. Por lo tanto\ las moléculas de agua se presentan de dos formas, siendo algunas de ellas más pesadas que la variedad común. Las moléculas de agua más pesadas se evaporan menos fácilmente que las del agua común, pero se congelan con mayor facilidad a partir del vapor de agua contenido en el aire; la proporción exacta de cada variedad dentro de la nieve que cae depende de la temperatura media que hubiera en esa parte del mundo en aquella época, por lo que la temperatura que hubo en tiempos remotos puede deducirse de las mediciones de la proporción de agua pesada presente en el hielo desde la parte más profunda del testigo.

Otras técnicas similares que utilizan distintos isótopos del oxígeno, obtenidos, tanto a partir de testigos de hielo, como de conchas de organismos muertos hace mucho tiempo y enterrados en profundos sedimentos oceánicos, dan una imagen más

amplia del modo en que las temperaturas globales fueron cambiando durante el mismo período. Los registros muestran que durante el período más frío, tanto de la última glaciación como de la penúltima, las temperaturas de todo el mundo dieron una media inferior en 9 °C a la actual, mientras que, durante los años más calientes del período interglacial comprendido entre las dos glaciaciones mencionadas, las temperaturas fueron superiores a las actuales en 2 °C. Al final de cada una de las dos glaciaciones últimas, la concentración de> dióxido de carbono en la atmósfera aumentó desde 190 ppm (partes por millón) a 280 ppm, lo que da un incremento del 47 por ciento, con una disminución comparable a principios de la última glaciación. Y, como puede verse en la ilustración, las fluctuaciones del dióxido de carbono y las de la temperatura marchaban al unísono por todo el intervalo cubierto por el testigo de Vostok.

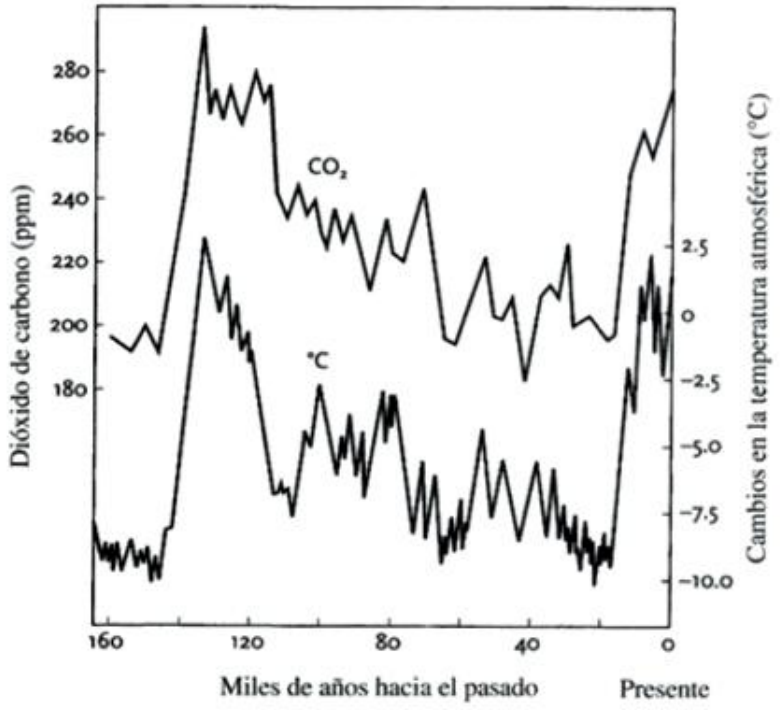

Figura 7.4. Cuando la Tierra entra en un período interglacial o sale de él, las variaciones de temperatura se producen de manera acompasada con los cambios en la concentración del dióxido de carbono. (Datos del testigo de Vostok.)

La pregunta es ¿cuál fue el detonante del cambio en la concentración de dióxido de carbono que tanto contribuyó a que sobreviniera una glaciación?

Una posibilidad obvia es que un aumento de la actividad biológica de los océanos, que cubren dos tercios de la superficie del planeta, extrajera dióxido de carbono de la atmósfera y lo fijara en forma de compuestos carbonados en las conchas de aquellos organismo fotosintetizadores (formas de plancton) que caen al fondo del océano cuando mueren. En la década de 1980 John Martin y Steve Fitzwater, de los Moss Landing Marine Laboratories, en California, siguieron un presentimiento de Martin, quien sospechaba que el hierro podía desempeñar un papel fundamental en la fertilización de los océanos.

Sabían que las aguas del océano Antártico y del océano Pacífico subártico eran ricas en nutrientes vegetales importantes, como los fosfatos y los nitratos. Estos nutrientes son esenciales para el crecimiento de las plantas y, en general, son absorbidos rápidamente por éstas; pero el plancton que vivía en esas zonas ni se expandía, ni los absorbía. Estaba claro que al plancton le faltaba alguna otra cosa, lo cual le impedía crecer y aprovechar aquellos nutrientes. El hierro parecía un buen candidato, porque es un componente crucial de la clorofila, el pigmento verde responsable de la absorción de la luz que se utiliza en la fotosíntesis.

Efectivamente, cuando Martin y Fitzwater tomaron muestras de agua del nordeste del Pacífico y le añadieron hierro (en forma de compuestos ferruginosos disueltos), hubo un rápido crecimiento del plancton, que absorbió el hierro y otros nutrientes que se encontraban en el agua. Los estudios mostraron que lo que impide que aumente la cantidad de plancton en las aguas frías, y ricas en nutrientes, de latitudes altas es precisamente la escasez de hierro. Desde entonces, a estos experimentos a escala de laboratorio les han seguido una serie de ensayos a gran escala en el mar, consistentes en añadir compuestos de hierro a los océanos, produciendo en ocasiones drásticos aumentos de la cantidad de plancton. Incluso se ha llegado a sugerir que esos aumentos de la cantidad de plancton podrían impulsarse a gran escala, para absorber dióxido de carbono del aire y reducir así el impacto del efecto invernadero causado por la actividad humana[102], pero esto no tiene nada que ver con el tema que estamos tratando ahora.

La cuestión fundamental es que el polvo arrastrado por el viento desde los

[102] Aunque esta idea resulta sumamente obvia, yo fui el primero que la hizo pública en un medio impreso, concretamente en Nature, en 1988 (vol. 331, p. 570). Como son tan escasas las ideas brillantes que se me han ocurrido, es pero que se me perdonará la inmodestia de mencionar ésta aquí.

continentes podría proporcionar una fuente vital de hierro para fertilizar los océanos en esas zonas, y que, dado que hay menos precipitaciones de lluvia cuando la Tierra se enfría, la cantidad de ese polvo que transporta hierro tiene probabilidades de aumentar cuando alguna influencia externa (como los ritmos astronómicos) inclina el equilibrio térmico de la Tierra a favor de una glaciación.

La situación es tal que el comienzo de un enfriamiento, desencadenado por influencias astronómicas (o quizá por alguna otra causa), conduce a unas condiciones de mayor sequía en tierra firme a grandes altitudes, porque cuando el planeta se enfría hay menos evaporación en los océanos y, por lo tanto, menos precipitaciones de lluvia. La situación de sequía permite que el polvo que lleva el viento desde los continentes y que transporta compuestos ferruginosos se extienda por los océanos a altas latitudes, donde (gracias al DMS de las algas marinas) cae con la lluvia y fertiliza los mares.

El aumento de actividad biológica resultante absorbe dióxido de carbono del aire, reduciendo el efecto invernadero y favoreciendo un enfriamiento aún mayor, lo cual permite que el viento levante y se lleve más polvo de la tierra, para fertilizar los océanos. Este proceso se detiene cuando todos los fosfatos y nitratos libres se han agotado. Al mismo tiempo, dado que se produce más cantidad de DMS, habrá más nubes sobre los océanos, y éstas reflejarán una mayor cantidad de energía solar, impidiendo su entrada, con lo que se favorecerá el enfriamiento. Con esta descripción es fácil ver cómo puede pasar el planeta de un período interglacial a una plena glaciación como resultado de la acción de un desencadenante externo de muy poca importancia. La salida de una glaciación puede producirse cuando las influencias astronómicas conspiran para ocasionar un calentamiento máximo en latitudes septentrionales altas, fundiendo parte del hielo e inundando la tierra seca situada en torno a los bordes del casquete polar. Como el resultado es que el viento levantará menos polvo, habrá menos hierro para el plancton, y la retroalimentación se invertirá, con un declive de la vida oceánica a medida que los casquetes polares retroceden cada vez más, pero con un florecimiento de la vida en tierra firme a medida que los continentes se vuelven más húmedos.

Esto es sólo una sencilla versión caricaturesca de lo que sucede, pero es suficiente para ilustrar la predicción que realiza el modelo del DMS. Si las glaciaciones están

asociadas de este modo con un aumento de la actividad biológica de los océanos, durante la glaciación debe ser mucho mayor la cantidad de DMS que pasa a la atmósfera. Uno de los productos del DMS que está en el aire es el ácido metanosulfónico, o AMS. Por lo tanto, una de las predicciones del modelo de Gaia desarrollado por Lovelock y sus colegas es que durante una glaciación debería haber más AMS en la atmósfera y, por consiguiente, tendría que caer una cantidad mayor de AMS con la lluvia o la nieve.

Los testigos de hielo muestran que, en efecto, durante la última glaciación, cayó cada año sobre la Antártida, con la nieve, entre dos y cinco veces más AMS que en los tiempos actuales. Los mismos testigos de hielo muestran también que con la nieve caía entonces más polvo que en el presente. Todo encaja. Es una prueba indiscutible de que los componentes biológicos y físicos de nuestro planeta forman parte de una única red que funciona de un modo autoorganizado (Lovelock utiliza la expresión autorregulado) para mantener unas condiciones que son ampliamente adecuadas para la existencia de vida, pero que sufren fluctuaciones a todas las escalas (incluidos los ritmos de alternancia de glaciaciones y períodos interglaciales, así como las extinciones masivas), siendo estas fluctuaciones análogas a las que se producen en el modelo del montón de arena.

En un sentido real, la Tierra es el lugar que alberga una red de vida única, y la existencia de esta red (Gaia) sería visible para cualquier forma de vida inteligente que hubiera en Marte y que fuera capaz de aplicar la prueba de Lovelock y buscar señales de reducción de la entropía.

Ni la NASA, ni nadie, tomó nunca la prueba de Lovelock lo suficientemente en serio como para aplicarla a la búsqueda de vida en el sistema solar; pero sí se está tomando en serio ahora para buscar vida más allá del sistema solar. Esta búsqueda ha cobrado ímpetu porque se piensa que, si hay vida más allá, es muy probable que sea una clase de vida más o menos como la que conocemos, constituida por los mismos bloques de construcción fundamentales (y sencillos) que funcionan en el mismo tipo de redes que vemos aquí en la Tierra.

La historia de la vida en el universo es otro ejemplo de complejidad superficial construida sobre cimientos de una profunda sencillez. Actualmente hay pruebas irrefutables de que el universo tal como lo conocemos surgió a partir de un estado

denso y caliente (el big bang) hace unos 14.000 millones de años[103]. Los bloques de construcción básicos que emergieron del big bang fueron el hidrógeno y el helio, casi exactamente en una proporción 3:1.

Todos los demás elementos químicos (excepto unos leves vestigios de unos pocos elementos muy ligeros, como el litio) han sido fabricados en el interior de las estrellas y dispersados por el espacio cuando éstas se dilataron y expulsaron materiales (en algunos casos explotaron) en las últimas etapas de sus vidas. Una estrella como el Sol genera calor convirtiendo hidrógeno en helio dentro de su núcleo; en otras estrellas los procesos cruciales incluyen fusiones sucesivas de núcleos de helio. Dado que cada núcleo de helio es una unidad que contiene cuatro «nucleones» (dos protones y dos neutrones), y este elemento se denomina abreviadamente helio-4, esto significa que los elementos cuyos núcleos contienen un número de nucleones que es múltiplo de cuatro son relativamente comunes en el universo, excepto el berilio-8, que es inestable. Concretamente, en las primeras etapas de este proceso se produce carbono-12 y oxígeno-16, y resulta que el nitrógeno-14, aunque no contiene un número entero de núcleos de helio-4, se obtiene como subproducto de una serie de interacciones en las que participan núcleos de oxígeno y de carbono que operan en estrellas de masa un poco mayor que la de nuestro Sol.

Como consecuencia, éstos son, con gran diferencia, los elementos más comunes, aparte del hidrógeno y del helio. Dado que este último es un gas inerte que no reacciona químicamente, se deduce que los cuatro elementos reactivos más comunes en el universo son el carbono, el hidrógeno, el oxígeno y el nitrógeno, conocidos en conjunto por el acrónimo CHON[104]. No es casualidad que los cuatro elementos químicos que participan con una aplastante mayoría en la composición de los seres vivos de la Tierra sean el carbono, el hidrógeno, el oxígeno y el nitrógeno. El carbono desempeña el papel clave en el desarrollo de la vida, porque un solo átomo de este elemento es capaz de combinarse químicamente nada menos que

[103] Los detalles sobre estas pruebas irrefutables pueden verse en The Birth of Time de John Gribbin (Phoenix, 1999) [Hay traducción al castellano: El nacimiento del tiempo. Cómo medimos la edad del universo, Ediciones Paidós Ibérica, S. A. Barcelona, 2000, trad. de Carlos Sánchez Rodrigo.]

[104] Aparte de la abrumadora cantidad de hidrógeno que hay en nuestro sistema solar, por cada 100 átomos de oxígeno hay 57 átomos de carbono y 13 átomos de nitrógeno, situándose por detrás el silicio, con una cantidad que es la mitad de la de nitrógeno. Pero todo ello, dejando a un lado el hidrógeno y el helio, constituye en conjunto sólo un 0,9 por ciento de la masa del sistema solar.

con otros cuatro átomos al mismo tiempo (incluidos otros átomos de carbono, que pueden estar unidos a su vez a más átomos de carbono, formando anillos y cadenas), de tal modo que este elemento tiene una química excepcional mente rica. En el ámbito de la ciencia-ficción, en obras relacionadas con el espacio como *Star Trek,* es frecuente que al referirse a nuestro tipo de vida se hable de «formas de vida basadas en el carbono», lo cual implica que podrían existir otras. *Podrían* existir; pero todas las pruebas que aporta la astronomía sugieren que es mucho mayor la probabilidad de que la vida más allá de nuestro universo este basada también en el CHON.

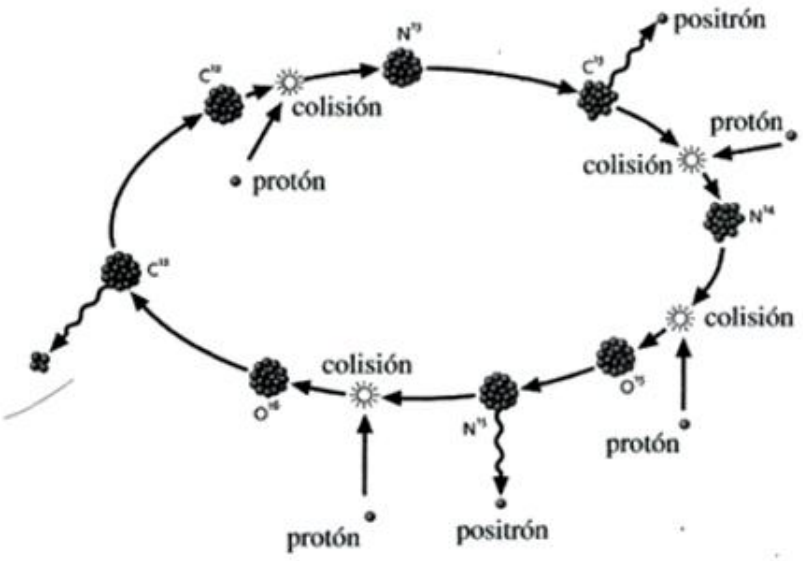

Figura 7.5. Ciclo de interacciones nucleares que se produce en el interior de las estrellas dotadas de una masa ligeramente mayor que la de nuestro Sol. En estas interacciones se combinan el carbono, el nitrógeno y el oxígeno, que, junto con el hidrógeno, son los elementos más importantes para el desarrollo de la vida.

Gran parte de estas pruebas proceden del análisis espectroscópico del material que está presente en nubes de gas y polvo que se encuentran en el espacio, el tipo de nubes a partir de las cuales se forman sistemas planetarios como nuestro sistema solar, y que contienen ellas mismas materiales expulsados de estrellas pertenecientes a generaciones anteriores. En estas nubes hay muchos compuestos construidos en torno a átomos de carbono, y este elemento es tan importante para

la vida que sus compuestos reciben en general el nombre de compuestos «orgánicos». Entre los compuestos detectados en nubes interestelares hay sustancias muy sencillas, como metano y dióxido de carbono, pero también materiales orgánicos mucho más complejos, entre los que cabe citar el formaldehido, el alcohol etílico, e incluso al menos un aminoácido, la glicina.

Esto constituye un descubrimiento muy esclarecedor, porque es probable que todos los materiales que existen en las nubes interestelares hayan estado presentes en la nube a partir de la cual se formó nuestro sistema solar, hace unos cinco mil millones de años. Inspirándose en estas observaciones, dos equipos de científicos llevaron a cabo aquí en la Tierra unos experimentos en los que unas materias primas del tipo de las que se sabe que existen en el espacio interestelar, en forma de mezclas de diversos hielos, se mantuvieron en recipientes sallados en unas condiciones como las que existen en el espacio, expuestas a la radiación ultravioleta a temperaturas inferiores a 15 grados Kelvin (es decir, inferiores a *menos* 258 grados de la escala Celsius). Los resultados se anunciaron en la primavera de 2002. Partiendo de una mezcla de agua, metanol, amoniaco y cianuro de hidrógeno, un equipo descubrió que en los recipientes surgían de manera espontánea tres aminoácidos (glicina, serina y alanina).

En el otro experimento, utilizando una mezcla de ingredientes ligeramente distinta, se producían no menos de dieciséis aminoácidos y otros compuestos orgánicos diversos en unas condiciones que eran las existentes en el espacio interestelar[105].

Para hacemos una idea, las proteínas de todos los seres vivos de la Tierra están compuestas por diversas combinaciones de tan sólo veinte aminoácidos. Todas las evidencias sugieren que este tipo de materia habría caído sobre los jóvenes planetas durante las primeras etapas de formación del sistema planetario, depositada por cometas que habrían sido barridos por la influencia gravitatoria de unos planetas que estaban aumentando de tamaño. Como vimos anteriormente, una sopa de aminoácidos posee la capacidad de organizarse por sí sola, formando una red con todas las propiedades que ha de tener la vida. De esto se deduce que los aminoácidos que estuvieron formándose durante largos períodos de tiempo en las profundidades del espacio (utilizando la energía proporcionada por la luz de las

[105] Véase *Nature* (2002), vol. 416, pp. 401 y 403.

estrellas) serían transportados a la superficie de *cualquier* planeta joven, como la Tierra. Algunos planetas pueden resultar demasiado calientes para que se desarrolle la vida, y otros demasiado fríos. Pero ciertos planetas, como la propia Tierra (y como la papilla de Baby Bear en la serie Goldilock) estarían justo a la temperatura adecuada. Allí, utilizando la expresión de Charles Darwin, en alguna «pequeña charca caliente», tendrían la oportunidad de organizarse en sistemas vivos.

Hasta que supimos que existían otros sistemas planetarios en nuestra galaxia, ni siquiera se podía considerar esta posibilidad como una prueba de que la vida planetaria fuera algo común en la Vía Láctea. Pero ahora se sabe que más de cien estrellas de nuestra zona de la galaxia tienen planetas que describen órbitas alrededor de ellas. Casi todos los planetas descubiertos hasta ahora son gigantes de gas, como Júpiter y Saturno (como era de esperar, los planetas grandes se descubrieron primero, por ser más fáciles de detectar que los planetas pequeños), sin embargo resulta difícil no llegar a la conclusión de que, si existen otros planetas como Júpiter, probablemente existen también otros planetas como la Tierra.

Lee Smolin, de la Universidad de Waterloo, Ontario, ha investigado la relación existente entre, por una parte, las estrellas, que convierten unos elementos más sencillos en algo como el CHON y arrojan estos materiales al espacio, y, por otra parte, las nubes de gas y polvo que hay en éste, que se contraen para formar nuevas estrellas. Nuestro hogar dentro del espacio, la Vía Láctea, es una entre los cientos de miles de millones de «islas» similares dispersas por todo el universo visible, y parece ser una más, con todas las características típicas, de tipo medio en cuanto a tamaño, composición química, etc. La Vía Láctea tiene forma de disco plano, con alrededor de cien mil años luz de diámetro, y está formada por unos pocos cientos de miles de millones de estrellas que describen órbitas en torno al centro del disco. El Sol (que no destaca por nada entre esa multitud de cientos de miles de millones de estrellas) recorre su órbita a una distancia del centro que viene a ser más o menos dos tercios del diámetro.

En el centro de la Vía Láctea las estrellas forman una protuberancia, de tal modo que desde el exterior esta galaxia parecería un enorme huevo frito, en el que la protuberancia es la yema. Sin embargo, el modo en que este disco rota revela que todo el material brillante que compone la parte visible de la Vía Láctea queda sujeto

por el tirón gravitatorio de una cantidad de materia negra que es más o menos diez veces la cantidad de materia de la galaxia y está diseminada en un halo situado alrededor de ella, extendiéndose mucho más allá del borde de este disco de estrellas brillantes.

Descubrir qué es realmente esta materia negra constituye un lema de crucial interés para los astrónomos, pero no tiene importancia para nuestra historia. Muchas galaxias en forma de disco se caracterizan por una especie de serpentinas que se alejan en espiral desde el centro, lo que hace que se les aplique el nombre de galaxias espirales. Es fácil estudiar las pautas que siguen los llamados «brazos espirales», porque las galaxias se encuentran relativamente cerca unas de otras, si comparamos estas distancias con sus tamaños. Andrómeda, la galaxia espiral más cercana comparable con la Vía Láctea, se encuentra con respecto a nosotros a una distancia de un par de millones de años luz; parece una gran distancia, pero la galaxia de Andrómeda es tan grande (un poco mayor que la Vía Láctea) que, incluso a esa distancia, vista desde la Tierra cubre un trozo de cielo del tamaño de la Luna, y puede observarse a simple vista en una noche despejada y sin luz lunar, si nos situamos lejos de las ciudades y de otras fuentes de emisión de luz[106].

Tras haber obtenido un mapa con las posiciones de las estrellas dentro de la Vía Láctea, los astrónomos saben que nuestra galaxia es también una espiral, y, como ya hemos señalado, tienen constancia de que hay una gran cantidad de gas y polvo formando unas nubes situadas entre las estrellas, por lo que, para entender correctamente qué es una galaxia espiral y cómo funciona, es necesario conocer el modo en que se produce el intercambio de la energía y la materia, en un proceso de doble vía entre las estrellas y el medio interestelar que las rodea.

También necesitamos calcular las escalas de tiempo que son importantes para una galaxia. Nuestra Vía Láctea tiene ya una edad de, más o menos, diez mil millones de años, y el Sol tarda unos 250 millones de años en describir una órbita completa alrededor del centro. Los-procesos de intercambio de materia y energía entre nubes interestelares y estrellas pueden parecer lentos según la escala humana del tiempo, pero sin embargo son rápidos según la escala de la propia galaxia.

[106] H ablando en términos aproximados, la distancia hasta la galaxia de Andrómeda es más o menos veinte veces su diámetro. Si la estrella más cercana al Sol estuviera a una distancia de veinte diámetros solares, describiría su órbita alrededor de este astro a una distancia de tan sólo 30 millones de kilómetros, con lo que se movería claramente dentro de la órbita de Mercurio.

Otro aspecto clave es que las estrellas tienen distintos tamaños, o, lo que es más importante, distintas masas. Cuanto más grande es una estrella (cuanta más cantidad de masa tiene), mayor es la cantidad de su reserva de combustible nuclear que ha de consumir (convirtiendo hidrógeno en helio, por ejemplo) con el fin de mantener su forma, resistiendo su propio peso. Esto hace que la estrella sea muy brillante, pero también que esté sometida a una combustión temprana. Aunque el Sol tiene una esperanza de vida de unos diez mil millones de años en esta fase estable, una estrella que tenga el doble de su masa sólo podrá mantenerse durante una cuarta parte de este tiempo, y una estrella que tenga treinta veces la masa del Sol vivirá sólo unos diez millones de años, luciendo un brillo igual al de 30.000 soles, hasta que su combustible nuclear se agote.

Entonces se colapsará hacia adentro, liberando una enorme cantidad de energía gravitatoria, que le dará la vuelta de dentro hacia afuera, con lo que toda la estrella explotará, convirtiéndose en una supernova. Sin embargo, estos sucesos tan drásticos son raros, porque son muy escasas las estrellas que tienen una masa tan enorme; los procesos mediante los cuales se forman las estrellas parecen favorecer la producción de grandes números de estrellas pequeñas, más o menos como el Sol. De hecho, como media sólo hay un par de explosiones de supernovas cada siglo en una galaxia como la Vía Láctea, pero, si esto lo situamos en un contexto adecuado a los procesos galácticos, la suma es de-unas veinte mil explosiones cada millón de años.

Los brazos espirales, que son una característica tan llamativa en galaxias como la nuestra, son visibles porque están bordeados por estrellas calientes de gran masa que relucen con mucho brillo. Esto significa que también son estrellas jóvenes, ya que no hay estrellas viejas que tengan gran cantidad de masa. Dado que una estrella suele tardar algo así como cien millones de años en recorrer una sola órbita por la galaxia, y las estrellas brillantes que bordean los brazos espirales brillan intensamente durante sólo unos pocos millones de años, quizá diez, se deduce que estas estrellas se formaron más o menos donde las vemos ahora. Los brillantes brazos espirales no indican unas zonas donde las estrellas sean especialmente abundantes, sino sólo zonas donde son especialmente brillantes, Smolin dice que están allí «como las luces en un árbol de Navidad».

No hay misterio alguno en cuanto al modo en que se mantiene esa forma en espiral. Se debe exclusivamente a un fenómeno de retroalimentación. Las nubes gigantescas a partir de las cuales se forman las estrellas pueden contener hasta un millón de veces la masa del Sol cuando empiezan a contraerse gravitatoriamente para formar estrellas. Cada nube que se contrae produce, no una sola estrella de gran tamaño, sino todo un conglomerado de estrellas, así como muchas estrellas menores. Cuando las estrellas brillantes emiten su luz, la energía de esta luz estelar (especialmente en la parte ultravioleta del espectro) forma una burbuja dentro de la nube, y tiende a frenar la formación de más estrellas. Sin embargo, una vez que las estrellas de gran masa han recorrido sus ciclos vitales y han explotado, sembrando además el material interestelar con elementos de distintos tipos, la onda expansiva ejerce presión sobre las nubes interestelares cercanas y hace que éstas comiencen a contraerse.

Las ondas procedentes de distintas supernovas, al entrecruzarse unas con otras, actúan mutuamente para barrer el material interestelar y formar nuevas nubes de gas y polvo que se contraen produciendo más estrellas y supernovas, en un ejemplo clásico de interacción que se mantiene por sí sola y en la que intervienen una absorción de energía (procedente de las supernovas) y una retroalimentación. Las simulaciones realizadas mediante ordenador ponen de manifiesto la existencia de una densidad en la nube que es la ideal para que este proceso autosostenido continúe, y que las retroalimentaciones impulsan el proceso de forma natural hacia estas condiciones óptimas. Sida nube es demasiado densa, su parte interna se contraerá gravitatoriamente de manera muy rápida, formando unas pocas estrellas grandes que recorren sus ciclos vitales rápidamente y revientan la nube en pedazos antes de que puedan formarse muchas estrellas.

Esto significa que la generación siguiente de estrellas nace de una nube más delgada, porque ha habido pocas supernovas que barrieran el material formando pedazos densos. Si la nube es tan delgada que su densidad queda por debajo de la densidad óptima, nacerán muchas estrellas, y habrá gran cantidad de explosiones de supernovas, lo cual producirá gran número de ondas de choque que barrerán el material interestelar, acumulándolo en nubes más densas. Por ambas partes, las retroalimentaciones operan para mantener un equilibrio aproximadamente

constante entre la densidad de las nubes y el número de supernovas (y estrellas del tipo del Sol) que se producen en cada generación. La propia pauta espiral resulta del hecho de que la galaxia realiza un movimiento de rotación y está sometida al tirón gravitatorio de la materia oscura, de manera análoga a como se produce una pauta espiral al verter la nata en el café por el hecho de que la nata se mueve en una rotación y está sometida a la sujeción que ejerce el café negro. De hecho, la pauta espiral se mueve alrededor de la galaxia a una velocidad de unos 30 kilómetros por segundo, mientras que las estrellas y las nubes de gas y polvo que constituyen el núcleo del disco de la galaxia se mueven a velocidades de unos 250 kilómetros por segundo, adelantando a los brazos espirales y recibiendo presiones al pasar a través de ellos. Más o menos cada cien millones de años, *todo* lo que hay en el disco de la galaxia se ve sometido a estas presiones dos veces, una por cada lado de la galaxia, mientras órbita alrededor del centro[107].

Al margen de toda esta actividad, en el resto de cualquier galaxia del tipo de la Vía Láctea, existen estrellas más tranquilas y de vida más larga, como el Sol, y, además, en el disco de la galaxia hay material en forma de gas y polvo en una cantidad que es al menos un 15 por ciento de todo el material que se encuentra en forma de estrellas. Este material interestelar se presenta en diversas variedades. Existen nubes de gas y polvo fríos, que son ricas en interesantes moléculas y se llaman nubes moleculares gigantes; a partir de estas nubes se forman nuevas estrellas (y planetas). Hay nubes de lo que consideraríamos gas «normal», formadas por átomos y moléculas de sustancias tales como el hidrógeno, y quizá tan calientes como la habitación en la que está usted ahora.

Además, hay regiones que se han calentado hasta temperaturas extremas mediante la energía procedente de explosiones estelares, de tal modo que los electrones han sido arrancados de sus átomos para formar un plasma cargado de electricidad. También existe una amplia variedad de densidades dentro del medio interestelar. En la modalidad más ligera, la materia que está entre las estrellas es tan escasa que sólo hay un átomo por cada mil centímetros cúbicos de espacio; en la modalidad más densa, las nubes que están a punto de producir nuevas estrellas y nuevos planetas contienen un millón de átomos por centímetro cúbico.

[107] Al menos, en una galaxia espiral sencilla que conste de dos brazos. Algunas galaxias poseen estructuras más complicadas, pero la explicación de estas sutilezas queda fuera del alcance de este libro.

Sin embargo, esto es algo muy diluido si se compara con el aire que respiramos, donde cada centímetro cúbico contiene más de diez trillones de moléculas, pero incluso una diferencia de mil millones de veces en la densidad sigue siendo un contraste espectacular. La cuestión que Smolin y unos pocos investigadores más destacaron a finales de la década de 1990 es que en todos estos aspectos, composición, temperatura y densidad, el medio interestelar dista mucho de ser uniforme. Por decirlo de una forma más categórica, no está en equilibrio, y parece que lo que lo mantiene lejos del equilibrio son unos procesos asociados con la generación de las pautas espirales.

Esto significa que la Vía Láctea (como otras galaxias espirales) es una-zona de reducción de la entropía. Es un sistema autoorganizador al que mantienen lejos del equilibrio, por una parte, un flujo de energía que atraviesa el sistema y, por otra, como ya hemos visto, la retroalimentación. En este sentido, nuestra galaxia supera el test de Lovelock para la vida, y además Smolin ha argumentado que las galaxias deben ser consideradas como sistemas vivos. Un planteamiento más prudente señala que el test de Lovelock constituye lo que se llama una condición «necesaria, pero no suficiente» para la existencia de vida. Si un sistema se encuentra en equilibrio termodinámico, si no supera el test de Lovelock—, podemos tener la seguridad de que está muerto. Si está vivo, debe producir una reducción de la entropía y superar el test de Lovelock. Pero un sistema podría producir entropía negativa sin estar vivo, como en el caso de contracción por efecto de la gravedad que hemos comentado anteriormente.

Desde este punto de vista, no hay una frontera claramente definida entre los objetos vivos y los de materia inerte. Un montón de arena sobre el que se deja caer granos desde arriba se mantiene en un estado crítico, alimentándose de energía externa, pero, por supuesto, no está vivo; un ser humano está ciertamente vivo; todavía se discute si se puede considerar que Gaia es un solo sistema vivo; en cuanto a las galaxias, este tipo de investigación no ha hecho más que comenzar, y prefiero no apostar sobre cuál será el consenso al que se llegará.

El mero hecho de que la frontera entre la vida y la ausencia de vida sea difusa, y que el lugar en que hay que trazar la línea sea un tema de discusión, es, sin embargo, un descubrimiento importante. Contribuye a dejar claro que en relación

con la vida no hay nada insólito en el contexto del modo en que funciona el universo. Como ya hemos visto, es natural que los sistemas simples se organicen en redes al borde del caos y, una vez que lo hacen, es natural que la vida surja allí donde hay «una pequeña charca caliente» que sea adecuada para ello.

Esto es parle de un proceso más o menos continuo, sin que haya un salto repentino en el que comience la vida. Desde este punto de vista, lo más importante que la ciencia podría lograr sería el descubrimiento de, al menos, otro planeta en el que haya surgido la vida. Gracias a la teoría de Lovelock sobre la naturaleza de la vida, estamos a punto de poder conseguirlo, y es posible que dentro de veinte o treinta años se lancen al espacio unos telescopios capaces de detectar otros sistemas como Gaia. .

Hay dos etapas en el descubrimiento de estas otras Gaias. En primer lugar, hemos de detectar planetas del tamaño de la Tierra que describan órbitas alrededor de otras estrellas; luego, tenemos que analizar las atmósferas de esos planetas para buscar pruebas de que los procesos de reducción de la entropía están en marcha. I .os primeros planetas «extrasolares» (o planetas que se encuentran «fuera del sistema solar») se detectaron utilizando técnicas Doppler, que ponían de manifiesto unos cambios pequeñísimos en el movimiento de las estrellas alrededor de las cuales orbitaban dichos planetas. Este efecto, que lleva el nombre del físico del siglo XIX Christian Doppler, modifica la posición de las líneas en el espectro de luz de un objeto, desplazándolas en una cantidad que depende de lo rápido que el objeto se mueva con respecto al observador. Para hacernos una idea de lo que es este tipo de observaciones, pensemos que el tirón gravitatorio que Júpiter ejerce sobre el Sol produce en éste un cambio de velocidad de unos 12,5 metros por segundo, y lo desplaza (con respecto al centro de masa del sistema solar) a una distancia de 800.000 kilómetros, más de la mitad del diámetro de este astro, cuando el Sol y Júpiter orbitan en torno a sus recíprocos centros de masa.

La velocidad de este movimiento es comparable a la de un corredor olímpico de los cien metros lisos y, para un observador situado fuera del sistema solar, esto, por el efecto Doppler, produce un pequeñísimo desplazamiento de vaivén en la posición exacta de las líneas del espectro de la luz emitida por el Sol. Se trata del tipo de desplazamiento que se ha detectado en la luz a partir de los datos de algunas

estrellas de nuestro entorno, y demuestra que en torno a ellas orbitan cuerpos celestes similares a Júpiter. Para ilustrar esto con una comparación, diremos que la Tierra induce en el Sol, mientras órbita alrededor de él, un cambio de velocidad de tan sólo 1 metro por segundo (la velocidad de un agradable paseo), y desplaza al Sol únicamente 450 kilómetros, con respecto al centro de masa del sistema solar. No disponemos aún de la tecnología necesaria para medir un efecto tan pequeño a distancias tales como las de nuestras estrellas, razón por la cual no se ha detectado mediante este método ningún planeta similar a la Tierra.

Hay otras técnicas que podrían servir para identificar planetas pequeños, y que a veces se mencionan en los informativos. Por ejemplo, si resulta que el planeta pasa directamente por delante de su estrella (una ocultación o un tránsito), se produce un empalidecimiento regular de la luz procedente de dicha estrella. Según las estadísticas, dado que las órbitas de los planetas extrasolares podrían estar inclinadas en cualquier dirección con respecto a nuestra posición, sólo el 1 por ciento de estos planetas estará en órbitas tales que podríamos ver ocultaciones y, en cualquier caso, cada tránsito dura sólo unas pocas horas (una vez al año para un planeta que tenga una órbita como la de la Tierra; una vez cada once años para uno cuya órbita sea como la de Júpiter). Sin embargo, hay proyectos para lanzar al espacio en los próximos años (quizá ya en-2005) unos satélites que controlen cada uno de ellos un gran número de estrellas con el fin de buscar esas ocultaciones^ Si se estudian 100.000 estrellas, y 1.000 de ellas muestran tránsitos, la estadística resultante implicaría que prácticamente toda estrella similar al Sol está acompañada por planetas. Sin embargo, aunque todas las búsquedas de este tipo son de un valor inestimable, la técnica Doppler es la que se puede aplicar de manera más general a la búsqueda de planetas similares a la Tierra. De cualquier manera, independientemente de los planetas de este tipo que se descubran, la siguiente fase en la búsqueda de otros sistemas como Gaia será igual.

La mejor perspectiva inmediata de hallar un gran número de planetas como la Tierra (al menos, de sus mismas dimensiones) es la que nos ofrece el satélite de la NASA llamado SIM (Space Interferometry Mission). Este satélite podría ser lanzado durante los próximos dos años (hacia 2005), y utilizará una técnica conocida como interferometría, que consiste en combinar los datos obtenidos por varios telescopios

pequeños para imitar la capacidad de observación de un solo telescopio (mucho mayor. Si todo va bien, el SIM podrá medir la posición de las estrellas con tanta exactitud que detectará las oscilaciones ocasionadas por cualquiera de los planetas similares a la Tierra que describen órbitas alrededor de cualquiera de las 200 estrellas más cercanas al Sol, así como por cualquiera de los planetas similares a Júpiter hasta una distancia del Sol que podría llegar hasta los 3.000 años luz.

Hacia el final de la primera década del siglo XXI (de nuevo, si todo va bien), la Agencia Espacial Europea lanzará un satélite cuyo nombre, algo equívoco, será GAIA y que tendrá como misión principal, no precisamente buscar otras Gaias, sino trazar un mapa con las posiciones de los mil millones de objetos celestes más brillantes. Dado que GAIA tendrá que observar tantas estrellas, no mirará cada una muchas veces ni durante mucho tiempo, por lo que no podrá detectar las oscilaciones ocasionadas por planetas similares a la Tierra; pero sí que podrá detectar planetas del tamaño de Júpiter que recorren órbitas cuyos períodos duran algo más de un par de años.

Si estos planetas son tan abundantes como sugieren las primeras indicaciones obtenidas desde telescopios terrestres, dentro de diez años deberíamos tener identificados decenas de miles de sistemas planetarios extrasolares en las zonas de la Vía Láctea próximas a nosotros. Sin embargo, seguiría tratándose de observaciones indirectas y, para captar los espectros de algunos de esos planetas, se necesita dar un salto más en cuanto a tecnología.

Tanto la NASA como la Agencia Espacial Europea están trabajando en una nueva generación de proyectos que podrían usar esa tecnología para realizar las observaciones hacia el año 2030. Sin embargo, a causa del elevado coste de cualquier misión de este tipo, parece probable que en alguna etapa todos los esfuerzos se fusionen en un proyecto verdaderamente global. Sería una colaboración entre todos los expertos de renombre que hay en la Tierra para buscar la prueba de que no están (y no estamos) solos en el universo, Gaia en su conjunto buscando otras Gaias. El proyecto de la Agencia Espacial Europea se conoce como proyecto Darwin, pero también se denomina, de una manera más prosaica, Interferómetro Espacial de Infrarrojos (IRSI = Infrared Space Interferometer); el proyecto equivalente de la NASA es el Terrestrial Planet Finder (TPF). No obstante,

ambos proyectos funcionarán según los mismos principios.

Sin embargo, por sorprendente que pueda parecer, especialmente después de ver las imágenes de la Tierra tomadas desde el espacio, en las cuales ésta aparece como una brillante bola azul y blanca sobre un fondo oscuro, la luz visible no ofrece las mejores perspectivas para detectar directamente otros planetas similares a la Tierra. Esto es así por dos razones.

En primer lugar, la luz visible que se recibe desde un planeta como la Tierra es en esencia el reflejo de la luz procedente de su estrella progenitora, por lo que no sólo es relativamente débil, sino que resulta muy difícil de captar a distancias astronómicas sobre un fondo iluminado por el resplandor de dicha estrella.

En segundo lugar, los planetas del tipo de la Tierra alcanzan en realidad su brillo máximo en la parte de rayos infrarrojos del espectro electromagnético, por el modo en que la energía absorbida procedente del Sol vuelve a irradiarse en la zona de infrarrojos de dicho espectro, con longitudes de onda más largas que las de la luz visible. En una longitud de onda de unas pocas mieras, la Tierra es el planeta más brillante del sistema solar y destacaría como un objeto impactante si se utiliza cualquier telescopio de infrarrojos suficientemente sensible situado en nuestra proximidad estelar.

El problema es que, dado que la radiación de infrarrojos es absorbida por los propios gases de la atmósfera terrestre, como el dióxido de carbono y el vapor de agua, que son lo que nos interesa descubrir, el telescopio que se utilice para buscar otros planetas como la Tierra tendrá que ser colocado en las profundidades del espacio, lejos de cualquier fuente potencial de contaminación. También tendrá que ser muy sensible, lo cual significa muy grande. Por esto es por lo que estamos hablando de un proyecto internacional muy caro que tardará décadas en realizarse.

Este plan precisa de un telescopio basado en la técnica del interferómetro y que sea mucho más potente que el SIM. Como todos los interferómetros, éste necesitará que las señales procedentes de los diferentes telescopios se sumen con una gran exactitud. Cuando esta técnica se desarrolló inicialmente en tierra, utilizando radiotelescopios, las señales se reunieron mediante cables; ahora, lo habitual es usar rayos láser en este tipo de tareas, y esta tecnología, independientemente del nombre que reciba en cada caso, es esencial para el telescopio con el que se va a

realizar la búsqueda de planetas.

Cada una de las dos posibilidades que actualmente permanecen guardadas en el cajón está pensada para un grupo de seis satélites (seis telescopios) que volarán en formación, separados por una distancia de al menos 100 metros, en los vértices de un hexágono. Las posiciones relativas de estos vehículos espaciales habrán de medirse con una precisión de al menos un milímetro (una vez más, utilizando rayos láser), mientras que la información que proporcionen los telescopios se combinará para dar una sola señal en un satélite central, y se enviará también mediante rayos láser a la Tierra, tal vez desde una distancia de 600 millones de kilómetros.

Este instrumento será capaz de detectar las emisiones de infrarrojos procedentes de planetas similares a la Tierra (o, más bien, de planetas similares a la Tierra, a Venus o a Marte que describan órbitas comparables a las de estos planetas) que orbiten alrededor de unos pocos cientos de estrellas situadas a una distancia de unos cincuenta años luz. El paso siguiente es captar los espectros de esos planetas, concretamente, el paso que dio Jim Lovelock fue el mejor modo de buscar vida en Marte.

Todo esto lleva su tiempo y, cuantos más planetas examinemos en cada etapa de la búsqueda, menos tiempo pasaremos contemplando cada planeta. Así, en el caso de la misión planeada para seis años, los dos primeros se dedicarán a la observación de tantas estrellas como sea posible, con el fin de encontrar cualquier signo de emisión de rayos infrarrojos desde los planetas. Durante los dos años siguientes, se hará un estudio más detallado de los ochenta mejores candidatos, dedicando unas 200 horas de observación a cada objetivo, para buscar las características espectrales más destacadas en la zona de infrarrojos del espectro asociada con el dióxido de carbono, y las características, ligeramente menos destacadas, relacionadas con el vapor de agua.

La sola presencia de estos gases no es un signo de que exista vida, pero sí de la existencia de planetas que serían del tipo de la Tierra en el sentido de que tendrían una atmósfera como Venus y Marte, mientras que, en particular, la presencia de agua indicaría la probabilidad de que existiera un lugar adecuado para la vida. En la última etapa de la misión se seleccionarían entre los ochenta candidatos sólo veinte objetivos, con el fin de dedicar al estudio de cada uno de ellos unas 800 horas, el

tiempo necesario para obtener vestigios de la característica potencial más interesante de esta zona del espectro, las características que produce la presencia de oxígeno en el espectro de infrarrojos. No se trataría del tipo de moléculas de oxígeno diatómicas ordinarias (O_2) que respiramos en el aire, porque el oxígeno diatómico no se emite ni se absorbe en esa zona del espectro.

Sin embargo, cualquier planeta del tipo terrestre que, como la propia Tierra, tenga una atmósfera rica en oxígeno, tendrá asimismo una capa de ozono, producida por la acción de la luz procedente de su estrella progenitora sobre el oxígeno diatómico para fabricar oxígeno triatómico (O_3). Además el ozono produce una marcada huella en el espectro de infrarrojos, justo entre las características asociadas con el dióxido de carbono y el vapor de agua. La sola presencia de la huella característica producida por el ozono (o las huellas características que indiquen la presencia de otros compuestos activos, como el metano) sería suficiente para decirnos que la atmósfera de un planeta qué gire en órbita alrededor de una estrella a una distancia de decenas de años luz no se encuentra en equilibrio termodinámico, y que los procesos de reducción de la entropía, en otras palabras, la vida, están funcionando en la superficie de ese planeta. Y todo esto sin recurrir a sondas espaciales, y mucho menos a una tripulación humana, incluso fuera del sistema solar.

Este es el ejemplo más impactante que conocemos de la profunda sencillez sobre la que se asienta el universo. Los objetos más complejos del universo conocido son los seres vivos, como, por ejemplo, nosotros mismos. Estos sistemas complejos están hechos de las materias primas más comunes que existen en galaxias como la Vía Láctea.

En forma de aminoácidos, estas materias primas se ensamblan de manera natural, dando lugar a sistemas autoorganizadores, donde unas causas subyacentes muy sencillas pueden producir complejidad en la superficie, como en el caso del leopardo y sus manchas. Finalmente, con el fin de detectar la presencia tic esta complejidad máxima de unos sistemas universales no necesitamos ninguna prueba sofisticada para distinguir la materia viva de la materia inerte, sino únicamente las técnicas más sencillas (aunque asistidas por tecnologías altamente avanzadas) para identificar la presencia de uno de los compuestos más simples del universo: el oxígeno.

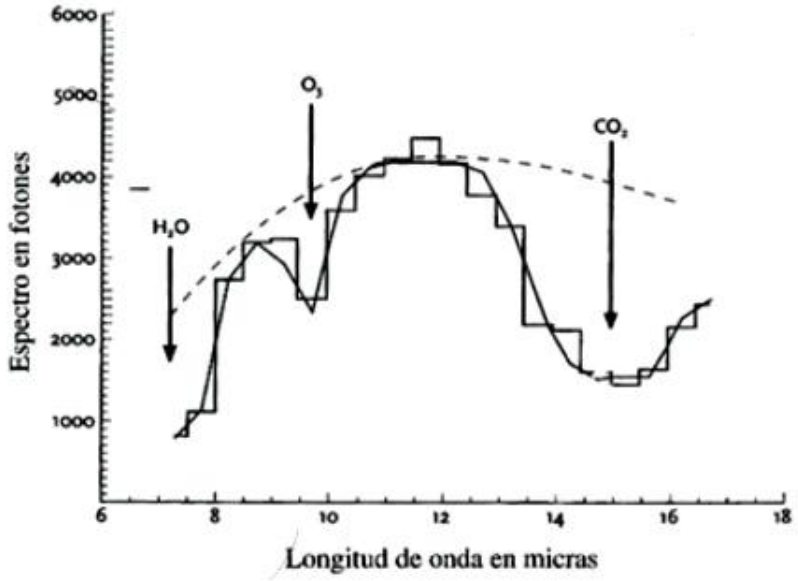

Figura 7.6. Simulación de lo que sería el espectro de la Tierra en el telescopio de Darwin a una distancia de treinta años luz. Si instrumentos como el de Darwin obtuvieran espectros como éste para planetas que describen órbitas alrededor de otras estrellas, se demostraría que en dichos planetas tienen lugar procesos de reducción de la entropía, y se tendría un indicio de que es probable la presencia de vida más allá del sistema solar.

El caos y la complejidad se combinan para hacer del universo un lugar muy ordenado que es justo el entorno adecuado para formas vivas como nosotros mismos. Como dijo Stuart Kauffman, «en el universo estamos en nuestra propia casa». Sin embargo, no es que el universo se haya diseñado así para beneficiarnos a nosotros. Por el contrario, lo que sucede es que estamos hechos a imagen y semejanza del universo.

Lecturas complementarias

En esta lista aparecen varios de mis propios libros, no porque sean necesariamente los mejores sobre cada uno de los temas tratados aquí, sino porque se refieren de forma complementaria a algunas de las cuestiones que he examinado en esta obra.

- Alexrod, Robert, The Evolution of Cooperation, Basic Books, Nueva York, 1984.
- Bak, Per, How Nature Works, Oxford University Press, Oxford, 1997.
- Ball, Philip, The Self-Made Tapestry, Oxford University Press, Oxford, 1999.
- Boulter, Michael, Extinction, Fourth Estate, Londres, 2002.
- Buchanan, Mark, Ubiquity, Weidenfeld & Nicolson, Londres, 2000.
- Carroll, Lewis, Though the Looking-Glass and What Atice Found There, Macmillan, Londres, 1871 (hay trad. cast., A través del espejo y lo que Alicia encontró al otro lado, Alianza, Madrid, 1997).
- Darwin, Charles, On the Origin of Species by Means of Natural Selection, Penguin, Londres, 1968; reimpresión de la primera edición de 1859 (hay trad. cast., El origen de las especies, Planeta De Agostini, Barcelona, 1996).
- Davies, Paul, The Fifth Miracle: The Search for the Origin and Meaning of Life, Alien Lane/Penguin Press, Londres, 1998 (hay trad. cast. El quinto milagro: la búsqueda del origen y significado de la vida. Crítica, Barcelona, 2000).
- Drake, Slillman, Galilea at Work, Dover, Nueva York, 1978.
- Ekeland, Ivar. The Broken Dice, And Other Mathematical Tales of Chance, University of Chicago Press, Illinois, 1993.
- Fauvel, J. K. R. Flood, M. Shortland y R. Wilson, eds., Let Newton be! Oxford University Press, Oxford, 1988.
- Feder, J. Fractals, Plenum, Nueva York, 1988.
- Field, Michael y Martin Golubitsky, Symmetry in Chaos, Oxford University Press, Oxford, 1992.
- Fisher, Ronald Aylmer, The Genetical Theory of Natural Selection, Clarendon Press, Oxford, 1929.
- Gamow, George, My World Line, Viking, Nueva York, 1970.
- Glass, León y Michael Mackey, From Clocks to Chaos: The Rhythms of life,

Princeton University Press, Princeton, 1988.

- Gleick, James, Chaos: Making a New Science, Heinemann, Londres, 1988 (hay trad. cast., Caos: la creación de una ciencia, Seix Barral, Barcelona, 1988).
- Gribbin, John, In Search of the Big Bang, Penguin, Londres, 1998 (hay trad. cast., En busca del big bang, Pirámide, Madrid, 1988).
- Almost Everyone's Guide to Science, Weidenfeld & Nicolson, Londres, 1998 (hay trad. cast., Introducción a la ciencia: una guía para todos, Crítica, Barcelona, 2000); Stardust, Alien Lane, Penguin Press, Londres, 2000; , Science: A History, Alien Lane, Penguin Press, Londres, 2002 (hay trad. cast., Historia de la ciencia, 1543-2001, Crítica, Barcelona, 2003).
- Gribbin, John y Mary Gribbin, Fire on Earth, Simon & Schuster, Londres, 1996; Chaos and Uncertainly, Hodder Children's Books, Londres, 1999.
- Hall, Nina, ed., The New Scientist Guide to Chaos, Penguin, Londres, 1992.
- Hao, Bai Lin, ed., Chaos, World Scientific, Singapur, 1984.
- Hodges, Andrew, Alan Turing, Bumett Books, Londres, 1983.
- Holland, John, Hidden Order: How Adaptation Builds Complexity, Addison-Wesley, Reading, Mass., 1995 (hay trad. cast., El orden oculto: de cómo la adaptación crea la complejidad, F. C. E., México, 20()4); Emergence, Oxford University Press, Oxford, 1998.
- Kauffman, Stuart, The Origins of Order, Oxford University Press, Nueva York, 1993; At Home in the Universe, Oxford University Press, Nueva York, 1995.
- Laplace, Pierre, A Philosophical Essay on Probabilities, Dover, Nueva York, 1951. Traducción del original Essai philosophique sur les probabilités, publicado inicialmente en inglés en 1819 (hay trad. cast., Ensayo filosófico sobre las probabilidades, Alianza, Madrid, 1985).
- Lewin, Roger, Complexity, Macmillan, Nueva York, 1993 (hay trad. cast., Complejidad, Tusquets, Barcelona, 2002).
- Lorenz, Edward, The Essence of Chaos, University of Washington Press, 1993 (hay trad. cast., La esencia del caos, Debate, Madrid, 1995).
- Lotka, Alfred, Elements of Mathematical Biology, Dover, Nueva York, 1956; publicado originalmente en 1925 como Elements of Physical Biology.

- Lovelock, James, The Ages of Gaia, Oxford University Press, Oxford, 1988 (hay trad. cast., Las edades de Gaia: una biografía de nuestro planeta vivo, Tusquets, Barcelona, 1995); Homage to Gaia, Oxford University Press, Oxford, 2000.
- Mandelbrot, Benoit, The Fractal Geometry of Nature, W. H. Freeman, Nueva York, 1983; la edición original es de 1977 (hay trad. cast., La geometría fractal de la naturaleza, Tusquets, Barcelona, 1997).
- Margulis, Lynn, The Symbiotic Planet: A New Look of Evolution, Weidenfeld & Nicolson, Londres, 1998 (hay trad. cast., Planeta simbiótico: un nuevo punto de vista sobre la evolución, Debate, Madrid, 2002).
- Monod, Jacques, Chance and Necessity, Knopf, Nueva York, 1971 (hay trad. cast., El azar y la necesidad: ensayo sobre la filosofía natural de la biología moderna, Tusquets, Barcelona, 20006).
- Murray, James, Mathematical Biology, Springer, Nueva York, 1993; la segunda edición corregida del primer volumen fue publicada inicialmente en 1989.
- Pagels, Heinz, The Dreams of Reason, Simón & Schuster, Nueva York, 1988 (hay trad. cast., Los sueños de la razón, Gedisa, Barcelona, 1991).
- Peitgen, Heinz-Otto y Peter Richter, The Beauty of Fractals, Springer, Nueva York, 1986.
- Peitgen, Heinz-Otto, Hartmtrt Jürgens y Dietmar Saupe, Chaos and fractals, Springer, Nueva York, 1992.
- Pelerson, Ivars, Newton's Clock, W. H. Freeman, Nueva York, 1993 (hay trad. cast.. El reloj de Newton: caos en el Sistema Solar, Alianza, Madrid, 1995).
- Poincaré, Henri, Science et Méthode, Flammarion, París, 1908 (hay trad. cast., Ciencia y método, Espasa-Calpe, Madrid, 19633).
- Prigogine, Ilya, The End of Certainty, The Free Press, Nueva York, 1997 (hay trad. cast., El fin de las certidumbres, Taurus, Madrid, 1997).
- Prigogine, Ilya e Isabelle Stengers, Order out of Chaos, Heinemann, Londres, 1984.
- Paup, David, Extinction: Bad Genes or Bad Luck?, Norton, Nueva York, 1991.
- Richardson, Lewis Fry, Weather Prediction by Numerical Process, Dover,

Nueva York, 1965; reimpresión de la edición original de 1922.

- Ruelle, David, Chaotic Evolution and Strange Attractors, Cambridge University Press, Cambridge, 1989; Chance and Chaos, Princeton University Press, Princeton, 1991 (hay trad. cast., Azar y caos, Alianza, Madrid, 1995).
- Schroeder, M., Fractals, Chaos, Power Laws, W. H. Freeman, Nueva York, 1991.
- Scoll, Stephen K., Chemical Chaos, Oxford University Press, Oxford, 1991.
- Smith, John Maynard, Evolution and the Theory of Games, Cambridge University Press, Cambridge, 1982.
- Smolin, Lee, The Life of the Cosmos, Oxford University Press, Nueva York, 1997.
- Stewart, lan, Does God Play Dice?, Penguin, Londres, 1990 (hay trad. cast., ¿Juega Dios a los dados?, Crítica, Barcelona, 2001).
- Thompson, D'Arcy, On Growth and Form, Cambridge University Press, Cambridge, 1917 (hay trad. cast., Sobre el crecimiento y la forma, Cambridge University Press, Madrid, 2003).
- Waldrop, Mitchell, Complexity, Simón & Schuster, Nueva York, 1992.
- Weiner, Jonalhan, The Beak of the Finch, Jonathan Cape, Londres, 1994 (hay trad. cast., El pico del pinzón: una historia de la evolución en nuestros días, Círculo de Lectores, Barcelona, 2002).
- White, Michael y John Gribbin, Darwin: A Life in Science, Simón & Schuster, Londres, 1995.
- Zhang, Shu-yu, Bibliography on Chaos, World Scientific, Singapur, 1991.
- Zimmer, Carl, Evolution, Heinemann, Londres, 2002.
- Zipf, George, Human behavior and the Principie of Least Effort, Addison-Wesley, Cambridge, Mass., 1949.